A simulated weather-driven bio-economic optimization model for agricultural planning.

A Thesis Submitted to the Committee on Graduate Studies in Partial Fulfillment of the Requirements for the Degree of Masters of Science in the Faculty of Arts and Science

TRENT UNIVERSITY

Peterborough, Ontario, Canada

Applied Modelling and Quantitative Methods

M.Sc. Graduate Program

September 2024

ABSTRACT

A weather-drive bio-economic optimization model for agricultural planning

Bunnel Bernard

This thesis introduces a weather-driven bio-economic optimization model for agricultural planning and decision-making. The model integrates weather simulations—including precipitation, temperature, relative humidity, and reference evapotranspiration (ETo)—to estimate crop yields using the AquaCrop simulator. These yield estimates are then incorporated into a multi-objective optimization (MOO) model that aims to maximize gross profit and economic water productivity (ET), while minimizing land use. The MOO model's results provide insights into key agricultural planning questions, such as what, where, when, and how much to plant. The findings demonstrate the model's potential to enhance agricultural decision-making by offering optimized crop combinations that improve both economic returns and land use efficiency. This research contributes to the development of a dynamic agricultural planning model by integrating weather forecasting, crop simulation, and multi-objective optimization.

Key Words: Markov chains, Stochastic differential equation, Reference evapotranspiration, Artificial neural network, AquaCrop, Multi-objective optimization

Acknowledgements

Firstly, I would like to start by extending praise to the most high Elohim, Yahuah, whose grace and blessing made this possible. I wish to express my sincere gratitude to my supervisors, Prof. Kenzu Abdella and Prof. Suresh Narine, for allowing me the opportunity to work with them. Their expert tutelage and words of encouragement allowed me to successfully complete my research. I would like to thank David Riegert for assisting me with the programming of several models and for always providing me with great feedback whenever I had any questions or simply wished to talk (I know he probably got tired of me after a while, but Dave's the best). I would also like to thank my mother, father, and sister, Nikki, for their prayers and words of encouragement. My brother, Chevey, for managing my assets while I was away, and my sister, Nolly, for always calling to check up on me and providing great advice. Sincerest thanks are also extended to Cindy Gagnon for being a wonderful and amazing person and for treating me so well while I was here. She is the best landlady in Canada. Lou Gagnon, for being the coolest dude in Canada and providing me with great health tips; Jean-Paul Stit, for being a great listener and having a heart of gold; and Dennis Nephin, the local legend, for the trips, generosity, and kindness. Lastly, I would like to show gratitude to my friends Hamant France and Rondel Ramkissoon, whose help was instrumental in my settling into life in Canada and to professor Laziz Bouzidi for his help and Guidance.

Contents

List of Figures

List of Tables

Chapter 1

General introduction

1.1 Research background and objective

Agricultural crop production is a cornerstone of food security, sustainable agriculture, and economic stability. In tropical countries where climatic conditions favour year-round crop production, farming is often the primary livelihood for many. In Guyana, agriculture plays a significant role in the country's economy, employing 17% of the labour force and contributing approximately 25% to non-oil GDP. Historically, traditional farming practices dominated crop production in Guyana. However, over the past decade or more, there has been a significant shift towards modern farming practices. These include the incorporation of advanced technology and mechanization, increased large-scale crop production for export, and the implementation of some precision agricultural techniques. Precision agriculture, which focuses on the

use of technology and data analysis, is essential for optimized crop production. One fundamental aspect of precision agriculture that is important to optimize crop production is agricultural planning, which often seeks to answer the questions of what, when, where, and how much to plant. In Guyana, as in many other tropical countries, these questions of agricultural planning pose significant challenges to farmers and policymakers alike, who are urgently seeking answers to optimize economic gain or to effectively utilize land resources.

Optimizing profit and economic productivity from water sources while minimizing land usage are critical objectives for many farmers and agricultural decision-makers who are interested in financial gain or improvements in food security and sustainable practices. However, these economic and land usage objectives conflict. Answering the questions of agricultural planning to satisfy these conflicting objectives leads to a multi-objective optimization problem.

When trying to optimize profit, economic water productivity, and land usage, certain factors need to be considered. These factors include crop yield, unit price of the yield, production costs, and the crop's evapotranspiration rate during growth. Among these factors listed, determining the crop yield is the most challenging and significant task, demanding consideration of various other supplemental factors such as soil type, field management practices, crop characteristics, cultivar, irrigation method, planting date, and weather conditions. Weather, being beyond our control and prone to variation, stands out among these supplemental factors and presents uncertainties that

should be considered.

These challenges and uncertainties posed by the weather, however, can be addressed using statistical, mathematical, and machine-learning techniques to model its behaviour. If the behaviour of certain weather variables is accurately modelled, then combining its results with the other supplemental factors in a crop growth simulator can allow us to estimate crop yields.

Crop growth simulators, used to simulate crop yield and biomass development, can determine crop yields over different planting periods and varying climates and estimate their corresponding total evapotranspiration. By harnessing the power of the estimated yield from the crop growth simulator, along with simulated weather data, pricing, and production costs, we can answer the agricultural planning questions of what, when, where, and how much to plant. This approach holds immense promise, as it can serve as a valuable tool to aid policy and decision-making in the agricultural sector. Consequently, the primary objective of this thesis focuses on developing a weather-driven, bioeconomic optimization model that can be used as a tool to aid policy and decision-making in the agricultural sector. This model structure has three primary parts: Weather models, crop simulators and multi-objective optimization.

1.2 Outline of thesis

The organizational structure of the thesis is as follows.

1. In Chapter 2, we focus on modelling precipitation using a two-step process. The first step of the modelling uses first-order Markov chains to model precipitation occurrence, while the second step uses a distribution fit for each month to model precipitation amount. We use the Akaike information criterion (AIC) and Probability-Probability plots (PP plots) to determine the appropriate distribution fit.

2. In Chapter 3, we propose and use a stochastic differential (SDE) equation using an Ornstein-Uhlenbeck process to model daily minimum and maximum temperature and relative humidity. To assess the effectiveness of the model, we compare it to observed minimum and maximum temperature and relative humidity using summary statistics, RSME, MAE and MAPE.

3. In Chapter 4, we compare the performance of an artificial neural network (ANN) model for estimating reference evapotranspiration to 14 common empirical methods. The ANN model was calibrated against the FAO Penman-Monteith equation and uses minimum and maximum temperature, relative humidity and solar radiation as input variables. Solar radiation is estimated using the Hargreaves equation. This ANN model will assimilate simulated data from

the temperature and humidity models to obtain reference evapotranspiration for use in AquaCrop.

4. In Chapter 5, we use the simulated data obtained from Chapters 2, 3, and 4 and integrate it into AquaCrop's climate file along with corresponding observed weather data. In addition to weather data, AquaCrop is also integrated with crop-specific data as well as soil and field management practices data. After this data integration is completed, calibration takes place. We then simulate crop yields under observed and simulated weather conditions to assess how well the simulated data performs in estimating crop yields. The crops used in this analysis are rice, soy, and corn.

5. In Chapter 6, the yield data obtained from AquaCrop and crop evapotranspiration estimated from AquaCrop, dynamic production costs, and dynamic price per unit of each crop are integrated into a multi-objective optimization model. This model is then used to determine the best crop combinations and corresponding planting intervals to satisfy user objectives. The optimization is solved using an NSGA-II algorithm. A description of how the model can be used to aid in planning is provided at the end.

6. In Chapter 7, we present the general conclusion and discuss future work. We summarize each model's and AquaCrop's results for the general conclusion. We then briefly discuss how the model can be used as a tool to aid farmers and

policymakers, and we examine some limitations of the work. For future work, we explore improvements that can be made to enhance the model's accuracy.

1.3 Schematic diagram of model structure

Figure 1.1 depicts the schematic structure of the model. The first level of the model shows the first four climatic variables modelled. The next level of the model depicts an ANN for estimating ETo. After the ANN model has been trained and calibrated, it is fed with simulated data from the minimum and maximum temperatures and relative humidity model, together with solar radiation data estimated from the minimum and maximum temperatures. This integration of simulated data in the ANN model is used to estimate simulated ETo, which is then fed into AquaCrop. The simulated data from the precipitation and minimum and maximum temperature models are also fed into AquaCrop to create a climate file. In addition to simulated climatic data, AquaCrop is provided with field management, soil type, and crop characteristic data to obtain crop yields over various planting intervals and their corresponding crop evapotranspiration. The yield data obtained from AquaCrop, its corresponding evapotranspiration, and dynamic pricing and production cost data are then fed into the multi-objective optimization (MOO) model. The integration of these datasets into the MOO model allows for the computation of the optimal crop combinations that can be used to satisfy user requirements.

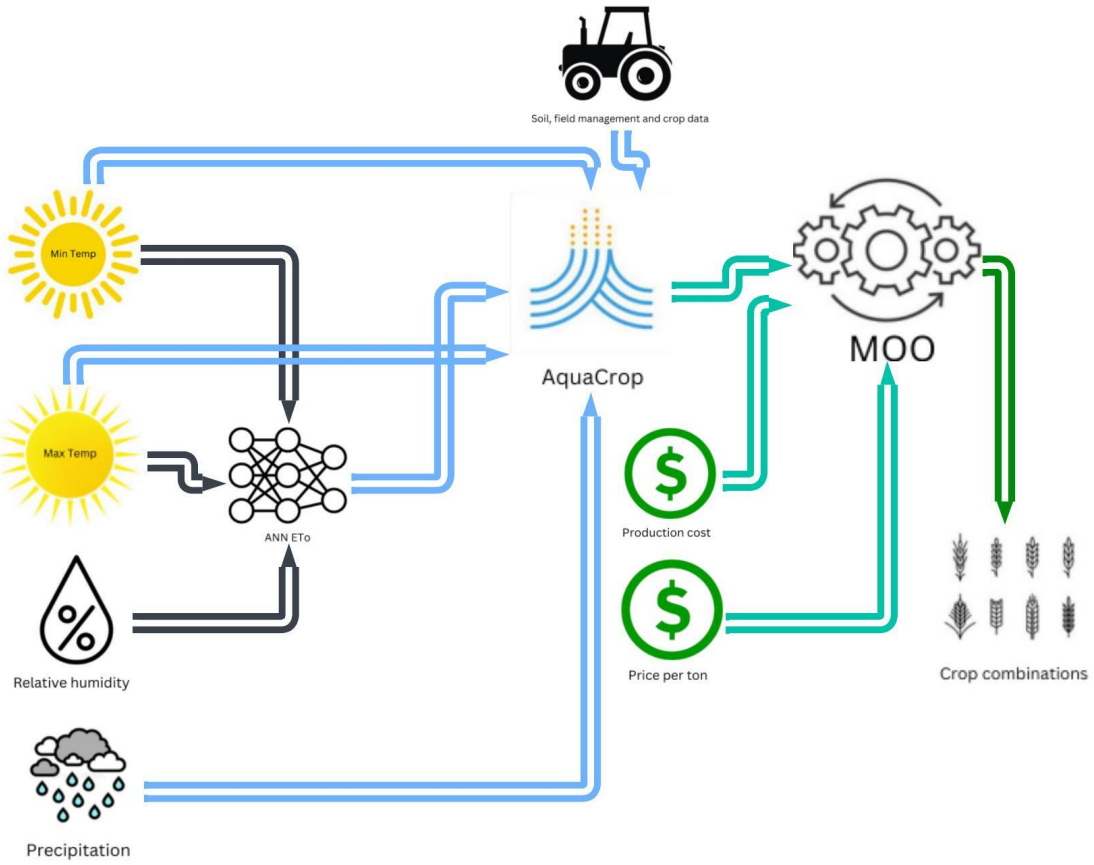


Fig. 1.1. Schematic diagram of the weather-driven bio-economic optimization model

1.4 Crops used in simulation and description of model testing locations

This research uses rice, corn, and soybeans to demonstrate the model's utility and effectiveness. Simulations and testing are done using the land in Rose Hall estate and Ebini. The Rose Hall Estate, situated in East Canje Berbice Corentyne and 79

miles away from the capital city Georgetown, is a former sugar plantation with 3,808 hectares of fertile Whittaker 37 clay soil series variety. Approximately 51 miles south of Rose Hall estate and east of the Berbice River lies the region of Ebini, an area encompassed by the intermediate savannahs of Guyana. Ebini, a mixed woodland grassland ecosystem with soil classified as Ebini sandy loam, possesses thousands of hectares of potential farmland for planting.

For this research, Simulations will be done using approximately 2500 hectares of land in both areas. An overview of the research locations can be seen below.

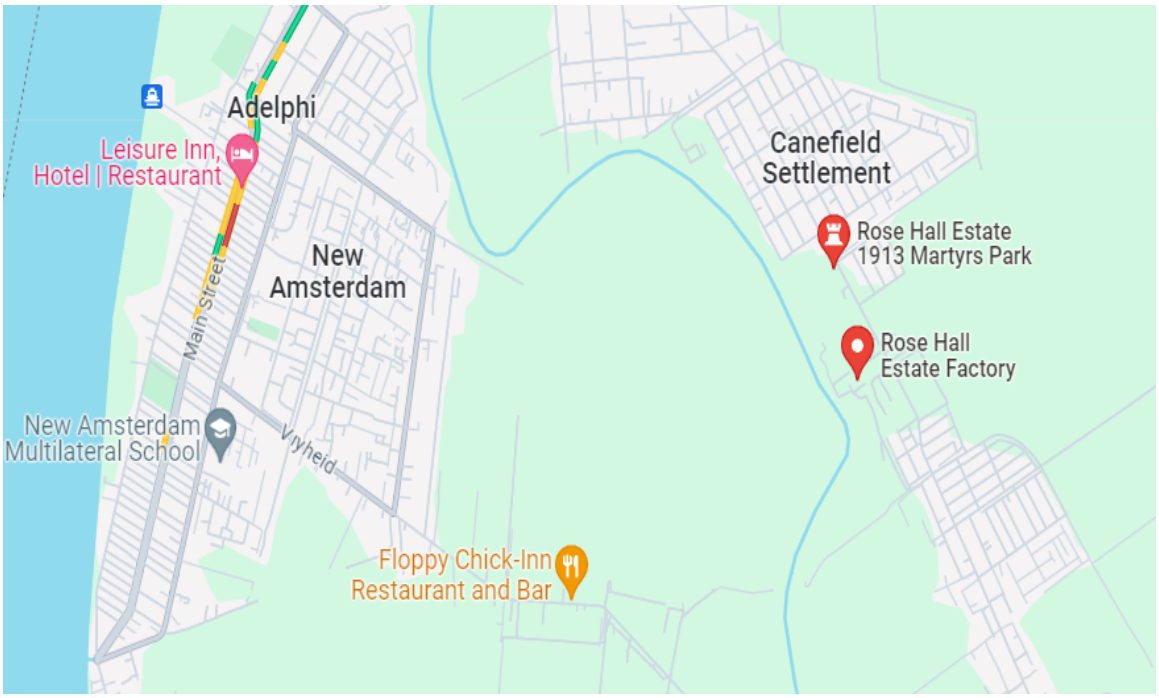


Fig. 1.2. Overhead map of the Rose Hall estate lands

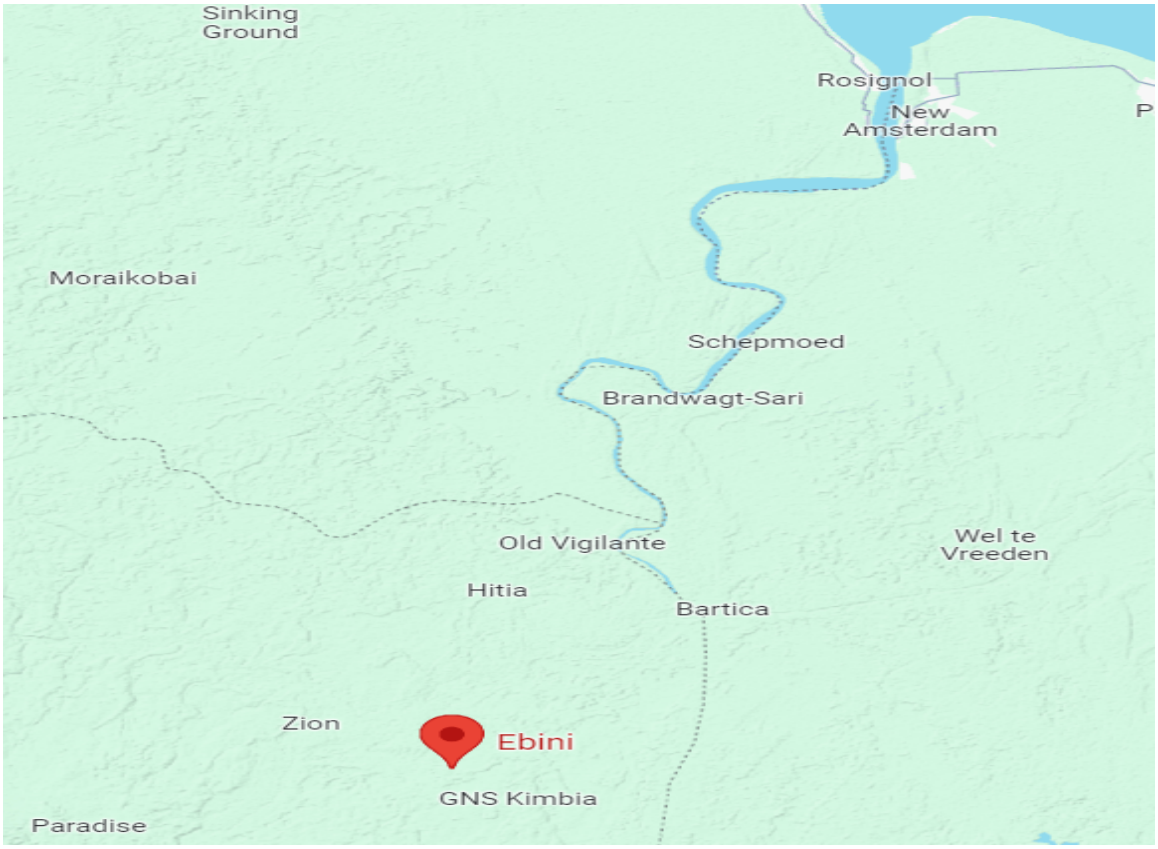


Fig. 1.3. Overhead map of the lands in Ebini

1.5 The climatic data used

The climatic simulations for this study were conducted at three locations in Guyana: New Amsterdam (temperatures, relative humidity, and ETo), Rose Hall estate (precipitation), and Ebini (temperature, relative humidity, ETo, and precipitation). Since Rose Hall estate is close to New Amsterdam, the weather data from the New Amsterdam weather station, as well as simulation data, was used to estimate temperatures,

relative humidity, and ETo at Rose Hall. Climate Engine provided weather data for Ebini (minimum and maximum temperatures, dew point, and wind speed) and precipitation data for Rose Hall due to the absence of weather station data. The Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS)-4.8km-daily dataset in Climate Engine was utilized for precipitation, while the European Centre for Medium-Range Weather Forecasts reanalysis v5 for agriculture (ECMWF Reanalysis v5, also called ERA5 Ag)-9.6km-daily dataset was used to obtain all weather data for use at Ebini.

1.6 Climate at locations used in the study

The climate in coastal areas such as New Amsterdam and Rose Hall estate is one of high humidity, high temperatures (with minimal temperature fluctuations), and heavy rainfall, which typically occurs during rainy seasons. Nighttime temperatures along the coast range from the mid to low 20s, while daytime temperatures reach the mid to low 30s. The average relative humidity typically hovers around 75%–80% and, along with temperatures, is mitigated by the trade winds. The rainfall amount in these areas averages about 2300 mm per year and is influenced by the inter-tropical front or doldrums. There are two rainy seasons: one from April to July and a shorter one from November to early December. The rainfall pattern, as well as temperatures and humidity in this area, greatly influences the rate of evapotranspiration, which ranges from about 4 to 5 mm day^{-1}.

The climate in Ebini is similar to New Amsterdam and other coastal locations in Guyana, being classified as tropical humid and tropical wet and dry. The mean annual rainfall in Ebini is about 2250 mm with a bimodal distribution. Approximately $40 - 60\%$ of the rainfall occurs during the long rainy season from mid-April to mid-August, and less than 20% falls during the shorter, less predictable rainy season. The mean annual temperature is 26°C, with diurnal temperature fluctuations up to 10°C, which are more pronounced than seasonal changes of about 1.5°C. The highest temperatures are recorded from August to November.

1.7 Softwares used

The following software packages were used to build the models and run yield simulations.

1. R studio: R was used to build and run simulations for the weather and multi-objective optimization model. This software was used for its simplicity and ease of use.
2. AquaCrop: Aqucrop is a water-based crop growth simulator developed by the Food and Agricultural Organization in 2009. It balances robustness, simplicity, and accuracy for broad use. This model was chosen based on its ease of use and smaller data requirements compared to other growth simulators.
3. SPSS: This package was used to compute the summary statistics for most of

the models used in the research. It was chosen because of its ease of use.

4. SPC for Excel: This software is an Excel add-in used to obtain the various distribution fits for the precipitation model.
5. Climate Engine: Originating from a Google Faculty Research grant for the White House Climate Initiative, Climate Engine is a collaboration between Google, the Desert Research Institute, and the University of Idaho. It processes satellite and climate data on-demand using Google Earth Engine in a web browser. It also offers on-demand mapping of environmental monitoring datasets, including gridded meteorological observations and remote sensing.

Bridging text

Precipitation, the condensation of water vapour in the atmosphere that falls under the influence of gravity, is an important part of crop yield development. To simulate crop yields, precipitation data is required for assimilation into AquaCrop. As such, the following chapter focuses on modelling precipitation.

Precipitation

Chapter 2

Stochastic precipitation model

2.1 Abstract

Precipitation simulations are important in various sectors, including transportation, farming, and agriculture. This study uses a stochastic model for simulating precipitation over Guyana's Rose Hall estate and the Ebini region. The precipitation model utilizes a two-step process. Firstly, it employs a first-order Markov chain to classify days as wet or dry. In the second step, the model fits a statistical distribution to the wet days to estimate the precipitation amount. Given Guyana's seasonal rainfall pattern, the dataset is divided into twelve strata, each corresponding to a specific month. This stratification aids in identifying rainfall seasonality and facilitates the formulation of customized distribution patterns for the wet days of each month. The results indicate that the model is able to effectively replicate the statistical properties

of observed data, with similarities seen between the mean, median and upper and lower quartiles over both study locations.

2.2 Introduction

Accurate simulations of meteorological variables, especially precipitation, are essential for hydrological, ecological, and agricultural models and decision-making [80]. Examination of daily precipitation time series consistently reveals a distinct stochastic pattern related to the amount and occurrence of rainfall. In these models, the occurrence of rainfall is well characterized by a Markov chain, selected based on the appropriate order that aligns with the conditional dependence of physical events [26]. The fundamental premise of the Markov Model asserts that the likelihood of rainfall on a given day depends on previous rainfall occurrences, with the number of preceding days considered determining the Markov chain's order. In many cases, the daily occurrence of precipitation can be effectively determined by using a basic first-order Markov chain model, which is widely used in conjunction with other lower-order Markov models in the existing literature. Utilizing first-order and other lower-order chains in Markov models for precipitation serves several purposes, including reducing the required parameters for estimation and increasing the model's capacity to compute additional values, such as the likelihood of extended dry spells [29]. The second part of the two-step precipitation model focuses on determining the amount of rainfall that falls on days considered wet, which was determined from the Markov occurrence

models [102]. Determining this rainfall amount involves fitting statistical distributions to these identified wet days. Within the literature, the gamma, lognormal, and Weibull distributions are commonly employed for this purpose [29, 75]. Considering the seasonal variations in rainfall patterns observed in Guyana and the coastal regions, this study focuses on simulating precipitation over the Rose Hall estate and Ebini-Savannah regions. The precipitation simulation involves fitting distributions to historical data from wet days, segmented into twelve parts corresponding to each month of the year. This approach allows for a more accurate model that is capable of simulating the seasonal rainfall patterns observed in the study locations.

2.3 Literature review

Many of the existing stochastic models employed for precipitation simulation typically involve a two-process formulation that encompasses the modelling of both precipitation occurrence and precipitation amount. This two-process model typically employs a Markov chain to represent the occurrence of precipitation. At the same time, a probability distribution is used to determine the amount of precipitation that falls on a wet day [29, 80, 86, 102, 130, 152, 153]. [153] and [86] initiated an investigation on analyzing precipitation patterns in Tel Aviv, Israel. Utilizing the Markov chain method, their research revealed a significant finding – the occurrence of wet days displayed a level of independence across various months, and frequency distributions of sequences of fine days aligned to a logarithmic series. [78] used a stochastic weather

simulation model utilizing a first-order, two-state Markov chain to generate daily sequences of precipitation, temperature, and solar radiation. Their validation across diverse climates demonstrated the model's capability to produce climatic data consistent with long-term observations, particularly emphasizing its utility in locations with scarce historical data. [55] identified limitations in standard first-order Markov models for replicating area-average daily precipitation series, proposing higher-order models with many states to better capture seasonal variance and heavy precipitation spells. In South Korea, [93] investigated the Markov chain properties of daily precipitation occurrences, revealing spatial and temporal dependencies in precipitation patterns and providing valuable insights into the spatial heterogeneity and temporal dynamics of daily precipitation occurrences. Studies by [31], [72], [125], and [152] simulated precipitation by taking higher-order Markov chains into account when modelling precipitation occurrence. This higher-order method was shown to lengthen the Markov model's memory of previous rainy and dry days compared to first-order approaches. [117] further investigated the optimal order of Markov chain models for daily precipitation occurrence, highlighting discrepancies between model orders identified by Bayesian information criteria (BIC) and those suggested by Kolmogorov–Smirnov (K-S) tests, especially concerning wet-spell length distributions. [75] introduced a spatiotemporal precipitation generator based on latent and transformed Gaussian processes, demonstrating robust performance across diverse spatiotemporal scales and geographic contexts. [126] affirmed the adequacy of a first-order Markov chain model

to describe wet and dry weather patterns in Colombo, Sri Lanka, particularly in capturing observed wet spells. [73] proposed a Markov chain-mixed exponential model for stochastic precipitation simulation, showcasing its suitability for reproducing the statistical properties of observed precipitation data. [64] applied a first-order Markov chain model to predict annual rainfall intervals, demonstrating its efficacy in providing accurate forecasts based on historical records. In Qatar [102] used a two-step Markov chain model to simulate precipitation. Their results showed that a lower-order Markov chain model was sufficient for simulating precipitation occurrence in Qatar. For simulating precipitation amounts, a four-kappa distribution was used.

2.4 Materials and methods

The study areas for this research are The Rose Hall estate and Ebini in Guyana. Because both areas lack rainfall data, reanalysis data from the climate engine from 1981-2022 was used for the precipitation simulations. Data from the climate engines' CHIRPS dataset, which offers the highest resolution among the available options (4.8 km), was used to extract data from 1981 to 2022 to build both precipitation models.

2.4.1 Precipitation model

The precipitation model used for simulating rainfall from 1981-2030 draws upon historical precipitation data over the Rose Hall estate and Ebini. The simulation uses a first-order Markov chain to determine precipitation occurrence, while a probability distribution determines the precipitation amount on wet days. For this research, four distinct probability distributions are employed. These distributions are the Lognormal, Weibull, and Gamma (2 and 3 parameters). Given the seasonal nature of rainfall patterns observed in both locations, we stratify the data into 12 parts corresponding to each month. After this, a distribution fit is carried out on each stratum. This approach helps identify the most appropriate distribution that accurately aligns with the precipitation characteristics for that specific month. The selection of appropriate distribution models for each month is determined using AIC [7] and PP plots.

2.4.2 Precipitation occurrence model

In the precipitation simulation model, rainfall occurrence is modelled using a first-order Markov chain. This is done by categorizing the data into two states: "ones," representing wet days, and "zeroes," representing dry days. The transitions between these states are then counted to determine transition probabilities and final states.

2.4.3 Markov chains

For precipitation simulation, using higher-order Markov chains can enhance the model's performance in replicating precipitation occurrence [102]. However, the estimation of parameters becomes more computationally intensive as the order grows. When examining the results, the improvement in model performance in estimating rainfall occurrence achieved by higher-order methods is minimal. As a result, it is more advantageous to utilize a first-order approach [102]. A first-order Markov chain is a type of stochastic process characterized by the fact that the value of the process at time t, X_t, is solely determined by its value at the previous time step $t-1, X_{t-1}$. The value of X_t does not explicitly depend on the specific sequence of values that the process went through to reach X_{t-1}. This means that the likelihood of precipitation occurring tomorrow is contingent solely upon today's weather conditions, namely whether it is currently raining or not. The explanation can be attributed to the Markov property, which can be described as follows:

$$P(X_{t+1} = s_{t+1}|X_t = s_t, X_{t-1} = s_{t-1}, \cdots, X_0 = s_0) = P(X_{t+1} = s_{t+1}|X_t = s_t)\,, \quad (2.1)$$

where the time variable, denoted as t, takes on values from the set $\{0, 1, 2, \cdots, T\}$. The state space, denoted by s, consists of elements from the set $\{w, d\}$.

For a Markov chain, it is essential to have a transition matrix to establish the probability associated with the occurrence of each event. In this context, the elements of the transition probability matrix give the conditional probability of future state j

given the current state i. The transition matrix $\mathbf{P}$ is given as

$$\mathbf{P} = \begin{bmatrix} p_{11} & \cdots & p_{1j} \\ \vdots & \ddots & \vdots \\ p_{i1} & \cdots & p_{ij} \end{bmatrix} \qquad \text{for } i, j \in S\,, \tag{2.2}$$

where $p_{ij} = P(X_{t+1} = j | X_t = i)$. The transition matrix has the property that the sum of each row must be equal to one [102]

$$\sum_{j=1}^{S} P_{ij} = \sum_{j=1}^{S} P(X_{t+1} = j | X_t = i) = \sum_{j=1}^{S} P_{\{X_t = i\}}(X_{t+1} = j) = 1. \tag{2.3}$$

The incidence of precipitation may be classified into two states: dry and wet. Hence, the transition matrix is determined by four conditional probabilities. These are given by

$$P_{dd} = P(\text{dry on day t+1} | \text{dry on day t})\,, \tag{2.4}$$

$$P_{dw} = P(\text{wet on day t+1} | \text{dry on day t})\,, \tag{2.5}$$

$$P_{wd} = P(\text{dry on day t+1} | \text{wet on day t})\,, \tag{2.6}$$

$$P_{ww} = P(\text{wet on day t+1} | \text{wet on day t})\,. \tag{2.7}$$

Given that only two states exist, the transition probabilities inside a given state are complimentary. It is optional to estimate all four transition probabilities; instead, it suffices to estimate a single probability for each pair of transitions [102]. Given this, we can define the transition probability matrix as follows

$$\mathbf{P} = \begin{bmatrix} p_{dd} & p_{dw} \\ p_{wd} & p_{ww} \end{bmatrix} \qquad 0 \leq p_{ij} \leq 1 \qquad i, j = \{d, w\}\,, \tag{2.8}$$

where $p_{ww} = 1 - p_{wd}$.

Employing the transition matrix makes it possible to compute the stationary state vector. The stationary state vector π, has the unique property of $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}$. It gives the long-run relative frequency of being in a wet and dry state and has the property: $\sum_{i=1}^{S} \pi_i = 1$, where $\pi_i \geq 0$ for all $i.\pi_i$ denotes the probability of being in state i [51, 102, 118].

$$\boldsymbol{\pi} = \begin{pmatrix} \pi_d & \pi_w \end{pmatrix}, \tag{2.9}$$

$$\begin{pmatrix} \pi_d & \pi_w \end{pmatrix} = \begin{pmatrix} \pi_d & \pi_w \end{pmatrix} \begin{bmatrix} 1 - p_{dw} & p_{dw} \\ p_{wd} & 1 - p_{wd} \end{bmatrix}. \tag{2.10}$$

For a two-state Markov chain, the stationary probabilities π_d and π_w are given as follows:

$$\pi_d = \frac{p_{wd}}{p_{wd} + p_{dw}}, \tag{2.11}$$

$$\pi_w = \frac{p_{dw}}{p_{wd} + p_{dw}}. \tag{2.12}$$

2.4.4 Determining precipitation amount

The second part of the stochastic precipitation model seeks to determine the precipitation amount. This is determined by finding a suitable generalized probability distribution that fits the wet days of the precipitation data stratified into months. According to existing literature findings, the gamma distribution emerges as one

of the most widely used distributions for fitting precipitation data on wet days [29, 51, 75, 98, 118]. Considering this and drawing insights from the density plot and bar charts presented below, distributions skewed to the right will be considered for modelling precipitation quantities. This study will test the following distributions: the Weibull, two-parameter Gamma, three-parameter Gamma, and Lognormal. Log-likelihood estimation was used to estimate the parameters of these distributions.

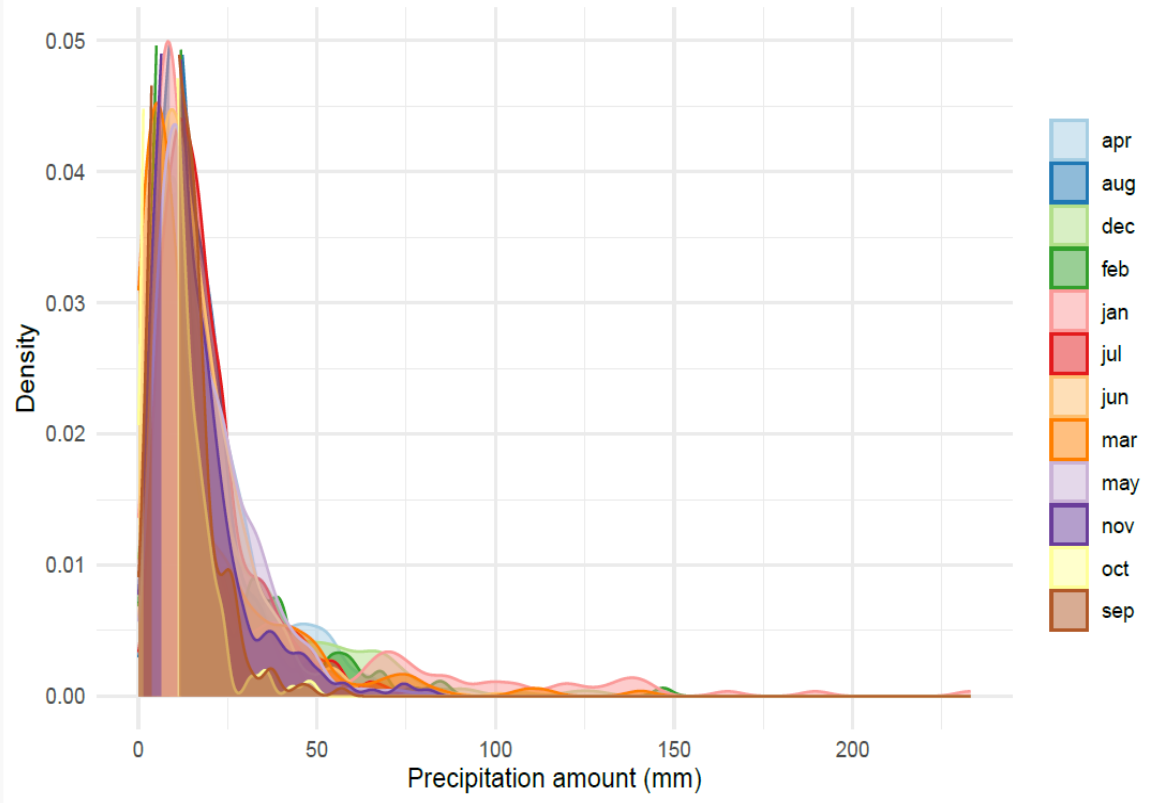


Fig. 2.1. Density plot of monthly precipitation at Rose Hall (1981-2022)

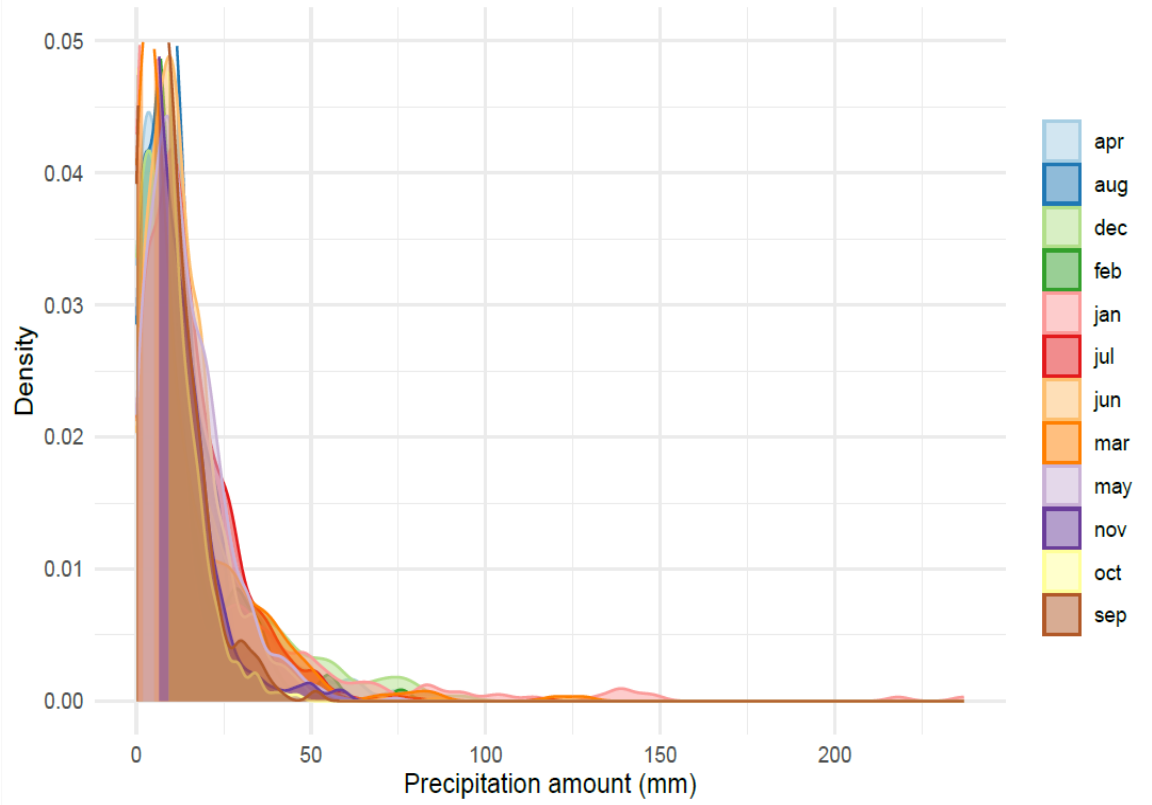


Fig. 2.2. Density plot of month precipitation at Ebini (1981-2022)

2.4.5 Probability distributions

Four distributions will be tested to determine the best fit for the monthly precipitation data. The choices of distribution were made based on the prevalence in the literature.

1) Lognormal distribution

The lognormal distribution is a continuous probability distribution that displays right skewness, characterized by elongated tails in the positive direction. It can be

generated through a variable method transformation from the normal distribution. As a result, if the random variable X follows a lognormal distribution, then Y = ln(X) follows a normal distribution. This distribution finds common application in modelling various natural phenomena, encompassing but not limited to revenue distributions and precipitation levels.

The standard lognormal distribution comprises two parameters representing location and scale.

$$f_{LN}(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(\ln x - \mu)^2}{2\sigma^2}\right), \tag{2.13}$$

where $0 \leq x < \infty, -\infty < \mu < \infty, \sigma > 0$. μ and σ are the parameters of the distribution, called the location and scale parameters.

The mean of the lognormal distribution is given by $E(f_{LN}) = \exp\left(\mu + \frac{1}{2}\sigma^2\right)$ and its variance is given by $V(f_{LN}) = \left(e^{\sigma^2} - 1\right)\exp\left(2\mu + \sigma^2\right)$. The maximum likelihood estimate of the two lognormal distribution parameters is given below [102]

$$\hat{\mu} = \frac{\sum_k \ln x_k}{n}, \tag{2.14}$$

$$\hat{\sigma}^2 = \frac{\sum_k (\ln x_k - \hat{\mu})^2}{n}. \tag{2.15}$$

2) Gamma distribution (Two and Three parameters)

The Gamma distribution is one of the most widely used distributions for determining the precipitation amount [51, 118]. The two-parameter gamma distributions have two parameters corresponding to α (shape) and β (scale). whereas the three parameters distributions have α (shape), β (scale) and λ (threshold). The probability density function of the Two and three gamma distributions are given as

$$f_{G2}(x;\alpha,\beta) = \frac{\beta^{-\alpha}x^{\alpha-1}e^{-x/\beta}}{\Gamma(\alpha)}. \tag{2.16}$$

$\Gamma(\alpha)$ is a gamma function evaluated at α and given by

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}e^{-t}dt. \tag{2.17}$$

The two-parameter gamma distribution has a mean of $E(f_{G2}) = \alpha\beta$ and a variance of $V(f_{G2}) = \alpha\beta^2$ [74].

$$f_{G3}(y;\gamma,\alpha,\beta) = \begin{cases} \dfrac{1}{\Gamma(\alpha)\beta^\alpha}(y-\gamma)^{\alpha-1}e^{-\dfrac{y-\gamma}{\beta}}, & \gamma < y < \infty, \beta > 0, \alpha > 0 \\ 0, & \text{otherwise.} \end{cases} \tag{2.18}$$

The three-parameter distribution (Eq 2.18) has a mean of $E(f_{G3}) = \lambda + \alpha\beta$ and a variance of $V(f_{G3}) = \alpha\beta^2$. [17]

3) Weibull distribution

The Weibull distribution is a widely applied continuous probability distribution that has found extensive use in examining product dependability, analyzing life statis-

tics, and modelling failure times in several fields, such as biology, economics, and engineering [161]. The probability density function for the Weibull distribution is given as

$$f_W(x;\lambda,k) = \begin{cases} \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k}, & x \geq 0, \\ 0, & x < 0, \end{cases} \tag{2.19}$$

where $k > 0$ is the shape parameter and $\lambda > 0$ is the scale parameter of the distribution. The mean and variance of the Weibull distribution are

$$E(f_W) = \lambda\Gamma(1 + 1/k) \tag{2.20}$$

$$V(f_W) == \lambda^2 \left[\Gamma\left(1 + \frac{2}{k}\right) - \left(\Gamma\left(1 + \frac{1}{k}\right)\right)^2\right]. \tag{2.21}$$

2.4.6 Validation of distribution fit

When modelling precipitation amounts, choosing the distribution that best fits the data on the amount of precipitation that falls on wet days is essential. This process is commonly known as goodness-of-fit testing and is commonly used in the literature [29, 51, 75, 102, 118]. This research utilizes two procedures: the AIC and PP-plots.

1) Akaike information criterion

The AIC was one of the first model selection criterion that gained widespread recognition in the statistical community. Formulated by [7], AIC was introduced as a method to evaluate the effectiveness of linear regression models. The fundamental concept of the AIC is to identify a model that minimizes the loss of information from the given data. The formulation for the AIC is given as $AIC = -2\ln(L) + 2K$. L represents the log-likelihood function in this equation, commonly used in statistical modelling. K, conversely, denotes the number of parameters included in the model being analyzed. A penalty term, represented as K, is associated with the number of parameters. Smaller values of AIC are preferred when deciding on distribution fit.

2) PP-Plots

PP-plots are graphical methods used to determine how closely two datasets agree or how well a particular dataset fits a specific distribution. The data is plotted against a theoretical distribution, expecting the plotted points to exhibit a linear pattern. Deviation from the linear trend suggests a departure from the designated distribution.

2.4.7 Precipitation simulation

The subsequent steps are used to simulate precipitation from 1981 to 2022 and 2023 to 2030. The initial three steps encompass the processes necessary for generating precipitation occurrences. The final step involves the procedure required for generating precipitation amounts. The combination of these steps is essential to the formulation of the precipitation model.

Step 1: Transform your data into a binary representation of wet and dry states (0 for dry, 1 for wet). Employ this binary state to construct a transition probability matrix where the probability of transitioning from a wet day to another wet day is calculated by dividing the number of wet-to-wet transitions by the total number of wet days. Similarly, the probability of a wet-to-dry transition is determined by dividing the number of wet-to-dry transitions by the total number of wet days. This process is repeated for all transitions, including dry-to-wet and dry-to-dry. From this, the final state vectors, (π_d, π_w), can be computed.

Step 2: For generating precipitation occurrences throughout the projected period, employ a random number generator sourced from a uniform distribution with a range of 0 to 1. If the generated random number is smaller than π_d, the state is considered dry; otherwise, it is deemed wet. Next, produce another random number within the range of 0 to 1 and compare it to the transition probability for transitioning from the current state to an alternative state. If the number is equal to or less than that

probability of transitioning to the next state, the subsequent state will be that state. Otherwise, you will be in another state.

Step 3: Repeat step 2 until you have generated all the days required.

Step 4: Once the precipitation occurrence has been determined, random values are generated from the probability distribution fitted to the precipitation amount on wet days. These values correspond to the projected amount of rain simulated to occur on days identified as wet days by the precipitation occurrence model.

2.5 Results and discussions

May and June exhibited the lowest transition probability from wet to dry days in both study locations (Tables 2.1 and 2.2) and aligned with the traditional rainfall patterns observed in these areas [105]. During the observed periods, the transition probabilities remained below 0.4. Moreover, May and June, along with July, showed the highest probability of transitioning from a wet day to another wet day, with probabilities exceeding 0.5 and 0.7 in Rose Hall estate and Ebini, respectively. The highest likelihood of transitioning from a wet to a dry day was observed in January and February, with transition probabilities ranging from 0.51 to 0.61. Results also showed that the likelihood of moving from a wet day to another wet day was always larger than the probability of transitioning from a dry day to a wet day. This can be expressed as $p_{dw} < \pi_w < p_{ww}$ [46].

Table 2.1. Transition probability matrix for months at Rose Hall from 1981-2022

State	Dry	Wet	**Month**	State	Dry	Wet	**Month**
Dry	0.836	0.164	**January**	Dry	0.565	0.435	**July**
Wet	0.607	0.393		Wet	0.463	0.537	
Dry	0.829	0.171	**February**	Dry	0.666	0.334	**August**
Wet	0.619	0.381		Wet	0.502	0.498	
Dry	0.858	0.142	**March**	Dry	0.776	0.225	**September**
Wet	0.573	0.427		Wet	0.612	0.388	
Dry	0.827	0.173	**April**	Dry	0.774	0.226	**October**
Wet	0.513	0.487		Wet	0.594	0.406	
Dry	0.669	0.331	**May**	Dry	0.791	0.209	**November**
Wet	0.396	0.604		Wet	0.549	0.451	
Dry	0.526	0.474	**June**	Dry	0.796	0.204	**December**
Wet	0.392	0.608		Wet	0.557	0.443	

Table 2.2. Transition probability matrix for months at Ebini from 1981-2022

State	**Dry**	**Wet**	**Month**	**State**	**Dry**	**Wet**	**Month**
Dry	0.754	0.246	**January**	Dry	0.442	0.558	**July**
Wet	0.509	0.491		Wet	0.287	0.713	
Dry	0.780	0.220	**February**	Dry	0.510	0.490	**August**
Wet	0.518	0.482		Wet	0.380	0.620	
Dry	0.811	0.189	**March**	Dry	0.666	0.334	**September**
Wet	0.503	0.497		Wet	0.445	0.555	
Dry	0.761	0.239	**April**	Dry	0.681	0.319	**October**
Wet	0.389	0.611		Wet	0.471	0.529	
Dry	0.570	0.430	**May**	Dry	0.708	0.292	**November**
Wet	0.293	0.707		Wet	0.424	0.576	
Dry	0.376	0.624	**June**	Dry	0.749	0.251	**December**
Wet	0.261	0.739		Wet	0.444	0.556	

2.5.1 Steady state vector

The transition probability matrix provides information for computing the steady-state vector, which determines the long-term average occurrence of precipitation. Analysis of the steady-state vectors reveals that May, June, and July consistently showed probabilities exceeding 0.45 for encountering wet conditions across both study locations. These values suggest that these months experienced wet days at least 45% of the time over an extended period. This observation also aligns with Guyana's established seasonal rainfall pattern of May-June/July.

Table 2.3. Final state probability vectors at Rose Hall

January	**February**	**March**	**April**
$[0.788, 0.212]$	$[0.783, 0.217]$	$[0.801, 0.199]$	$[0.748, 0.252]$
May	**June**	**July**	**August**
$[0.545, 0.456]$	$[0.392, 0.608]$	$[0.516, 0.484]$	$[0.600, 0.400]$
September	**October**	**November**	**December**
$[0.732, 0.268]$	$[0.725, 0.275]$	$[0.724, 0.276]$	$[0.732, 0.268]$

Table 2.4. Final state probability vectors at Ebini

January	**February**	**March**	**April**
$[0.674, 0.325]$	$[0.702, 0.297]$	$[0.726, 0.273]$	$[0.619, 0.380]$
May	**June**	**July**	**August**
$[0.405, 0.594]$	$[0.294, 0.705]$	$[0.339, 0.660]$	$[0.436, 0.563]$
September	**October**	**November**	**December**
$[0.570, 0.429]$	$[0.596, 0.403]$	$[0.591, 0.408]$	$[0.639, 0.360]$

2.5.2 Summary statistics at Rose Hall and Ebini

The rainfall patterns recorded in Ebini from 1981 to 2022 showed variability across each month. This area experienced a prolonged rainfall season from April to mid-August and a shorter, drier period from September to November. September, specifically, had the lowest average precipitation amount on rainy days, with an average of 8.28mm. These rainfall observations, drawn from four decades of reanalysis data, reveal notable fluctuations in monthly rainfall values. January had the highest recorded maximum rainfall at 236.903mm, while October had the lowest at 45.910mm (see Appendix A). The mean rainfall amount also saw notable variation, ranging from 8.279mm in October to 17.2042mm in January (see Appendix A).

Variations in the number of rainy days were also observed each month, with February recording the fewest at 353. February's lower count of rainy days contrasts with June and July, which both exceeded 800 rainy days. Months such as December and January, with fewer rainy days compared to peak months, saw higher levels of average precipitation on rainy days than in all other months. These differences and other findings in the descriptive measures further emphasize the importance of fitting a separate distribution for each month.

Data gathered over the Rose Hall estate showed a rainfall season that ran from May to August. The results revealed a more varied rainfall pattern than the Ebini, evidenced by the larger standard deviations. Despite May-August having the highest

number of rainy days, January, like in Ebini, recorded the highest maximum monthly rainfall at 233.016 mm. Other months saw significantly less maximum single-day rainfall, with September and October having the lowest. From 1981 to 2022, mean rainfall on rainy days showed significant fluctuations, ranging from 8.958mm in October to 25.715 mm in January (see Appendix A). Over this location, the rainy days ranged from 257 days in February to 691 days in June. The results further show that despite having fewer monthly wet days, the average precipitation amount on those wet days was greater than that of Ebini.

2.5.3 Distribution parameters estimation and selection

The Maximum Likelihood Estimation (MLE) was employed to estimate the distribution parameters. The parameter estimation for each month and associated AIC values is shown in Table 2.5 & 2.6. The determination of the most suitable distribution for fitting the data of wet days in each month was achieved by utilizing the AIC method and PP plots. Lower AIC values, along with data points that closely align with the reference line on the PP-Plot, are considered optimal. The table below presents the selected distribution for each month. The PP plots for the best-fitting distributions corresponding to each month can be found in the Appendix. Good fits were obtained for almost all months in Ebini. However, this was not the case in Rose Hall estate, where some months had moderately fit (January, February). As a result, multiple simulations had to be run to achieve alignment with observed data.

Table 2.5. Best fitting distributions by months

(a) Rose Hall

Month	Distribution	Location	Shape	Scale	Threshold	AIC
January	Lognormal	2.667		1.002		2260.7
February	Lognormal	2.491		0.784		1889.0
March	Gamma		0.687	23.62		1944.7
April	Gamma		2.043	9.687		2460.8
May	Gamma (3p)		1.879	9	1.562	4436.6
June	Gamma		2.297	7.577		5115.2
July	Lognormal	2.73		0.66		4708.2
August	Lognormal	2.624		0.636		3738.8
September	Gamma		2.432	4.736		2212.4
October	Gamma (3p)		1.746	4.966	0.285	2211.9
November	Lognormal	2.482		0.718		2488.5
December	Lognormal	2.746		0.828		2780.7

(b) Ebini

Month	Distribution	Location	Shape	Scale	Threshold	AIC
January	Lognormal	1.846		1.541		3131.8
February	Gamma (3p)		0.721	14.17	0.141	2324.4
March	Weibull		0.787	11.87		2544.4
April	Gamma (3p)		0.848	16.84	0.0346	3516.9
May	Weibull		1.263	15.22		5607.4
June	Weibull		1.307	14.7		6333.6
July	Weibull		1.195	16.41		6388.2
August	Weibull		1.214	12.6		5061.5
September	Weibull		1.119	10.06		3536.9
October	Gamma		1.088	7.612		3271.2
November	Gamma (3p)		0.889	10.68	0.045	3343.1
December	Gamma (3p)		0.736	22.07	0.008	3540.7

2.5.4 Simulation results

The precipitation model generated simulation data spanning 50 years (1981-2030) across both study locations. This data was divided into two periods: simulated data from 1981 to 2022 and 2023 to 2030. Tables 2 and 3 show that the first-order Markov chain model effectively generates precipitation occurrences in Rose Hall and Ebini. Remarkably, the simulation of precipitation occurrences demonstrated a high degree of accuracy for most months, with May, June, July, and November exhibiting the highest accuracy for Rose Hall and June, July, August, and October showing high accuracy for Ebini.

Table 2.6. Simulated and observed precipitation occurrence by month in Rose Hall and Ebini

(a) Rose Hall

Month	Observed	Simulated
Jan	276	282
Feb	257	265
Mar	260	261
Apr	318	315
May	594	603
Jun	691	699
Jul	631	626
Aug	520	527
Sep	338	335
Oct	358	345
Nov	348	340
Dec	350	360

(b) Ebini

Month	Observed	Simulated
Jan	423	437
Feb	353	366
Mar	357	340
Apr	481	502
May	775	744
Jun	890	887
Jul	859	845
Aug	734	723
Sep	542	567
Oct	525	523
Nov	514	502
Dec	471	473

2.5.5 Simulation of precipitation amount

The comparison between observed and simulated precipitation showed fairly close statistical agreement between the mean and median values in both research locations. The overall similarities in average precipitation between observed and simulated data show the precipitation model's ability to represent average precipitation accurately. This suggests that, on average, the simulated data captures the central tendency of precipitation for both regions reasonably well. Moreover, in the Ebini, the simulated data displays closer proximity to observed mean and median values across the months,

highlighting a potentially more accurate representation than the Rose Hall.

Analysis of the first, second, and third quartiles demonstrates consistent data spreads in both observed and simulated data, underlining the model's ability to reliably capture the central precipitation values. Despite occasional variations in extreme values, which might indicate outliers or model limitations, the overall consistency in quartile ranges implies a certain level of reliability in representing typical precipitation patterns. These results underscore the model's ability to replicate the observed data's statistical properties. This supports its potential utility as a good model for use in various sectors. However, continued validation and refinement are needed to enhance accuracy.

2.5.6 Simulation projections for 2023-2030

The model was further used to simulate precipitation from 2023 to 2030 over both study locations. The descriptive measures obtained are presented in the Appendix. As observed in the previous simulation and consistent with historical data, the months of May, June, and July had the highest number of wet days across both locations, recording 138, 150, and 126 wet days for Rose Hall and 139,180 and 177 days for Ebini. Over this period, September and October once again had the fewest wet days for Rose Hall, while February and March had the lowest number of wet days for Ebini. However, several months saw lower average precipitation than previously simulated months (see Appendix A). In Rose Hall estate, the simulated precipitation

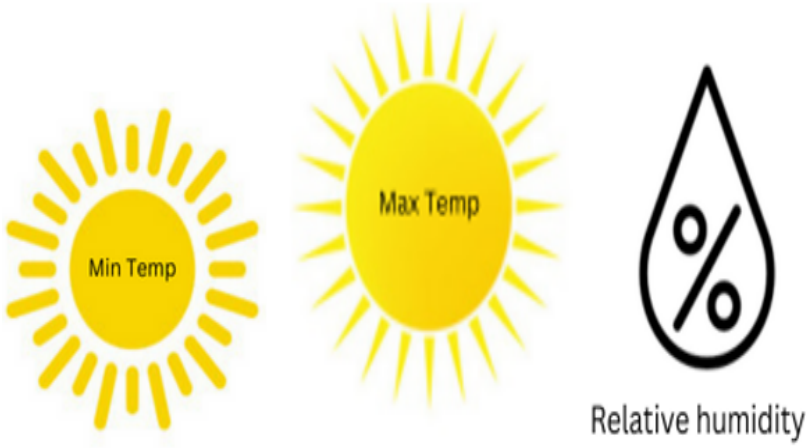
from 2023-2030 shows notable variations across months. December, January, and April had relatively higher mean precipitation values on rainy days; similar trends were seen in the observed data from 1981-2022. Moreover, similar to what was observed from 1981-2022, January again had substantial variability compared to other months, with a standard deviation of 24.581. Notable differences were seen in the mean precipitation on wet days for January, February and April. Simulations produced lower mean precipitations for January and February and a higher mean for April. On the other hand, Ebini had slightly different precipitation patterns over the 2023-2030 simulation period. Over this time frame, January, March, and December stood out with higher mean precipitation levels. While January and March demonstrate higher variability, October and February showcase lower fluctuations in simulated precipitation. Quartile analyses emphasize differences in data spreads between these regions. In Ebini narrower interquartile ranges across the various months were observed, indicating more significant precipitation variability than the Rose Hall estate.

2.6 Conclusion

Stochastic precipitation models are often designed to simulate the occurrence and amount of precipitation on wet/rainy days. This study used a two-step first-order Markov chain model to simulate precipitation over Rose Hall and Ebini, Guyana. The model simulates precipitation occurrence using transitional probabilities and corresponding state vectors, while the amount of precipitation is determined using fitted probability distributions for historical days classified as wet (rainfall occurred). The data used in the model spanned from 1982 to 2022. The results of the first part of the model, which estimates precipitation occurrence, showed similar results when compared to observed data in both research locations. When simulations were conducted, the descriptive measures indicated similar statistical characteristics between the observed and simulated precipitation summary statistics. These combined results suggest that the two-step Markov chain model performs well and can be relied upon to simulate precipitation occurrence and amount in Ebini, Rose Hall, and other possible locations in Guyana.

Bridging text

In Chapter 2, we modelled precipitation, which is crucial for simulating crop yield development and estimation. In the forthcoming chapter, we will focus on modelling minimum and maximum temperatures as well as relative humidity, which are essential weather components needed for estimating and simulating ETo and crop yield development.



Chapter 3

Temperature and relative humidity model

3.1 Abstract

This study uses a stochastic weather model to simulate minimum and maximum temperatures and relative humidity over New Amsterdam and Ebini, Guyana. The proposed model integrates a stochastic differential equation (SDE) with an adjusted Ornstein-Uhlenbeck process. Parameter estimation for the equation is conducted using the least squares method, assuming linear relationships between consecutive observations and normally distributed errors. Twenty and forty years of data were collected from the New Amsterdam weather station and climate engine to develop and validate the models. Comparative analysis of model projections for the years

2021 and 2022 in New Amsterdam and 2019-2022 for Ebini against observed data revealed mean absolute percentage errors of 4.73%, 4.46% and 7.024% for maximum and minimum temperatures and relative humidity, respectively, in New Amsterdam. Comparisons over Ebini had mean absolute percentage errors of 4.625%, 2.723% and 5.565% for maximum and minimum temperatures and relative humidity, respectively. Descriptive measures showed statistical similarities between observed and simulated data. These combined results underscore the SDE model's effectiveness in modelling and replicating the statistical properties of temperature and relative humidity in New Amsterdam and Ebini, Guyana.

3.2 Introduction

Modelling meteorological variables is important to many industries and sectors. Among these variables, maximum and minimum temperatures and relative humidity play pivotal roles due to their impact on agricultural productivity, the weather, and human comfort. Temperature, resulting from sensible heat transfer, conduction, and diffusion processes, is responsible for the growth and development of most living organisms and is a key determining factor of weather conditions [12]. Temperature changes can affect the formation of clouds, precipitation and atmospheric pressures.

Relative humidity, or RH, on the other hand, represents the quantity of water vapour or vapour pressure in the air, expressed as a percentage of the total moisture that the atmosphere can hold at constant pressure and temperature [11]. Relative

humidity is important since it also influences weather patterns, precipitation, and human comfort. Given the combined importance of these two weather variables, accurately modelling them is essential.

Within the literature, deterministic approaches (like time series analysis) are often employed to model temperature and relative humidity. However, while these approaches are good at capturing seasonal changes and linear trends, they often struggle to account for the observed fluctuations and variability in temperature and humidity data. Since temperature and humidity exhibit changing and evolving patterns over time, their behaviour displays stochastic (random) characteristics. Therefore, a model that incorporates both stochasticity and seasonality would be better suited to reflect the behaviour of these variables. When examining temperature and relative humidity data, a tendency to return to a long-term average while also displaying random fluctuations can be seen [102]. For this reason, these variables can be modelled using an Ornstein-Uhlenbeck process, which describes the continuous evolution of a particle or variable as it experiences random fluctuations around its mean. This incorporation allows for the representation of uncertainty, which is often seen in atmospheric processes [102]. Taking seasonality into consideration, an adjustment to the Ornstein-Uhlenbeck is recommended [102].

In the following work, we apply an adjusted Ornstein-Uhlenbeck (O-U) Stochastic Differential Equation (SDE) model to simulate relative humidity and minimum and maximum temperatures in New Amsterdam and Ebini. Leveraging the mean-

reverting behaviour of the O-U stochastic process, the model captures the temporal dynamics of meteorological variables, considering their inherent volatility. Adjusting the model and segmenting the data into months facilitates the estimation of SDE parameters, allowing the incorporation of seasonality. This and the adjusted O-U process enhance the model's potential accuracy, providing a good approach to modelling temperature and humidity in these regions.

3.3 Literature review

Temperature simulation and forecasting models have caught the attention of several researchers and have been a tool in many studies. [22] examined the distinctive features of temperature in developing a trend model for daily temperature. These include cyclic patterns, fluctuations around seasonal averages, the influence of global warming and urbanization, autoregressive shifts, and notably higher volatility during the winter compared to the summer. Their finding suggested a k-lag autocorrelation model for daily residuals. [113] examined global temperature modelling for managerial decision-making by developing a parsimonious forecasting model using ARIMA and GARCH techniques. His results found that the forecasting model performs well compared to rival models and that the estimation results confirm the findings of previous studies. In Southeastern Nigeria, [71] employed SARIMA models to analyze monthly temperature data, observing stable temperature values within forecasted years. Their study emphasized the practical application of mathematical modelling

in tracking temperature fluctuations over time. Expanding the scope to European sites, [96] used various methodologies, including ARIMA, regression, and seasonal models, to forecast daily meteorological time series. Their results showcased the ability of these models to accurately capture temperature dynamics, thereby facilitating sensible forecasts applicable across different climatic zones. Other time series-based model approaches included [24], who used SARIMA techniques to predict monthly mean temperatures in Nanjing, China, yielding promising results with high forecasting accuracy. Their study emphasized the practical utility of SARIMA models in capturing temperature variations over time. Similarly, [100] applied ARIMA models to forecast temperatures in the Sylhet Division of Bangladesh. Their findings highlighted the usefulness of time series analysis in meteorology, revealing suitable ARIMA models for short-term temperature predictions in specific districts. [151] conducted a comparative analysis of ARIMA and ARIMAX models for temperature forecasting in Thailand's southwestern region. Their research underscored the importance of considering dynamic climatic variables, with the optimal ARIMAX model demonstrating superior performance in Ranong, Thailand.

Although time series models have been extensively used to simulate and forecast temperature, they often lack the ability to replicate stochasticity and continuous-time dynamics. As such, continuous models and models that incorporate stochasticity are often used. [37] modelled temperature fluctuations as a regression between desonalized daily temperature. [21] modelled temperature dynamics using a stochastic

process called fractional Brownian motion. The model indicated that the change in temperature is regressed on the previous day's de-seasonalized temperature. Following the elimination of the seasonal average from Central England's daily average temperature dataset, distinct indications of fractional behaviour in the temperature variations were identified. [9] modified the model proposed by [37], incorporating monthly variations in volatility by modelling volatility as a piece-wise constant function. The quadratic variation of the volatility remained stable throughout each month in the dataset, confirming the appropriateness of their volatility selection. In later years, [158] expanded the mean-reverting model proposed by [37] to include seasonality in both the level and volatility, validating this extension using temperature data gathered in Paris over more than 100 years. [18] propose a mean-reverting model driven by a Levy process derived from empirical rejection of normality tests. However, assumptions of constant mean reversion speed and including a Levy process complicate the model. Furthermore, the inclusion of AR(1) inadequately captures slow auto-correlation decay in temperature data. Research by [150] devised a practical pricing model for weather derivatives using the daily average temperature as a key variable, applying the mean-reverting O-U process to a dataset of 62 years' worth of daily recordings to demonstrate how the temperature changes over time. Similar to [150], [102] employed a stochastic differential equation with an O-U process to model daily temperature in Qatar to aid in pricing weather derivatives. A linear assumption between temperature observations was used to determine the model

parameters. Their results showed the Adjusted O-U process's ability to replicate seasonality and the statistical properties of observed temperature. [133] proposed a temperature model for calculating temperature indices. Their model employed a mean-reverting process driven by a Levy process, accounting for jumps and other temperature features. By measuring temperature indices as deviations from a base temperature, the model accommodated the impact of jumps, which can significantly influence temperature variations in certain locations. In a more Novel approach, [40] developed a Levy process-based O-U temperature model with a time-varying speed of mean reversion. Unlike conventional models driven by Wiener processes, their model used an inverse Gaussian distribution to capture the non-normality of temperature differences. The incorporation of cyclic functions for seasonal mean and volatility added to the model's sophistication.

Modelling of humidity using discrete methods has been achieved by using time series models. [116] used an ARIMA model to forecast relative humidity in Ahwaz Station, Iran, setting an initial approach for subsequent research. [83] developed ARIMA and long short-term memory (LSTM) models for daily average relative humidity prediction in Gansu Linxia, China, signifying a shift toward more sophisticated forecasting techniques. The literature on the use of continuous approaches for simulating relative humidity over time is limited. Research by [102] used a stochastic differential equation model with mean reversion to simulate relative humidity over Doha. Their use of an adjusted O-U process represented, at the time, a more novel

approach to estimating relative humidity. The results obtained showed close statistical agreement between observed and simulated humidity.

3.4 Materials and methods

The research locations used in this study are New Amsterdam and Ebini. To conduct this study, 22 years of weather data over New Amsterdam and 43 years of data over Ebini were collected from the New Amsterdam weather station and Climate Engine (for Ebini), respectively. For model validation, data from 2001-2020 for New Amsterdam and data from 1979-2018 for Ebini were used to build the models. Data from 2021-2022 and 2019-2022 were used for validation testing over New Amsterdam and Ebini, respectively.

3.4.1 Relative humidity

Relative humidity (RH) represents the air's saturation level, represented as the ratio between the actual ($e_a(T)$) and the saturation ($e^\circ(T)$) vapour pressure at the equivalent temperature (T) [44]. $RH = 100\frac{e_a(T)}{e^\circ(T)}$. Relative humidity has no dimensions and is often expressed as a percentage [44]. The temperature and absolute water vapour concentration can also affect relative humidity. As a result, a change in relative humidity may be explained by a change in dew point temperature (Tdew) [79].

Since humidity is linked to temperature and dew point, building a model that considers this relationship would be reasonable. However, constructing a model that

considers this would be difficult [102]. Given this intricacy, the humidity model will be created using only time and no additional climatic factors.

3.4.2 Trend and seasonality of temperature and humidity

Since incorporating seasonality is important to the development of a temperature and humidity model, summary statistical measures were computed from the temperature and humidity data collected from both the weather station and Climate Engine. The statistics computed were from 2001-2020 for New Amsterdam and 1979-2018 for Ebini. These descriptive measures were taken to gain a better insight into the data.

If a clear linear trend and seasonality exist within the measured temperature and humidity data, a sinusoid can be used to describe it. This sinusoidal function can be obtained by using a Fourier series with linear trend $a_0 + a_1 t$ as

$$\begin{aligned} DT(t)/\text{DRH(t)} &= a_0 + a_1 t + \sum_{i=1}^{2} b_i \cos(iwt) + \sum_{i=1}^{2} c_i \sin(iwt) \\ &= a_0 + a_1 t + b_1 \cos(wt) + b_2 \cos(2wt) + c_1 \sin(wt) + c_2 \sin(2wt), \end{aligned} \tag{3.1}$$

where w is $2\pi/365$ [102]. The coefficients of this Fourier model can be solved using least squares.

3.4.3 Temperature and relative humidity model with mean reversion

For modelling and simulating daily temperature and relative humidity over New Amsterdam and Ebini, we used a mean reverting O-U process. As suggested by [102], this process is defined as follows

$$dX_t = \left[\lambda\left(\mu - X_t\right) + \frac{d\mu}{dt}\right]dt + \gamma dW_t, \tag{3.2}$$

where X_t represents the daily temperature (minimum and maximum) and relative humidity, μ is the mean where the process reverts to, λ is the speed of mean reversion, γ is the volatility of the model, and W_t is Wiener process. This process is normally distributed with a mean of 0 and standard deviation of $\sqrt{t}$. A large value of λ, means that the values are quickly pulled back to their mean value μ.

This adjusted Ornstein Uhlenbeck (O-U) process can be solved using Itô's Lemma [19, 102] (see Appendix B)

3.5 Parameter estimation using least squares

The least squares approach assumes that successive observations demonstrate a linear relationship with normally distributed errors. This relationship is described by the equation: $X_t = a + bX_{t-1} + \epsilon$, where X_t is the temperature/ relative humidity at time t [19, 102]. The derivation and description of the link between the parameters of the linear equation and the solution of the stochastic differential equation (SDE) are shown below

$$a = \mu\left(1 - e^{-\lambda\delta}\right),\ b = e^{-\lambda\delta},\ \text{ and } sd(\epsilon) = \gamma\sqrt{\frac{1 - e^{-2\lambda\delta}}{2\lambda}}. \tag{3.3}$$

In the above equations, δ represents the time step between time t and time $t-1$; for estimating and simulating daily temperature and humidity, δ is set to 1. Transposing with respect to parameters in the SDE, we have

$$\lambda = -\frac{\ln b}{\delta}, \mu = \frac{a}{1-b},\ \text{ and } \gamma = sd(\epsilon)\sqrt{\frac{-2\ln b}{\delta\left(1-b^2\right)}}. \tag{3.4}$$

The parameters of the least square fit were calculated as follows

$$a = \frac{S_y - bS_x}{n}, b = \frac{nS_{xy} - S_xS_y}{nS_{xx} - S_x^2} \text{ and } sd(\epsilon) = \sqrt{\frac{nS_{yy} - S_y^2 - b\left(nS_{xy} - S_xS_y\right)}{n(n-2)}}, \tag{3.5}$$

where $S_x = \sum_{t=1}^{n} X_{t-1}$, $S_y = \sum_{t=1}^{n} X_t$, $S_{xx} = \sum_{t=1}^{n} X_{t-1}^2$, $S_{xy} = \sum_{t=1}^{n} X_{t-1}X_t$ and $S_{yy} = \sum_{t=1}^{n} X_{\mathrm{t}}^2$ [19, 102].

3.5.1 Equation for simulation of temperature and relative humidity

The solution of the SDE, $X_t = \mu - (\mu - X_{t-1})\, e^{-\lambda\delta} + \gamma\sqrt{\frac{1-e^{-2\lambda\delta}}{2\lambda}} N_{0,1}$, was used to conduct simulations for minimum and maximum temperatures and relative humidity.

3.5.2 Performance metrics

To evaluate the accuracy of the SDE model's simulations, the simulated temperatures and relative humidity for the years 2021 to 2022 for New Amsterdam and 2019 to 2022 for The Ebini were compared with the observed data for the same period. This comparison used the root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE), along with summary statistics.

$$\begin{aligned} \text{R M S E} &= \sqrt{\frac{1}{n}\sum_{i=1}^{n}(S_i - O_i)^2}, \\ \text{MAE} &= \frac{1}{n}\sum_{i=1}^{n}\left|S_i - \text{O}_i\right|, \\ \text{MAPE} &= \frac{1}{n}\sum_{i=1}^{n}\left|\frac{o_i - S_i}{o_i}\right| X100, \end{aligned} \tag{3.6}$$

where S_i represent Simulated temperature and O_i represent observed temperature.

3.6 Results and discussion

3.6.1 Temperature and humidity trends and seasonality at New Amsterdam.

From September to November, temperatures in New Amsterdam consistently exceeded 32°C and 23.9°C for average minimum and maximum temperatures, respectively. February emerged as the warmest month, having the lowest maximum (30.18°C) and minimum (23.05°C) temperatures (see appendix B). While monthly temperature variations are evident, Guyana's tropical climate contributes to limited temperature variability, which is evident in low standard deviation values compared to mean temperatures (0.834 to 1.47 for maximum temperatures and 0.98 to 1.317 for minimum temperatures).

Summary statistics for relative humidity from 2001 to 2020 show noteworthy monthly variations, with several months experiencing maximum humidity levels exceeding 100% (indicating supersaturation, which may have been due to equipment error). May recorded the highest relative humidity level of 103.91%, followed by June at 103.91%, and July at 95.84%. Minimum humidity fluctuated from 56.24% (March) to 70.12% (June), with maximum values ranging from 90.47% (August) to 103.91% (June). Mean humidity varied from 74.02% (October) to 81.52% (June), highlighting seasonal shifts and increased moisture content over New Amsterdam. The range and

standard deviation for humidity were notably higher than for temperatures, emphasizing greater fluctuations (see Appendix B)

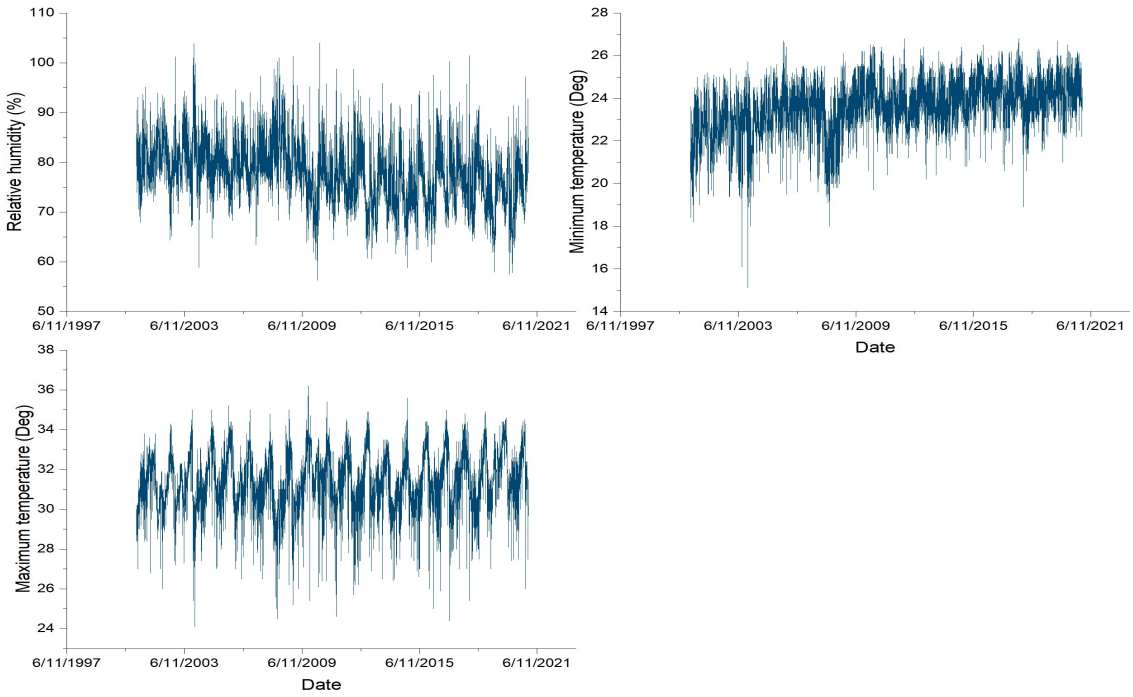


Fig. 3.1. Daily temperatures and relative humidity from 2001-2020

Analyzing the plots in Figure 3.1 and the descriptive measures (see Appendix B), it's evident that minimum temperature and relative humidity have noticeable linear trends—minimum temperatures undergo slight increases, while relative humidity exhibits a slight decrease over time. Meanwhile, a very subtle linear trend is observed in maximum temperature.

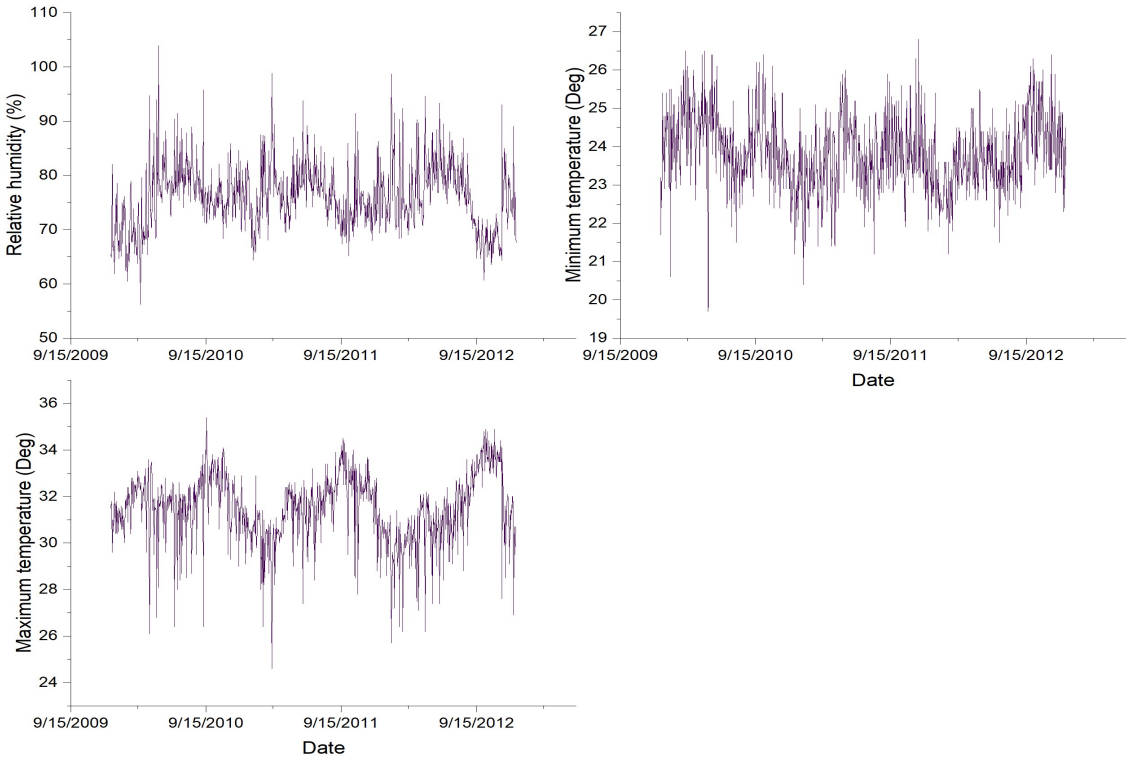


Fig. 3.2. Daily temperatures and relative humidity from 2010-2012

Figure 2 shows the three weather variables on a reduced time scale (2010-2012). Over this period, the seasonal and periodic patterns can be seen.

3.6.2 Fourier series representation

The Fourier series representation of minimum and maximum temperature and relative humidity with coefficients obtained through the method of least squares is given as

$$\begin{aligned}\text{Min}\, T(t) &= 22.63 + 0.0027t - 0.138\cos(wt) - 0.329\cos(2wt) \\ &\quad - 0.204\sin(wt) - 0.28\sin(2wt), \end{aligned} \tag{3.7}$$

$$\begin{aligned}\text{Max}(t) &= 31.198 + 0.000054t - 0.175\cos(wt) - 0.635\cos(2wt) \\ &\quad - 1.131\sin(wt) - 0.268\sin(2wt), \end{aligned} \tag{3.8}$$

$$\begin{aligned}RHT(t) &= 81.648 - 0.0011t - 1.742\cos(wt) + 2.628\cos(2wt) \\ &\quad + 0.595\sin(wt) - 0.25\sin(2wt). \end{aligned} \tag{3.9}$$

The linear trend coefficients for the Fourier series are 0.0027, -0.0011, 0.0000547, and for minimum temperature, relative humidity and maximum temperatures, respectively. These values suggest a slight rise and slight drop in minimum temperature and relative humidity and a very minor increase in maximum temperatures. The intercept values of 22.63, 31.198 and 81.648 align with the average minimum and maximum temperatures and relative humidity observed from 2001-2020. Figure 3.3 compares the daily Fourier series estimates and observed values of minimum and maximum temperature, along with relative humidity. Significant differences are apparent between the two plots, with the random fluctuations observed in the data absent in the

Fourier series. Additionally, the Fourier series tends to generate more distinct upward or downward trends in minimum temperature and Relative humidity compared to the observed data across the region.

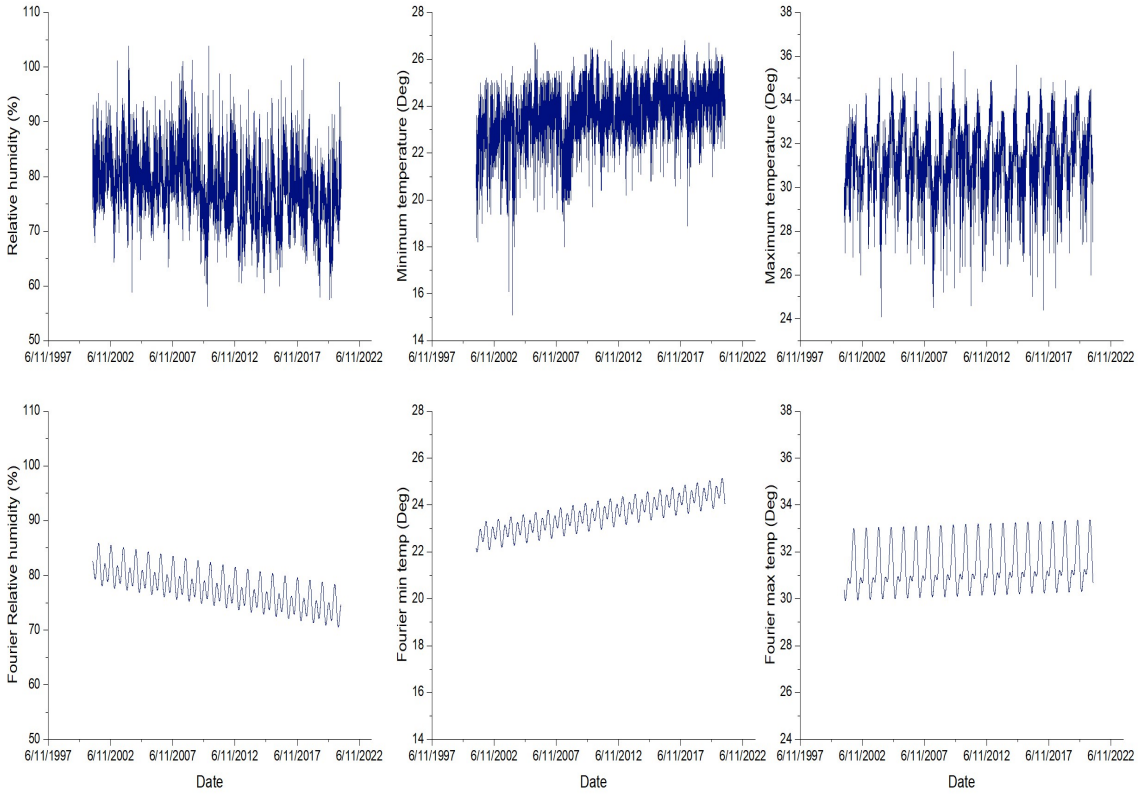


Fig. 3.3. Fourier representation and observed temperature and relative humidity from 2001-2020

3.6.3 Temperature and relative humidity trends and seasonality in Ebini

Temperature data (1979-2018) from Guyana's Ebini shows small seasonal variations. February recorded the lowest average minimum temperature of 22.111°C, while January saw the lowest average maximum of 29.144°C. Standard deviations indicate stable fluctuations, with minimum and maximum temperature deviations ranging from 0.377 to 0.803°C and 0.894 to 1.233°C, respectively. September and October had the highest average maximum temperatures, ranging from 31.518°C to 31.821°C, similar to New Amsterdam (see Appendix B). Relative humidity data (1979-2018) reveals nuanced seasonal shifts. October had the lowest minimum humidity at 56.119%, while June had the highest at 70.473%. May recorded the highest maximum humidity of 93.589%, while October had the lowest at 90.169%. Mean humidity ranged from 73.439% in March to 81.860% in June, reflecting varying moisture content. Standard deviations ranged from $3.222\% - 5.234$. These variations were notably less than those in New Amsterdam.

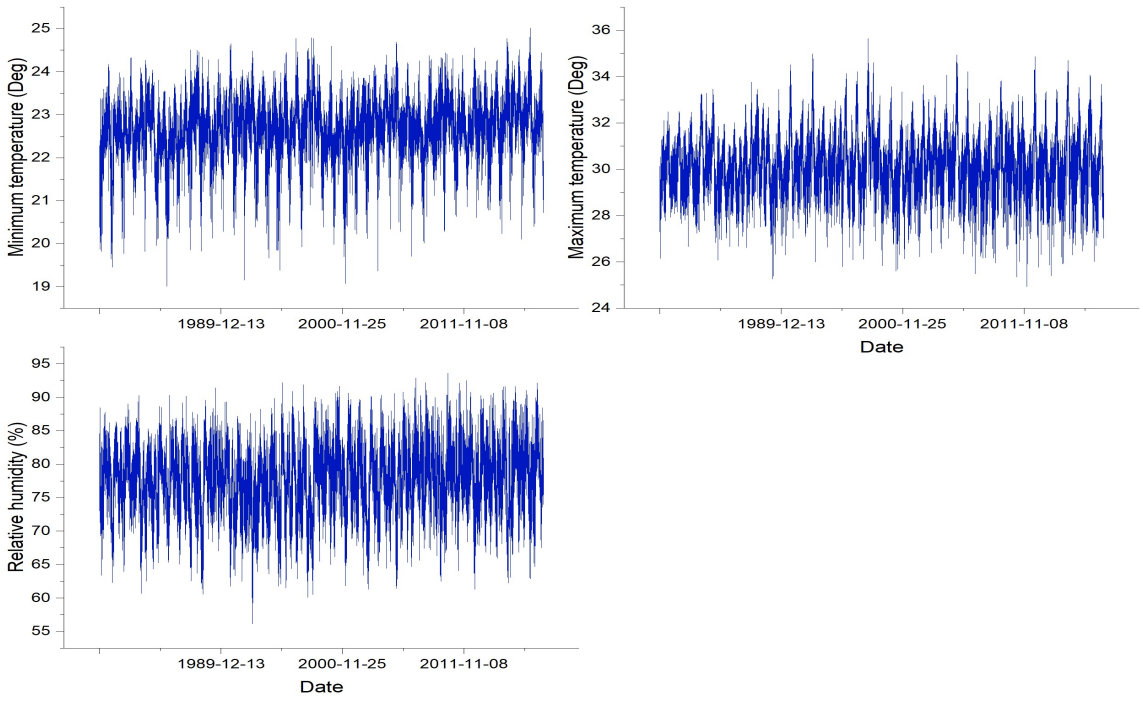


Fig. 3.4. Daily temperature and relative humidity at Ebini

Based on the plots seen in Fig.3.4 and summary statistics (see Appendix B), no noticeable linear trends are seen in temperature and humidity.

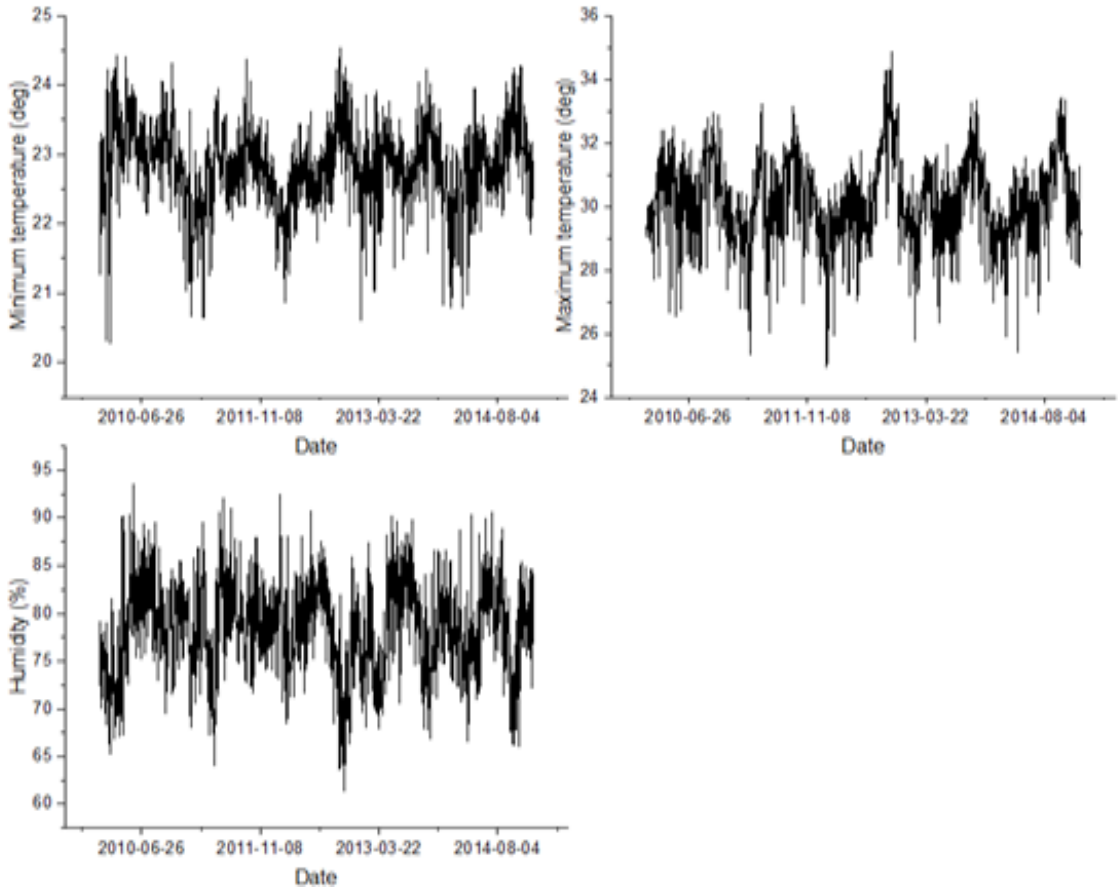


Fig. 3.5. Daily temperature and relative humidity in Ebini from 2010-2014

Figure 3.5 shows the three meteorological variables observed in Ebini from 2010 and 2014. It shows the seasonal and periodic patterns of minimum and maximum temperature and relative humidity.

3.6.4 Fourier series representation

The Fourier series representation of minimum and maximum temperature and relative humidity are given as

$$\begin{aligned}\text{Min}\, T(t) &= 22.703 - 0.114\cos(wt) - 0.164\cos(2wt) \\ &\quad - 0.344\sin(wt) - 0.341\sin(2wt), \end{aligned} \tag{3.10}$$

$$\begin{aligned}\text{Max}\, T(t) &= 30.189 - 0.042\cos(wt) - 0.781\cos(2wt) \\ &\quad - 0.79\sin(wt) - 0.333\sin(2wt), \end{aligned} \tag{3.11}$$

$$\begin{aligned}RHT(t) &= 77.515 - 2.301\cos(wt) + 2.94\cos(2wt) \\ &\quad - 0.793\sin(wt) + 0.101\sin(2wt). \end{aligned} \tag{3.12}$$

The constant values of 22.703, 30.189, and 77.515 align with the minimum and maximum temperatures and relative humidity observed in Ebini from 1978 to 2018 (see Appendix B). A comparison between the daily Fourier series estimates and observed values of minimum and maximum temperature, along with relative humidity, is depicted in Figure 3.6 . Similar to the results seen in New Amsterdam, the Fourier series models do not exhibit the random variations that were seen in the observed data.

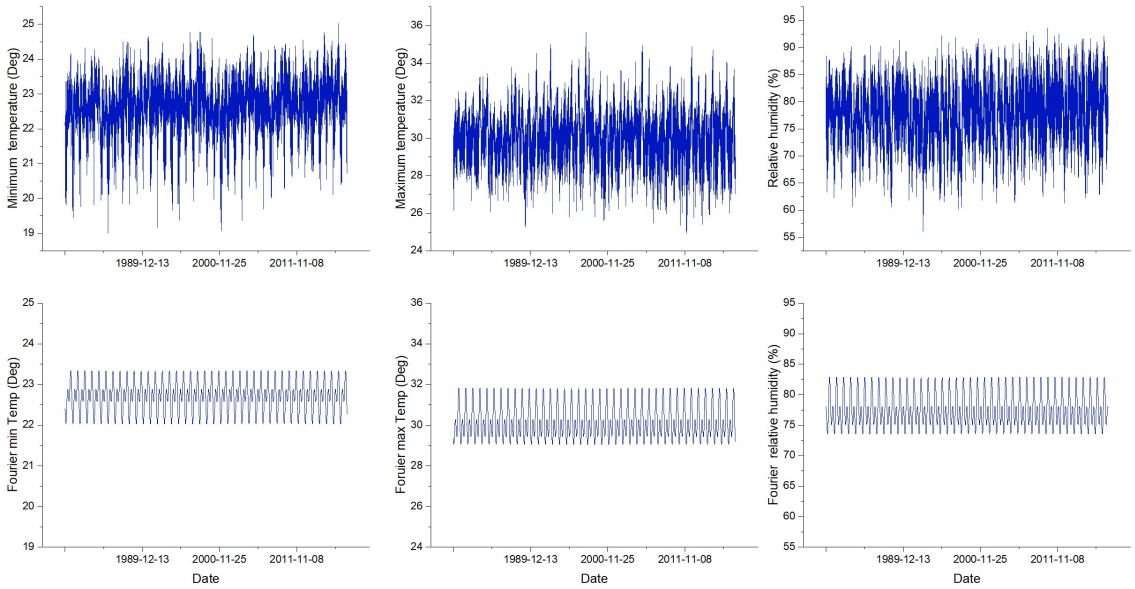


Fig. 3.6. Daily temperature and relative humidity in Ebini from 1978-2022

3.6.5 SDE parameter estimation for temperatures at New Amsterdam and Ebini

Estimating the parameters of the SDE by months was unsuccessful. Consequently, instead of obtaining the parameters for each month, the parameters were calculated using the entire dataset for New Amsterdam (2001-2020) and Ebini (1979-2018). From the results obtained (see Table 3), the minimum temperature in New Amsterdam had the highest volatility (γ=1.0563) and the fastest speed of mean reversion (λ=0.5764) when compared to the maximum temperature. This higher speed of mean reversion suggests that the minimum temperature reverts to its long-term average faster than the maximum temperature in this area. Similar to New Amsterdam, minimum

temperatures in Ebini also had higher speeds of means reversion when compared to maximum temperature (Table 3.1). However, the volatility of the minimum temperature was lower than that of the maximum temperature. Seasonal fluctuations, seen in observed temperatures, were lost in the simulations due to the inability to compute the parameters of the SDE model for each month.

Table 3.1. SDE parameters for temperatures at Ebini and New Amsterdam

Temperature	μ	λ	γ	**location**
Maximum	31.406	0.4573	1.0066	New Amsterdam
Minimum	23.611	0.5764	1.0563	
Maximum	30.195	0.393	0.834	Ebini
Minimum	22.707	0.485	0.498	

3.6.6 SDE parameter estimation For relative humidity at New Amsterdam and Ebini

SDE parameters were estimated for the relative humidity in New Amsterdam (2001 - 2020) and Ebini (1979 - 2018) (Table 3.2). In New Amsterdam, the mean relative humidity (μ) ranged from 74.17% in October to 81.66% in June, while the speed of mean reversion (λ) varied from 0.58 in March to 13.82 in June. In June, July, and August, the highest speeds of mean reversion were observed, with an average across these months of 13.07. Moreover, volatility (γ) peaked during this period, ranging from 9.62 to 15.97.

In Ebini, similar patterns were observed, with mean reversion levels ranging from 73.53% in March to 81.93% in June. The speed of mean reversion varied from 0.74 in October to 13.82 in June. Similar to New Amsterdam, the speed of mean reversion was highest in June, July, and August. This time frame also saw the highest levels of volatility compared to other months, ranging from 12.57 to 13.9.

Table 3.2. SDE parameters for relative humidity at New Amsterdam and Ebini

	New Amsterdam			Ebini		
Month	$\boldsymbol{\mu}$	$\boldsymbol{\lambda}$	$\boldsymbol{\gamma}$	$\boldsymbol{\mu}$	$\boldsymbol{\lambda}$	$\boldsymbol{\gamma}$
Jan	77.820	0.624	5.419	77.112	1.460	5.783
Feb	77.144	0.788	6.143	74.981	0.988	5.138
Mar	75.527	0.577	5.021	73.527	1.052	5.491
Apr	77.067	0.704	5.665	75.733	0.769	5.080
May	80.611	1.995	8.494	79.605	1.131	5.777
Jun	81.656	13.816	15.971	81.928	13.816	13.918
Jul	80.690	13.185	12.290	81.744	13.816	11.758
Aug	78.492	12.215	9.625	80.015	13.816	12.567
Sep	75.710	2.053	6.602	76.360	1.141	5.396
Oct	74.173	1.454	6.829	74.473	0.736	4.905
Nov	76.065	1.267	6.937	76.786	0.910	4.694
Dec	79.360	0.828	6.652	78.780	2.145	5.351

3.6.7 Simulation of temperature

Summary statistics for the simulated temperatures and humidity are juxtaposed with observed data. This was done to evaluate the statistical performance of the SDE

model (see Table 5). Observed daily mean minimum and maximum temperatures closely match simulated values in both study locations. The Descriptive metrics, including standard deviations and quartiles, showed statistical consistency between observed and simulated data. Based on the results seen in the table below, the mean-reverting O-U process in the temperature model effectively replicated the statistical patterns of the observed temperature data.

Table 3.3. Simulated vs observed temperatures at New Amsterdam and Ebini

	New Amsterdamn				Ebini			
	Sim Max	Obs Max	Sim Min	Obs Min	Sim Max	Obs Max	Sim Min	Obs Min
Mean	31.370	31.394	23.583	23.604	30.196	30.200	22.696	22.700
Median	31.367	31.400	23.595	23.700	30.194	30.100	22.698	22.700
STD	1.061	1.440	0.989	1.235	0.952	1.300	0.517	0.700
Minimum	27.506	24.100	19.151	15.100	26.658	24.900	20.409	19.000
Maximum	34.946	36.200	28.093	26.800	33.709	35.600	25.079	25.000
Q1	30.637	30.500	22.918	22.900	29.196	28.900	22.164	22.100
Q3	31.367	31.400	23.595	23.700	30.194	30.100	22.698	22.700

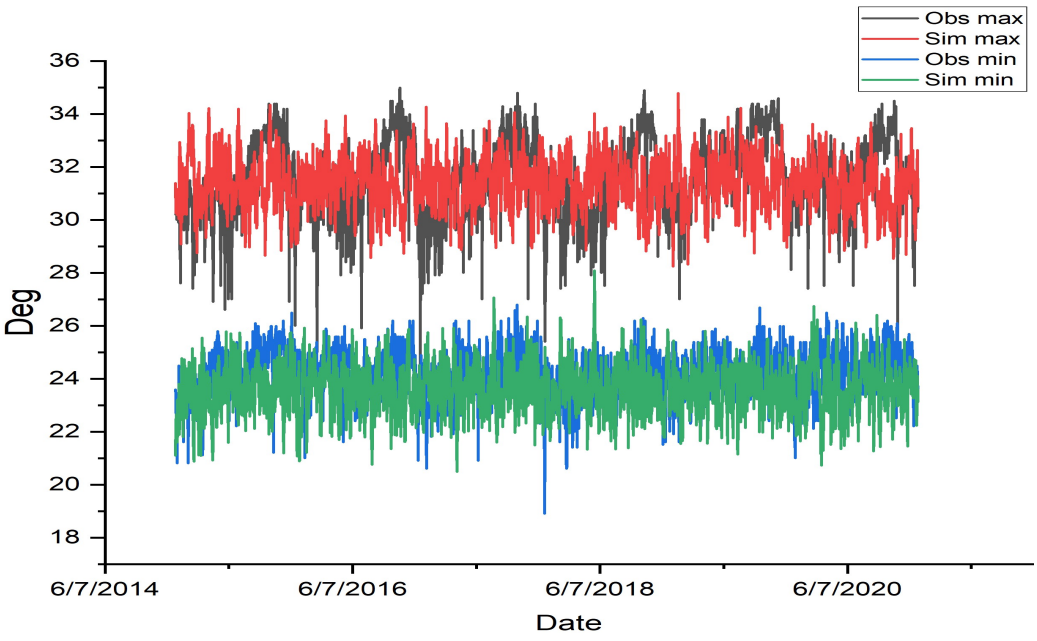


Fig. 3.7. Simulated and observed minimum and maximum temperatures at New Amsterdam from 2015-2020

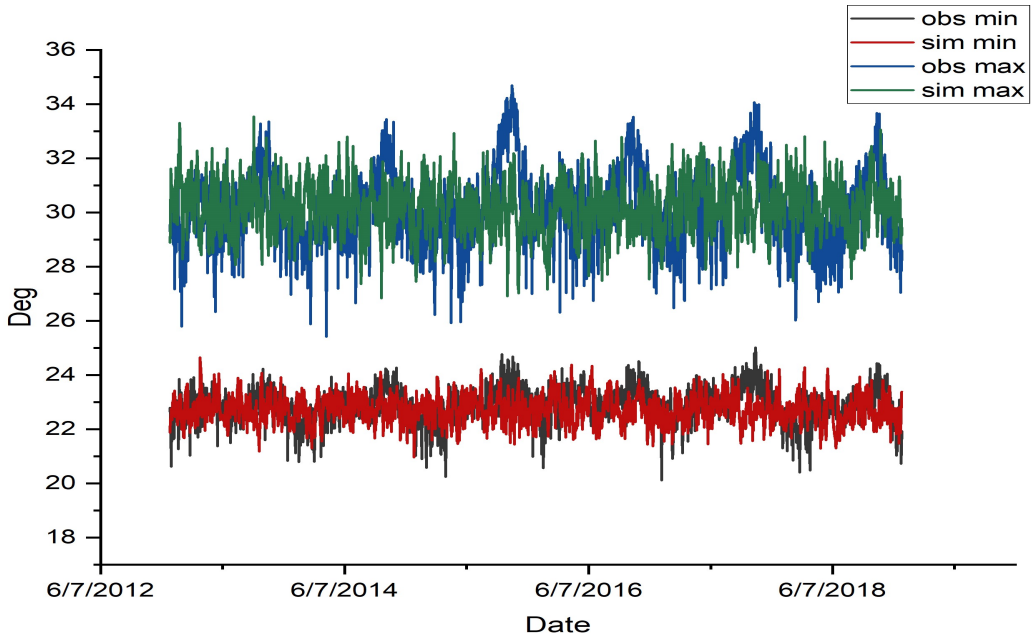


Fig. 3.8. Simulated and observed minimum and maximum temperatures at Ebini from 2013-2018

Figures 3.7 and 3.8 show the SDE model effectively replicating observed temperature trends. The O-U process applied in the Stochastic Differential Equation (SDE) model successfully introduced stochasticity into the simulations, which is evident from the observed random fluctuations within the simulated temperatures. The Simulations of maximum temperatures exhibited more variability than minimum temperatures across both study locations, a discrepancy attributed to the higher variability in maximum temperatures observed in the study areas. Overall, the simulated fluctuations and stochastic behaviour of the SDE model showed similar behaviour with observed minimum temperatures than with maximum temperatures. This pattern

was observed in both locations (see Appendix B).

3.6.8 Simulation of relative humidity

The SDE model was used to simulate humidity in New Amsterdam and Ebini. Table 6 shows similarities in mean, median, and percentiles between simulated and observed humidity across both locations. However, differences were seen in the simulated minimum and maximum values compared to observed values. Both locations had lower standard deviations in simulated data. Although some differences were seen between simulated and observed data, the mean-reverting O-U model performed well in replicating the statistical properties of relative humidity.

Table 3.4. Summary statistics of simulated and observed relative humidity (2001-2020; 1979-2018)

	New Amsterdam		Ebini	
Descriptixes	Obs	Sim	Obs	Sim
Mean	77.679	77.944	77.516	77.635
Median	77.812	78.000	77.905	78.000
STD	6.060	4.605	5.070	4.328
Minimum	56.245	60.000	56.119	61.000
Maximum	103.909	96.000	93.589	92.000
Q1	73.746	75.000	74.317	75.000
Q3	81.420	81.00	80.980	81.000

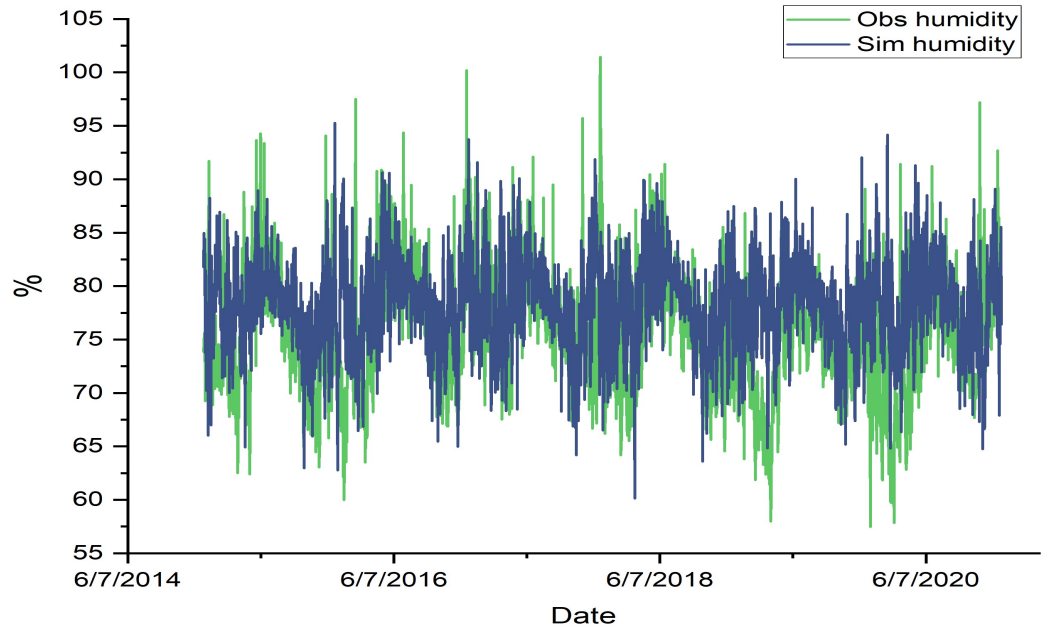


Fig. 3.9. Simulated and observed relative humidity at New Amsterdam 2015-2020

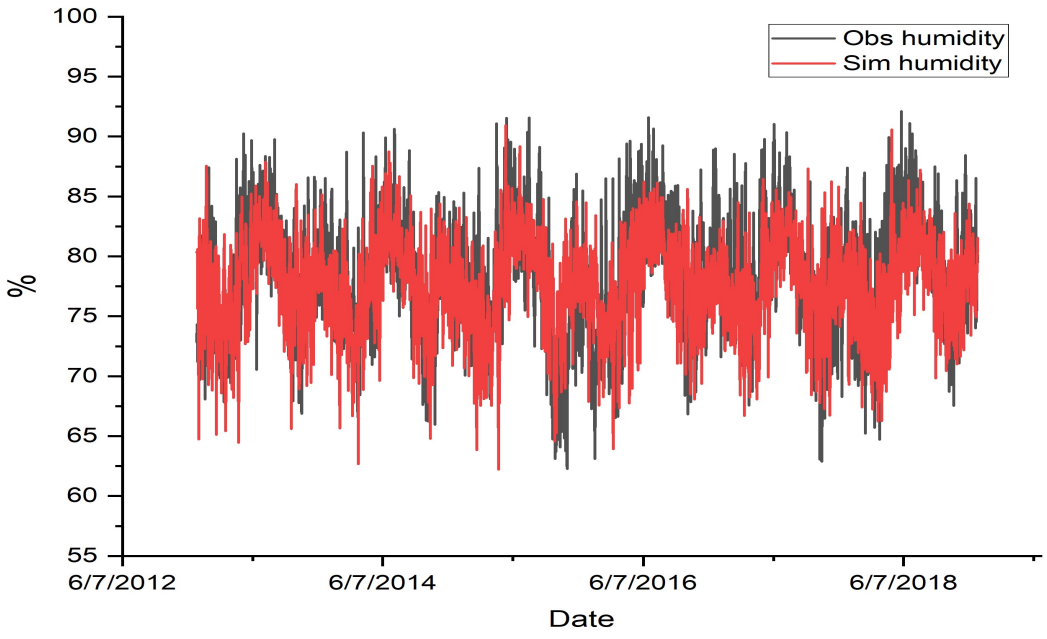


Fig. 3.10. Simulated and observed relative humidity at Ebini 2013-2018

The figures above (3.9 and 3.10) show that the adjusted O-U process effectively introduced stochasticity to the model. The diagrams further highlight the SDE model's effective replication of observed relative humidity trends, showing similar seasonality with the observed data. This trend and seasonal replication are seen in both study locations. Unlike temperature, SDE parameters were estimated for each month, enabling the incorporation of seasonality into the model's simulation. This aspect further contributed to the improved accuracy of the simulations.

3.6.9 Simulations for temperature and relative humidity from 2021-2022; 2019-2022

The SDE model was used to simulate minimum and maximum temperatures from 2021 to 2022 and 2019 to 2022 for New Amsterdam and Ebini, respectively. The results revealed statistical similarities between observed and simulated means. In both locations, the RMSE and MAE remained below 1.9 and 1.5 °C day^{-1}, respectively (see Appendix B). Furthermore, notable low values of MAPE were observed across both areas. Simulations for relative humidity yielded statistical results similar to observed data for all measures of central tendency. The RMSE for both locations was less than 7% day^{-1}; however, New Amsterdam exhibited a higher RMSE and MAPE compared to Ebini.

3.6.10 Temperature and humidity simulations from 2023-2030

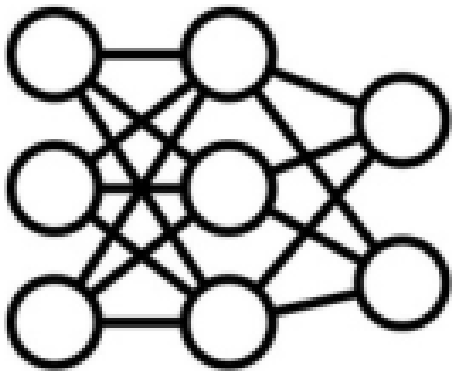
Relative humidity and temperature simulations were done from 2023 to 2030 for both study locations. The results indicated estimates for average temperature and relative humidity similar to those in previous years. Consistent with observed data, simulated temperatures in New Amsterdam were higher than in Ebini. Relative humidity across both locations was found to be similar. However, differences in maximum and minimum relative humidity were observed (see Appendix B)

3.7 Conclusion

Modelling temperature and Relative humidity plays a pivotal role in many industries. This paper used an adjusted Ornstein-Uhlenbeck Stochastic differential equation model to simulate maximum and minimum temperatures and relative humidity in New Amsterdam and Ebini of Guyana. This Model was constructed using data from 2001 to 2020 for New Amsterdam and from 1979 to 2018 for Ebini. Overall, the model performed reasonably well when compared to observed data, producing similar statistical values over this period. Minimal temperature fluctuations over the study regions resulted in the inability to calculate the SDE parameters for minimum and maximum temperatures. As a workaround, temperatures were simulated by using the entire dataset. No such issues occurred during the simulation of relative humidity. The results demonstrated that the model effectively replicated the statistical properties, seasonality, and random fluctuations observed in the temperature and humidity data. This suggests that the model can be used in weather simulations across New Amsterdam, Ebini, and other parts of Guyana.

Bridging text

Chapter 3 focused on modelling temperature and relative humidity, which are important components of not only crop yield development but also the estimation of reference evapotranspiration (ETo). ETo is crucial as it provides a baseline for estimating crop water requirements, a key factor in crop yield development. Given the importance of ETo, the upcoming chapter focuses on ETo estimation using an ANN and subsequently compares its performance with several other methods for ETo estimation. The developed ANN model will later be integrated with data from the temperature and humidity models in Chapter 3 to estimate ETo under simulated conditions. This simulated ETo will then be used in AquaCrop to aid in estimating crop yields.



ANN ETo

Chapter 4

Estimating reference evapotranspiration

4.1 Abstract

Knowledge of reference evapotranspiration, which is the combined loss of water to the atmosphere from the processes of transportation and evaporation, is important for accurate water management, irrigation strategies, and crop yield estimation and development. In this study, we compared the performance of an ANN to multiple temperature-humidity and radiation-based methods for estimating daily reference evapotranspiration. The analysis used 22 years of weather data over New Amsterdam and 43 years of data over the Ebini. The ANN model followed a 4-8-1 architectural structure and used a rectified linear activation function (ReLU). Minimum

and maximum temperature, relative humidity, and solar radiation were used as input variables. Based on several computed statistical metrics, the ANN model outperformed all empirical methods tested for ETo estimation. Performance metrics for the ANN models across the study locations were: RSME = 0.1373 mm day^{-1}, MAE = 0.123 mm day^{-1}, $R^2 = 0.9876$, A = 0.99, and B = 0.161 for New Amsterdam and RSME = 0.1017 mm day^{-1}, MAE = 0.0783 mm day^{-1}, $R^2 = 0.9711$, A = 0.989, and B = 0.044 For Ebini. These results underscore the ANN model's ability to estimate ETo accurately. Consequently, this model shows good potential for accurate ETo estimation over New Amsterdam, Ebini and several other possible regions in Guyana.

4.2 Introduction

Accurately estimating water loss associated with reference evapotranspiration (Eto) is crucial for efficient water management and planning. Eto plays a pivotal role in the hydrology cycle as it represents water lost to the atmosphere through evaporation and transpiration [139]. Estimating ETo can be done using some direct methods such as water budget measurement using lysimeters and water vapour transfer methods using eddy covariance and Bowen ratios [60]. However, these methods are expensive, time-consuming, and have spatial and temporal limitations [52, 65, 104]. Given these drawbacks, other direct methods and indirect process-based approaches relying on meteorological data are frequently used to calculate ETo [132]. Among direct meth-

ods, the FAO Penman-Monteith (PM-56) equation is recommended as the standard method for the definition and computation of the reference evapotranspiration [**?**]. This method is effective since it can accurately estimate ETo over diverse climates and geographic regions [25, 48, 68, 134, 140]. However, despite its widespread application, the PM-56 equation suffers from a notable limitation associated with its extensive meteorological variable requirements [23]. Acquiring reliable daily values for wind speed, humidity, radiation and even temperature, essential inputs for ETo calculation, can be challenging, particularly in developing countries [85]. As a result, indirect methods that require significantly less data for estimating ETo have been used [2, 58, 65, 70, 112, 117, 136, 137]. While these simplified methods are more accurate for estimating monthly and weekly ETo, their accuracy for daily ETo is lower [135]. Additionally, certain empirical methods perform better in specific climatic conditions [66].

Given the PM-56 disadvantages and inefficiency of empirical methods, machine learning approaches such as ANNs are often used to estimate ETo. Given their ability to learn from examples, detect patterns in data, adapt to solutions over time, and process information efficiently [69], ANNs are often able to accurately estimate ETo using less climatic data than process-based and empirical methods.

[77] compared the effectiveness of ANNs with the widely used Penman-Monteith technique for estimating ETo. [157] tested an ANN in Campos dos Goytacazes county, State of Rio de Janeiro to calculate reference evapotranspiration (ETo) based on the

highest and lowest air temperatures. [27] demonstrated that ANN can reliably predict ETo despite insufficient climatological data for all months. Antonopoulos et al. (2016) used an ANN to estimate daily evaporation from Lake Vegoritis in Northern Greece, comparing the results to traditional empirical techniques such as Penman, Priestley-Taylor, and the mass transfer method.

Within the literature, a limited number of studies focus on estimating or modelling ETo in Guyana. Consequently, this study aims to assess the performance of an ANN for estimating ETo in New Amsterdam and Ebini. Daily historical weather data from the New Amsterdam weather station and reanalysis data from the climate engine were used for the analysis.

4.3 Literature review

Many empirical techniques are often used, tested or investigated in the literature to assess their performance in estimating ETo from limited data. Studies like [39] explored the modifications of methods such as Hargreaves to enhance accuracy under different climatic conditions, highlighting Modified-Hargreaves' effectiveness in arid regions. [48] focused on the dynamics of ETo in the Bolivian highlands. Their work found the temperature-based hargreaves-samani method to be able to estimate ETo on the northern parts of the Altiplano but not on the southern end due to the exclusion of aerodynamic factors. Similar to [39], [141] explored the modifications of methods like Hargreaves to enhance accuracy under varying climatic conditions, high-

lighting its effectiveness in arid areas and showed the varying performance of other methods across Iranian provinces. Studies like [34] in Senegal's Senegal River Valley and [54] in the Brazilian Cerrado highlighted the challenges of selecting suitable ETo equations under specific climatic conditions and locations with [34] favouring Romanenko, Schendel, Trabert, Mahringer, and Valiantzas in the Senegal river valley while [54] favoured ASCE-Penman monteith, and Blaney-Criddle. [6] introduced optimized ETo equations using the Hargreaves-Samani method. Their approach and choice of empirical method displayed superior performance in Australia against several other tested approaches. [35] examined mass transfer-based ETo equations in Tanzania and Kenya, favouring Mahringer, Trabert, and Romanenko among empirical methods tested for enhanced water management. In Bosnia and Herzegovina, [164] sought to determine the most accurate empirical approach for estimating ETo from limited data. Their results indicated that an adjusted Trajkovic and an adjusted Hargreaves-Samani are suitable alternatives to Penman-Monteith. In Brazil, [92] conducted a comprehensive evaluation of 29 methods for ET estimation. Using principal component analysis (PCA), they found that the behaviour of ETo was mainly influenced by global solar radiation, maximum air temperature, and relative humidity. Their results showed the effectiveness of Turc and Abtew methods across all regions and seasons tested.

The quest to accurately Estimate ETo while using limited data has spurred the exploration of machine-learning models, particularly ANNS. Early studies, such as

that by [77] in California, showcased the superiority of ANNs over traditional methods. Adopting a 6-7-1 architectural structure and calibrating against lysimeter data, their results showed the superior performance of the ANN model against methods like Penman-Monteith. As researchers delved deeper into ETo estimation, studies like that of [135] and [157] shed light on forecasting methodologies and key climatic variables influencing ETo estimation, respectively. [135] employed historical ETo and weather parameters through the Hargreaves Equation and Multivariate Relevance Vector Machine (MVRVM), yielding favourable results. On the other hand, [157] demonstrated successful ETo estimation in Campos dos Goytacazes, highlighting the significance of maximum and minimum air temperatures, extraterrestrial radiation, and daylight hours as inputs for the ANN model. Subsequent research by [69] and [32], [62] elucidated the importance of temperature, radiation, and relative humidity in accurate ETo estimation, particularly in arid regions. [69] identified temperature and radiation as crucial inputs for accurate estimation. [62] and [32] further underscored the significance of temperature but showed that relative humidity is important in arid climates, boosting the understanding of climatic variables in ETo modelling. Advancements in ANN-based ETo estimation continued with studies like those by [13], [16] and [43], showcasing the potential of ANNs in forecasting ETo across different regions. [13] used air temperature data as input variables, showcasing the potential of ANNs in accurate ETo forecasting in Brazil. [16] achieved high accuracy with root mean square error (RMSE) values ranging from 0.574 to 1.33 mm/day and

correlation coefficients (r) from 0.955 to 0.986 using ANN models in Greece. Further research explored the implications of climate change on ETo dynamics [88] and the comparative performance of various artificial intelligence models, which was carried out by [87]. Makwana compared various artificial intelligence models for ETo modelling, emphasizing the superior performance of ANNs over other models and underscoring the importance of selecting appropriate input variables. Novel approaches to ETo estimation using ANN have been done by [146] and [139]. In the case of [146], who used a ReLu activation function which was not covered much in the literature. Their results showed the effectiveness of this approach versus the sigmoid activation function.

4.4 Materials and methods

The research areas used for estimating ETo are New Amsterdam, located at Latitude 6°24 N and Longitude 57°51′W, and Ebini, located at Latitude 5°38′39.26″N Longitude 57°46′17.51″W. The elevation of these areas is estimated to be around 5.22 m and 25.52 m above sea level, respectively. The study used weather data from the New Amsterdam weather station for 2001-2022 and Climate Engine at Ebini for 1979-2022. The variables in the datasets used to estimate ETo using the PM-56 method were maximum temperature, minimum temperature, average dew point, and 2-meter wind speed. The remaining data needed for estimating ETo were calculated using [10] (see Appendix C).

The daily extraterrestrial radiation used for ETo and solar radiation estimation was obtained from the Sirad package in R [20]. This function requires the Julian day and Latitude in radians as input parameters. Our study used latitudes of 6.243 (radians) and 5.6442 (radians) for the New Amsterdam and Ebini regions, respectively. About 0.5 percent of the data in the New Amsterdam dataset was missing. This was addressed using a K-nearest neighbour algorithm with K set to 5.

4.4.1 Estimation of reference evapotranspiration

The PM-56 equation was used to estimate daily ETo (mm day^{-1}) per FAO's accepted practice [10] for training and testing the ANN. This approach combines energy balance and aerodynamics and was applied in daily time steps. The PM-56 is given as follows:

$$\text{ETo} = \frac{0.408 \cdot \Delta \cdot (R_n - G) + \gamma \cdot \frac{900}{T+273} \cdot u_2 \cdot (e_s - e_a)}{\Delta + \gamma \cdot (1 + 0.34 \cdot u_2)}, \tag{4.1}$$

where, Δ = slope of the vapor pressure curve (kPa°C^{-1}), R_n = net radiation (MJ m^{-2} day^{-1}), G = soil heat flux density (MJ m^{-2} day^{-1}), γ = psychrometric constant (kPa°C^{-1}), T = mean daily air temperature (°C), u_2 = wind speed at 2 meters above ground (m/s), e_s = saturation vapor pressure (kPa), and e_a = actual vapor pressure (kPa). For the computation of daily ETo, the value of G in the PM-56 formula can be treated as zero.

4.4.2 Solar radiation

The following equation is used to approximate solar radiation, R_s, from available weather data [59].

$$R_s = \mathrm{K_{Rs}}\sqrt{(T_{\max} - T_{\min})} \cdot R_a, \tag{4.2}$$

where $\mathrm{K_{Rs}}$ is set at 0.19 for coastal regions (New Amsterdam) and 0.16 for inland locations (Ebini). R_a = Extraterrestrial radiaition was calculated using the R sirad package in MJ $\mathrm{m^{-2}}$ $\mathrm{day^{-1}}$ [20].

4.4.3 ETo estimates using empirical methods

In this research, 12 commonly used empirical methods for estimating ETo were tested to assess the performance of the ANN model on observed data. These techniques were divided into two categories depending on their required inputs: temperature-based and radiation-based. (Table 4.1 and Table 4.2)

Table 4.1. Temperature and humidity-based ETo methods

(a)

Empirical method	Equation	References
Ivanov (IV)	$ET_0 = 0.00006(25 + T_{\text{mean}})^2(100 - RH)$	[67]
Hargreaves (HS1)	$ET_0 = 0.0023R_a(T_{\text{max}} - T_{\text{min}})^{0.5}(T_{\text{mean}} + 17.8)$	[58]
Schendel (SC)	$ET_0 = 16\frac{T_{\text{mean}}}{RH}$	[117]
Hargreaves (HS3)	$ET_0 = 0.408 \cdot 0.003R_a(T_{\text{max}} - T_{\text{min}})^{0.4}(T_{\text{mean}} + 20)$	[58]
Hargreaves (HS4)	$ET_0 = 0.0448 \cdot 0.0025R_a(T_{\text{max}} - T_{\text{min}})^{0.5}(T_{\text{mean}} + 16.8)$	[58]
Hargreaves (HS2)	$ET_0 = 0.0023\frac{R_a}{2.45}(T_{\text{max}} - T_{\text{min}})^{0.5}(T_{\text{mean}} + 17.8)$	[58]
Trajkovic (TR)	$ET_0 = 0.0023R_a(T_{\text{max}} - T_{\text{min}})^{0.24}(T_{\text{mean}} + 17.8)$	[136]

Table 4.2. Radiation-based ETo methods

(a)

Empirical method	Equation	References
Modified Turc (AB)	$\text{ET}_o = (0.2868R_s + 0.6) \cdot \frac{T_{\text{max}}}{T_{\text{max}} + 15}$	[2]
Linacre (LI)	$\text{ET}_o = \frac{\frac{700 \cdot (T_{\text{mean}} + 0.006 \cdot \text{Z})}{100 - L} + 15 \cdot (T_{\text{mean}} - T_d)}{80 - T_{\text{mean}}}$	[84]
Turc (TU)	$\text{ET}_o = 0.013 \left(\frac{T_{\text{mean}}}{T_{\text{mean}} + 15}\right)(R_s + 50)$	[137]
Jensen & Haise (JH)	$\text{ET}_o = \frac{1}{38 - (2 \cdot \frac{Z}{305}) + 7.6 \cdot \frac{50}{[e_s(T_{\text{max}}) - e_s(T_{\text{min}})]}} \cdot \text{A}$	[70]
Irmak (IR3)	$\text{ET}_o = -0.478 + 0.156R_s - 0.0112T_{\text{max}} + 0.0733$	[65]
Irmak (IR2)	$\text{ET}_o = -0.642 + 0.174R_s + 0.0353$	[65]
Irmak (IR1)	$\text{ET}_o = -0.611 + 0.149R_s + 0.079$	[65]

Here, $\mathrm{ET}_o\left(\mathrm{mm\,day^{-1}}\right)$ = reference evapotranspiration; $T_{\mathrm{mean}}\,(^{\circ}\mathrm{C})$ = mean air temperature; $T_{\mathrm{min}}\,(^{\circ}\mathrm{C})$ and $T_{\mathrm{max}}\,(^{\circ}\mathrm{C})$ = minimum and maximum air temperature, respectively; $R_a\left(\mathrm{MJ\,m^{-2}\,day^{-1}}\right)$ = extraterrestrial solar radiation; RH(%) = average relative humidity; $R_s\left(\mathrm{MJ\,m^{-2}\,day^{-1}}\right)$ = solar radiation; z(m) = elevation; L (degrees) = latitude; $T_d\,(^{\circ}\mathrm{C})$ = mean dew point; e_s(mm) = saturated vapor pressure, and $\mathrm{A} = \left(T_{\mathrm{mean}} - \left\{-2.5 - 0.14\left[e_s\left(T_{\mathrm{max}}\right) - e_s\left(T_{\mathrm{min}}\right)\right] - \frac{Z}{550}\right\}\right) \times R_a$

4.4.4 Artificial neural networks

ANNs serve as mathematical representations of biological neural networks, enabling them to learn from examples and detect patterns in data, adapt solutions over time and process information efficiently [69]. In mathematical terms, ANNs function as nonlinear regression models, utilizing a set of weights (w_{ij}) or coefficients to establish input-output mappings. The diagram below illustrates the fundamental architecture of an ANN. This structure consists of three layers: an input layer, a hidden layer, and an output layer. Each layer consists of interconnected neurons (or nodes), with the connections governed by the assigned weights. The number of neurons in each layer defines the model's architectural structure. For instance, (i−j−k) is an ANN with i, j, and k neurons in the input, hidden, and output layers, respectively [16]. The number of neurons in the input layer corresponds to the number of input variables used for estimating ETo. In contrast, the output layer contains a single neuron that estimates the ETo value. Putting too few neurons in a model may result in underfitting and

limited generalization, whereas excess neurons may lead to overfitting [62].

4.4.5 Architectural structure

Predefining an optimal network architecture often necessitated a trial-and-error approach to determine the appropriate number of hidden neurons [28, 69]. However, [77] demonstrated that a single hidden layer with at least n+1 (n = number of input variables) neurons adequately models ETo. [120] demonstrated that using the formula $(4n^2 + 3) / (n^2 - 8)$, where n is the number of input variables can be used to estimate the number of hidden neurons in a neural network. Using these two approaches with n set 4 gave possible hidden neuron values of 5 and 8. Based on testing, the ANN model that used 8 hidden neurons produced the best results. As a result, this study uses a (4-8-1) architectural structure.

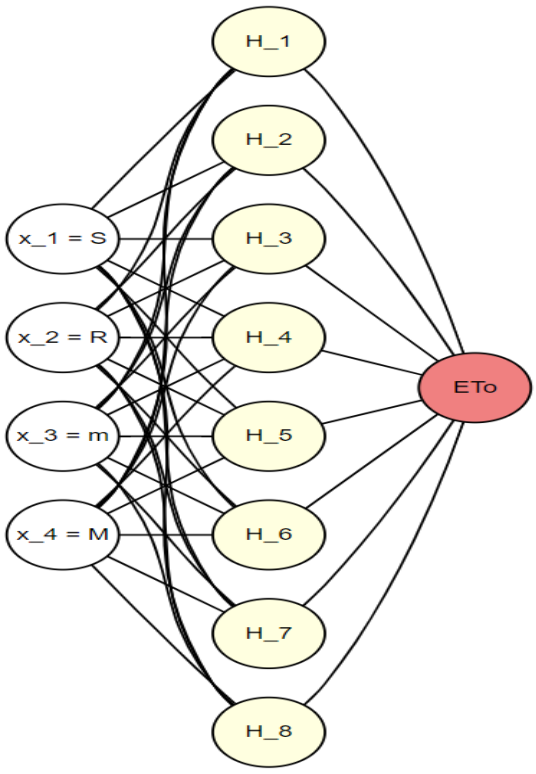


Fig. 4.1. Layout of ANN model with 4-8-1 architectural structure.

where the four input variables are S = solar radiation, m = minimum temperature, M = maximum temperature, and R = relative humidity;

$$H_j = \sum_{i=1}^{n} x_i w_{ij} + b_j, \quad f(H_j) = \max(0, H_j),$$

$$ETo = w_{p0} + \sum_{j=1}^{m} w_{Rj} f\left(\sum_{i=1}^{n} x_i w_{ij} + b_j\right)$$

In the diagram above, each neuron in a given layer is connected to neurons in the subsequent layer through weights denoted by w_{ij}. These weights are adjusted during the training process. For each neuron in the hidden layer, the output is calculated by first taking the weighted sum of the input variables and adding a bias term (b_j) [16].

This sum is then transformed into an output value by applying an activation function. Specifically, each hidden neuron h_j produces an output that is transmitted through the weights w_{pj} to the output neuron. The weighted sum of these contributions is given by $\sum_{j=1}^{m} w_{pj} f\left(\sum_{i=1}^{n} x_i w_{ij} + b_j\right)$. Finally, a bias term w_{p0} is added to this sum to produce the final output, ETo.

4.4.6 Input variables

Obtaining certain climatic variables for estimating ETo can be challenging. Consequently, this research aims to develop a model that can effectively estimate ETo using a limited set of inputs. [69] identified temperature, radiation, and humidity as crucial inputs for accurate ETo estimation. [43] and [16] demonstrated the effectiveness of minimum temperature, maximum temperature, and solar radiation in achieving accurate ETo estimation. In light of these findings, minimum and maximum temperatures, solar radiation, and relative humidity will be used as inputs in the ANN.

4.4.7 Batch size, epochs and learning algorithm.

The number of epochs indicates the total number of times the complete training dataset undergoes forward and backward passes within the model. Insufficient epochs may yield an ANN that inadequately captures system intricacies, while excessive epochs can lead to over-fitting. Conversely, the batch size determines the subset or number of samples propagated through the network at each iteration [146]. For

this research, a consistent epoch value of 500 was chosen to ensure adequate training duration and model convergence, while a small fixed batch size of 10 was employed [146].

The learning algorithm optimizes the objective function and is the primary parameter responsible for the model's learning capacity, assuming sufficient data is available (Walls et al., 2020). Due to its efficiency, stability and faster convergence speed [121], this study uses a Root mean square propagation (RMSprop) Learning Algorithm.

The RMSprop algorithm is given by:

$$w_{\text{new}} = w_{\text{old}} - \frac{\alpha}{\sqrt{\text{MS}(w,t)}} \nabla E\left(w_{\text{old}}\right), \tag{4.3}$$

where w_{new} and w_{old} are the new and old weights, α is the learning rate which was set to 0.0001, ∇E is the gradient of the cost function, and $\text{MS}(w,t)$ is given by

$$\text{MS}(w,t) = \gamma_d \,\text{MS}(w,t-1) + (1-\gamma_d)\left(\nabla E(w)\right)^2, \tag{4.4}$$

and γ_d is the level of decay in the rms term $= 0.9; t =$ the target output and MS $=$ mean square.

4.4.8 Activation function

In the literature, the sigmoid activation function has been commonly used in ANN models estimating ETo [13, 16, 50, 163, 87, 146]. However, this function results in extended training periods, especially with large datasets. Consequently, this study adopts the Rectified Linear Unit (ReLU) activation function, which has demonstrated improved performance and faster training than the sigmoid function [88, 146]. The ReLU activation function is defined as: $f(\mathbf{x}) = \max(0, \mathbf{x})$.

4.4.9 Standardization

The input and output variables used have been standardized to expedite the convergence process. The standardization procedure was carried out using the following equation

$$\mathrm{x_{ni}} = \frac{\mathrm{x_i} - \mathrm{x_{min}}}{\mathrm{x_{max}} - \mathrm{x_{min}}}, \tag{4.5}$$

where: x_{ni} is a normalized dimensionless variable; x_i is the physical value of the variable; $\mathrm{x_{max}}$ is the maximum value of the physical variable, and $\mathrm{x_{min}}$ is the minimum value of the physical variable [16].

4.4.10 Loss function

The loss or error function computes the model's output error relative to the desired output and is the function to be minimized during training; the mean square error (MSE) was used as the performance measure in this study.

$$\text{MSE} = \frac{1}{n}\sum_{i=0}^{n}\left(y_i - \widehat{y}_i\right)^2, \tag{4.6}$$

where y_i = the observed values of ETo; $\hat{y}_i$ = the predicted values of ETo, and n = the number of observations.

4.4.11 Training and testing

The ANN model will be trained using observed data from 2001-2018 for New Amsterdam and from 1979-2012 for Ebini. The remaining observed data will be used for testing. This split gives a train-to-test ratio of 81:19 and 77:23.

4.4.12 Statistical analysis of model performance

The performance of the ANN and empirical methods were assessed from the RMSE (root mean square), MAE (mean absolute error), R^2 (coefficient of determination) and a regression coefficient-based method.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}\left(\text{ET}_{ES} - \text{ET}_{\text{PM56}}\right)^2}{N}}, \tag{4.7}$$

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|\mathrm{ET}_{ES} - \mathrm{ET}_{\mathrm{PM56}}\right|, \tag{4.8}$$

$$R^2 = \left[\frac{\sum_{i=1}^{N}\left(\mathrm{ET}_{ES} - \bar{E}T_{ES}\right)\left(\mathrm{ET}_{\mathrm{PM56}} - \bar{E}T_{\mathrm{PM56}}\right)}{\sqrt{\sum_{i=1}^{N}\left(\mathrm{ET}_{ES} - \bar{E}T_{ES}\right)^2 \sum_{i=1}^{N}\left(\mathrm{ET}_{\mathrm{PM56}} - \bar{E}T_{\mathrm{PM56}}\right)^2}}\right]^2, \tag{4.9}$$

where $\mathrm{ET_{ES}}$ is the reference crop evapotranspiration of the empirical methods and the ANN model tested in the research and $\mathrm{ET_{PM-56}}$ is the reference evapotranspiration estimated by the Penman-Monteith Method. R^2 solely quantifies variability. Thus, R^2 may not be a good match for consistently overpredicted or underpredicted data. Therefore, equation (10) [48, 66] was used to define the calibration parameters for quantitative evaluation.

$$\mathrm{ET}_{ES} = \mathrm{A} \cdot \mathrm{ET}_{\mathrm{PM56}} + \mathrm{B}, \tag{4.10}$$

where A is the regression gradient, and B is the intercept. A and B values closer to unity and zero are deemed optimal [66]

4.5 Results and discussion

4.5.1 Estimating daily ETo using PM-56

Daily ETo was estimated from 2019-2022 in New Amsterdam, Guyana, using the PM-56 method. Throughout this timeframe, the daily ETo values ranged from 1.71 mm day^{-1} to 6.34 mm day^{-1}, with the highest recorded ETo occurring on March 3, 2020, and the lowest on January 15, 2021 (see Appendix C). The average ETo over this period was estimated to be 4.57mm day^{-1} with a standard deviation of 0.54 mm day^{-1}. Throughout these four years, August and September, which were recognized as two of the hottest periods in Guyana, consistently displayed the highest average ETo values over two months. In contrast, January and February consistently registered the lowest average ETo values over two months among the warmest months.

In Ebini, however, daily ETo values were lower and ranged from 2.08 mm day^{-1} to 5.82 mm day^{-1}, reaching their highest point on October 10, 2015, and their lowest on December 12, 2021 (see Appendix C). The average ETo for this period was estimated to be 3.79 mm day^{-1}, with a standard deviation of 0.54 mm. Over the ten years, March, September, and October consistently displayed the highest average ETo values, while May and June consistently registered the lowest daily average ETo values.

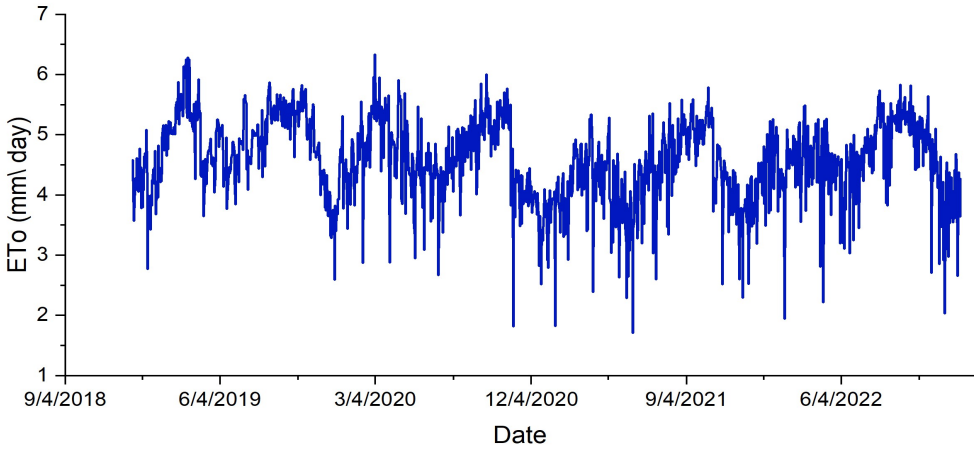


Fig. 4.2. Plot of estimated ETo at New Amsterdam from 2019-2022 using PM-56

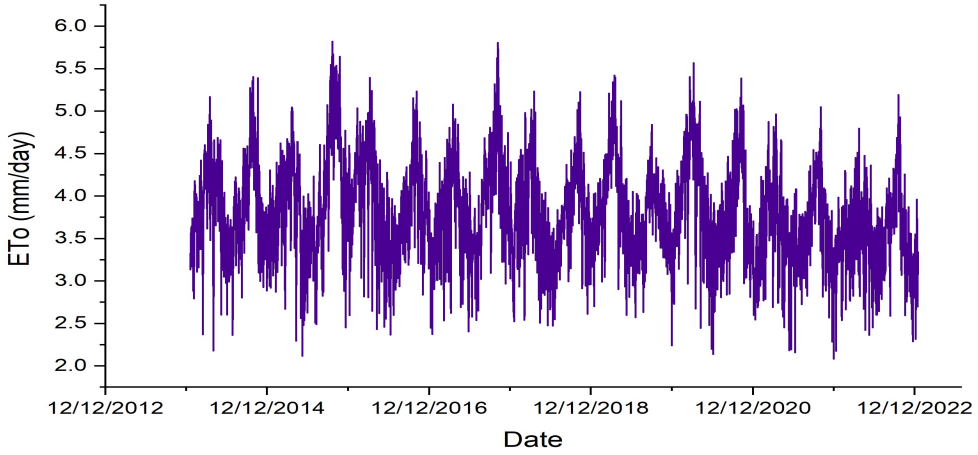


Fig. 4.3. Plot of estimated ETo at Ebini from 2013-2022 using PM-56

4.5.2 Comparison of empirical methods for ETo estimation

In New Amsterdam, the best-performing empirical method, based on RMSE, MAE, R^2, and regression coefficients A and B, was HS3. HS3 exhibited a high correlation

with PM-56 and demonstrated lower MAE and RMSE values compared to all other empirical methods. Moreover, it displayed R^2 values that were higher than almost all others. Additionally, values of A close to 1 and B close to zero were observed (see Appendix C). Overall, HS3 outperformed every other method in at least 3 out of 5 metrics. Daily ETo values computed by this method ranged from 2.18 mm day^{-1} to 5.78 mm day^{-1} (see Appendix C).

In Ebini, HS2 emerged as the best-performing empirical method, outperforming the others in 3 out of 5 metrics. Its RMSE and MAE were 0.3844 mm day^{-1} and 0.3289 mm day^{-1}, respectively, lower than those of all other methods tested (see Appendix C).

4.5.3 Comparison between ANN and empirical methods

The performance of the 12 empirical methods was compared to that of the ANN models in both study locations. Results indicate that the ANN model outperformed the empirical methods in 4 out of 5 performance metrics in New Amsterdam and 5 out of 5 performance metrics in Ebini. In New Amsterdam, it exhibited RMSE and MAE values of 0.1373 mm day^{-1} and 0.123 mm day^{-1}, respectively. The ANN model also demonstrated a strong correlation and a high coefficient of determination value of 98.76%. Regression coefficient values A and B were 0.99 and 0.161, respectively. In Ebini, it showcased RMSE, MAE, and R^2 of 0.1017 mm day^{-1}, 0.0783 mm day^{-1}, and 97.11%, respectively, with regression parameter values for A and B of 0.989 and

0.044 (see Appendix C). In addition to the strong statistical performance of the ANN models, time series plots reveal that they align more closely with PM-56 and exhibit similar trends and seasonal fluctuations.

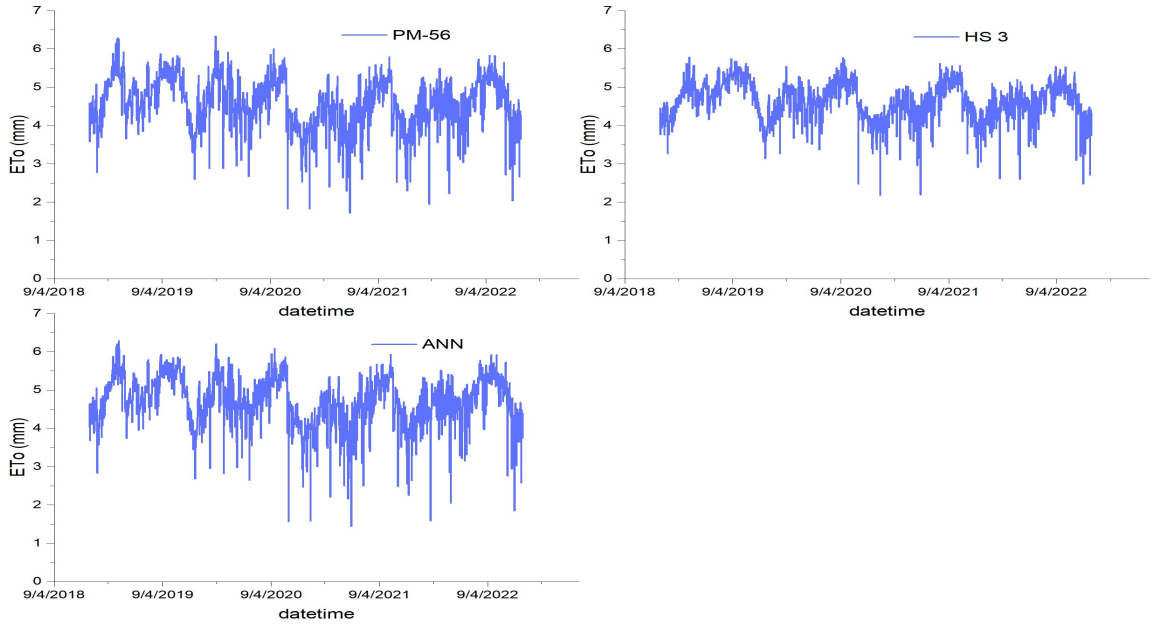


Fig. 4.4. Time series plot of ANN, HS3 and PM-56 methods at New Amsterdam

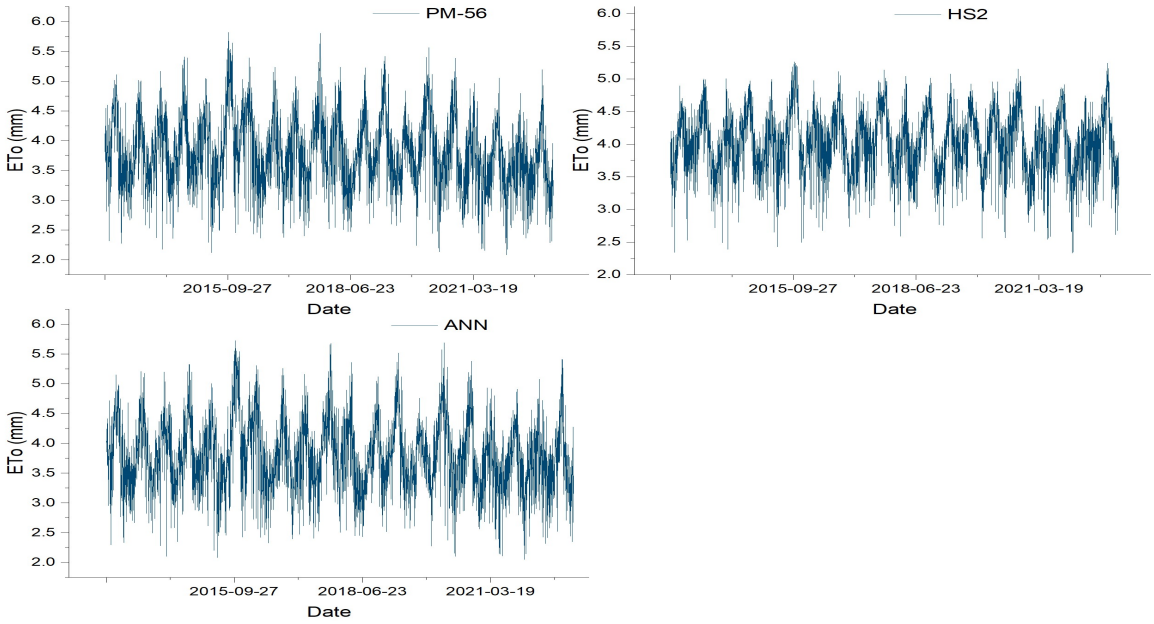


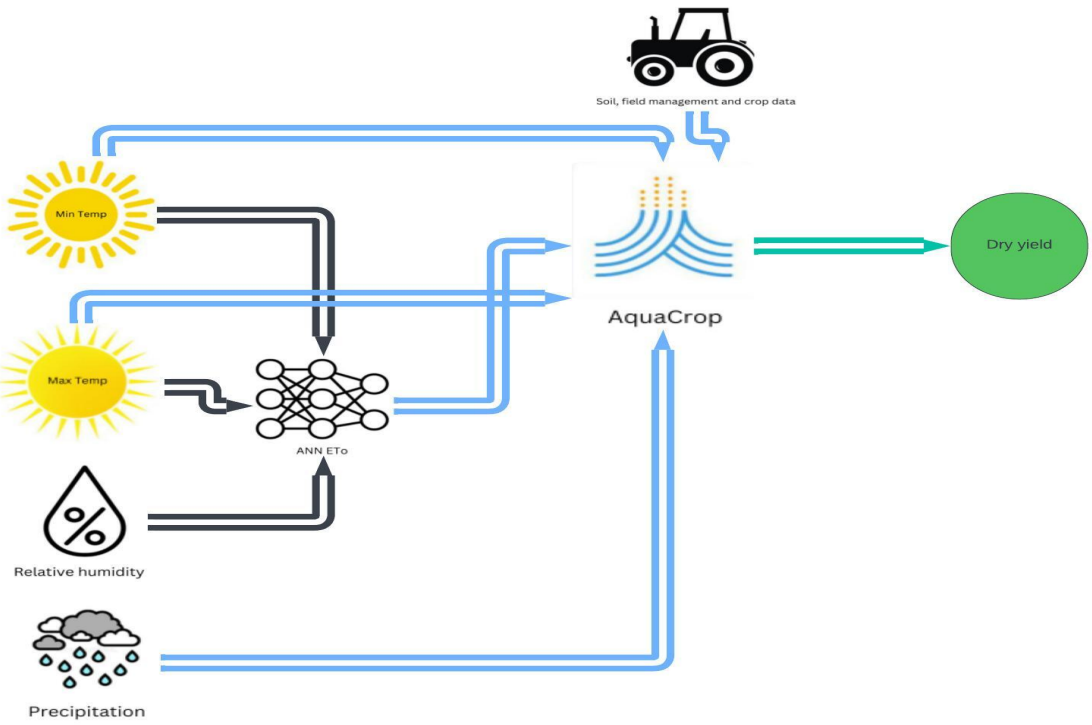
Fig. 4.5. Time series plot of ANN, HS2 and PM-56 methods at Ebini

4.6 Conclusion

This study undertook a comparison of the performance of 14 empirical methods for estimating reference evapotranspiration to that of an ANN. Both the empirical methods and the ANN were calibrated against the Penman-Monteith equation. The ANN model, which used minimum and maximum temperatures, relative humidity, and solar radiation as inputs, was trained on data spanning from 2001 to 2018 in New Amsterdam and from 1979 to 2012 in Ebini. Following training, it underwent testing using data from 2019 to 2022 and 2013 to 2022 in New Amsterdam and Ebini, respectively. The statistical analysis revealed that the ANN model outperformed all other tested empirical methods across both locations in four out of five metrics in New Amsterdam and all five metrics in Ebini. These findings underscore the applicability and effectiveness of employing an ANN to estimate reference evapotranspiration in New Amsterdam and Ebini and possibly other parts of Guyana.

Bridging text

AquaCrop is a powerful tool for estimating crop yields under various soil types, field management practices, and climatic conditions. In the preceding chapters, precipitation, temperature, and ETo, all important climatic variables for yield estimation, were modelled. Using the data from these models to create a climate file in AquaCrop, the upcoming chapter centers on estimating crop yields using observed and simulated weather data. AquaCrop's climate file database necessitates minimum and maximum temperatures, precipitation, and ETo at daily, 10-day, or 30-day intervals. For this research, daily time steps were used. Information regarding crops, field management practices, and soil types was sourced from the literature.



Chapter 5

Simulating yields using AquaCrop

5.1 Abstract

This study used and fine-tuned the FAO's AquaCrop model, using both actual and simulated weather data. The objective was to determine the yields of paddy rice, maize, and soybeans across different soil types and irrigation methods (for rice) using simulated and observed weather data. Yield simulations used the May to June and October to November planting seasons from 2014 to 2025 for these crops. To evaluate the AquaCrop model's performance and the effectiveness of using simulated weather data to simulate crop yields, various metrics such as Root Mean Square Error (RSME), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) were employed. Due to the absence of historical yield data, validation was only conducted for irrigated rice simulated at the Rose Hall estate, Model calibration

resulted in RSME, MAE, and MAPE values of 0.642 tons,0.478 tons, and 7.518% for rice using observed weather data and 0.618 tons,0.456 tons, and 7.237% for simulated weather data. Validation yielded values of 0.32 tons,0.282 tons, and 4.35% for RSME, MAE, and MAPE for rice using observed weather data and 0.312 tons,0.29 tons, and 4.39% for simulated weather data. Yield simulations for rainfed rice, soy, and maize demonstrated favourable to acceptable results when comparing observed and simulated weather conditions.

The study's findings highlight AquaCrop's effectiveness in simulating rice, corn, and soybean yields across coastal and inland in Guyana. It also underscores the effectiveness of using simulated weather data to estimate crop yields.

5.2 Introduction

Estimating potential and attainable crop yields across varying soil types and climates is important for agricultural planning, decision-making, food security, and sustainable agriculture. Tackling this challenge necessitates understanding how a crop's yield reacts to water availability [128] and various soil types. Traditionally, the link between agricultural output and climatic conditions, water supply, and soil types has relied on empirical production functions [36]. However, these functions may only be reliably applicable within the specific locations in which they were developed [45] In response to this limitation, calibrated crop growth simulation models are progressively supplanting empirical methods to evaluate crop yields across diverse environmental conditions

and agricultural practices [45]. One such crop simulation model is AquaCrop, which was designed by the Food and Agricultural Organization's Land and Water Division (FAO) of the United Nations. This model blends simplicity and robustness into a simulator that requires fewer inputs than other simulators, like DSSAT, WOFOST, and APSIM. Once properly calibrated, the AquaCrop simulator can accurately simulate how field crops (such as cereals, roots, tubers, legumes), vegetables, and tree crops respond to water availability and is well suited to locations where water is a key limiting factor in crop production [1, 107]. AquaCrop's accuracy in modelling crop responses across diverse conditions enables customized agricultural planning. Its accurate yield predictions support farmers, policymakers, and researchers in making informed decisions to enhance agricultural productivity and sustainability [1, 107].

In Guyana, where the government aims to bolster rice production and increase soybean and corn production, AquaCrop can hold significant potential in providing local policymakers and farmers with a tool for estimating past and projected crop yields of these and other crops in varied soil types and climatic conditions across the country.

With the literature, AquaCrop has primarily been used for water management and crop productivity analysis, typically relying on observed weather data. However, the incorporation of simulated weather data into AquaCrop is seldom explored despite its potential benefits, such as yield projections and scenario analysis. As a result, this research aims to evaluate the utility of assimilating simulated weather data into

AquaCrop to estimate crop yields in the Rose Hall estate farmlands and Ebini. The simulations in AquaCrop are done for rice, corn, and soybean (yields).

5.3 Literature review

AquaCrop is a crop water productivity model developed by the Food and Agriculture Organization (FAO) to simulate crop yield, water requirements, and water use efficiency under water-limited conditions [128]. The model blends simplicity, precision, and robustness by evolving from [36] conceptual framework regarding agricultural production response to water. Extensively cited in the literature, AquaCrop finds application primarily in two areas: water management and crop productivity analysis.

5.3.1 Water management

AquaCrop is a tool often used for optimizing crop irrigation strategies and managing water resources. [103] identified optimal irrigation strategies for cabbage, showcasing significant water savings with minimal yield reduction. [111] examined AquaCrop's efficacy in modelling rice growth, acknowledging its accuracy while noting the model's tendencies to overestimate early-season dry matter. Similarly, [155] evaluated AquaCrop's performance in simulating rice yields in paddy soil, noting acceptable accuracy in simulating canopy cover, biomass and yield, but faced challenges in simulating soil moisture content during drying-wetting cycles, suggesting refinements are needed. Further rice simulation studies included [106], who explored

AquaCrop's utility in assessing rice irrigation requirements concerning climatic variations, demonstrating its ability to correlate irrigation needs with the reconnaissance drought index for improved water allocation strategies based on seasonal forecasts. [33] calibrated AquaCrop for canola under varied irrigation levels, showcasing its effectiveness in simulating biomass, canopy cover, and the yield response to differing water regimes. Further studies by [81], [53], [14], [159] and [99] expanded AquaCrop's applications to cotton, wheat, and maize showcasing its adaptability in predicting yield, optimizing irrigation, and improving water use efficiency under diverse conditions and irrigation strategies. In addition to the studies by [81], [53], [14], [159] and [99], studies by [122] and [147] also looked at AquaCrop's application to wheat and maize. [122] investigated the performance of the AquaCrop model in evaluating winter wheat and summer maize under various irrigation scenarios in the North China Plain (NCP). They found that under full irrigation, the model accurately predicted grain yield and biomass for both crops, with deviations below $\pm 11.00\%$. However, under water-saving irrigation, the accuracy for winter wheat decreased, emphasizing the need for caution when simulating water-stressed conditions. Meanwhile, [147] evaluated the AquaCrop model's efficacy for winter wheat under different irrigation strategies on the NCP (The North China Plain). Their study highlighted the importance of irrigation scheduling, with optimum results observed at 60% field capacity. Recent investigations by [156] and [145] further expanded AquaCrop's application to potatoes, highlighting its potential in optimizing irrigation schemes, forecasting

yield under water scarcity, and offering insights for future food security in varying agro-climatic zones.

5.3.2 Crop productivity analysis

Research using AquaCrop often covers applications such as assessing and predicting crop yields, canopy cover, and biomass under varying soil types, fertilization applications and environmental conditions. [91] and [114] used AquaCrop to model rice yield responses to soil salinity, showing decent accuracy in simulating rice productivity under saline water levels. studies like [8] assessed AquaCrop's performance in simulating rainfed maize yields under varying nitrogen fertilizer levels in Nigeria. Their findings indicated moderate to very good agreement between simulated and actual maize yields, with some overestimation suggesting the need for further calibration. [154] proposed a joint simulation-optimization framework for irrigation and fertilizer scheduling, integrating the improved AquaCrop model with multi-objective programming. Their study demonstrated that the model accurately simulated evapotranspiration under nitrogen stress, enabling precise irrigation and fertilization management to maximize the efficiency of maize yield and water use. Similarly, [110] assessed AquaCrop's performance in simulating maize response to nitrogen stress in a semi-arid environment, affirming its utility for managing maize production under varying nitrogen levels. Conversely, [61] investigated the integrated strategies of deficit irrigation, nitrogen fertilization, and biochar addition for winter wheat using AquaCrop,

highlighting its capability to simulate soil water content, canopy cover, biomass, and grain yield. However, they noted limitations in simulating certain aspects, particularly under biochar-amended treatments and drought conditions, suggesting areas for further model refinement. While some studies emphasized AquaCrop's precision in simulating specific crops like lettuce [15] or wheat under stress conditions [90](Mohammadi et al., 2017), others projected AquaCrop's potential in predicting alfalfa yields [107] and managing rainfed maize cultivation in changing climates [138].

5.4 Materials and methods

5.4.1 Study locations

The research locations selected for this study are the Rose Hall Sugar Estate farmlands and Ebini in Guyana.

5.4.2 Model description

AquaCrop simulates crop growth and yields based on atmospheric conditions (rainfall, temperature, reference evapotranspiration and carbon dioxide concentration), crop characteristics, soil composition, and management practices. This multi-crop and water model aims to balance robustness, simplicity, and accuracy for broad use. It only needs a relatively limited number of explicit parameter values and mostly intuitive input variables, which may be obtained using straightforward methods [128],

unlike other simulation models, which require a lot of data.

Fig. 5.1. AquaCrop simulator user interface

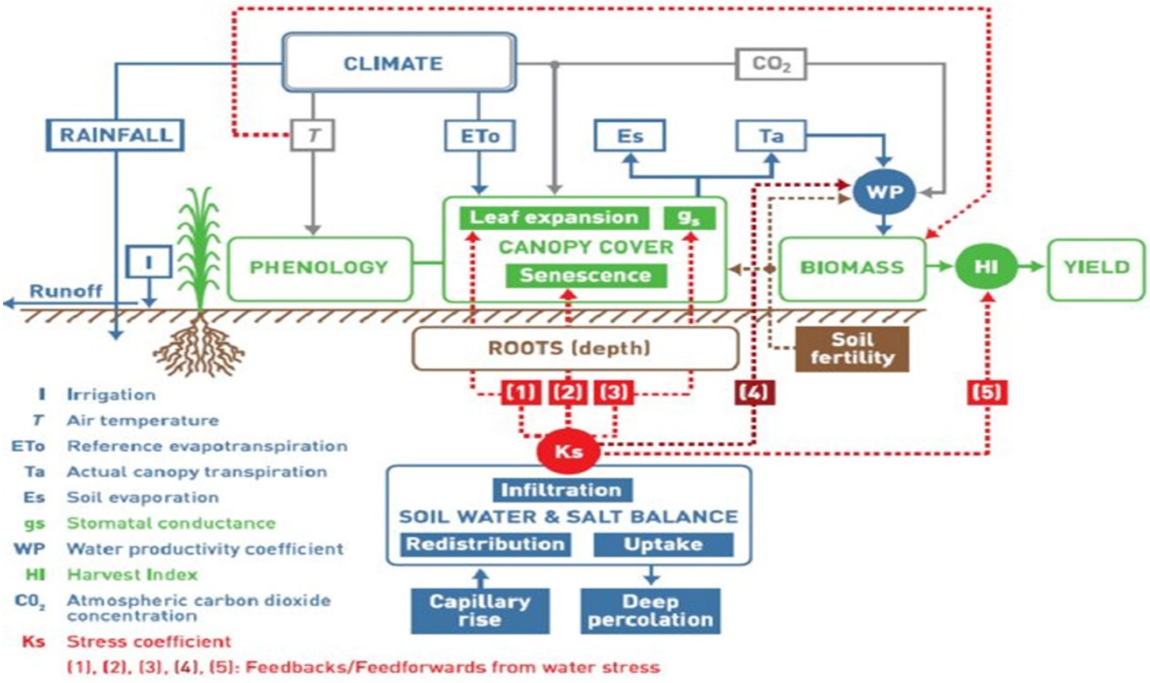


Fig. 5.2. Schematic diagram of AquaCrop.

Source: [1]

5.4.3 Climatic data

The climate component of AquaCrop utilizes five weather input variables: rainfall, maximum and minimum air temperatures, reference crop evapotranspiration (ETo), and mean annual atmospheric carbon dioxide concentration (CO2) to define the crop's atmospheric conditions [128]. Except for CO2 concentration, these parameters can be integrated at daily, ten-day, or monthly time intervals. Temperature influences crop phenology, while ETo and rainfall determine the soil root zone water balance. Additionally, CO2 affects leaf growth and water productivity (WP) [128]. The observed

climatic data used in this study were obtained from The New Amsterdam weather station and Climate Engine. Given the Rose Hall estates' proximity to New Amsterdam, the temperature and ETo data collected from the weather station were used to approximate conditions over the Rose Hall estate during the study period. The data obtained from Climate Engine were used as the observed weather conditions over Ebini. Simulated data for temperature, precipitation, and ETo were obtained from the weather models. In particular, simulated ETo was derived from an ANN that used simulated temperature, simulated relative humidity, and simulated solar radiation as inputs. The ANN model was trained and calibrated against the Penman-Monteith equation. Observed weather data ranged from 2014 to 2022, while the simulated data covered the period from 2014 to 2025.

5.4.4 AquaCrop yield development

The normalized biomass water productivity (WP*) and the ratio of daily crop transpiration (Tr) over the reference evapotranspiration for that day (ETo) is used to calculate the daily (m) and cumulative (B) dry aboveground biomass production

$$\begin{aligned} m &= WP^* \left(\frac{Tr}{ET_o}\right), \\ B &= WP^* \sum \left(\frac{Tr}{ET_o}\right), \end{aligned} \tag{5.1}$$

$$Y_{dry} = HI \cdot B,$$

where B is the total dry above-ground biomass generated at crop maturity (Eq.2), HI is the yield percentage of B, and Ydry is the dry yield. When water and/or temperature stress occurs during the crop cycle, the Harvest Index is adjusted to the stresses at run time for fruit/grain-producing crops as well as roots and tuber crops and may deviate from the reference harvest index (HI_o) [107, 108].

5.4.5 Soil data

Distinct soil types and characteristics define the local soil profiles across the research sites of Rose Hall estate and Ebini. The soil type at Rose Hall estate Estate corresponds to the Whittaker Series, 37, as documented by [129]. This type covers approximately 45,800 acres (6.4% of the total area) and spans the coast from Rose Hall Sugar Estate in the northwest to number 51 Village in Leeds in the southeast. Detailed characteristics of the Whittaker soil are outlined in Table 5.1. Conversely, the soil in the Ebini area is classified as Ebini sandy loam characterized by well-drained properties and featuring a dark grey loamy sand layer over light yellowish-brown sandy soil, as described by ESDC. Due to the unavailability of precise soil characteristic breakdowns for Ebini, the default sandy loam soil type from AquaCrop was used to conduct simulations over this region.

It should be noted that due to insufficient data and time limitations, we did not model or account for the future changes in soil nutrient dynamics, as well as the potential impact that mechanization can have on soil compaction and its effects on

yield.

Table 5.1. Measured soil properties of Whittaker 37 soil series (n= 12 samples)

Layer	Depth (cm)	Particle Size Distribution (%)			USDA Texture	Bulk density (gcm^3)	Ksat (mm/day)
		Clay	Silt	Sand			
1	0-22	63	35	2	Clay	1.23	237.6
2	22-65	65	32	3	Clay	1.29	129.6
3	65-100	66	32	2	Clay	0.94	79.2
4	100-141	45	45	10	Silty-Clay	1.11	984

Source: [47]

5.4.6 Irrigation sources and methods

The lands in Rose Hall estate Estate rely on the Canje River as their primary irrigation source, ensuring an adequate water supply for agricultural activities. The irrigation method employed for simulating rice yields in Rose Hall estate was basin irrigation, with a schedule of 50mm every eight days. In contrast, the Ebini area predominantly practices rainfed agriculture due to its location, depending solely on natural rainfall for crop irrigation. Without proper irrigation sources in the Ebini area, all simulations conducted there will solely consider rainfed conditions. Additionally, it should be noted that simulations for soy and corn at the Rose Hall estate are also conducted under rainfed conditions.

5.4.7 Crops used in simulations

The GRDB 10, a local rice variety with a yield potential of 6.8-7 tons per hectare, will be used along with corn and soybeans for the simulation. As there is no available information on the varieties of corn and soybeans currently under mass cultivation in Guyana, information concerning these crops will be sourced from the literature. Specifically for simulation purposes, the open-field Pollinated Yellow corn variety and the MSOY 9144 variety will be used for Maize and Soybeans, respectively. The parameters employed for the AquaCrop simulations are detailed in Table 5.3

5.4.8 Calibration and validation

AquaCrop was calibrated and validated using weather data from 2014 to 2022. The calibration process relied on observed weather data from 2014 and 2015, while the validation phase used data from 2016 and 2022. However, due to the recent large-scale production of soybeans and corn in Guyana, relevant historical yield data for these crops during the simulation period was unavailable for validation. Consequently, validation was exclusively performed for rice simulations in the Rose Hall estate area. The calibration of AquaCrop focused on refining non-conservative parameters within the system. Its primary goal was to minimize key performance metrics, such as root mean square error, mean absolute error, and mean absolute percentage error between simulated and observed yields. Additionally, the calibration and validation process

expanded to specific crop cultivation locations. In the case of rice simulations at the Rose Hall estate, calibration used yield data obtained from the Blackbush Polder rice region (an area that is close to the Rose Hall estate), covering the May-June planting season from 2014 to 2022.

In the absence of corresponding historical yield data for corn and soy within this timeframe, calibration relied on research experiments conducted on these crops on clay soils by [143], [144], [124] and [94]. [144], focused their research on soybean yields across Guyana's clay (Inki and Whittaker 37 series) and peat soils, demonstrating that optimal conditions, including adequate drainage, fertilizer application, and effective crop and soil management, led to favourable yields ranging from 2 to 2.4 tons/ha and reaching as high as 4.9 tons/ha. Leveraging this valuable insight, AquaCrop underwent calibration in clay soils against these obtained yields. Additionally, research by [94] (investigating the short-term effects of soil property amendments on agronomic productivity on clay soils) and [124] (Performance of corn genotypes at coastal and savannah regions and cost of production of cultivation) showcased yields ranging from 5.6 to 10.5 tons/ha under optimal conditions, encompassing proper spacing, drainage, fertilizer application, and effective soil management. Using the results of the experiments by [124] and [94] as a guide, corn and maize were calibrated specifically on the Whittaker clay soils in the Rose Hall estate. This calibration subsequently served as a foundational framework for simulation across the sandy loam soil type found in Ebini.

Given that rice is not conventionally cultivated in Ebini, yield data for calibrating against this area's sandy, loamy soil type was unavailable. No local data was sourced either. To address this challenge, research by [38] was used to aid calibration. Their findings revealed a 46% average yield decline for rice on Sandy Loamy soils compared to clay soils, regardless of irrigation strategies. Calibration of Rice simulation over Ebini was conducted based on this insight.

5.4.9 Planting windows and intervals

Table 5.2. Planting windows

Crop	Planting -window	location	# of intervals
Rice (irrigated)	May15th-Jun30th	Rose Hall	5
Rice(rainfed)	May 15th-Jun15th	Rose Hall	3
Soy	May30th-Jun30th	Rose Hall	3
Corn	Oct15th - Nov30th	Rose Hall	5
Rice	May 15th-June 15th	Ebini	3
Soy	May 15th-June 30th	Ebini	5
Corn	May 15th-June 30th	Ebini	5

Table 5.3. Planting intervals

Planting intervals	Planting dates in interval
May15th-May30th	M15, M23, M30
May23rd-Jun7th	M23, M30, J7
May30th-Jun15th	M30, J7, J15
Jun7th-Jun23rd	J7, J15, J23
Jun15th-Jun30th	J15, J23, J30
Oct15th-Oct30th	O15, O23, O30
Oct23th-Nov7th	O23, O30, O7
Oct30th-Nov15th	O30, N7, N15
Nov7th-Nov23rd	N7, N15, N23
Nov15th-Nov30th	N15, N23, N30

Tables 5.2 and 5.3 display the planting windows, planting intervals, and corresponding planting dates for the crops used in the simulations. As the time during which these crops are planted can affect yields, the planting windows were subdivided into smaller intervals spanning 14–16 days. These smaller windows also facilitate better comparisons of yields obtained under observed and simulated weather conditions. To obtain the yields over the smaller intervals, we divide the yields obtained from the planting dates in each interval by three. For example, the average for May15-May30th interval would be (M15 + M23 + M30)/3.

5.4.10 Performance metrics

To assess the performance of AquaCrop's yield simulations, the following three performance metrics will be used: Root mean square error (RSME), Mean absolute percentage error (MAPE), and Mean absolute error (MAE).

$$\begin{aligned} \text{RMSE} &= \sqrt{\frac{1}{\mathrm{n}} \sum_{i=1}^{\mathrm{n}} (\mathbf{S}_i - \mathbf{M}_i)^2}, \\ \text{MAE} \ &= \frac{1}{\mathbf{n}} \sum_{i=1}^{n} \left| \mathbf{S}_i - \mathbf{M}_i \right|, \\ \text{MAPE} \ &= \frac{1}{n} \sum_{i=1}^{n} \left| \frac{S_i - M_i}{M_i} \right| \times 100. \end{aligned} \tag{5.2}$$

5.5 Results and discussion

5.5.1 Observed and simulated climatic conditions

The monthly minimum and maximum temperatures observed in Ebini and Rose Hall estate remained relatively stable. In Ebini, temperatures ranged from 29.49°C to 32.11°C for maximum and 23.01°C to 23.52°C for minimum throughout the growing period (May-October). Over in Rose Hall estate, slightly higher temperatures, ranging from 30.73°C to 33.28°C for maximum and 24.41°C to 24.74°C for minimum temperatures were seen. The observed average monthly precipitation during the growing season in Ebini was slightly higher than that in the Rose Hall estate area,

averaging 218.12 mm compared to 204.147 mm. Despite the slightly higher rainfall in Ebini, its daily evapotranspiration rate was lower than Rose Hall, averaging about 3.75 mm compared to 4.69 mm.

The simulated monthly minimum and maximum temperatures in Ebini and Rose Hall estate also remained relatively stable. In Ebini, maximum temperatures ranged from 30.01°C to 30.6°C, and minimum temperatures ranged from 22.6°C to 22.82°C. In Rose Hall estate, simulated temperatures were somewhat higher, ranging from 31.17°C to 31.6°C for maximum and 23.49°C to 23.74°C for minimum temperatures. The simulated average monthly precipitation in Ebini was slightly higher than at the Rose Hall estate, averaging 218.82 mm compared to 206 mm. In Ebini, the daily evapotranspiration rate was lower than in Rose Hall, averaging about 3.81 mm compared to 4.61 mm.

5.5.2 Calibration and validation

The parameters used in the AquaCrop simulations after calibration are outlined in Appendix A. Some of the data used for calibration was sourced from the literature [95, 119, 123] and a local farmer. After calibration, each crop was found to have a minimum of 90% maximum canopy cover, with a harvest index of at least 36%. Following calibration, using both observed and simulated weather, the simulated yield for rice in Rose Hall demonstrated a close match with the actual yield under irrigated conditions. In both instances, the Root Mean Square Error (RMSE) was less than

0.65, and the Mean Absolute Percentage Error (MAPE) was under 8% for both weather conditions. Average yearly yields under observed and simulated conditions were found by finding the average yield across all planting dating in the planting window.

Table 5.4. Performance metrics of yield simulations after calibration for May 15th -June 30th window

Year	Yield (ton/ha)				RMSE (ton/ha)	MAE	MAPE %
	2014	2015	2016	2017			
Actual Yield	6.1	6	5.2	5.9			
Obs weather	6.11	6.22	6.35	6.45	0.642	0.478	7.518
Sim weather	6.27	6.27	6.38	6.445	0.618	0.456	7.237

Following calibration, validation occurred using simulated and observed weather data from 2018 to 2022. The results strongly suggest that the calibrated parameters are acceptable and can be used for simulating rice yields. Additionally, there was close agreement between yield simulations based on observed weather patterns and those from simulated climatic conditions. This indicates the viability of using simulated weather data to predict rice yields across the Rose Hall estate. Since no historical data was available for the yields of rainfed rice, corn and soybeans, validation for these crops could not be done.

Table 5.5. Performance metrics of rice yield simulations after validation

	Yield (ton/ha)					RMSE	MAE	MAPE
Year	2018	2019	2020	2021	2022			
Actual Yield	6.55	6.8	6.1	6	6.3			
Obs weather	6.45	6.42	6.39	6.55	6.59	0.32	0.282	4.35
Sim weather	6.35	6.49	6.52	6.55	6.53	0.312	0.29	4.49

5.5.3 Yields obtained under observed weather data

After calibrating AquaCrop, the average yields over the corresponding planting intervals for irrigated rice, rainfed rice, soybeans, and maize were simulated in at the Rose Hall estate. These yield simulations assume good agricultural practices such as adequate drainage, fertilizer application, and effective crop and soil management are done. After simulations, crop yields ranged from 5.872 to 6.59 (irrigated rice), 5.161 to 6.649 (rainfed rice), 2.123 to 3.33 (soy), and 7.71 to 10.019 (corn) tons/ha (see Appendix D). In Ebini, the yields for corn and soybeans were higher than those in Rose Hall and ranged from 9.784 to 9.909 (corn) and 2.74 to 3.785 (soy) tons/ha (see Appendix D). The simulated yields over this area from 2014-2022 were notably higher than at the Rose Hall estate. This increase can be attributed to the sandy, loamy soil texture in the Ebini location, a soil type much more suited for corn and soybeans. However, rice yields in this location were lower than in the other study locations, ranging from 3.367 to 4.096 tons/ha (see Appendix D). This lower rice yield

is attributed to the soil type.

5.5.4 Performance assessment of yields obtained under simulated weather data

At the Rose Hall estate, from 2014-2022, the yields obtained for rice (both irrigated and rainfed) and soy were slightly higher than those obtained under observed weather conditions, ranging from 5.87 to 6.59 tons/ha for irrigated rice, 5.162 to 6.649 tons/ha for rainfed rice, and 2.132 to 3.328 tons/ha for soy. However, corn saw slightly lower yields under simulated weather conditions, ranging from 7.53-10.07 tons/ha (see Appendix D). When examining the performance metrics, the yields obtained via simulated weather proved to be good estimates of the yields of both irrigated and rainfed rice and corn under observed weather conditions. Further inspection of the performance metrics shows that rice and corn had MAPE values below 10%, with irrigated rice performing the best under simulated conditions with MAPE values ranging from 0.97% to 2.34%. RMSE values for irrigated rice were also low, ranging from 0.08 tons/ha to 0.21 tons/ha. Rainfed rice, which was the second best-performing crop under simulation at the Rose Hall estate, saw yields comparable to that of irrigated rice. This parity was also seen under observed weather conditions (see Appendix D). Soybeans, on the other hand, performed moderately well, with MAPE values ranging from 9.99% to 13.320% and RMSE ranging from 0.366 tons/ha to 0.421 tons/ha (see Appendix D). In Ebini, higher yields for corn and soybeans were observed, similar to

those simulated under observed weather conditions. Better performance metrics for each crop were observed in this area as well. This can be attributed to the soil type favouring corn and soybean growth and the increased rainfall under both simulated and observed weather conditions compared to Rose Hall. Simulated yields for each crop remained stable over the years, with soybeans experiencing the most significant deviations, ranging from 2.74 to 3.79 tons/ha. Soybeans also saw the largest MAPE value at 7.7% (see Appendix D). This yield stability can be attributed to the observed and simulated climatic conditions used and the lack of detailed soil data. The overall good results highlight the utility of using simulated data in AquaCrop to estimate crop yields.

5.5.5 Yield projections

Simulated weather data was used to project crop yields across the intervals of 2023–2025 at both study locations. Results suggest that with the maintenance of good agricultural practices, yields for irrigated rice and rainfed rice at the Rose Hall estate should remain relatively stable. Throughout this timeframe, the average yields for rainfed irrigated rice varied between 6.62 tons/ha (May 15th-May 30th, 2024) and 6.08 tons/ha (May 30th-Jun 15th, 2025). Meanwhile, yields for irrigated rice ranged from 6.51 tons/ha (June 7th-June 22nd, 2025) to 6.64 tons/ha (May 15th-May 30th, 2025) (see Appendix D). Corn and soybeans exhibited minor fluctuations in average yearly yields, with 2025 witnessing the highest soybean yields, while 2024/25 saw peak corn

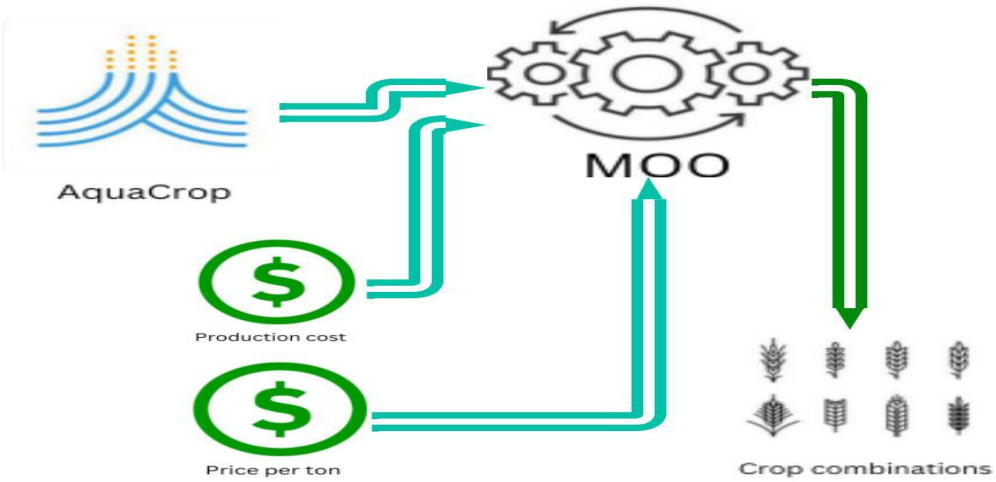
yields. In Ebini, corn yields remained relatively stable from year to year. However, soybean and rice yields experienced reductions in 2025. (see Appendix D).

5.6 Conclusion

AquaCrop underwent calibration and subsequent validation to evaluate its performance in simulating rice yields under irrigated conditions at the Rose Hall estate, using both observed and simulated weather data. The results demonstrated a strong agreement between the yields simulated under both observed and simulated weather conditions, showcasing the effectiveness of AquaCrop post-calibration in simulating rice yields. It also highlighted the efficacy of using simulated weather data for estimating crop yields. As historical yield data for rainfed rice, soy, and corn were unavailable, no validation was conducted for these crops in the study locations. To assess the effectiveness of using simulation data in AquaCrop for estimating yields of these crops, yields under both simulated and observed weather conditions were generated and compared using various performance metrics. The overall performance metrics showed a good statistical agreement between the yields obtained under observed and simulated weather. However, increased accuracy could be achieved with a more comprehensive soil analysis for each study region, as well as detailed field management practices for corn and soybean.

Bridging text

The preceding four chapters modelled and simulated climatic variables and yield development. Crop yields, which are dependent upon climatic conditions, are pivotal for agricultural planning. When these yields are combined with dynamic production cost and pricing data, an optimization model can be created to aid farmers and policymakers in answering the questions of where, when, what, and how much to plant. Consequently, the upcoming chapter focuses on developing a multi-objective optimization model (MOO) that integrates the yield data from Chapter 5. In addition to the yield data, the model integrates time series pricing data and dynamic production cost data to create a more realistic crop selection optimization model.



Chapter 6

Crop interval mix planning model

6.1 Abstract

The decision of what to plant, how much to plant, where, and when to plant poses challenges for many farmers and agricultural producers. Given that some crops have wider planting periods, determining the optimal planting intervals is crucial as it can lead to improved economic benefits and better utilization of water resources. Consequently, this paper introduces a constrained multi-objective optimization model that utilizes various planting interval combinations for mixed crop planning. The central objective of this work is to maximize gross profit, ET economic water productivity, and minimize land usage to determine the optimal crop and planting interval combinations. Simulation experiments were conducted using the Non-dominated Sorting Genetic Algorithm II (NSGA II). The results indicated that rice, whether irrigated

or rainfed, double-cropped with corn, yielded the highest projected economic value at the Rose Hall estate for 2024/25, under both 50% and 25% export scenarios, with profits ranging from $\$4,216,054$ to $\$4,788,175$ and economic water productivity ranging from \$0.595 per m^3 to \$0.616 per m^3 for the double cropping combination that included rainfed rice. These projected values were achieved when rice was planted between May 15th and May 30th, 2024, and between June 15th and June 30th, 2024. In Ebini, where polyculture is practiced, corn and soybeans poly-cropped from May 30th - June 15th for soy and June 15th - June 30th for corn consistently yielded the highest projected gross profit, with rice cropped with corn performing the worst. This was observed under both 25% and 50% export scenarios. Profits for this combination ranged from $\$1,940,002$ to $\$2,392,755$ with a land use of 1793 ha for soy and 707 ha for corn. The results underscore the utility of the model and its potential to be used as a tool in agricultural planning and decision-making for food security and sustainable agriculture.

6.2 Introduction

The agricultural sector, particularly crop production, is essential since it sustains the lives of many people and animals on the planet. Since producing enough food to cater to the ever-increasing global demand is vital for food security and sustainable agriculture, efficient planning is important. Within the literature, much of the crop production planning research typically focuses on crop rotation or mixed/ploy-

cropping methodologies, which are used to maintain continuous production across planting areas [3]. The practice of mixed cropping depends on several factors, such as weather conditions, labour, production cost, yield rates, land availability, fertilizer, and farming preferences. It is a common approach farmers use to ensure food security and sustainable agriculture. Since many mixed cropping models found in the literature focus on determining the optimal mix of crops to plant to satisfy one or several objectives under certain constraints, they inherently answer the question of what to plant, which is one of the central aims of agricultural planning. However, the literature on crop production models has little, if any, work that addresses the question of what to plant, when to plant and where to plant. The success of a "crop interval mix" planting technique, which answers the question of what to plant among those considered and when to plant, depends on models that can manage the variables and parameters important to its production [5, 115]. Since agricultural planning often requires the simultaneous satisfaction of several objectives, choosing which objectives to satisfy is essential.

Globally, agriculture is the primary consumer of water and is responsible for around 70% of global freshwater withdrawals [101]. Given this importance, an objective that focuses on some form of water use efficiency would be crucial for a crop interval mix planting model. Profit optimization is another critical aspect of agricultural planning, sustainable agriculture and food security. Profit is essential for the agricultural industry as it ensures the economic viability of production and drives in-

novation and investment. These investments further improve productivity, efficiency, and overall quality. Furthermore, profit incentivizes farmers and policymakers to adopt sustainable practices and technologies that reduce the environmental impact of crop production. Given its importance, obtaining optimal profits should be a key objective of the crop interval mix model. Finally, maximizing productivity while conserving land use is crucial for sustainable agriculture as it ensures a reliable food supply, mitigates environmental degradation and promotes resilience in the agricultural sector. Given this, minimizing land use should be another important objective the model satisfies.

Given the above objectives, this study focuses on developing a constrained multi-objective optimization model using simulated crop yield and time-series pricing data to determine the optimal planting period of considered crops and crop combinations. The objectives of this model are to maximize gross profit, maximize ET economic water productivity, and minimize land use.

6.3 Literature review

Several works in the literature have addressed the use of multi-objective optimization in agricultural production and crop planning issues. These optimization approaches vary primarily in the types of objectives and constraints used, as well as their areas of focus. Among the pursued objectives, profit optimization is one of the most commonly targeted goals. Considering the widespread use of profit optimization in the literature,

the review was categorized into two sections focusing on land use and water usage objectives.

6.3.1 Water usage

Several studies have delved into multi-objective optimization techniques to tackle the intricate challenges of water management in agriculture. For instance, [162] introduced an Interactive Fuzzy Multi-Objective Optimization (IFMOO) approach for crop planning and water resource allocation, demonstrating its effectiveness in addressing vagueness and imprecision in decision-making processes. By applying IFMOO, they optimized crop patterns for maximum net returns, expected grain yield, and environmental returns and showcased its utility in water resources management through trade-off procedures. [42] investigated integration among Egyptian governorates to optimize irrigation water use, emphasizing the benefits of multi-objective optimization (MOO) in achieving optimal water utilization. Their study illustrated that integrated cropping patterns, guided by MOO, saved water and reduced cultivated area, lessened fertilizer usage, and increased net returns from crop production. [148] explored MOO of water and fertilizer management for potato production, identifying optimal combinations to maximize yield, quality, and water/fertilizer productivity. Their results revealed significant effects of irrigation frequency, water amount, and fertilizer rate on potato growth and quality attributes. [76] proposed a multi-objective approach to optimize irrigation and fertilizer scheduling at the farm level, aiming to increase crop

yield while minimizing water and nitrogen usage and environmental impact. Their optimization model found irrigation and nitrogen schemes that reduced water usage by 48%, nitrogen usage by 26%, and nitrogen leaching by 51%.

6.3.2 Land usage

Several studies have investigated the impact of multi-objective optimization techniques on various aspects of land use planning related to agriculture. [109] investigated the impact of integrated watershed land use plans on forest fodder dependency in a Central Himalayan watershed. Through MOO, alternate land use plans were developed to optimize farm income, employment, and nutritional security and minimize soil loss. Similarly, [89] proposed a multi-objective linear fractional programming approach for land use planning in agricultural production systems, demonstrating its suitability in optimizing conflicting goals in agricultural production planning. [160] developed a Multi-objective Land Use Optimization Allocation (MOLUOA) model using a multi-agent system integrated with particle swarm optimization. Applied to land use optimization in Changsha, China, their model demonstrated the ability to achieve multi-objective land use optimization in terms of quantity, space, and time. [127] proposed a fuzzy multi-objective-based goal programming model for optimal land allocation in a canal command area, simultaneously addressing multiple socio-economic and environmental objectives. That same year, [49] introduced MAUSS, a multi-objective model for sustainable land use allocation, applied to the Plains of

San Juan, Puebla, Mexico. By maximizing income, minimizing negative environmental pressure, and reducing food deficit, MAUSS provided a more sustainable land use allocation compared to models at the time. [131] introduced CoMOLA, a tool for constrained multi-objective optimization of land use allocation, demonstrating its effectiveness in exploring feasible solution spaces for complex land use allocation problems. [82] developed a model for evaluating cropland use sustainability, integrating the ecological footprint model with multi-objective optimization. Their model facilitated the generation of sustainable cropland use patterns, promoting water use efficiency, economic benefit, and reduction of greenhouse gas emissions.

6.4 Materials and methods

The multi-objective optimization problem examined in this study has three objectives: maximization of gross profit, maximization of ET economic water productivity, and minimization of land usage. The crop yield data over varying planting intervals and their corresponding crop evapotranspiration in m^3 were obtained by integrating simulated weather and crop-specific data into the AquaCrop simulator. The selected locations for running simulations are the Rose Hall estate farmlands and Ebini, Guyana.

6.4.1 Model parameters and variables

Index and parameters

1. i: A crop which can be considered for production
2. j: A crop combination made up from i
3. l: A crop planting interval combination dependent upon j
4. n: The number of crops in a single combination determined by j
5. Y_{ijl}: The average yield in metric tons of crop i in crop combination j for crop planting interval combination l
6. P_{il} : Export price per metric ton of crop i in crop planting interval combination l
7. Q_{il} : Local price per metric ton of crop i in crop planting interval combination l
8. X_i : The percentage of crop i to be exported
9. PC_{ijl}: Cost to produce 1 hectare of crop i in crop combination j for crop planting interval combination l
10. $EPRT_{ijl}$: Export profit obtained from producing 1 hectare of crop i in crop combination j for crop planting interval combination l . $= (X_i \times P_{il} \times Y_{ijl} - PC_{ijl})$
11. $DPRT_{ijl}$: Domestic profit obtained from producing 1 hectare of crop i in crop combination j for crop planting interval combination $l = ((1 - X_i) \times P_{il} \times Y_{ijl} - PC_{ijl})$
12. ET_{ijl}: Average evapotranspiration in cubic meters per hectare of crop i, in crop combination j for crop planting interval combination l

13. D_i : Demand for crop i in metric tons of crop land

14. L: The amount of land in hectares

15. TI: Total investment

16. TD: Total demand

Decision variable

A_{ijl} : The area of land in hectares to be cultivated for crop i in crop combination j for crop planting interval combination l

6.4.2 Objectives

Maximize gross profit

Stakeholders, policymakers, and farmers grapple with questions about how much to produce when to produce, and which production methods to employ. However, the underlying foundation upon which these choices are made is rooted in profit maximization. Therefore, the decisions ultimately made are aimed at achieving maximum profits. Gross Profit, which can be seen as the difference between revenue and expenditure, can be expressed as follows.

$$\text{Profit } = \sum_{i=1}^{k} \text{EPRT}_{ijl} \cdot A_{ijl} + \sum_{i=1}^{k} \text{DPRT}_{ijl} \cdot A_{ijl}. \tag{6.1}$$

In this equation, $EPRT_{ijl}$ and $DPRT_{ijl}$ represent the export and domestic profits obtained from selling the commodities (milled rice, dried corn and soybeans). The

equations used to compute these values considered production cost, which was divided into fixed and variable costs;

$$\text{cost} = \text{variable cost} + \text{fixed cost}$$

The fixed and variable costs considered were

Variable cost: Seed, Inoculants, Fertilizer and manure, Herbicides, Insecticides, Other chemicals, Crop insurance, customs work, drying, irrigation, equipment repair, labour cost, miscellaneous, and interest.

Fixed costs: Land lease, and taxes.

In tables 6.1 and 6.2, a breakdown of the production cost for rice and corn can be seen.

Table 6.1. Rice production cost per hectare in coastal regions

Particulars	Expenses (GYD per ha)
Power tiller cost	$\$50,000$
Labor cost	$\$24,000$
Seed cost	4 bags / ha ($\$20,000$)
Fertilizer cost	$\$104,760$
Manure cost	\$0
Insecticides	$\$3,000$
Irrigation charge	\$80000
Land use cost	$\$48,000$

Table 6.2. Corn production cost per hectare in coastal regions in 2013

Inputs and operations per ha	Total expenses (GYD)
Ploughing, sowing, threshing etc	50000
Sowing, irrigation, spraying, weeding, harvesting etc	150000
Basal dressing (NPK) 200 kg	34000
Top dressing (Urea) 152kg	20064
Herbicide cost (Round-up @3 lit/ha)	6000
Insecticide cost (2 sprayings)	6000
Seed cost @ 25 kg/ha	2125
Land annual lease	1000
Total cost inclusive land cost	$\$269,188$

Maximize ET economic water productivity

Water productivity quantifies the yield obtained relative to the amount of water used by the crop. Therefore, its optimization focuses on achieving higher yields for each unit of water used by the crop. However, this dimension of water productivity can be broadened to include the economic value associated with water use. This form of water productivity is known as ET economic water productivity, and its maximization can be expressed mathematically as follows

$$\text{ET economic water productivity } = \sum_{i=1}^{n} \frac{Profit}{ET_{ijl} \cdot A_{ijl}}. \tag{6.2}$$

Optimal land use

Efficient land use is vital for sustainable agriculture as it can boost productivity while minimizing environmental impact. Optimal land allocation conserves resources, mitigates land degradation and ensures sufficient food production. Its minimization can be expressed as

$$\text{Land use} = \sum_{i=1}^{n} A_{ijl}. \tag{6.3}$$

6.4.3 Constraints

Investment constraint:

Total amount of money spent for production must be less than the amount available for investment.

$$\sum_{i=1}^{n} PC_{ijl} \cdot A_{ijl} \leq \text{TI}. \tag{6.4}$$

Land constraint:

The total land area used for the selected crops in each region must not exceed the available land

$$\sum_{i=1}^{n} W_k \cdot A_{ijl} \leq L_k, \forall k, \tag{6.5}$$

where $W_1 = 1$ for the polyculture scenario and $W_2 = 1/2$ for double cropping because two crops are planted on the same land.

Demand constraint:

The total crop yields must meet the economic demand.

$$\sum_{i=1}^{k} Y_{ijl} \cdot A_{ijl} \geq T. \tag{6.6}$$

6.4.4 Domestic and export prices

The export prices used in the simulation study were obtained from the World Bank commodity prices database. To forecast prices for 2024 to 2025, the R package auto. Arima was employed to derive ARIMA models for forecasting international rice, corn, and soy prices on Monthly data from January 1988 to March 2024. The auto. arima function in R applies a modified version of the Hyndman-Khandakar method, which incorporates unit root testing, AICc minimization, and maximum likelihood estimation to formulate an ARIMA model [63]. Following data assimilation, the optimal ARIMA models for forecasting were determined to be ARIMA (1,1,2), ARIMA (2,1,2), and ARIMA (2,1,3) for rice, soy, and corn, respectively. Price projections were made for 14 months, extending to May 2025.

Only two years' worth of domestic pricing data were available for each crop, which was derived from local market data (rice) and feasibility studies conducted by orga-

nizations such as NAREI for soy and corn (see Appendix E). A conversion rate of 204 GYD = 1 USD was consistently applied throughout the study. As domestic prices (like international prices) fluctuate over time, it's imperative to integrate fluctuating costs into the optimization model for a more realistic representation. Due to limited data, achieving dynamic domestic pricing resulted in an assumption of fixed price changes. This involved calculating the difference between prices and dividing the result by the number of months between data collection periods to achieve a fixed value. This fixed value was added to domestic prices on a monthly basis to create non-stationary price changes. To maintain consistency in how the dynamic pricing data for export and production costs are determined, all pricing data obtained are assumed to represent year-end prices (see Appendix E).

6.4.5 Production cost estimates

Local production cost data were only available for rice (four years) and corn (one year), respectively. To establish dynamic production cost data for rice over time, a similar approach to that used for domestic pricing was employed. Given that only one year of corn and no years' worth of soybean production data were available, Iowa State's historical cost of production data was used. To convert these prices to ones suitable for use in Guyana, the cost of producing 1 hectare of corn along Guyana's coast (of which Rose Hall is a part) in 2013 was compared to the cost of producing 1 hectare of corn in Iowa in 2013. It was then deduced that the cost of

producing 1 hectare of corn on the coast was approximately 67.72% of the cost of producing the same hectare in Iowa. Since the Iowa dataset provided historical yearly average production costs from 1975 to 2024 for both corn and soybeans, the estimated 67.72% cost of production was used to estimate costs for both crops (see Appendix E). Since the data were collected on a yearly time frame, the same procedure used for determining domestic pricing was applied. Because of its location, the cost of production in Ebini differs from that of coastal locations like the Rose Hall estate. Through consultations with local farmers and agronomists, it was determined that the cost of production in Ebini would be approximately 10% higher than at the Rose Hall estate (and other coastal locations). The milling cost for rice was estimated to be about 49 USD per ton, with a paddy-to-mill rice conversion rate of 0.65%. The cost for drying corn and soybeans was estimated to be 25 USD and 20 USD, respectively.

6.5 Multi-objective optimization

Optimization is the systematic process of identifying the most favourable or optimal solution for an optimization problem. Multi-objective optimization (MOO) is a type of optimization that includes searching for both maximum and minimum values for several objectives [56, 57]. Decision-making in MOO allows for compromises or trade-offs in objectives. The exploration of the trade-offs between objectives in an MOO

can be expressed as follows.

$$\text{Max / Min } \mathbf{U}(\mathbf{x}) = (U_1(\mathbf{x}), U_2(\mathbf{x}), \ldots, U_k(\mathbf{x}))^T,$$

$$\mathbf{x} = (x_1, x_2, \ldots, x_n)^T,$$

subject to i constraints:

$$g_i(\mathbf{x}) \leqslant h_i,$$

where $U_1(\mathbf{x}), \ldots, U_k(\mathbf{x})$ are the objective functions that are simultaneously maximized or minimized, and $(x_1, \ldots, x_n)$ are the decision variables [56]. The decision variables can take on a prescribed array of values, $\mathbf{x} \in S$, where S is the solution or parameter space. The Constraints can arise from the specific problem formulation [56, 57].

Pareto method

An MOO problem using the Pareto method can be expressed as follows [41, 57].

$$\begin{aligned} f_{1,\text{ opt}} &= \min f_1(x) \\ f_{2,\text{ opt}} &= \min f_2(x) \\ &\vdots \\ f_{n,\text{ opt}} &= \min f_n(x). \end{aligned}$$

The Pareto method maintains the separation (independence) of the elements in the solution vectors throughout optimization, and the idea of dominance is used to distinguish between dominated and non-dominated solutions [41, 57]. In MOO, the optimal value and dominance solution are typically attained when one objective

function cannot be increased without causing the other objective function to decrease. The term for this state is Pareto optimality, and the Pareto optimal solution refers to the set of optimal solutions in the MOO [41, 57].

6.5.1 Model description

This multi-objective optimization model aims to determine the optimal amount of land to allocate for various crop-planting interval combinations for polyculture, double cropping, triple cropping, and N-cropping. Considering the study locations' climatic conditions, soil types, and the crops used for modelling, simulations are limited to polyculture and double cropping. Specifically, double cropping is being simulated for the Rose Hall estate lands, while polyculture is being simulated for Ebini. In this study, two crops will be planted during the same growing season in the polyculture scenario in Ebini. The growing season for Ebini will be the May-June. The crops used for simulations have several potential planting intervals, each approximately two weeks wide (see Appendix E). Upon initiation, the model first determines whether simulations will be conducted for polyculture or double-cropping. Following this determination, the appropriate cropping combination mix will be identified based on the chosen cropping method. For instance, the model identifies suitable cropping combinations if a double-cropping is selected. For this research, double cropping combinations are rice and corn and soy and corn. Each of these crops has M, N, and O planting intervals, which generate $M \times O$ and $N \times O$ possible combinations. The

model finds these individual combinations and runs the multi-objective optimization for each, generating $(M \times O) + (N \times O)$ different Pareto fronts. Each Pareto front comprises the optimal land sizes (the decision variable) necessary to optimize all three objectives for a specific crop combination, subject to several constraints. This separation allows for the comparison of different crop combinations, which enables specific objectives to be isolated. For example, finding the crop combination interval with the highest profit or the highest ET economic water productivity.

6.5.2 NSGA II

The NSGA II is a second-generation multi-objective evolutionary algorithm (MOEA) created by [30]. This algorithm enhanced the original NSGA by employing a more efficient non-domination sorting scheme, removing the sharing parameter, and introducing an implicitly elitist selection mechanism that substantially facilitates the capture of Pareto surfaces [4, 30]. The NSGA II algorithm uses a fast non-domination sorting technique to rank each solution based on the number of solutions that dominate [30, 73, 142]. After assigning fitness, a two-step, crowded binary tournament selection is carried out. When two solutions have different rankings, the one with the lower rank is selected [4, 30, 142]. The NSGA II algorithm was chosen for this study because it was successfully used in other crop planning research [5, 4].

6.5.3 NSGA II parameter settings

In this study, the NSGA II parameters were adjusted based on values from the literature [5, 4, 149]. Cross over probability $P_c = 0.9$, mutation probability $P_m = \frac{1}{Dv}$, where DV is the number of decision variables. Population size $N = 200$ (100-300 was suggested), cross-over distribution index $CD_i = 20$, mutation distribution index $MD_i = 20$.

6.5.4 Simulation scenarios

This study will encompass six simulated projections for the forthcoming 2024/25 crop season. Four of these simulations will be concentrated on the agricultural fields of the New Amsterdam estate, while two will be conducted on the Ebini farmlands. The simulations will be carried out on 2500 hectares in both areas, with a specific emphasis on exporting approximately 25% and 50% of the total crop production.

Interval overlaps

For planting intervals that overlap, the results that occur in the prior interval should be used. For example, the two planting windows May 15^{th} – May 30^{th} and may 22^{nd} – June 7^{th} over lap between May 22^{nd} – May 30^{th} . If planting occurs and is completed in this overlapped timeframe, the results obtained for the May 15^{th} - May 30^{th} window should be used.

6.6 Results and discussion

6.6.1 Constraint settings

The demand limitations used in the simulation projections remained constant at 4000,2000 and 7000 tonnes for rice, soybeans, and maize, respectively. The investment limitation was defined as the sum of the variable costs of all crop combinations plus an additional $\$10,000$.

6.6.2 Projections under 25 percent export for 2024/25 in New Amsterdam

Multiple simulations were conducted in New Amsterdam, where a 25% export rate of milled rice, dried corn, and dried soybeans was used. In this region, double cropping simulations were conducted separately for rainfed rice, corn, and soybeans and for irrigated rice, corn, and soybeans. It should be noted that soybeans and corn were grown under rainfed conditions only.

The double cropping combinations based on the given planting time and crops employed were as follows: 1) irrigated rice and corn, and soy and corn; and 2) rainfed rice and corn, and soy and corn. The total number of combinations for the double cropping scenarios was 40 (for irrigated rice and corn and soy and corn) and 30 (for rainfed rice and corn and soy and corn). These combinations produced 40 and 30

unique Pareto optimum fronts. Two of these Pareto optimal fronts can be seen in Figures 6.1 & 6.2

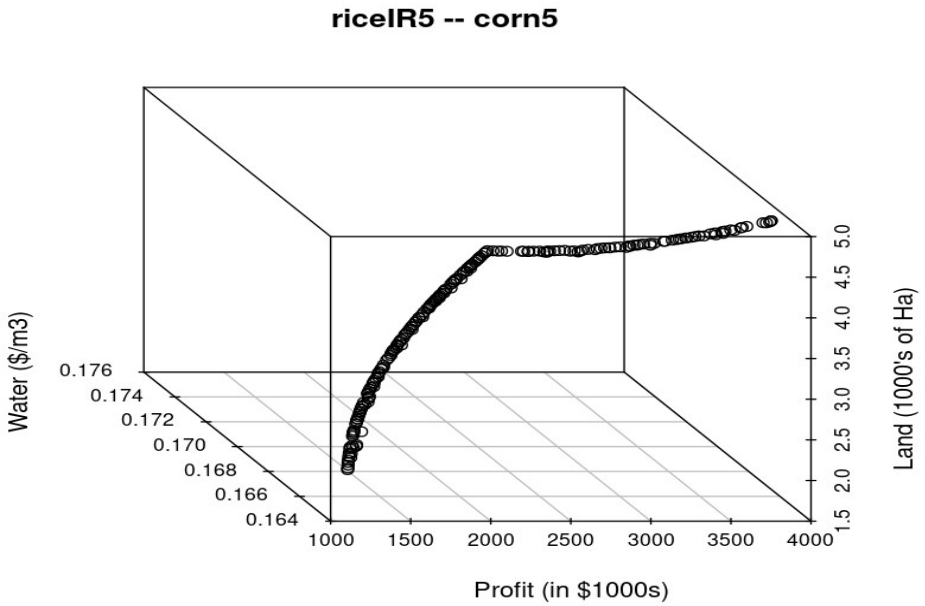


Fig. 6.1. The Pareto optimal front for double crop combination irrigated rice interval 1- corn interval 5 for 2024 under 25% export

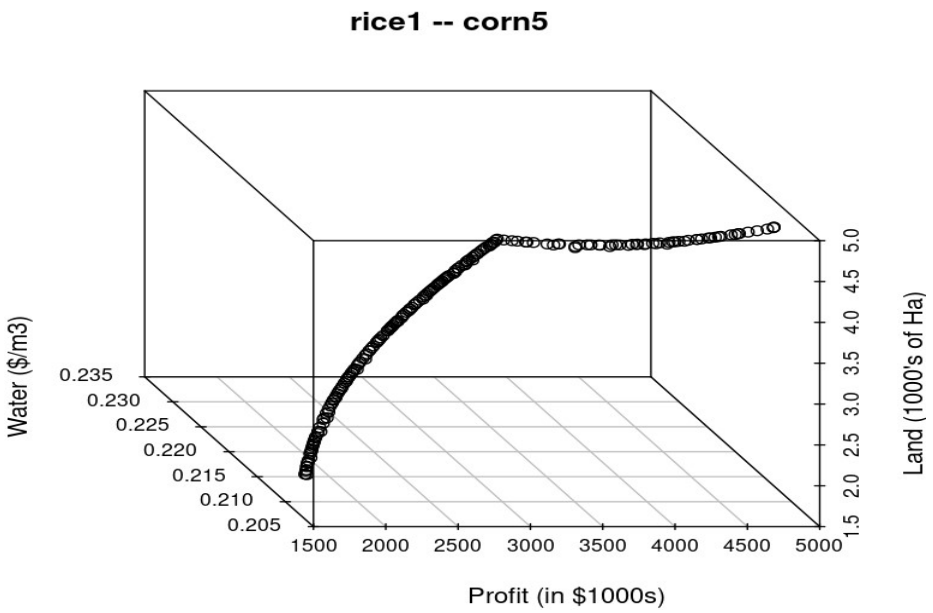


Fig. 6.2. The Pareto optimal front for double crop combination rainfed rice interval 1- corn interval 5 for 2024 under 25% export

Within the Pareto optimum solution set, no solution is better than any other solution for all objectives. Ultimately, the decision maker must choose one option from this collection for system implementation based on their preferred needs [5, 97]. In this research, each crop-planting interval combination has a distinct Pareto optimal solution set, making it easier to compare the different combinations. To analyze and compare the simulation results and determine the model's effectiveness, the solution in each Pareto front that yielded the highest and lowest profits was used to determine the crop planting interval combination that was the best and worst for obtaining the optimal profit. (Tables 6.1 and 6.5).

Table 6.3. The five best and five worst projected crop combinations for rainfed rice for 2024/25 growing season under 25% export

Crop 1	Crop 2	Area 1	Area 2	profit	EWP
rice1	corn5	2500	2500	4788.175	0.208
rice2	corn5	2500	2500	4732.885	0.205
rice3	corn5	2500	2500	4648.74	0.202
rice1	corn4	2500	2500	4621.205	0.199
rice2	corn4	2500	2500	4565.914	0.97
soy2	corn1	2500	2500	2454.721	0.109
soy1	corn4	2500	2500	2427.223	0.114
soy1	corn3	2500	2500	2353.372	0.109
soy1	corn2	2500	2500	2251.659	0.103
soy1	corn1	2500	2500	2214.882	0.101

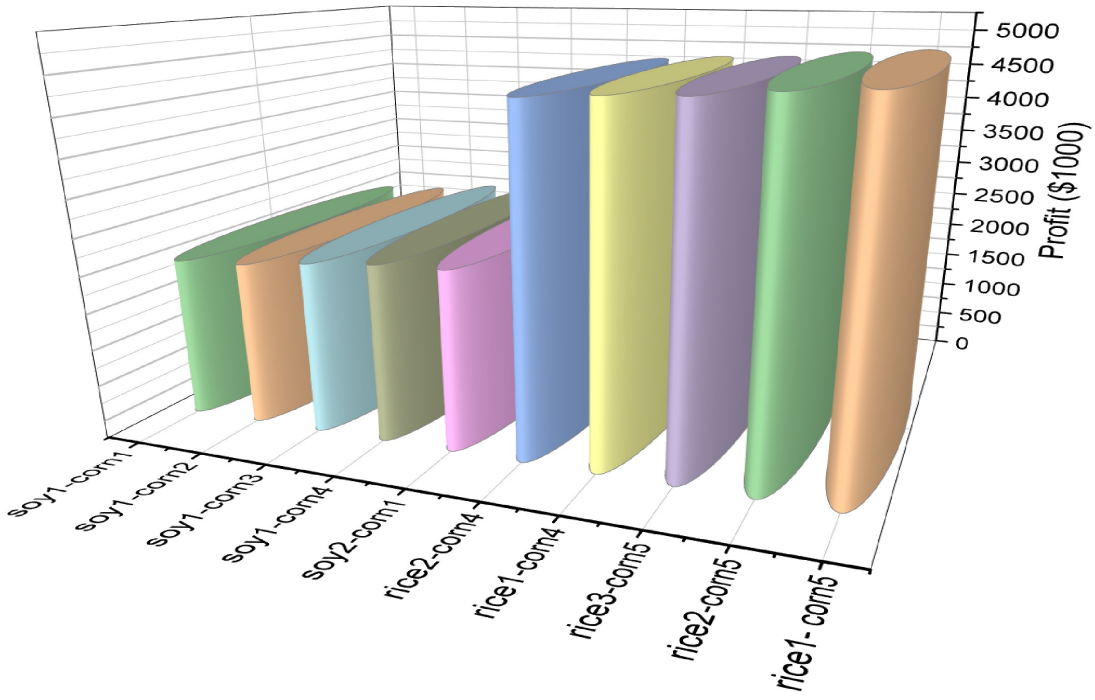


Fig. 6.3. Bar chart of the best and worst performing crop-planting interval combinations with rainfed rice for 2024/25

Table 6.4. The five best and five worst projected crop combinations for irrigated rice for 2024/25 growing season under 25% export

Crop 1	Crop 2	Area 1	Area 2	profit	EWP
riceIR5	corn5	2500	2500	3874.28	0.165
riceIR1	corn5	2500	2500	3834	0.165
riceIR4	corn5	2500	2500	3803.716	0.163
riceIR3	corn5	2500	2500	3784.442	0.165
riceIR2	corn5	2500	2500	3783.737	0.437
soy2	corn1	2500	2500	2454.721	0.109
soy1	corn4	2500	2500	2427.223	0.114
soy1	corn3	2500	2500	2353.372	0.109
soy1	corn2	2500	2500	2251.659	0.103
soy1	corn1	2500	2500	2214.882	0.101

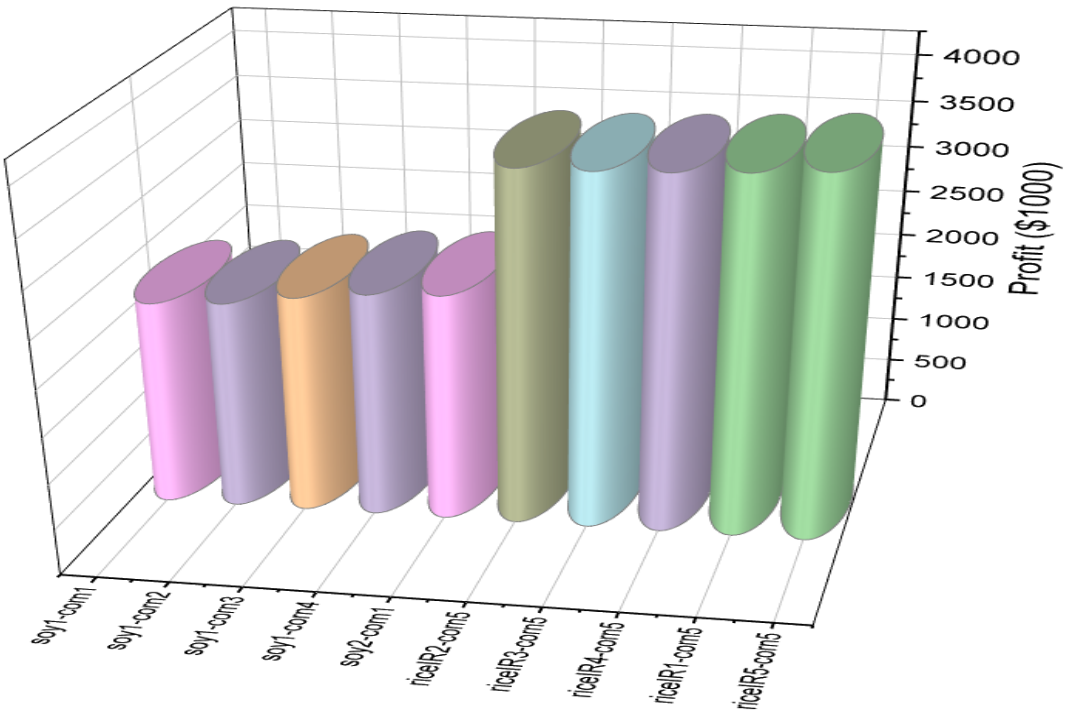


Fig. 6.4. Bar chart of the best and worst performing crop-planting interval combinations with irrigated rice for 2024/25

In Table 6.1, the results demonstrate that double-cropping rice and corn, with production timed for planting between May 15th and May 30th of 2024 for rice and between November 15th and November 30th, 2024 for corn, is projected to generate $\$4,788,175$ (USD) of gross profit. This combination is projected to have an ET economic water productivity (EEWP) value of approximately \$0.208 per m^3 of water effectively used by the crop through evapotranspiration. On the other hand, if soy is planted between June 1st and June 15th, 2024, and corn between October 15th and October 30th, 2024, the projected outcome is a gross profit of $\$2,214,882$ (USD) with an EEWP value of \$0.101 per m^3 of water consumed. This combination of soy and corn decreases earnings by 53.74% compared to the most successful combination. Alternatively, if soy and corn are selected for production, to achieve optimal profits, it is recommended to schedule soy planting between June 15th and June 30th, 2024, and corn planting between November 15th and November 30th, 2024. With this planting interval combination, gross profit is projected to be $\$3,017,783$.

For double-crop combinations featuring irrigated rice, planting between June 15th and June 30th, 2024, for rice and between November 15th and November 30th, 2024, for corn is projected to generate the highest gross profit of $\$3,874,280$, with an EEWP of \$0.165 per m^3. Conversely, planting soy between June 1st and June 15th, 2024, and corn between October 15th and October 30th, 2024, resulted in a gross profit of $\$2,214,882$, the lowest of all combinations (Table 6.2). Interestingly, the highest projected profit earned for the irrigated rice and corn combination was approximately

19.09% less than the combination featuring rainfed rice.

6.6.3 Projections under 50 percent export for 2024/25 in New Amsterdam

Based on simulated projections, if 50% of the commodities produced through double cropping featuring rainfed rice are exported, planting rice and corn during the May 15th- May 30th and November 15th- November 30th, Window will once again yield the highest gross profit. Under this level of export, however, gross profit is projected to decrease by $\$572,121$ to $\$4,216,054$, with the lowest-performing combination seeing a significant reduction of $\$709,480$ to $\$1,505,402$ (see Appendix E). Like results under 25% export, soy planted between June 1st- June 15th and corn between October 15th-October 30th saw the lowest gross profits. The EEWP for the best-performing and worst-performing combinations were \$0.183 and \$0.069 per m^3, respectively. This overall reduction in profits can be attributed to higher projected domestic prices versus export prices for each crop considered in the study.

Double cropping featuring irrigated rice in the combinations is also projected to see a reduction in gross profits. This reduction in profit was approximately 15.27% and amounted to $\$592,120$. Like exporting 25% of the commodities, the same crop planting interval of June 15 th-June 30th (rice) and November 15th-November 30th (corn) generated the highest gross profits (see Appendix E).

6.6.4 Simulation projections in Ebini

Ebini, located approximately 53 miles from the Rose Hall estate, has a different soil texture (sandy loam) and lacks irrigation sources. Consequently, crops grown in this area primarily rely on rainfall to meet water demands. Due to the rainfall pattern in this region, double cropping is not commonly practiced. To demonstrate the flexibility of the optimization model, simulated projections for this area in 2024 will focus on polyculture (Specifically, when two crops are planted in the same area during the May 15th to June 30th) growing season. Although Ebini boasts thousands of hectares of farmland, simulations will use 2500 hectares to maintain consistency with other projections. Since we are using polyculture with two crops grown in the same area, the model will generate possible combinations of crops that simultaneously optimize the research objectives. As discussed previously, each of these combinations has a unique Pareto front, in which each solution in these distinct fronts is considered optimal. This gives users a wide variety of choices and allows them to decide what they consider optimal. However, for demonstration purposes, we will once again organize the solutions of the Pareto fronts by maximum profit to determine which combination generates the highest gross profit and their corresponding land allocations. The cropping combinations used were rice-soy, rice-corn and soy-corn. These combinations resulted in 55 distinct Pareto fronts. The Pareto front for one of the combinations can be seen below.

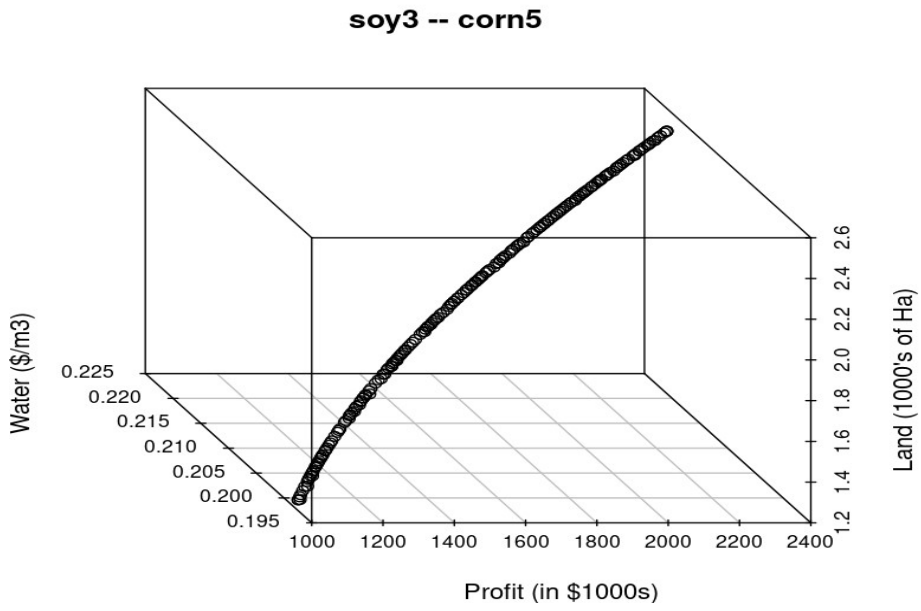


Fig. 6.5. The Pareto optimal front for crop combination soy interval 3- corn interval 5 for 2024 under 25% export

6.6.5 Projections with 25 percent export for 2024 in Ebini

Over Ebini, results indicate that a polyculture of soybeans and corn is projected to yield the highest gross profit for 2024. Specifically, if production is timed so that planting is done and completed between May 30th and June 15th for soybeans and between June 15th and June 30th for corn. In that case, gross profit is projected to be $\$2,392,912$, with an EEWP of \$0.221 per m^3 of water. It should be noted that EEWP in this region is higher than in Rose Hall. This implies that rainfall water in Ebini is being used more effectively by the crops than it is in New Amsterdam.

For the above-mentioned koy-corn combination, 71.76% of the land (1794 ha) is al-

located to soybean production, while 28.24% (706 ha) is allocated to corn production. Planting rice in combination with soy or corn is projected to be the least profitable venture, generating maximum profits of $1,641,885 (rice-soy) and $1,178,101 (rice-corn), with EEWP values of $0.149 (rice-soy) and $0.113 (rice-corn) per $ m^3$ (see Appendix E). Land allocations were 1497 ha to rice and 1003 ha to both corn and soy. These values are projected if planting is done May 30th -June 15th , June 15th-June 30th (soy-corn), May 22nd- June 7th , May 30th -June 15th (rice-soy) and May 22nd- June 7th, June 15 th -june 30th (rice-corn). The least profitable combination was rice and corn, both planted from May 22nd- June 7th, which is projected to gross $1,118,419 with the same land allocation as the other mentioned combinations (see Appendix E).

6.6.6 Projections under 50 percent export for 2024 in Ebini

When projections were made considering the 50% exportation rate, minimal changes were seen in the order of the crop planting interval combinations compared to those observed under 25% exportation. However, Significant changes were seen in the gross profits of combinations vs 25% export. This increase in exports resulted in a profit reduction of approximately 18.92%, 13.04%, and 21.02% for soy-corn, rice-soy and rice-corn combinations that yielded the highest profits. The projected gross profits of these combinations are $1,940,024.1, $1,427,705 and $881121.2, respectively (see Appendix E). Moderate reductions in EEWP were also observed.

6.6.7 Using model results for agricultural planning, policy and decision making

The results produced by the model under various scenarios can be used by farmers and decision-makers to answer questions related to agricultural planning. The first question that can be addressed is what to plant to satisfy specific objectives. For this research, these objectives were maximum gross profit, ET economic water productivity (EEWP), and minimum land usage. The model produces distinct Pareto fronts for each crop combination and their corresponding planting intervals. The results of each Pareto front can then be filtered based on user interests, allowing the user to decide on what to plant based on the objectives of the optimization. The second question that the model can answer is where to plant. Due to the integration of location-specific climatic data and supplemental data such as soil type, irrigation strategy, field management practices, and crop characteristics into the AquaCrop simulator, location-specific crop yields can be obtained. When these yield data are assimilated into the optimization model along with location-specific production cost data, comparisons between different planting locations can be made. This enables policymakers to see the outcomes of planting certain crops in one location versus another, such as comparing the results of planting soy and corn in Rose Hall estate versus planting in Ebini. The third question concerns how much to plant. The decision variable of the optimization model is A_{ijl}, which represents the number of hectares of each crop

that needs to be planted to simultaneously satisfy the optimization model objectives. This value inherently provides users with the amount of land that needs to be planted. Lastly, the model can answer when to plant. The integration of climatic data and AquaCrop allows for the generation of yields over several varying planting intervals. When this is combined with dynamic production cost data for each of those intervals, the optimization model can produce the optimal crop combinations and their corresponding planting intervals. This enables comparisons between a crop combination planted over different intervals from which the choice of the best period to plant can be made.

6.7 Conclusion

This study aimed to develop a constrained multi-objective optimization model (MOO) that can be used to aid farmers and policymakers in decision-making regarding crop selection and planting timing. The constrained MOO model incorporated dynamic yield, ET water usage, and pricing data obtained from AquaCrop and time series forecasting. Its objectives were to i) maximize combined gross profit, ii) maximize ET economic water productivity, and iii) minimize land usage, subject to constraints on land, investment, and demand. Scenario analysis involved testing the model under assumptions of 25% and 50% export of agricultural commodities, considering both double cropping and polyculture. Simulations were conducted on 2500 hectares of land in Rose Hall Estate and Ebini.

The model produced diverse and promising results, identifying combinations that yielded the highest profit for simulation and scenario analysis after the results were adjusted. Consistently, rice, whether irrigated or rainfed, double-cropped with corn, yielded the highest projected economic value at the Rose Hall estate under 25% and 50% export scenarios. Conversely, soybeans and corn, mainly when soybeans were planted between June 1st -June 15th, 2024, and corn between October 15th-October 30th, 2024, consistently had the lowest projected gross profit under both export scenarios.

In Ebini, where simulations were done under polyculture, corn and soybeans

cropped from May 30th- June 15th for soy and June 15th–June 30th for corn generated the highest projected gross profit, with rice cropped with corn performing the worst. The overall results show the potential utility the model can provide farmers and decision-makers in scenario simulations, policy formulation, and planning.

Chapter 7

General conclusion and future work

7.0.1 Conclusion

The research focused on developing a weather-driven bio-economic optimization model to aid policy and decision-making. This model is comprised of three major components: weather simulations, crop yield simulations, and multi-objective optimization, to create a more dynamic and realistic optimization model. The first component of the model involved developing and simulating models for precipitation, minimum and maximum temperatures, relative humidity, and reference evapotranspiration. The precipitation model showed similarities with summary statistics and seasonal variations similar to those observed in the data. Close statistical similarities and similar seasonal variations were also observed in the model for relative humidity. However,

although the model for minimum and maximum temperatures showed close similarities for summary statistics to observed data, lacked similar seasonal variations. This lack of seasonal variations was due to the inability to calculate the parameters of the SDE model. The final weather model was an ANN used to estimate ETo. This model yielded excellent results when tested with observed weather data and was later integrated with simulated minimum and maximum temperature and relative humidity to obtain simulated ETo data for use in AquaCrop. For the second component of the model, the AquaCrop simulator was used to assess the performance of simulated weather data for yield estimation. The results showed that the simulated weather data performed well in simulating the yields of all crops in both study locations. Very good performances were observed in the simulation of rice and corn yields, and good results were obtained for soybeans. However, improvements in yield simulations over Ebini can be obtained with a more precise soil data analysis. The Multi-objective optimization was the final piece of the overall model structure. It used the yields obtained from AquaCrop under simulated weather conditions together with dynamic production and price data to generate optimal crop combinations based on user needs. Results were organized for maximum profit under double cropping and polyculture practices to demonstrate the model's utility. The obtained results not only present potential users with a single value but also with a variety of choices based on their needs and preferences.

7.0.2 Future work

1. Use of an improved temperature model: The SDE parameter for each month could not be calculated under the assumptions of constant speeds of mean and volatility. Consequently, the seasonal variations observed in the minimum and maximum temperatures were not prevalent in the simulations. This is probably due to the lack of significant fluctuations in the monthly data. As a result, seasonal fluctuations in the simulations were lost. Future work in this area includes the development of a model that does not consider constant speeds of mean reversion and volatility, enabling the addition of seasonal fluctuations to simulations.
2. Acquisition of more detailed soil data for yield simulations: Locations-specific and updated soil analysis was not available for simulations of crop yields. As a result, data from literature and default AquaCrop soil types were used over Rose Hall estate and Ebini, respectively. Future work for the yield simulations includes acquiring better soil data over research locations to improve the accuracy of yield simulations and more specific field management data and irrigation strategies.
3. Acquisition of accurate and historical production cost data for developing a production cost model: Acquiring accurate production data is important to the accuracy of the projections of the multi-objective optimization model. So,

further research should include acquiring accurate production cost data and using a model that can forecast the cost of production.

4. Since the changing nature of soil nutrients was not considered, future work should include a model or models that can effectively replicate the nutrient evolution of the soil on a yearly or semi-yearly time scale. Additional work can be further conducted on the effects of mechanization on soil compaction and its potential impact on soil yield.
5. While this research used simulated weather data to obtain crop yields, the models used did not account for the potential impacts of climate change. Future studies could focus on updating these weather models to incorporate climate change projections, allowing for a more comprehensive analysis of climate change's effects on crop yields and crop yield projections and its potential impact on the economic aspects of crop production.

Bibliography

[1] Shaibu Abdul-Ganiyu, Nicholas Kyei-Baffour, Wilson Agyei Agyare, and Wilson Dogbe. *Evaluating the Effect of Irrigation on Paddy Rice Yield by Applying the AquaCrop Model in Northern Ghana.* 2018.

[2] Wossenu Abtew. Evapotranspiration measurements and modeling for three wetland systems in South Florida. *Water Resources Bulletin*, 32(3), June 1996.

[3] O Adekanmbi and O Olugbara. Multiobjective Optimization of Crop-mix Planning Using Generalized Differential Evolution Algorithm. *J. Agr. Sci. Tech*, 17:1103–1114, 2015.

[4] Oluwole A. Adekanmbi, Oludayo O. Olugbara, and Josiah Adeyemo. A comparative study of state-of-the-art evolutionary multi-objective algorithms for optimal crop-mix planning. *International Journal of Agricultural Science and Technology*, 2:8, 2014.

[5] Josiah Adeyemo and Fred Otieno. Differential evolution algorithm for solving multi-objective crop planning model. *Agricultural Water Management*, 97:848–

856, 6 2010.

[6] Matin Ahooghalandari, Mehdi Khiadani, and Mina Esmi Jahromi. Developing equations for estimating reference evapotranspiration in australia. *Water Resources Management*, 30:3815–3828, 9 2016.

[7] Hirotugu Ai. Ac-19, 1974.

[8] Uvirkaa Akumaga, Aondover Tarhule, and Ado A. Yusuf. Validation and testing of the FAO AquaCrop model under different levels of nitrogen fertilizer on rainfed maize in Nigeria, West Africa. *Agricultural and Forest Meteorology*, 232:225–234, 1 2017.

[9] Peter Alaton, Boualem Djehiche, and David Stillberger. On modelling and pricing weather derivatives. *Applied Mathematical Finance*, 9:1–20, 2002.

[10] Richard G Allen. Crop evapotranspiration. *FAO irrigation and drainage paper*, 56:60–64, 1998.

[11] Samer Alsadi and Tamer Khatib. Modeling of relative humidity using artificial neural network. *Journal of Asian Scientific Research*, 2:81–86, 2012.

[12] Clayton Alcarde Alvares, José Luiz Stape, Paulo Cesar Sentelhas, and José Leonardo de Moraes Gonçalves. Modeling monthly mean air temperature for Brazil. *Theoretical and Applied Climatology*, 113:407–427, 8 2013.

[13] Walison B. Alves, Glauco De S. Rolim, and Lucas E.De O. Aparecido. Reference evapotranspiration forecasting by artificial neural networks. *Engenharia Agricola*, 37:1116–1125, 2017.

[14] E. Amiri, A. Bahrani, S. Irmak, and N. Mohammadiyan Roshan. Evaluation of irrigation scheduling and yield response for wheat cultivars using the AquaCrop model in an arid climate. *Water Supply*, 22:602–614, 1 2022.

[15] M. Amirouche, D. Smadhi, and L. Zella. Calibration and validation of the AquaCrop model for the culture lettuce (Lactuca sativa L.) under fertilization levels in pluvial condition. *AgricEngInt: CIGR Journal*, 23(1):36, March 2021. Open access at http://www.cigrjournal.org.

[16] Vassilis Z. Antonopoulos and Athanasios V. Antonopoulos. Daily reference evapotranspiration estimates by artificial neural networks technique and empirical equations using limited input climate variables. *Computers and Electronics in Agriculture*, 132:86–96, 1 2017.

[17] Indrani Basak and N. Balakrishnan. Estimation for the three-parameter Gamma distribution based on progressively censored data. *Statistical Methodology*, 9:305–319, 5 2012.

[18] Fred Espen Benth and Jurate Šaltyte Benth. Weather derivatives and stochastic modelling of temperature. *International Journal of Stochastic Analysis*, 2011, 2011.

[19] Thijs Van Den Berg. Calibrating the ornstein-uhlenbeck (vasicek) model, 2011.

[20] Jedrzej S. Bojanowski. sirad: Functions for calculating daily solar radiation and evapotranspiration. https://cran.r project.org/web/packages/sirad/sirad.pdf, 2016. R package version 2.3-3.

[21] Dorje C. Brody, Joanna Syroka, and Mihail Zervos. Dynamical pricing of weather derivatives. *Quantitative Finance*, 2:189–198, 2002.

[22] Melanie Cao and Jason Wei. Equilibrium valuation of weather derivatives. 2000. , Working Paper, Department of Economics, Queen's University, Kingston, Ontario.

[23] Sindikubwabo Celestin, Feng Qi, Ruolin Li, Tengfei Yu, and Wenju Cheng. Evaluation of 32 simple equations against the penman–monteith method to estimate the reference evapotranspiration in the hexi corridor, Northwest China. *Water (Switzerland)*, 12, 10 2020.

[24] Peng Chen, Aichen Niu, Duanyang Liu, Wei Jiang, and Bin Ma. Time series forecasting of temperatures using sarima: An example from Nanjing. *IOP Conference Series: Materials Science and Engineering*, 394, 8 2018.

[25] F H S Chiew, N N Kamaladasa, H M Malano, and T A Mcmahon. Agricultural water management elsevier Penman-monteith, FAO-24 reference crop

evapotranspiration and class-a pan data in australia. *Agricultural Water Management*, 28:9–11, 1995.

[26] Edwin H. Chin. Modeling daily precipitation occurrence process with Markov chain. *Water Resources Research*, 13(6):949–956, December 1977.

[27] Archana Chowdhary and R.K. Shrivastava. Reference crop evapotranspiration estimation using artificial neural networks. *International Journal of Engineering Science and Technology*, 2, 09 2010.

[28] Paulin Coulibaly, François Anctil, Ramon Aravena, and Bernard Bobée. Artificial neural network modeling of water table depth fluctuations. *Water Resources Research*, 37:885–896, 2001.

[29] Priyaranjan R Dash. A Markov chain modeling of daily precipitation occurrences in Odisha. *International Journal of Advanced Computer and Mathematical Sciences*, 3:482–486, 2012.

[30] Kalyanmoy Deb, Amrit Pratap, Sameer Agarwal, and T Meyarivan. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION*, 6, 2002.

[31] M D Dennett, J A Rodgers, and J D H Keatinge. Simulation of a rainfall record for the site of a new agricultural development: An example from Northern Syria. *Agricultural Meteorology*, 29:247–258, 1983.

[32] M J Diamantopoulou, P E Georgiou, and D M Papamichail. Performance evaluation of artificial neural networks in estimating reference evapotranspiration with minimal meteorological data. *Global NEST Journal*, 13:18–27, 2011.

[33] Tinashe Lindel Dirwai, Aidan Senzanje, and Tafadzwanashe Mabhaudhi. Calibration and evaluation of the FAO AquaCrop model for canola (Brassica napus) under varied moistube irrigation regimes. *Agriculture (Switzerland)*, 11, 2021.

[34] Koffi Djaman, Alpha B. Balde, Abdoulaye Sow, Bertrand Muller, Suat Irmak, Mamadou K. N'Diaye, Baboucarr Manneh, Yonnelle D. Moukoumbi, Koichi Futakuchi, and Kazuki Saito. Evaluation of sixteen reference evapotranspiration methods under Sahelian conditions in the Senegal River valley. *Journal of Hydrology: Regional Studies*, 3:139–159, 3 2015.

[35] Koffi Djaman, Komlan Koudahe, Mamadou Sall, Isa Kabenge, Daran Rudnick, and Suat Irmak. Performance of twelve mass transfer based reference evapotranspiration models under humid climate. *Journal of Water Resource and Protection*, 09:1347–1363, 2017.

[36] J. Doorenbos and A. H. Kassam. *Yield Response to Water*. FAO, 1979.

[37] Fabien Dornier and Michel Queruel. Caution to the wind. *Energy & power risk management*, 13(8):30–32, 2000.

[38] Fugen Dou, Junel Soriano, Rodante E. Tabien, and Kun Chen. Soil texture and cultivar effects on rice (Oryza sativa, l.) grain yield, yield components and water productivity in three water regimes. *PLoS ONE*, 11, 3 2016.

[39] Peter Droogers and Richard G Allen. Estimating reference evapotranspiration under inaccurate data conditions. *Irrigation and Drainage Systems*, 16:33–45, 2002.

[40] Nelson Christopher Dzupire, Philip Ngare, and Leo Odongo. Levy process based Ornstein-Uhlenbeck temperature model with time varying speed of mean reversion. *Advances and Applications in Statistics*, 53:199–224, 9 2018.

[41] Matthias Ehrgott. *Multicriteria Optimization*, volume 491 of *Lecture Notes in Economics and Mathematical Systems*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2005.

[42] Inas El-Gafy. Integration among Egyptian governorates to optimize irrigation water use utilizing weight goal programming technique. *Irrigation and Drainage*, 62(5):559–577, 2013.

[43] Ahmed Elbeltagi, Attila Nagy, Safwan Mohammed, Chaitanya B. Pande, Manish Kumar, Shakeel Ahmad Bhat, József Zsembeli, László Huzsvai, János Tamás, Elza Kovács, Endre Harsányi, and Csaba Juhász. Combination of limited meteorological data for predicting reference crop evapotranspiration using artificial neural network method. *Agronomy*, 12, 2 2022.

[44] Food and Agriculture Organization of the United Nations. Calculations of Meteorological Variables, 2024. [Accessed: 13-Jul-2024]. Available at: https://www.fao.org/4/x0490e/x0490e07.htm.

[45] T. Foster, N. Brozović, A. P. Butler, C. M.U. Neale, D. Raes, P. Steduto, E. Fereres, and T. C. Hsiao. AquaCrop-os: An open source version of FAO's crop water productivity model. *Agricultural Water Management*, 181:18–22, 2 2017.

[46] K. R. Gabriel and J. Neumann. A Markov chain model for daily rainfall occurrence at Tel Aviv. *Quarterly Journal of the Royal Meteorological Society*, 88(375):90–95, 1962.

[47] Naresh Gaj. An analysis of the hydrology of a sugarcane field in Guyana. Master's thesis, McGill University, Montreal, Canada, May 2014. A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Master of Science.

[48] Magali Garcia, Dirk Raes, Rick Allen, and Carlos Herbas. Dynamics of reference evapotranspiration in the Bolivian Highlands (Altiplano). *Agricultural and Forest Meteorology*, 125:67–82, 9 2004.

[49] Guadalupe Azuara García, Efrén Palacios Rosas, Alfonso García-Ferrer, and Pilar Montesinos Barrios. Multi-objective spatial optimization: Sustainable land use allocation at sub-regional scale. *Sustainability (Switzerland)*, 9, 2017.

[50] M. A. Genaidy. Estimating of evapotranspiration using artificial neural network. *Misr Journal of Agricultural Engineering*, 37(1):81–94, 2020.

[51] Shu Geng, Frits W T Penning De, and Iwan Supit. A simple method for generating daily rainfall data. *Agricultural and Forest Meteorology*, 36:363, 1986.

[52] Ikhlas Ghiat, Rajesh Govindan, and Tareq Al-Ansari. Evaluation of evapotranspiration models for cucumbers grown under CO2 enriched and hvac driven greenhouses: A step towards precision irrigation in hyper-arid regions. *Frontiers in Sustainable Food Systems*, 7, 2023.

[53] Mohiaddin Goosheh, Ebrahim Pazira, Ali Gholami, Bahram Andarzian, and Ebrahim Panahpour. Improving irrigation scheduling of wheat to increase water productivity in shallow groundwater conditions using AquaCrop. *Irrigation and Drainage*, 67:738–754, 12 2018.

[54] Jackeline T. Gotardo, Lineu N. Rodrigues, and Benedito M. Gomes. Comparison of methods for estimating reference evapotranspiration: An approach to the management of water resources within an experimental basin in the Brazilian Cerrado. *Engenharia Agricola*, 36:1016–1026, 2016.

[55] J M Gregory, Tml Wigley, and P D Jones. Application of Markov models to area-average daily precipitation series and interannual variability in seasonal totals. *Climate Dynamics*, 8:299–310, 1993.

[56] Jeroen C.J. Groot, Gerard J.M. Oomen, and Walter A.H. Rossing. Multi-objective optimization and design of farming systems. *Agricultural Systems*, 110:63–77, 7 2012.

[57] Nyoman Gunantara. A review of multi-objective optimization: Methods and its applications. *Cogent Engineering*, 5:1–16, 1 2018.

[58] George Hargreaves and Zohrab Samani. Reference crop evapotranspiration from temperature. *Applied Engineering in Agriculture*, 1, 01 1985.

[59] George H Hargreaves and Zohrab A Samani. Estimating potential evapotranspiration. *Journal of the irrigation and Drainage Division*, 108(3):225–230, 1982.

[60] Guomin Huang, Lifeng Wu, Xin Ma, Weiqiang Zhang, Junliang Fan, Xiang Yu, Wenzhi Zeng, and Hanmi Zhou. Evaluation of catboost method for prediction of reference evapotranspiration in humid regions. *Journal of Hydrology*, 574:1029–1041, 7 2019.

[61] Mingyi Huang, Ce Wang, Wei Qi, Zhanyu Zhang, and Hui Xu. Modelling the integrated strategies of deficit irrigation, nitrogen fertilization, and biochar addition for winter wheat by AquaCrop based on a two-year field study. *Field Crops Research*, 282, 6 2022.

[62] Z. Huo, S. Feng, S. Kang, and X. Dai. Artificial neural network models for reference evapotranspiration in an arid area of Northwest China. *Journal of Arid Environments*, 82:81–90, 7 2012.

[63] Rob J. Hyndman and Yeasmin Khandakar. Automatic time series forecasting: The forecast package for r. *Journal of Statistical Software*, 27(3):1–22, 2008.

[64] A. O. Ibeje, J. C. Osuagwu, and O. R. Onosakponome. A Markov model for prediction of annual rainfall. *International Journal of Scientific Engineering and Applied Science (IJSEAS)*, 3(11), 2018.

[65] S. Irmak, A. Irmak, R. G. Allen, and J. W. Jones. Solar and net radiation-based equations to estimate reference evapotranspiration in humid climates. *Journal of Irrigation and Drainage Engineering*, 129:336–347, 10 2003.

[66] A. I. Issaka, J. Paek, K. Abdella, M. Pollanen, A. K. S. Huda, S. Kaitibie, I. Goktepe, M. M. Haq, and A. T. Moustafa. Analysis and calibration of empirical relationships for estimating evapotranspiration in Qatar: Case study. *Journal of Irrigation and Drainage Engineering*, 143, 2 2017.

[67] NN Ivanov. On the determination of evaporation values. *Yzv. HHO, Moskow*, pages 189–196, 1954.

[68] Jennifer M Jacobs, Sudheer Reddy, Satti Graduate, and John M Fitzgerald. Evaluation of reference evapotranspiration methodologies and afsirs crop water use simulation model (final report), 2001.

[69] S. K. Jain, P. C. Nayak, and K. P. Sudheer. Models for estimating evapotranspiration using artificial neural networks, and their physical interpretation. *Hydrological Processes*, 22:2225–2234, 6 2008.

[70] Marvin E Jensen and Howard R Haise. Estimating evapotranspiration from solar radiation. *Journal of the Irrigation and Drainage Division*, 89(4):15–41, 1963.

[71] Chisimkwuo John, Uchechukwu George, and Okezie Sampson Chukwuemeka. Time series analysis and forecasting of monthly maximum temperatures in South Eastern Nigeria. *International Journal of Innovative Research & Development*, 3(1), January 2014.

[72] P G Jones and P K Thornton. Spatial and temporal variability of rainfall related to a third-order Markov model. *Agricultural and Forest Meteorology*, 86:127–138, 1997.

[73] K. Senthamarai Kannan and A. Jawahar Farook. Stochastic simulation of precipitation using Markov chain-mixed exponential model. *Applied Mathematical Sciences*, 9:3205–3212, 2015.

[74] Manad Khamkong. Parameter estimation method for the two parameter Gamma distribution based on transformation. *International Journal of Applied Engineering Research*, 13:1261–1264, 2018.

[75] William Kleiber, Richard W. Katz, and Balaji Rajagopalan. Daily spatiotemporal precipitation simulation using latent and transformed Gaussian processes. *Water Resources Research*, 48, 2012.

[76] Ian Kropp, A. Pouyan Nejadhashemi, Kalyanmoy Deb, Mohammad Abouali, Proteek C. Roy, Umesh Adhikari, and Gerrit Hoogenboom. A multi-objective approach to water and nutrient efficiency for sustainable agricultural intensification. *Agricultural Systems*, 173:289–302, 7 2019.

[77] M. Kumar, N. S. Raghuwanshi, R. Singh, W. W. Wallender, and W. O. Pruitt. Estimating evapotranspiration using artificial neural network. *Journal of Irrigation and Drainage Engineering*, 128(4):224, 2002.

[78] G. A. Larsen and R. B. Pense. Stochastic simulation of daily climatic data for agronomic models. *Agronomy Journal*, 74:510–514, 1982.

[79] Mark G. Lawrence. The relationship between relative humidity and the dewpoint temperature in moist air: A simple conversion and applications. *Bulletin of the American Meteorological Society*, 86:225–233, 2 2005.

[80] Jan Lennartsson, Anastassia Baxevani, and Deliang Chen. Modelling precipitation in Sweden using multiple step Markov chains and a composite model. *Journal of Hydrology*, 363:42–59, 12 2008.

[81] Fawen Li, Dong Yu, and Yong Zhao. Irrigation scheduling optimization for cotton based on the AquaCrop model. *Water Resources Management*, 33:39–55, 1 2019.

[82] Mo Li, Yan Zhou, Yukuan Wang, Vijay P. Singh, Zongxing Li, and Yuexin Li. An ecological footprint approach for cropland use sustainability based on multi-objective optimization modelling. *Journal of Environmental Management*, 273, 11 2020.

[83] ZhiQiang Li, HongXia Zou, and Bin Qi. Application of arima and lstm in relative humidity prediction. pages 1544–1549, 10 2019.

[84] Edward T Linacre. A simple formula for estimating evaporation rates in various climates, using temperature data alone. *Agricultural meteorology*, 18(6):409–424, 1977.

[85] P. Lobit, A. Gómez Tagle, F. Bautista, and J. P. Lhomme. Retrieving air humidity, global solar radiation, and reference evapotranspiration from daily temperatures: development and validation of new methods for Mexico. part iii: reference evapotranspiration. *Theoretical and Applied Climatology*, 133:787–797, 8 2018.

[86] R. W. Longley. The length of dry and wet periods. *Quarterly Journal of the Royal Meteorological Society*, 79(342):520–527, 1953.

[87] Jaydip J. Makwana, Mukesh K. Tiwari, and B. S. Deora. Development and comparison of artificial intelligence models for estimating daily reference evapotranspiration from limited input variables. *Smart Agricultural Technology*, 3, 2 2023.

[88] Junaid Maqsood, Aitazaz A. Farooque, Farhat Abbas, Travis Esau, Xander Wang, Bishnu Acharya, and Hassan Afzaal. Application of artificial neural networks to project reference evapotranspiration under climate change scenarios. *Water Resources Management*, 36:835–851, 2 2022.

[89] Babita Mishra, Anil Kumar Nishad, and S. R. Singh. Fuzzy multi-fractional programming for land use planning in agricultural production system. *Fuzzy Information and Engineering*, 6:245–262, 6 2014.

[90] M. Mohammadi, B. Ghahraman, K. Davary, H. Ansari, A. Shahidi, and M. Bannayan. Nested validation of AquaCrop model for simulation of winter wheat grain yield, soil moisture and salinity profiles under simultaneous salinity and water stress. *Irrigation and Drainage*, 65:112–128, 2 2016.

[91] M. Shahjahan Mondal, Abul Fazal M. Saleh, Md Abdur Razzaque Akanda, Sujit K. Biswas, Abu Zofar Abu, Sinora Zaman, Attila N. Lazar, and Derek

Clarke. Simulating yield response of rice to salinity stress with the AquaCrop model. *Environmental Science: Processes and Impacts*, 17:1118–1126, 6 2015.

[92] Ana Flávia Martins Monteiro, Fabrina Bolzan Martins, Roger Rodrigues Torres, Vitor Hugo Marrafon de Almeida, Marcel Carvalho Abreu, and Enrique Vieira Mattos. Intercomparison and uncertainty assessment of methods for estimating evapotranspiration using a high-resolution gridded weather dataset over brazil. *Theoretical and Applied Climatology*, 146:583–597, 10 2021.

[93] Sung-Euii Moon, Sang-Boom Ryoo, and Jae-Gi Kwon. A Markov chain model for daily precipitation occurrence in South Korea, 1994.

[94] Nall Moonilall, Oudho Homenauth, and Rattan Lal. Short-term effects of amendments on soil properties and agronomic productivity for a coastal Guyana soil. *Tropical Agriculture*, 97:9–31, 01 2020.

[95] Guia Marie M. Mortel. Remote sensing of crop inventories and crop model simulations for irrigation management along the Guyana Coastal Plains. Master's thesis, McGill University, Montreal, Canada, 2023. A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Master of Science.

[96] Małgorzata Murat, Iwona Malinowska, Magdalena Gos, and Jaromir Krzyszczak. Forecasting daily meteorological time series using arima and regression models. *International Agrophysics*, 32:253–264, 4 2018.

[97] Antonio L. Márquez, Raúl Baños, Consolación Gil, María G. Montoya, Francisco Manzano-Agugliaro, and Francisco G. Montoya. Multi-objective crop planning using pareto-based evolutionary algorithms. *Agricultural Economics*, 42:649–656, 11 2011.

[98] N. M. Neykov, P. N. Neytchev, and W. Zucchini. Stochastic daily precipitation model with a heavy-tailed component. *Natural Hazards and Earth System Sciences*, 14:2321–2335, 9 2014.

[99] Tangzhe Nie, Yi Tang, Yang Jiao, Na Li, Tianyi Wang, Chong Du, Zhongxue Zhang, Peng Chen, Tiecheng Li, Zhongyi Sun, and Shijiang Zhu. Effects of irrigation schedules on maize yield and water use efficiency under future climate scenarios in heilongjiang province based on the AquaCrop model. *Agronomy*, 12, 4 2022.

[100] Ah Nury, M Koch, and MJB Alam. Time series analysis and forecasting of temperatures in the sylhet division of Bangladesh. In *Proceedings of the Conference on Meteorology and Environmental Sciences*, August 2014.

[101] Oecd and Ocde. Sustainable management of water resources in agriculture, 2010.

[102] Jayeong Paek, Marco Pollanen, and Kenzu Abdella. A stochastic weather model for drought derivatives in arid regions: A case study in Qatar. *Mathematics*, 11, 4 2023.

[103] G. S. Pawar, M. U. Kale, and J. N. Lokhande. Response of AquaCrop model to different irrigation schedules for irrigated cabbage. *Agricultural Research*, 6:73–81, 3 2017.

[104] Lingling Peng, Yi Li, and Hao Feng. The best alternative for estimating reference crop evapotranspiration in different sub-regions of mainland China. *Scientific Reports*, 7, 12 2017.

[105] Ryan Pereira, C. Isabella Bovolo, Nathan Forsythe, Nikolai Pedentchouk, Geoff Parkin, and Thomas Wagner. Seasonal patterns of rainfall and river isotopic chemistry in Northern Amazonia (Guyana): From the headwater to the regional scale. *Journal of South American Earth Sciences*, 52:108–118, 2014.

[106] Nader Pirmoradian and Naser Davatgar. Simulating the effects of climatic fluctuations on rice irrigation water requirement using AquaCrop. *Agricultural Water Management*, 213:97–106, 3 2019.

[107] Dirk Raes, Elias Fereres, Margarita García Vila, Yannick Curnel, David Knoded, Sema Kale Çelik, Yusuf Ucar, Mevlüt Türk, and Joost Wellens. Simulation of alfalfa yield with AquaCrop. *Agricultural Water Management*, 284, 6 2023.

[108] Dirk Raes, Pasquale Steduto, Theodore C Hsiao, and Elias Fereres. *Chapter 3 Calculation procedures AquaCrop Reference manual August 2023*, 2023.

[109] A. Raizada, Pradeep Dogra, and B. L. Dhyani. Assessment of a multi-objective decision support system generated land use plan on forest fodder dependency in a Himalayan watershed. *Environmental Modelling and Software*, 23:1171–1181, 9 2008.

[110] Arash Ranjbar, Ali Rahimikhoob, Hamed Ebrahimian, and Maryam Varavipour. Assessment of the AquaCrop model for simulating maize response to different nitrogen stresses under semi-arid climate. *Communications in Soil Science and Plant Analysis*, 50:2899–2912, 12 2019.

[111] Roxana Seyed Raoufi, Saeid Soufizadeh, Bahman Amiri Larijani, Majid AghaAlikhani, and Jafar Kambouzia. Simulation of growth and yield of various irrigated rice (Oryza sativa l.) genotypes by AquaCrop under different seedling ages. *Natural Resource Modeling*, 31, 5 2018.

[112] VA Romanenko. Computation of the autumn soil moisture using a universal relationship for a large area. *Proc. of Ukrainian Hydrometeorological Research Institute*, 3:12–25, 1961.

[113] Peter Romilly. Time series modelling of global mean temperature for managerial decision-making. *Journal of Environmental Management*, 76:61–70, 7 2005.

[114] Debanjali Saha and M. Shahjahan Mondal. Simulation of boro rice yield in South-West Coastal Bangladesh using the AquaCrop model. In *Proceedings*

of the 5th International Conference on Water & Flood Management (ICWFM-2015), Dhaka, Bangladesh, March 2015.

[115] R.A. Sarker and M.A. Quaddus. Modelling a nationwide crop planning problem using a multiple criteria decision making tool. *Computers Industrial Engineering*, 42(2):541–553, 2002.

[116] Amirpouya Sarraf, Seyed Farnood Vahdat, and Azita Behbahaninia. Relative humidity and mean monthly temperature forecasts in ahwaz station with arima model in time series analysis. 2011.

[117] U Schendel. Vegetationswasserverbrauch und-wasserbedarf. *Habilitation, Kiel*, 137(1):1–11, 1967.

[118] J. T. Schoof and S. C. Pryor. On the proper order of Markov chain model for daily precipitation occurrence in the contiguous United States. *Journal of Applied Meteorology and Climatology*, 47:2477–2486, 2008.

[119] Yuyang Shan, Ge Li, Lijun Su, Jihong Zhang, Quanjiu Wang, Junhu Wu, Weiyi Mu, and Yan Sun. Performance of AquaCrop model for maize growth simulation under different soil conditioners in shandong coastal area, china. *Agronomy*, 12, 7 2022.

[120] K. Gnana Sheela and S. N. Deepa. Review on methods to fix number of hidden neurons in neural networks. *Mathematical Problems in Engineering*, 2013, 2013.

[121] Naichen Shi, Dawei Li, Mingyi Hong, and Ruoyu Sun. Rmsprop converges with proper hyper-parameter. In *International Conference on Learning Representations (ICLR)*, 2021.

[122] Sana Zeeshan Shirazi, Xurong Mei, Buchun Liu, and Yuan Liu. Assessment of the AquaCrop model under different irrigation scenarios in the North China Plain. *Agricultural Water Management*, 257, 11 2021.

[123] Da Silva, Roberta Araújo e Silva, Girlene Figueiredo Maciel, Célia Campos Braga, José Luiz Cabral da Silva Júnior, Enio Pereira de Souza, Rafaela Silveira Rodrigues Almeida, Madson Tavares Silva, and Romildo Morant de Holanda. Calibração e validação do modelo AquaCrop para a cultura de soja cultivada mediante diferentes níveis de irrigação na região de matopiba, brasil. *Ciencia Rural*, 48, 2018.

[124] D P Singh, O Homenauth, N Cumberbatch, V Persaud, and F Benjamin. Performance of corn (Zea mays) genotypes at coastal and savannah regions and cost of cultivation in Guyana. *Greener Journal of Agricultural Sciences*, 4:310–320, 2012.

[125] S V Singh and R H Kripalani. Analysis of persistence in daily monsoon rainfall over India. *JOURNAL OF CLIMATOLOGY*, 6:55, 1986.

[126] D. U.J. Sonnadara and D. R. Jayewardene. A Markov chain probability model to describe wet and dry patterns of weather at Colombo. *Theoretical and Applied Climatology*, 119:333–340, 1 2015.

[127] Paritosh Srivastava and Raj Mohan Singh. Agricultural land allocation for crop planning in a canal command area using fuzzy multiobjective goal programming. *Journal of Irrigation and Drainage Engineering*, 143, 6 2017.

[128] Pasquale Steduto, Theodore C. Hsiao, Dirk Raes, and Elias Fereres. AquaCrop-the FAO crop model to simulate yield response to water: I. concepts and underlying principles. *Agronomy Journal*, 101:426–437, 5 2009.

[129] J.G Steele. Soil surveys, report to the government of Guyana, 1966.

[130] R D Stern. Analysis of daily rainfall at Samaru, Nigeria, using a simple two-part model. *ARCHIV Arch. Met. Geoph. Biokl., Ser. B*, 28:123–135, 1981.

[131] Michael Strauch, Anna F. Cord, Carola Pätzold, Sven Lautenbach, Andrea Kaim, Christian Schweitzer, Ralf Seppelt, and Martin Volk. Constraints in multi-objective optimization of land use allocation – repair or penalize? *Environmental Modelling and Software*, 118:241–251, 8 2019.

[132] Hai Tao, Lamine Diop, Ansoumana Bodian, Koffi Djaman, Papa Malick Ndiaye, and Zaher Mundher Yaseen. Reference evapotranspiration prediction using

hybridized fuzzy model with firefly algorithm: Regional case study in burkina faso. *Agricultural Water Management*, 208:140–151, 9 2018.

[133] Birhan Taştan and Azize Hayfavi. Modeling temperature and pricing weather derivatives based on temperature. *Advances in Meteorology*, 2017, 2017.

[134] Bekele Temesgen, Simon Eching, Baryohay Davidoff, and Kent Frame. Comparison of some reference evapotranspiration equations for California. *Journal of Irrigation and Drainage Engineering*, 131:73–84, 2 2005.

[135] Alfonso F. Torres, Wynn R. Walker, and Mac McKee. Forecasting daily potential evapotranspiration using machine learning and limited climatic data. *Agricultural Water Management*, 98:553–562, 2 2011.

[136] Slavisa Trajkovic and Srdjan Kolakovic. Wind-adjusted turc equation for estimating reference evapotranspiration at humid European locations. *Hydrology research*, 40(1):45–52, 2009.

[137] L Turc. The estimation of irrigation-water requirements; potential evapotranspiration. development of a simplified climatic formula (etp, f. 60). 1961.

[138] Barikara Umesh, K. S. Reddy, B. S. Polisgowdar, V. Maruthi, U. Satishkumar, M. S. Ayyanagoudar, Sathyanarayan Rao, and H. Veeresh. Assessment of climate change impact on maize (Zea mays l.) through AquaCrop model in

semi-arid alfisol of Southern Telangana. *Agricultural Water Management*, 274, 12 2022.

[139] Fatih Üneş, Süreyya Doğan, Bestami Taşar, Y Kaya, Mustafa Demirci, et al. The evaluation and comparison of daily reference evapotranspiration with ann and empirical methods. *Natural and Engineering Sciences*, 3(3):54–64, 2018.

[140] John D. Valiantzas. Simplified forms for the standardized fao-56 Penman-monteith reference evapotranspiration using limited weather data. *Journal of Hydrology*, 505:13–23, 11 2013.

[141] Mohammad Valipour. Evaluation of radiation methods to study potential evapotranspiration of 31 provinces. *Meteorology and Atmospheric Physics*, 127:289–303, 6 2015.

[142] Shanu Verma, Millie Pant, and Vaclav Snasel. A comprehensive review on NSGA-II for multi-objective combinatorial optimization problems. *IEEE Access*, 9:57757–57791, 2021.

[143] Abdul H. Wahab. Performance of soybean in Guyana's Peats and Peaty Clays. *The Journal of Agriculture of the University of Puerto Rico*, 63(4):443–449, Oct. 1979.

[144] Abdul H. Wahab and Imran Hassan. Soybean variety evaluation under three edaphic conditions in Guyana. *The Journal of Agriculture of the University of Puerto Rico*, 63(4):450–459, Oct. 1979.

[145] Aemro Wale, Mekete Dessie, and Hailu Kendie. Evaluating the performance of AquaCrop model for potato production under deficit irrigation. *Air, Soil and Water Research*, 15, 2022.

[146] Spencer Walls, Andrew D. Binns, Jana Levison, and Scott MacRitchie. Prediction of actual evapotranspiration by artificial neural network models using data from a Bowen ratio energy balance station. *Neural Computing and Applications*, 32:14001–14018, 9 2020.

[147] Guangshuai Wang, Faisal Mehmood, Muhammad Zain, Abdoul Kader Mounkaila Hamani, Jingjie Xue, Yang Gao, and Aiwang Duan. AquaCrop model evaluation for winter wheat under different irrigation management strategies: A case study on the North China Plain. *Agronomy*, 12, 12 2022.

[148] Haidong Wang, Xiukang Wang, Lifei Bi, Ying Wang, Junliang Fan, Fucang Zhang, Xianghao Hou, Minghui Cheng, Wenhui Hu, Lifeng Wu, and Youzhen Xiang. Multi-objective optimization of water and fertilizer management for potato production in sandy areas of Northern China based on topsis. *Field Crops Research*, 240:55–68, 7 2019.

[149] Qi Wang, Libing Wang, Wen Huang, Zhihong Wang, Shuming Liu, and Dragan A. Savić. Parameterization of NSGA-II for the optimal design of water distribution systems. *Water (Switzerland)*, 11, 5 2019.

[150] Zhiliang Wang, Peng Li, Lingyong Li, Chunyan Huang, and Min Liu. Modeling and forecasting average temperature for weather derivative pricing. *Advances in Meteorology*, 2015, 2015.

[151] Wandee Wanishsakpong and Bright Emmanuel Owusu. Optimal time series model for forecasting monthly temperature in the southwestern region of Thailand. *Modeling Earth Systems and Environment*, 6:525–532, 3 2020.

[152] D. S. Wilks and R. L. Wilby. The weather generation game: a review of stochastic weather models. *Progress in Physical Geography: Earth and Environment*, 23(3):329–357, 1999.

[153] C. B. Williams. Sequences of wet and of dry days considered in relation to the logarithmic series. *Quarterly Journal of the Royal Meteorological Society*, 78(335):91–96, 1952.

[154] Hui Wu, Qiong Yue, Ping Guo, Xiaoyu Xu, and Xi Huang. Improving the AquaCrop model to achieve direct simulation of evapotranspiration under nitrogen stress and joint simulation-optimization of irrigation and fertilizer schedules. *Agricultural Water Management*, 266, 5 2022.

[155] Junzeng Xu, Wenhuan Bai, Yawei Li, Haiyu Wang, Shihong Yang, and Zheng Wei. Modeling rice development and field water balance using AquaCrop model under drying-wetting cycle condition inEastern China. *Agricultural Water Management*, 213:289–297, 3 2019.

[156] Juan Yin, Yingpan Yang, Rasu Eeswaran, Zhen Yang, Zhenghu Ma, and Fubin Sun. Irrigation scheduling for potatoes (Solanum tuberosum l.) under drip irrigation in an arid region using AquaCrop model. *Frontiers in Plant Science*, 14, 2023.

[157] S. S. Zanetti, E. F. Sousa, V. P. S. Oliveira, F. T. Almeida, and S. Bernardo. Estimating evapotranspiration using artificial neural network and minimum climatological data. *Journal of Irrigation and Drainage Engineering*, 133(2):83, March/April 2007.

[158] A. Zapranis and A. Alexandridis. Modelling the temperature time-dependent speed of mean reversion in the context of weather derivatives pricing. *Applied Mathematical Finance*, 15:355–386, 8 2008.

[159] Biying Zhai, Qiang Fu, Tianxiao Li, Dong Liu, Yi Ji, Mo Li, and Song Cui. Rice irrigation schedule optimization based on the AquaCrop model: Study of the Longtouqiao irrigation district. *Water (Switzerland)*, 11, 2019.

[160] Honghui Zhang, Yongnian Zeng, Xiaobin Jin, Bangrong Shu, Yinkang Zhou, and Xuhong Yang. Simulating multi-objective land use optimization allocation

using multi-agent system-a case study in Changsha, China. *Ecological Modelling*, 320:334–347, 1 2016.

[161] Zhuzhu Zhang, Haitao Mao, Yulin Liu, Peng Jiao, Wenlin Hu, and Pei Shen. A risk assessment method of aircraft structure damage maintenance interval indexed by: considering fatigue crack growth and detection rate. *Eksploatacja i Niezawodność – Maintenance and Reliability*, 1 2023.

[162] Huicheng Zhou, Hui Peng, and Chi Zhang. Lncs 4688 - an interactive fuzzy multi-objective optimization approach for crop planning and water resources allocation. *LNCS*, 4688:335–346, 2007.

[163] Özgür Kişi. Evapotranspiration estimation using feed-forward neural networks. *Nordic Hydrology*, 37:247–260, 2006.

[164] Sabrija Čadro, Mirza Uzunović, Jasminka Žurovec, and Ognjen Žurovec. Validation and calibration of various reference evapotranspiration alternative methods under the climate conditions of Bosnia and Herzegovina. *International Soil and Water Conservation Research*, 5:309–324, 12 2017.

Appendix A

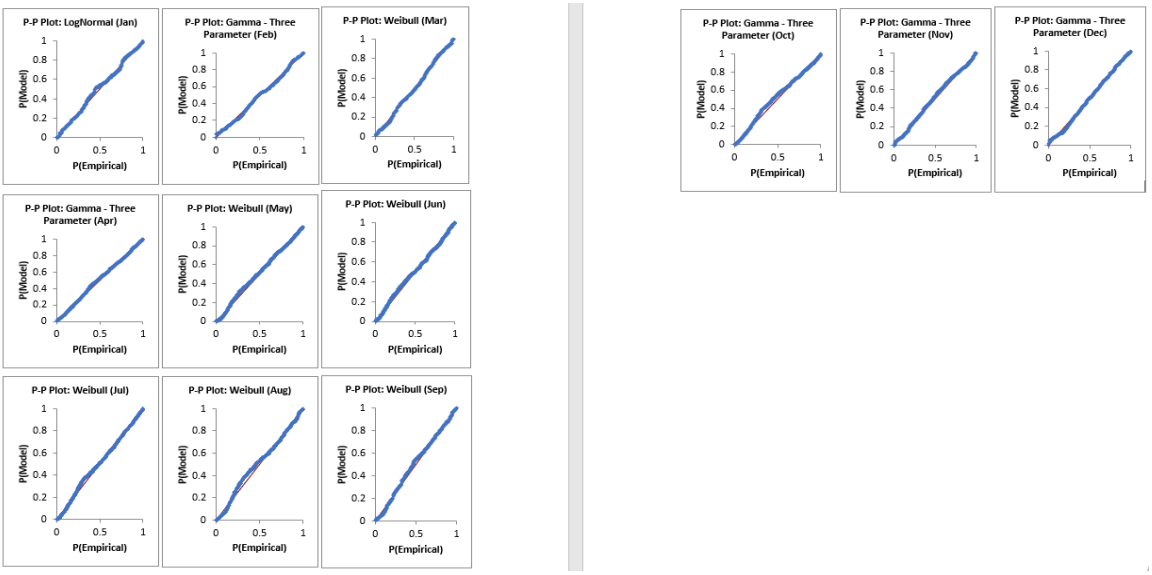


Fig. 7.1. PP-Plots of precipitation distribution at Rose Hall estate

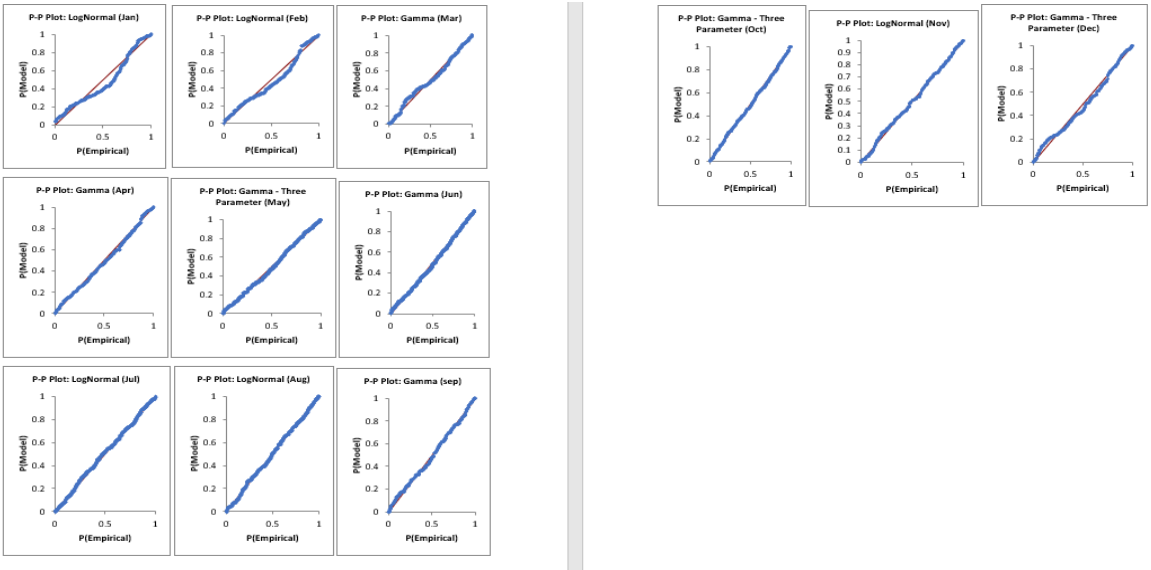


Fig. 7.2. PP-plots of precipitation distribution at Ebini

Table 7.1. Summary statistics for observed precipitation at Rose Hall from 1981-2022

Month	Maximum	Mean	STD	Q1	Q3	Rain days
Jan	233.016	25.715	34.651	7.291	26.127	276
Feb	146.951	17.066	17.655	7.462	18.801	257
Mar	140.665	16.222	20.211	4.057	21.762	260
Apr	77.883	19.786	14.344	9.528	25.667	318
May	81.007	18.469	12.491	9.143	24.940	594
Jun	100.732	17.406	12.318	8.650	23.113	691
Jul	77.907	18.876	12.616	10.369	23.584	631
Aug	77.797	16.874	11.691	9.346	21.069	520
Sept	56.857	11.517	10.541	6.218	15.351	338
Oct	48.931	8.955	7.758	4.305	11.776	358
Nov	81.027	15.462	12.285	7.644	18.915	348
Dec	128.347	22.040	20.323	8.064	28.906	350
Total						**4941**

Table 7.2. Summary statistics for observed precipitation at Ebini from 1981-2022

Month	Max	Mean	STD	Q1	Q3	wet days
Jan	236.903	17.204	28.440	2.165	17.712	423
Feb	76.817	10.358	12.438	1.662	13.005	353
Mar	127.811	13.614	17.169	2.422	19.290	357
Apr	83.892	14.318	14.790	3.496	19.523	481
May	74.105	14.200	10.840	6.316	19.834	775
Jun	86.009	13.587	10.535	6.230	18.250	890
Jul	116.766	15.482	12.840	6.539	21.411	859
Aug	61.088	11.852	9.517	5.321	15.963	734
Sep	51.493	9.670	8.360	3.469	13.756	542
Oct	45.910	8.279	7.354	2.732	11.366	525
Nov	59.239	9.540	9.585	2.548	13.719	514
Dec	94.116	16.245	18.349	2.677	23.124	471
Total						**6924**

Table 7.3. Summary statistics of simulated precipitation at Rose Hall 1981-2022

Month	Mean	Median	Max	Q1	Q3
Jan	24.518	14.708	281.124	7.508	29.381
Feb	15.586	12.036	100.412	6.893	20.647
Mar	17.884	10.170	137.700	3.593	23.158
Apr	20.981	17.332	78.211	10.131	27.970
May	16.789	14.355	74.523	7.571	22.781
Jun	17.940	15.595	103.541	9.155	24.045
Jul	20.117	16.445	115.240	10.395	25.220
Aug	17.586	13.887	85.091	9.195	23.115
Sep	10.754	9.431	44.594	5.910	14.276
Oct	9.028	7.736	40.478	4.142	12.276
Nov	15.250	11.414	78.272	6.991	17.626
Dec	21.398	15.954	126.095	9.008	28.485

Table 7.4. Summary statistics of simulated precipitation at Ebini 1981-2022

Month	Mean	Median	Max	Q1	Q3
Jan	17.231	6.060	208.231	2.107	17.720
Feb	10.523	5.583	95.092	1.618	13.998
Mar	14.642	7.368	127.379	2.237	20.220
Apr	14.080	9.169	117.002	3.240	19.297
May	14.498	11.769	71.631	5.634	20.835
Jun	13.600	10.863	62.079	5.349	19.220
Jul	15.907	12.154	95.380	5.906	21.537
Aug	12.098	9.565	60.719	4.078	17.217
Sep	9.504	7.121	48.934	3.410	13.138
Oct	8.216	5.959	46.483	2.720	11.682
Nov	9.953	6.798	61.544	2.242	13.914
Dec	16.519	10.013	126.347	3.357	23.432

Table 7.5. Summary statistics for simulated precipitation at Rose Hall from 2023-2030

Month	Mean	Median	STD	Max	Q1	Q3	Wet days
Jan	21.589	13.398	24.581	121.858	7.425	24.235	59
Feb	13.922	11.062	9.528	49.682	6.520	19.818	58
Mar	15.440	10.003	19.412	118.256	4.592	20.050	57
Apr	22.327	20.868	13.531	63.733	12.154	32.063	60
May	17.256	14.178	12.812	59.409	6.922	24.240	113
Jun	16.970	13.611	12.406	73.621	8.050	22.013	146
Jul	18.396	16.213	11.524	70.712	10.045	23.557	120
Aug	17.552	14.837	12.104	63.050	9.294	21.465	90
Sep	9.036	8.807	6.559	28.546	3.654	11.866	54
Oct	8.515	6.385	6.256	28.569	4.065	12.504	72
Nov	13.527	10.542	9.158	46.589	7.484	16.969	73
Dec	23.369	17.852	18.186	88.507	11.383	32.939	64

Table 7.6. Summary statistics for simulated precipitation at Ebini from 2023-2030

Month	Mean	Median	STD	Max	Q1	Q3	Wet days
Jan	16.859	5.283	29.464	169.149	2.288	15.526	95
Feb	8.868	4.726	11.339	43.386	1.443	10.465	64
Mar	15.248	8.314	19.945	95.994	2.036	20.830	58
Apr	14.577	8.817	17.176	83.699	3.009	19.213	85
May	13.846	10.105	11.368	66.679	5.328	19.778	139
Jun	14.627	12.619	9.677	58.276	7.719	20.203	180
Jul	15.001	12.221	11.782	63.587	5.841	22.232	177
Aug	12.321	10.245	9.709	46.013	4.382	17.589	130
Sep	9.676	7.079	8.508	38.110	2.591	14.177	105
Oct	6.867	4.295	6.980	29.253	1.948	9.171	97
Nov	9.640	6.253	8.842	35.548	2.644	16.270	101
Dec	15.726	8.314	18.602	86.662	3.556	21.414	90

Appendix B

Itô's Lemma Itô's Lemma is the chain rule of stochastic calculus and can be defined as follows: Suppose a stochastic variable X_t satisfies the following stochastic differential equation $dX_t = f(X_t, t)\,dt + g(X_t, t)\,dW_t$ Eq 6 and the given function $F = F(X_t, t)$ is a function with continuous partial derivatives $\frac{\partial F}{\partial t}$, $\frac{\partial F}{\partial X_t}$ and $\frac{\partial^2 F}{\partial X_t^2}$. Then by Itô's Lemma, $F = F(X_t, t)$ satisfies

$$\begin{aligned} dF &= \frac{\partial F}{\partial t}dt + \frac{\partial F}{\partial X_t}dX_t + \frac{1}{2}g^2(X_t, t)\frac{\partial^2 F}{\partial X_t^2}dt \\ &= \left(\frac{\partial F}{\partial t} + \frac{1}{2}g^2(X_t, t)\frac{\partial^2 F}{\partial X_t^2}\right)dt + \frac{\partial F}{\partial X_t}dX_t. \end{aligned} \tag{7.1}$$

Assuming

$$F = (\mu - X_t)\,e^{\lambda t}, \tag{7.2}$$

we have,

$$\frac{\partial F}{\partial t} = \lambda e^{\lambda t}(\mu - X_t) + e^{\lambda t}\frac{d\mu}{dt}, \frac{\partial F}{\partial X_t} = -e^{\lambda t}, \frac{\partial^2 F}{\partial X_t^2} = 0.$$

Using Itô's Lemma isometry, we can calculate a stochastic process's expected value and variance of a stochastic process.

$$E\left(\int_a^b \phi(t,\omega)dW_t\right) = 0, E\left[\left(\int_a^b \phi(t,\omega)dW_t\right)^2\right] = E\left[\int_a^b \phi^2(t,\omega)dW_t\right]. \tag{7.3}$$

By Itô's Lemma:

$$
\begin{aligned}
dF &= \left[\lambda e^{\lambda t}\left(\mu - X_t\right) + e^{\lambda t}\frac{d\mu}{dt}\right]dt - e^{\lambda t}dX_t \\
&= \lambda e^{\lambda t}\left(\mu - X_t\right)dt + e^{\lambda t}d\mu - e^{\lambda t}\left[\left\{\lambda\left(\mu - X_t\right) + \frac{d\mu}{dt}\right\}dt + \gamma dW_t\right] \qquad (7.4) \\
&= -e^{\lambda t}\gamma dW_t.
\end{aligned}
$$

Thus, $F(t) = F(0) + \int_0^t -e^{\lambda s}\gamma dW_s$, and the solution of the adjusted $\mathrm{O-U}$ process is

$$
X_t = \mu - e^{\lambda t}\left(\mu + X_0\right) + e^{-\lambda t}\int_0^t e^{\lambda s}\gamma dW_s
$$

From this solution, the expected value of X_t is

$$
E\left(X_t \mid X_0\right) = X_0 e^{-\lambda t} + \mu\left(1 - e^{-\lambda t}\right) = \mu + \left(X_0 - \mu\right)e^{-\lambda t}.
$$

Over some time, this conditional expectation will be approximately equal to the mean μ. In the above process, λ, μ, and γ need to be estimated from the data [102]. In a later section, we use the method of least squares to estimate the three parameters of the SDE.

Table 7.7. Summary statistics for min and max temperature in New Amsterdam

(a) Summary statistics for max temp

Months	Min	Max	Mean	STD
Jan	25.7	33.1	30.310	0.910
Feb	25	33	30.176	1.103
Mar	24.5	33.1	30.777	1.041
Apr	26	33.8	31.053	1.209
May	26.5	33.8	30.848	1.104
Jun	26.4	33.4	30.883	1.143
Jul	25.9	34.2	31.319	0.982
Aug	28.7	34.4	32.168	0.834
Sep	26.4	35.4	32.783	0.906
Oct	25.4	36.2	33.092	1.048
Nov	25.4	35	32.423	1.266
Dec	24.1	34.2	30.839	1.470

(b) Summary statistics for min temp

Min	Max	Mean	STD
18	25.5	22.976	1.304
18	25.6	23.051	1.246
19.4	26.5	23.522	1.217
19.6	26.5	23.770	1.210
19.7	26.4	23.892	1.174
19	26.1	23.534	1.154
19.3	25.9	23.322	1.086
16.1	26.2	23.620	0.980
20	26.7	24.071	1.106
20.9	26.8	24.142	1.136
19.5	26.8	23.910	1.231
15.1	26.5	23.412	1.317

Table 7.8. Summary statistics for relative humidity at New Amsterdam

Month	Min	Max	Mean	STD
Jan	57.446	98.785	77.556	6.822
Feb	60.475	97.53	76.911	6.576
Mar	56.245	100.144	75.256	6.723
Apr	57.957	101.004	76.82	6.569
May	62.39	103.909	80.461	5.388
Jun	70.119	95.56	81.519	4.507
Jul	69.191	95.835	80.56	3.994
Aug	66.315	90.475	78.365	3.644
Sep	61.308	95.781	75.566	4.524
Oct	58.764	95.195	74.018	5.126
Nov	63.037	100.977	75.884	5.569
Dec	64.526	103.857	79.131	6.743

Table 7.9. Summary statistics for min and max temperature at Ebini

(a) Summary statistics for max temp

Month	Min	Max	Mean	STD
Jan	24.933	31.972	29.144	0.887
Feb	25.781	33.333	29.418	1.012
Mar	25.258	34.141	29.996	1.059
Apr	25.402	34.608	30.192	1.152
May	25.941	32.825	29.885	1.148
Jun	25.488	32.799	29.571	1.054
Jul	26.39	31.816	29.763	0.921
Aug	26.318	32.783	30.568	0.894
Sep	27.358	34.619	31.518	0.989
Oct	26.626	35.647	31.821	1.233
Nov	27.443	34.569	30.847	1.151
Dec	26.246	32.59	29.519	0.945

(b) Summary statistics for min temp

Min	Max	Mean	STD
19.009	23.935	22.122	0.685
19.078	24.787	22.111	0.815
19.38	24.501	22.345	0.803
19.669	24.769	22.748	0.67
21.177	24.372	22.879	0.46
21.162	24.1	22.722	0.377
20.634	24.126	22.557	0.383
21.103	24.295	22.792	0.438
20.884	24.771	23.08	0.536
21.131	25.016	23.304	0.554
21.42	24.654	23.14	0.497
19.852	24.463	22.611	0.565

Table 7.10. Summary statistics for relative humidity at Ebini

Month	Min	Max	Mean	STD
Jan	63.084	92.500	77.033	4.141
Feb	61.784	88.740	74.877	4.611
Mar	60.520	89.635	73.439	4.609
Apr	60.566	91.117	75.607	5.234
May	61.954	93.589	79.508	4.704
Jun	70.473	92.891	81.860	3.549
Jul	68.172	91.601	81.678	3.222
Aug	67.729	91.665	79.951	3.297
Sep	61.810	88.910	76.268	4.427
Oct	56.119	90.169	74.360	5.193
Nov	62.251	88.447	76.678	4.504
Dec	66.667	90.285	78.708	3.439

Table 7.11. Temperature and humidity simulations from 2023-2030

(a) Temperature simulations

	New Amsterdam		Ebini	
	Sim Max Temp	Sim Min Temp	Sim Max temp	Sim Min temp
Mean	31.461	23.544	30.215	22.68
Median	31.468	23.551	30.216	22.674
STD	1.037	1.016	0.912	0.502
Min	27.658	20.073	26.711	20.998
Max	35.011	26.68	33.269	24.514

(b) Relative humidity simulations

	New Amsterdam	Ebini
	Sim humidity	Sim humidity
Mean	77.956	77.726
Median	78.262	78.064
STD	4.579	4.310
Min	61.473	61.247
Max	93.244	91.285

Table 7.12. Simulated vs observed temperatures 2021-2022 (New Amsterdam) and 2019-2022 (Ebini)

(a) New Amsterdam

	Sim Max	Obs Max	Sim Min	Obs Min
Mean	31.526	31.097	23.667	24.126
Q2	31.508	31.100	23.630	24.100
STD	1.060	1.439	1.001	0.845
Min	28.546	24.600	20.807	20.800
Max	34.569	34.900	27.171	26.800
RSME	1.858		1.351	
MAE	1.443		1.082	
MAPE	4.735		4.461	

(b) Ebini

	Stat	Sim Max	Obs Max	Sim Min
Obs Min				
Mean	30.251	30.068	22.702	22.849
Q2	30.236	30.070	22.714	22.893
STD	0.980	1.334	0.505	0.616
Min	26.968	25.998	21.049	19.182
Max	34.052	34.007	24.161	24.569
RSME	1.723		0.779	
MAE	1.381		0.6204	
MAPE	4.625		2.723	

(a) Simulated vs observed humidity at New Amsterdam and Ebini

	New Amsterdam		Ebini	
	Sim humidity	Obs humidity	Sim humidity	Obs humidity
Mean	77.2631	76.2731	77.7134	79.1692
Q2	77.616	75.7623	78.1617	79.4244
STD	4.637	5.4991	4.2773	4.9058
Min	62.1784	65.0859	63.4215	62.3673
Max	90.3773	95.9838	90.048	93.461
RSME	6.625		5.623	
MAE	5.324		4.417	
MAPE	7.024		5.565	

Appendix C

Table 7.14. Performance metrics of each method from 2019-2022 at New Amsterdam

Methods	RMSE	MAE	R^2	A	B
HS1	5.5712	5.5219	0.8303	1.738	2.152
HS2	0.5393	0.4624	0.8303	0.709	0.878
HS3	0.3008	0.2449	0.8416	0.685	0.008
HS4	0.3456	0.2739	0.8312	0.756	0.925
IV	0.9331	0.7244	0.5716	1.201	-1.504
JH	19.2129	19.0494	0.8695	4.252	4.191
LI	0.8275	0.7378	0.6877	0.612	2.502
IR1	0.4594	0.3717	0.8433	0.493	2.057
IR2	1.1197	1.0557	0.8074	0.518	1.146
IR3	0.8823	0.7938	0.8448	0.424	1.853
SC	1.3780	1.3164	0.6458	0.682	2.768
TR	1.4937	1.4629	0.84	0.673	2.957
TU	4.0473	3.9923	0.8604	0.031	0.437
AB	0.8548	0.7820	0.814	0.562	1.223
ANN	0.1373	0.123	0.9876	0.99	0.161

Table 7.15. The performance metrics of each method from 2013-2022 at Ebini

Methods	RMSE	MAE	R^2	A	B
HS1	6.0775	6.0319	0.7185	1.678	3.461
HS2	0.3844	0.3289	0.7185	0.685	1.413
HS3	0.7780	0.7134	0.7096	0.654	2.019
HS4	0.5637	0.4942	0.7187	0.73	1.492
IV	0.5232	0.4433	0.9139	1.201	-1.504
JH	18.5175	18.3688	0.7101	1.415	-1.955
LI	0.9832	0.9518	0.8777	0.669	2.207
IR1	0.3783	0.3087	0.7207	0.42	2.185
IR2	0.8987	0.8173	0.7125	0.435	1.234
IR3	0.6684	0.5454	0.7075	0.345	1.957
SC	1.6410	1.6285	0.9049	0.767	2.511
TR	2.104	2.0756	0.6664	0.626	0.3494
TU	3.2990	3.2478	0.7154	0.026	0.443
AB	0.6890	0.5859	0.7141	0.463	1.45
ANN	0.1017	0.0783	0.9711	0.989	0.044

Table 7.16. Descriptive measures of various methods

	Ebini				New Amsterdam			
method	Min	Max	Mean	STD	Min	Max	Mean	STD
PM 56	2.083	5.819	3.791	0.594	1.710	6.340	4.569	0.686
HS1	5.719	12.882	9.823	1.177	4.022	13.210	10.091	1.309
HS2	2.334	5.258	4.009	0.480	1.642	5.392	4.119	0.534
HS3	2.855	5.661	4.499	0.462	2.183	5.785	4.616	0.513
HS4	2.476	5.590	4.258	0.512	1.742	5.730	4.377	0.569
IV	0.973	6.601	3.408	0.880	0.441	6.831	3.983	1.090
JH	12.104	29.710	22.160	2.821	7.621	30.230	23.618	3.130
LI	3.593	6.535	4.743	0.424	3.288	7.062	5.300	0.507
IR1	2.724	4.529	3.778	0.294	2.508	5.151	4.307	0.368
IR2	1.859	3.747	2.973	0.306	1.591	4.498	3.513	0.396
IR3	2.379	3.855	3.264	0.244	2.222	4.503	3.790	0.317
SC	4.227	7.497	5.419	0.479	3.925	7.613	5.885	0.583
TR	4.247	6.943	5.866	0.456	3.657	7.090	6.032	0.504
TU	0.479	0.592	0.543	0.018	0.468	0.626	0.577	0.023
AB	2.045	4.038	3.205	0.326	1.768	4.844	3.790	0.427
ANN	2.051	5.722	3.794	0.597	1.441	6.284	4.683	0.684

Appendix D

Table 7.17. AquaCrop calibrated parameters used for simulations.

PARAMETER	Rice (GRDB10)	Corn (LUDAN510)	Soybean (MSOY9144)
type of planting method	direct sowing	direct sowing	direct sowing
Initial canopy cover	6.05%	1.50%	1.43%
Days to max Canopy Cover	64	50	65
Days to senescence	91	82	102
maximum canopy cover (CCx)	95%	90%	99%
Weed coverage	$10-25\%$	10	10%
Days to emergence	21	7	9
Days to maturity	112	105	115
Duration of flowering	18	14	62
Days to flowering	77 days	80	44
max effective rooting depth	0.5 m	1 m	0.6 m
Days to max root depth	44 days	78	50
Reference harvest	44%	40%	40%
Irrigation method	Basin and rainfed	rainfed	rainfed
Planting Density	201.8	22.9	286

Source: [95, 119, 123]

Table 7.18. Irrigated rice yields under simulated and observed weather at Rose Hall

year	INT1	obs	INT2	obs	INT3	obs	INT4	Obs	INT5	Obs	Stat
2014	6.35	6.30	6.30	6.18	6.23	5.99	6.22	5.89	6.28	5.87	
2015	6.38	6.22	6.25	6.19	6.15	6.19	6.07	6.18	6.11	6.16	
2016	6.38	6.42	6.39	6.43	6.39	6.35	6.38	6.21	6.35	6.10	
2017	6.45	6.45	6.44	6.44	6.44	6.44	6.45	6.44	6.44	6.43	
2018	6.40	6.48	6.38	6.46	6.32	6.42	6.30	6.32	6.26	6.12	
2019	6.51	6.39	6.50	6.36	6.48	6.39	6.46	6.42	6.48	6.47	
2020	6.52	6.54	6.52	6.51	6.52	6.37	6.51	6.23	6.50	6.13	
2021	6.55	6.55	6.55	6.55	6.56	6.55	6.56	6.56	6.52	6.56	
2022	6.50	6.59	6.48	6.58	6.51	6.58	6.57	6.57	6.57	6.57	
	0.08		**0.08**		**0.11**		**0.16**		**0.21**		RSME
	0.06		**0.06**		**0.08**		**0.11**		**0.14**		MAE
	0.97		**1.01**		**1.31**		**1.74**		**2.34**		MAPE
2023	6.58		6.57		6.55		6.55		6.54		
2024	6.62		6.56		6.55		6.58		6.64		
2025	6.64		6.56		6.52		6.51		6.59		

Table 7.19. Rainfed rice yields under simulated and observed weather Rose Hall

year	INT1	obs	INT2	obs	INT3	obs	Stat
2014	6.33	6.03	6.27	5.73	6.12	5.19	
2015	6.37	6.10	6.25	5.68	5.96	5.16	
2016	6.34	6.34	6.28	6.30	6.12	6.21	
2017	6.44	6.35	6.45	6.21	6.43	5.98	
2018	6.31	6.46	6.04	6.41	5.65	6.17	
2019	6.51	6.29	6.45	6.00	6.33	5.62	
2020	6.53	6.59	6.53	6.55	6.51	6.33	
2021	6.55	6.54	6.54	6.56	6.58	6.55	
2022	6.37	6.65	6.11	6.64	5.92	6.60	
	0.19		**0.38**		**0.58**		RSME
	0.15		**0.31**		**0.49**		MAE
	2.45		**5.07**		**8.57**		MAPE
2023	6.49		6.37		6.13		
2024	6.62		6.56		6.45		
2025	6.56		6.39		6.08		

Table 7.20. Soy yields under simulated and observed weather at Rose Hall

year	INT3	obs	INT4	Obs	INT5	Obs	Stat
2014	2.675	2.625	2.545	2.544	2.675	2.404	
2015	2.562	2.464	2.701	2.319	2.562	2.132	
2016	2.519	2.780	2.538	2.750	2.519	2.642	
2017	2.807	2.808	2.531	2.821	2.807	2.808	
2018	2.384	3.029	2.550	2.890	2.384	2.707	
2019	3.240	2.474	3.231	2.344	3.240	2.245	
2020	2.976	2.754	2.829	2.842	2.976	2.760	
2021	2.998	3.085	2.877	3.262	2.998	3.328	
2022	2.447	2.193	2.554	2.448	2.447	2.712	
	0.366		**0.385**		**0.421**		RSME
	0.242		**0.291**		**0.328**		MAE
	9.989		**11.194**		**13.320**		MAPE
2023	2.900		2.910		2.960		
2024	2.672		2.825		2.923		
2025	3.188		3.192		3.079		

Table 7.21. Corn yields under simulated and observed weather at Rose Hall

year	INT6	obs	INT7	obs	INT8	obs	INT9	Obs	INT10	Obs	Stat
2014	9.57	8.82	9.69	8.60	9.79	8.67	9.78	8.85	9.73	9.08	
2015	9.75	9.64	9.76	9.86	9.80	10.02	9.79	10.07	9.78	9.93	
2016	7.98	9.31	7.75	9.06	7.68	8.73	7.76	8.60	7.93	8.67	
2017	8.65	8.77	8.73	8.97	8.78	9.08	8.81	9.03	8.84	9.04	
2018	9.51	9.76	9.21	9.84	8.92	9.92	8.85	10.01	8.90	9.97	
2019	8.34	9.71	8.40	9.85	8.51	9.93	8.63	9.81	8.80	9.39	
2020	9.31	7.53	9.29	7.71	9.16	7.83	9.02	8.02	8.98	8.32	
2021	9.21	9.05	9.03	9.18	8.96	9.42	8.94	9.66	9.17	9.81	
	0.97		1.00		0.97		0.86		0.65		RSME
	0.73		0.81		0.86		0.79		0.59		MAE
	8.43		9.20		9.59		8.64		6.36		MAPE
2023	8.42		8.41		8.53		8.79		9.09		
2024	9.33		9.14		8.88		8.67		8.50		
2025	9.09		9.09		8.81		8.13		7.85		

Table 7.22. Soy and corn yields under simulated and observed weather at Ebini

year	INT1	obs	INT2	obs	INT3	obs	INT4	Obs	INT5	Obs	Stat	crop
2014	3.65	3.62	3.62	3.48	3.55	3.23	3.44	2.96	3.33	2.74		
2015	3.67	3.64	3.67	3.54	3.67	3.36	3.67	3.11	3.67	2.79		
2016	3.69	3.69	3.69	3.69	3.69	3.66	3.67	3.63	3.62	3.54		
2017	3.71	3.71	3.71	3.71	3.71	3.71	3.71	3.70	3.71	3.63		
2018	3.72	3.72	3.72	3.72	3.72	3.72	3.72	3.65	3.67	3.51		Soy
2019	3.74	3.74	3.74	3.74	3.74	3.74	3.74	3.74	3.74	3.74		
2020	3.76	3.76	3.76	3.76	3.76	3.74	3.75	3.68	3.66	3.50		
2021	3.77	3.77	3.77	3.77	3.74	3.77	3.71	3.77	3.68	3.77		
2022	3.79	3.78	3.79	3.78	3.79	3.78	3.79	3.78	3.78	3.79		
	0.01		0.06		0.15		0.25		0.36		RSME	
	0.01		0.03		0.08		0.14		0.23		MAE	
	0.20		0.86		2.34		4.55		7.70		MAPE	
2023	3.80		3.80		3.80		3.80		3.76			
2024	3.81		3.81		3.80		3.74		3.69			
2025	3.73		3.68		3.58		3.45		3.37			
2014	9.82	9.82	9.82	9.82	9.82	9.82	9.82	9.84	9.84	9.78		
2015	9.83	9.83	9.83	9.83	9.83	9.84	9.83	9.87	9.83	9.89		
2016	9.84	9.86	9.84	9.87	9.84	9.88	9.84	9.88	9.84	9.87		
2017	9.85	9.85	9.85	9.85	9.85	9.85	9.85	9.85	9.85	9.85		Corn
2018	9.86	9.86	9.86	9.86	9.86	9.86	9.86	9.86	9.86	9.86		
2019	9.87	9.87	9.87	9.87	9.87	9.87	9.87	9.87	9.87	9.88		
2020	9.88	9.88	9.88	9.88	9.88	9.88	9.88	9.88	9.88	9.89		
2021	9.89	9.89	9.89	9.89	9.90	9.89	9.92	9.89	9.94	9.89		
2022	9.90	9.90	9.90	9.90	9.90	9.90	9.90	9.91	9.90	9.91		
	0.004		0.009		0.014		0.021		0.033		RSME	
	0.002		0.003		0.007		0.014		0.023		MAE	
	0.016		0.034		0.070		0.141		0.236		MAPE	
2023	9.909		9.909		9.909		9.909		9.909			
2024	9.880		9.880		9.883		9.917		9.920			
2025	9.941		9.958		9.978		9.967		9.994			

Table 7.23. Rice yields under simulated and observed weather at Ebini

year	INT1	obs	INT2	obs	INT3	obs	Stat
2014	3.94	3.95	3.894	3.751	3.788	3.37	
2015	3.99	3.89	3.978	3.771	3.95	3.57	
2016	3.98	3.98	3.985	3.983	3.991	3.97	
2017	4	4.02	3.999	4.036	4.013	4.04	
2018	4.02	4.04	4.018	4.042	4.015	4.02	
2019	4.04	4.03	4.036	4.029	4.036	4.01	
2020	4.05	4.06	4.071	4.05	4.07	4.01	
2021	4.03	4.07	3.937	4.07	3.847	4.07	
2022	4.09	4.1	4.1	4.094	4.111	4.09	
	0.04		0.096		0.205		RSME
	0.02		0.064		0.131		MAE
	0.58		1.658		3.602		MAPE
2023	4.1		4.101		4.102		
2024	4.11		4.112		4.103		
2025	3.86		3.799		3.675		

Appendix E

Table 7.24. Planting windows

Crop	Planting -window	location	# of intervals
Rice (irrigated)	May15th-Jun30th	Rose Hall	5
Rice(rainfed)	May 15th-Jun15th	Rose Hall	3
Soy	May30th-Jun30th	Rose Hall	3
Corn	Oct15th - Nov30th	Rose Hall	5
Rice	May 15th-June 15th	Ebini	3
Soy	May 15th-June 30th	Ebini	5
Corn	May 15th-June 30th	Ebini	5

Table 7.25. Planting intervals

Planting intervals	Planting dates in interval
May15th-May30th	M15, M23, M30
May23rd-Jun7th	M23, M30, J7
May30th-Jun15th	M30, J7, J15
jun7th-jun23rd	J7, J15, J23
jun15th-Jun30th	J15, J23, J30
Oct15th-Oct30th	O15, O23, O30
Oct23th-Nov7th	O23, O30, O7
Oct30th-Nov15th	O30, N7, N15
Nov7th-Nov23rd	N7, N15, N23
Nov15th-Nov30th	N15, N23, N30

Table 7.26. Domestic price for rice, soy and corn per ton

Crop	cost per ton (GYD)	cost per ton (USD)
rice (2024)	130,000	637.25
rice (2019)	90,000	441.18
dried soy (2021)	110,000	539.22
dried soy (2024)	140,000	686.27
dried corn (2021)	40,000	196.07
dried corn (2024)	55,000	269.608

Table 7.27. Five best and worst projected crop combinations for rice 2024/25 growing season under 50% export at Rose Hall for irrigated and rainfed rice

Crop 1	Crop 2	Area 1	Area 2	profit	EEWP
rice1	corn5	2500	2500	4216.05	0.183
rice2	corn5	2500	2500	4161.84	0.181
rice3	corn5	2500	2500	4067.63	0.177
rice1	corn4	2500	2500	4059.15	0.175
rice2	corn4	2500	2500	4004.93	0.173
soy2	corn1	2500	2500	1733.67	0.077
soy1	corn4	2500	2500	1679.61	0.079
soy1	corn3	2500	2500	1612.37	0.075
soy1	corn2	2500	2500	1539.05	0.071
soy1	corn1	2500	2500	1505.4	0.069
riceIR5	**corn5**	**2500**	**2500**	**3282.16**	**0.14**
riceIR1	corn5	2500	2500	3261.88	0.141
riceIR4	corn5	2500	2500	3220.17	0.138
riceIR2	corn5	2500	2500	3212.59	0.140
riceIR3	corn5	2500	2500	3201.31	0.138
soy2	corn1	2500	2500	1733.67	0.077
soy1	corn4	2500	2500	1679.61	0.079
soy1	corn3	2500	2500	1612.37	0.075
soy1	corn2	2500	2500	1539.05	0.071
soy1	corn1	2500	2500	1505.4	0.069

Table 7.28. Five best and worst projected crop combinations for 2024 under 25% export in Ebini

Crop 1	Crop 2	Area 1	Area 2	profit	EEWP
soy3	corn5	1794	706	2392.912	0.221
soy3	corn4	1794	706	2369.911	0.219
soy1	corn5	1793	707	2367.939	0.221
rice2	soy3	1497	1003	1641.885	0.0.149
rice1	soy3	1498	1002	1639.796	0.150
rice3	soy3	1500	1000	1638.272	0.0.148
rice2	corn5	1497	1003	1178.101	0.113
rice1	corn5	1498	1002	1176.471	0.114
rice3	corn5	1500	1000	1176.113	0.110
rice2	corn2	1497	1003	1118.419	0.109

Table 7.29. Five best and worst projected crop combinations for 2024 under 50% export in Ebini

Crop 1	Crop 2	Crop 1	Crop 2	profit	EEWP
soy3	corn5	1793	707	1940.81	0.179
soy3	corn4	1794	706	1918.26	0.177
soy2	corn5	1794	706	1916.06	0.178
rice2	soy3	1498	1002	1427.75	0.129
rice1	soy3	1498	1002	1426.563	0.130
rice3	soy3	1500	1000	1420.49	0.128
rice2	corn5	1497	1003	928.495	0.089
rice1	corn5	1498	1002	927.05	0.09
rice3	corn5	1500	1000	922.48	0.088
rice2	corn1	1497	1003	881.152	0.086

Table 7.30. Datasheet used for double cropping simulation in Rose Hall with Rainfed rice

crop	export	domesticPrice	variableCost	yield	Etwater	demand	exportProp
rice1	589.52	631.81	1297.71	4.30	5425.56	4000	0.25
rice2	589.52	631.81	1294.66	4.26	5418.69	4000	0.25
rice3	587.69	634.53	1291.20	4.19	5418.44	4000	0.25
soy1	499.53	678.716	1196.57	2.67	4683.26	2000	0.25
soy2	502.66	678.716	1199.62	2.83	4875.04	2000	0.25
soy3	502.66	682.49	1196.81	2.92	4988.78	2000	0.25
corn1	197.03	275.729	1746.93	8.34	4093.47	7000	0.25
corn2	197.03	275.729	1748.52	8.40	4044.77	7000	0.25
corn3	194.67	277.769	1741.87	8.51	3949.11	7000	0.25
corn4	194.67	277.769	1745.05	8.63	3858.27	7000	0.25
corn5	196.55	279.89	1739.66	8.80	3795.97	7000	0.25

Table 7.31. Datasheet used for double cropping simulation in Rose Hall with irrigated rice

crop	export	domesticPrice	variableCost	yield	Etwater	demand	exportProp
riceIR1	589.52	631.81	1679.38	4.30	5479.37	4000	0.25
riceIR2	589.52	631.81	1676.60	4.27	5399.04	4000	0.25
riceIR3	587.69	634.53	1679.97	4.26	5477.59	4000	0.25
riceIR4	587.69	634.53	1681.03	4.28	5526.35	4000	0.25
riceIR5	587.69	637.25	1687.93	4.32	5580.98	4000	0.25
soy1	499.53	678.716	1196.57	2.67	4683.26	2000	0.25
soy2	502.66	678.716	1199.62	2.83	4875.04	2000	0.25
soy3	502.66	682.49	1196.81	2.92	4988.78	2000	0.25
corn1	197.03	275.729	1746.93	8.34	4093.47	7000	0.25
corn2	197.03	275.729	1748.52	8.40	4044.77	7000	0.25
corn3	194.67	277.769	1741.87	8.51	3949.11	7000	0.25
corn4	194.67	277.769	1745.05	8.63	3858.27	7000	0.25
corn5	196.55	279.89	1739.66	8.80	3795.97	7000	0.25

Table 7.32. Datasheet used for polyculture simulations in Ebini

crop	export	domesticPrice	variableCost	yield	Etwater	demand	exportProp
rice1	589.52	631.81	1292.19	2.67	4281.52	4000	0.25
rice2	589.52	631.81	1292.33	2.67	4343.77	4000	0.25
rice3	587.69	634.53	1293.90	2.67	4371.21	4000	0.25
soy1	499.53	678.716	1341.25	3.81	4448.00	2000	0.25
soy2	499.53	678.716	1341.25	3.81	4482.35	2000	0.25
soy3	502.66	682.49	1335.80	3.80	4526.44	2000	0.25
soy4	502.66	682.49	1334.44	3.74	4571.01	2000	0.25
soy5	504.49	685.89	1328.15	3.69	4588.00	2000	0.25
corn1	197.03	275.729	1963.99	9.88	3677.61	7000	0.25
corn2	197.03	275.729	1963.99	9.88	3718.58	7000	0.25
corn3	194.67	277.769	1953.94	9.88	3766.75	7000	0.25
corn4	194.67	277.769	1954.87	9.92	3828.15	7000	0.25
corn5	196.55	279.89	1944.40	9.92	3868.27	7000	0.25