A stochastic optimization model for efficient water management in cocoa crop

Juan David Márquez González

Degree study to obtain the title of Master in Industrial Engineering

First advisor
Leonardo Hernán Talero Sarmiento
M.Sc. Industrial Engineering

Co-Advisor
Henry Lamos Díaz
Ph.D. in Physics and Mathematics

Universidad Industrial de Santander
Facultad de Ingenierías Físico - mecánicas
Escuela de Estudios Industriales y Empresariales
Master in Industrial Engineering
Bucaramanga

2024

Dedication

I dedicate this research to my beloved and amazing parents Laura Mercedes Gonzalez and Gilberto Marquez whose unwavering love, support, and encouraging words guide me through this goal accomplishment. Definitely, words will never be enough to express how indispensable they are in every achievement of my life.

To Estefi, my steadfast support and personal pillar; to Andrés and Daniel, my dear and beloved brothers; to my cousin Miguel José for helping me clear my mind; to my aunt Merly and uncle Elkin for always being there for me; to my grandmother Julia, a voice of peace in the journey of my life; to my beloved nephew, my source of inspiration; to my dear sisters-in-law Viviana and Silvia for their constancy and support: thank you very much for being part of this process and standing by me in every step I take. I deeply love you all.

I wish to express my profound gratitude to Professor Leonardo Hernán Talero, whose wise guidance has been steadfast at every step of this process. As I conclude this stage, I can only convey my gratitude for encouraging me to embark on this journey, for guiding me diligently through this passage, and above all, for never doubting the outcome. His clarity in resolving doubts, skill in directing my frustrations and anxieties, stand as a testament to exceptional leadership. Infinite gratitude for his role as a teacher, colleague, friend, and family. Thank you. To Professor Henry Lamos for his wisdom and unwavering support at every stage of the process.

To my friends, especially Yessica Corzo, Juan Felipe Heredia, and Edgar Eduardo Córdoba, who were pivotal in bringing closure to this life project. With their words, faith, ideas, and above all, affection, they made the days lighter, allowing me to relish the journey. Thank you all; you are unique and continuously impact my life in unimaginable ways.

To my forever friends, who will always be there; to my best friend Aleja, my darling, you are and will always be a reason for self-improvement and personal growth, an unwavering support, and the epitome of friendship; to Lau, Dur, Martha, and Leo, because even as each of you strives and fights for your goals, I felt your support and love at every moment.

To the babies Gess and Pao, I never looked for them, but I always needed them, and now that they accompany me and are part of my life, I want nothing else but to continue admiring our achievements and building our lives together. And to all my loved ones, thank you; you are everything I have in life.

I also express my gratitude to the *Industrial University of Santander*, particularly to the *Escuela de Estudios Industriales y Empresariales* (EEIE), and to all those involved in the Master's in Industrial Engineering program, including all professors and colleagues.

Contents

List of Tables

List of figures

List of Appendices

Every appendix is available through: https://github.com/JuanDM0106/Master-research.git and they are attached.

Appendices	Description
Appendix A. Uncertain parameter modeling notebook	A notebook that relates the second objective calculations
Appendix B. Water Balance	A document that related the water balance theory
Appendix C. Yield, Ky, Swater factor	A document that relates the yield and factors estimation theory
Appendix D. Yield, Ky, Swater factor	A document that relates the yield and factors estimation
Appendix E. Evapotranspiration	A document that relates the evapotranspiration theory
Appendix F. Evapotranspiration estimation	A document that relates the document that relates the evapotranspiration estimation
Appendix G. Article	A document that relates the methodological framework built based on the SLR results
Appendix H. Data sources	A document that relates the main database used in this study
Appendix I. NASA POWER 1981-2020	A document that contains the parameter data from 1981 to 2020
Appendix J. Scenarios of each strategy	A file that contains the scenarios acquired in the simulation
Appendix K. Aggregate scenarios of each strategy	A file that contains every dataset of scenarios aggregated
Appendix L. Aggregation of water balance periods	A notebook that relates the procedure applied for data aggregation
Appendix M. Optimization models solutions	A file that contains the solution obtained after every model run
Appendix N. Deterministic and TSP models	A notebook that relates the models programmed and solved
Appendix O. Models parameters data	A file that contains the dataset required to run every model
Appendix P. NASA POWER 2021	A document that contains the parameter data in 2021
Appendix Q. Productive farm units	A document that relates the estimation of the productive farms average size in San Vicente
Appendix R. EVPI and VSS	A document that relates the results of metrics application
Appendix S. PAR and IPAR estimation	A document that relates the PAR and IPAR estimation
Appendix T. First conference paper	It relates the preliminary findings of the SLR
Appendix U. Second conference paper	It relates a TSP case study regarding price uncertainty

Acronyms

N	Name	Acronym
1	Artificial Neural Networks	ANN
2	Augmented Dickey-Fuller test	ADF
3	Autoregressive Integrated Moving Average	ARIMA
4	Capillary Rise	CR
5	Chance Constraint Programming	CCP
6	Cocoa Crop Water Management	CCWM
7	Crop Evapotranspiration	ETc
8	Decision-makers	DM
9	Drainage	D
10	Expected Value of Perfect Information	EVPI
11	Fuzzy Parameter	FPa
12	Fuzzy Programming	FP
13	Interval Parameter	IP
14	Interval Parameter Programming	IPP
15	Irrigation	I
16	Mathematical Programming	MP
17	Multistage Stochastic Programming	MSP
18	Partial Autocorrelation Function	PACF
19	Phillips-Perron	PP
20	Precipitation	P
21	Probability Distribution Function	PDF
22	Random Parameter	RP
23	Reference Evapotranspiration	ETo
24	Robust Programming	RP
25	Solar Radiation	Rs
26	Stochastic Programming	SP
27	Total Available Water	TAW
28	Two-Stage Stochastic Programming	TSP
29	Value of stochastic solution	VSS
30	Water Use Efficiency	WUE
31	Web of Science	WOS

Resumen

Título: Un modelo de optimización estocástica para la gestión eficiente del agua en el cultivo de cacao[*]

Autor: Juan David Márquez González[**]

Palabras clave: Programación Estocástica Bi-etapa, Incertidumbre climática, Generación de Escenarios, Gestión Hídrica Agrícola, Theobroma Cacao L

Descripción:

La producción agrícola desempeña un papel crucial en los países en desarrollo, especialmente a medida que los recursos se vuelven escasos, los hábitos alimentarios cambian y las poblaciones crecen. En Colombia, el cultivo de cacao se destaca como un producto principal, proporcionando una alternativa a los cultivos ilícitos y generando empleo para miles de familias. Sin embargo, soportar la producción de cacao en Colombia es desafiante debido a la incertidumbre asociada con diversos factores productivos, así como recursos naturales como el agua un recurso fundamental y limitado en la producción agrícola. La gestión adecuada del agua es esencial para garantizar la productividad del cultivo de cacao, pero la incertidumbre climática y los efectos del clima extremo plantean desafíos adicionales. Debido a lo anterior, es crucial desarrollar estudios que permitan apoyar la toma de decisiones en la gestión de recursos hídricos agrícolas. Estrategias como la programación matemática representan alternativas adecuadas para tomar la mejor decisión bajo restricciones. Por lo tanto, este estudio se centra en desarrollar un modelo de optimización estocástica para la gestión del agua en cultivos de cacao, considerando la incertidumbre climática. Se identifican brechas de investigación, incluida la falta de enfoque en la gestión del agua para cultivos específicos y la subutilización de técnicas de optimización. Se propone un marco metodológico para futuras investigaciones y se demuestra que los modelos estocásticos superan a los deterministas, proporcionando una base sólida para la toma de decisiones informadas en la gestión de recursos hídricos agrícolas, especialmente en el caso del cacao.

[*] Trabajo de grado

[**] Facultad de Ingenierías Fisicomecánicas. Escuela de Estudios Industriales y Empresariales (EEIE). Maestría en Ingeniería Industrial. Director: Leonardo Hernán Talero Sarmiento. Magister en Ingeniería Industrial. Codirector: Henry Lamos Diaz. Doctor en física y matemáticas.

Abstract

Title: A stochastic optimization model for efficient water management in cocoa crop[*]

Autor: Juan David Márquez González[**]

Palabras clave: Two-stage Stochastic Programming, Climate Uncertainty, Scenario Generation, Agricultural Water Management, Theobroma Cacao L

Description:

Agricultural production plays a crucial role in developing countries, especially as resources become scarce, eating habits change, and populations grow. In Colombia, cocoa cultivation stands out as a primary product, providing an alternative to illicit crops and creating employment for thousands of families. However, supporting cocoa production in Colombia is challenging due to the uncertainty associated with various production factors, as well as natural resources like water, which is fundamental and limited in agricultural production. Proper water management is essential to ensure cocoa crop productivity, but climatic uncertainty and the effects of extreme weather pose additional challenges. Therefore, it is crucial to develop studies that support decision-making in agricultural water resource management. Strategies such as mathematical programming represent suitable alternatives for making the best decision under constraints. Thus, this study focuses on developing a stochastic optimization model for water management in cocoa crops, considering climatic uncertainty. Research gaps are identified, including the lack of focus on water management for specific crops and the underutilization of optimization techniques. A methodological framework for future research is proposed, demonstrating that stochastic models outperform deterministic ones, providing a solid foundation for informed decision-making in agricultural water resource management, especially in the case of cocoa.

[*] Degree study

[**] Facultad de Ingenierías Físicomecánicas. Escuela de Estudios Industriales y Empresariales (EEIE). Master in Industrial Engineering. First advisor: Leonardo Hernán Talero Sarmiento. Master in Industrial Engineering. Co-advisor: Henry Lamos Diaz. Ph.D.in Physics and Mathematics.

Introduction

Agricultural production satisfies several needs in developing countries and is essential as resources become scarce, eating habits change, and populations increase. In Colombia, the cocoa crop represents a major product because it represents an alternative to illicit crops (Fedecacao, 2021; Minagricultura, 2016). It generates around 165,000 jobs and allows 52,000 families employability improving farmers' living conditions (Minagricultura, 2020). Furthermore, proper cocoa crop production meets several Agenda 2030 goals (United Nations, 2015). Nevertheless, improving cocoa crop production is complex due to multiple random processes related to productive factors. Within these factors, water represents a crucial production factor, limited and necessary to guarantee productivity in crop production (Ali & Talukder, 2008; Fedecacao, 2018; P. Guo et al., 2010). It is additionally related to high levels of uncertainty due mainly to extreme weather effects that influence the availability increasing the level of risk in decision-making for proper crop water management (Fu, Li, Li, et al., 2018; Kang et al., 2004; Y. Wang et al., 2020). Consequently, the importance of developing studies to support appropriate decision-making on water resource management is evident, for several decades through various optimization approaches (Y. Y. Wang et al., 2017).

Considering optimization approaches, Mathematical Programming (MP) offers a compilation of tools for agricultural water management. Such tools improve water management efficiency and irrigation activities, considering crop demands and resource availability constraints. In MP, deterministic modeling considers crisp parameters or invariant factors, which in most research areas (e.g., water management) is incorrect, providing improper decision-making in uncertain conditions (Muhammad & Pflug, 2014). Consequently, it is critical to integrate uncertainty analysis techniques into optimization methods (W. J. Zhang et al., 2021). Within these techniques, Interval Parameter Programming (IPP) addresses uncertainty using possible parameter bounds where the phenomenon is likely to exist or to be realized with unknown distribution (Robers & Ben-Israel, 1969). Fuzzy Programming (FP) supports uncertain optimization modeling regarding ambiguity or vagueness conditions, demanding expert contribution for defining parameters' behavior through fuzzy membership functions (Bellman & Zadeh, 1970). On the other hand, Stochastic Programming (SP), specifically multistage approaches, focuses on probability theory to model uncertain parameters as random variables setting probabilities considering the scenarios used (Dantzig, 1955). SP answers to more realistic situations and gives decision ability regarding the scenarios (Jamal et al., 2018). However, considering the mentioned advantages, the SP approach is not covered in the Cocoa Crop Water Management (CCWM) literature. Thus, it is pertinent to develop studies that use such techniques in agricultural water resource management to support a better-informed, well-structured, and data-driven decision process.

This study aims to develop a two-stage stochastic optimization model under climatic uncertainty to support decision-making regarding proper water resource management in cocoa crops. In this sense, this research supports water management decision-making considering climatic uncertainty modeled through scenarios in a stochastic optimization framework, reducing water overuse and improving the farmers' benefit. This study outlines the following. Section 1 presents the problem statement; later, Section 2 and Section 3 present the research objectives and hypothesis. Section 4 displays the study methodology. Section 5 relates the theoretical framework. Section 6 shows the developed literature review. Section 7 presents the case study addressed, displaying the related research results. Finally, Section 8 and Section 9 contain the discussion and conclusions of the study results, respectively.

1. Problem Statement

Cocoa is one of Colombia's most prominent agricultural products, also called "The Peace crop." It represents a relevant product for the nation since it supports as a substitute alternative for illicit crops in the post-conflict era (Fedecacao, 2021; Minagricultura, 2016), enabling the creation of 173,300 jobs, with the benefit and employment of around 65,341 families, and enhancing the overall quality of farmers life conditions (Organización Internacional del trabajo, 2023). Colombia is the fifth producer of cocoa in Latin America (CAF, 2020), going from producing 42,294 tons in 2010 to 63,048 tons in 2020, which represents a significant result in crop production (Finagro, 2020), where Santander is the highest productive department with 41% (26,315 tons) of the national production contribution (Fedecacao, 2020). Besides, cocoa allows for meeting the Agenda 2030 goals considering its potential to end poverty, end hunger, achieve food security, conserve and use sustainable resources, and promote peace and inclusive societies to ensure sustainable development (United Nations, 2015).

However, ensuring and improving cocoa crop production is increasingly complex due to several random factors and their interactions, causing high uncertainty, increasing the decision-making complexity in cocoa crop production, and decreasing the crop yield (Carr & Lockwood, 2011; Chapman et al., 2021a; Cilas & Bastide, 2020; Finagro, 2018; Plazas et al., 2017). Within these factors, natural resources play a central role in promoting sustainable agricultural development, in which water represents a renewable production factor, limited and necessary to guarantee productivity in crop production (Fedecacao, 2018; P. Guo et al., 2010). Nevertheless, water is related to high levels of uncertainty, considering the extreme weather effects that influence the availability of water resources, population growth, and the increase in human activities, which impacts the demand and supply relationship capacity of different users (Fu, Li, Li, et al., 2018;

Kang et al., 2004; Y. Wang et al., 2020). Such situations increase the level of risk in decision-making for proper crop water management.

In the last two decades, diverse irregular weather patterns, specifically extreme droughts, have affected cocoa crop production in Colombia (Caracol Radio, 2016, 2020; La opinión, 2015; Noticias Canal TRO, 2019; Portafolio, 2009; Semana, 2014) and internationally (Inter Press Service, 2020; Unión Europea, 2017). Such a situation worsens considering the change in the climatic conditions expected for future years, considering its direct relationship with the crop productivity (Agbenyo et al., 2022; Bomdzele & Molua, 2023; Caracol Radio, 2022; Ch & F, 2021; EL TIEMPO, 2021). In Colombia, few studies relate the impact of water resources scarcity on crop yield or crop stability, with only Naranjo-Merino et al. (2017) developing a study that determines the cocoa crop water footprint, the crop's water dependency, and yield. Internationally, numerous studies focused on explaining the crop sensitivity to the lack of water resources or water stress, considering the "El Niño" phenomenon or regular drought periods. Several studies analyze the effects of climatic factors such as temperature and lack of precipitation (i.e., water stress) on crop yield, growth, and development (Amfo et al., 2021; Chapman et al., 2021b; Dos Santos et al., 2014; Gateau-Rey et al., 2018; Läderach et al., 2013; Schwendenmann et al., 2010; Zuidema et al., 2005), indicating the significance of studying such parameters in cocoa crop production and resource management. Considering that agriculture is the primary user of water worldwide (FAO, 2017) and Colombia (Agronet, 2020), it is crucial to enable studies aimed at supporting water management decision-making to answer to uncertainty, thus enabling higher water use efficiency (WUE), assuring crop productivity, the farmers' benefit, and natural resources conservation.

Therefore, various studies focus on agricultural water management using simulation-optimization techniques. Specifically, such studies simulate parameters and subsequently use them as inputs for the optimization model (Linker, 2021), defining how to allocate water to the end users at a specific time. Such researches include deterministic optimization and optimization under uncertainty; however, the deterministic-based approaches do not guarantee adequate decision-making in uncertain and random climatic conditions (Pannell et al., 2000). As alternative, uncertain optimization approaches address weather randomness and related parameters based on IPP, FP, and SP frameworks. Under the SP, there are three main approaches. The approaches are the implicit stochastic approach (a solution for each considered scenario), the explicit one-stage approach (a solution considering all the scenarios at the same time), and the multistage stochastic approach (which includes uncertainty in the decision framework regarding the parameter future realization) that support facing more realistic problems and granting the decision adaptability regarding the scenario realization (Jamal et al., 2018). However, this research has not retrieved works in CCWM using uncertain optimization approaches. Therefore, if there is the availability of historical data related to the parameters, and if possible, to model the hydric factor in the cocoa production process, a stochastic programming model will lead to flexible and time-adjustable decision-making, allowing preserving and maintaining water resources, ensuring a proper development cocoa crop production.

Consequently, this research aims to include uncertain precipitation patterns that affect the water balance and impacts support water management decision-making in cocoa crop production. This study addresses resource decision-making through a stochastic optimization model based on a multistage framework, answering the supply crop water requirement in the different production stages of the time considered. Thus, the research conducts efforts to guarantee the resource in the amount and at the right time for reducing the environmental impact, protecting the crop conditions from water stress scenarios due to improper use and uncertain climatic patterns, reducing the associated costs, and consequently improving the farmers' benefit. Based on the implications of supporting water management decision-making, the proposed research question relates to How to allocate or supply limited water resources adequately under weather uncertainty in cocoa crops, considering stages in stochastic optimization approaches?

2. Objectives

2.1. Overall Objective

To develop a stochastic base optimization model for water allocation considering weather uncertainty.

2.2. Specific Objectives

- To identify the main stochastic modeling strategies for agricultural water management under conditions of uncertainty through a literature review.
- To build a dataset containing customs instances developed for the stochastic optimization model.
- To formulate a mathematical programming model for cocoa crop irrigation management under weather uncertainty.
- To evaluate the stochastic model performance (i.e., supply just the required quantity of water resources) against the deterministic equivalent model.

3. Hypothesis

A stochastic optimization model allows better performance than the equivalent expected model for farmer agricultural water resources management under uncertain weather.

4. Theoretical framework

4.1. Two-Stage Stochastic Programming

Two-stage Stochastic Programming (TSP) is a decision-making framework used to build optimization models that support situations where decision-makers face uncertain parameters (modeled through probability theory) and must make decisions in multiple stages or periods. TSP establishes an optimization model construction and solution into a two-stage fragmented decision scheme (Sahinidis, 2004). The first-stage decisions, commonly called “here and now” decisions, represent decisions made at the beginning of the planning horizon regarding the available information and assumptions of the input parameters. The second-stage decisions, usually called “wait and see,” are made after the uncertain parameter realizations, meaning their resolution is scenario dependent. Therefore, second-stage decisions represent corrective actions or recourse to adjust the problem solution according to the first-stage decisions (Huang & Loucks, 2000). A TSP model with two variables like x (first-stage variable) and y (second-stage variable), has the following resolution, according to Conejo et al. (2010):

- In the first-stage, the decision x is made.
- Just after the decision x is made, the uncertain parameter realized as ω
- Now, the second-stage decision that depends on the first-stage decision and the scenario realization must be made $y(x, \omega)$

A general TSP model has the following mathematical representation addressing both kinds of stage decisions (Beale, 1955; Dantzig, 1955):

$$\min \mathrm{z} = c^T x + E_\omega[Q(y, \omega)]$$

Subject to: [1]

$$Ax = b$$
$$x \in X$$

Where:

$$E_{\varepsilon}[Q(y, \omega)] = \min q(\omega)^T \, y(\omega)$$

Subject to: [2]

$$T(w)x + W(\omega)y(\omega) = h(\omega)$$
$$y(\omega) \in Y, \forall \; \omega \in \Omega$$

In this mathematical representation, x represents the first-stage decision and y the second-stage decision variables vector (a ω-scenario dependent variable), where $E_{\varepsilon}[Q(y, \omega)]$ relates a recourse problem containing the second-stage decisions. c, b, A, $q(\omega)$, $h(\omega)$, $T\ (\omega)$, and $W(\omega)$ are known vectors or estimated parameters. Considering a finite number of scenarios or well-established parameter scenarios, the general representation of the model can be reformulated as follows:

$$\min_{x, y(\omega)} \mathrm{z} = c^T x + \sum_{\omega \in \Omega} \pi(\omega) q(\omega)^T y(\omega)$$

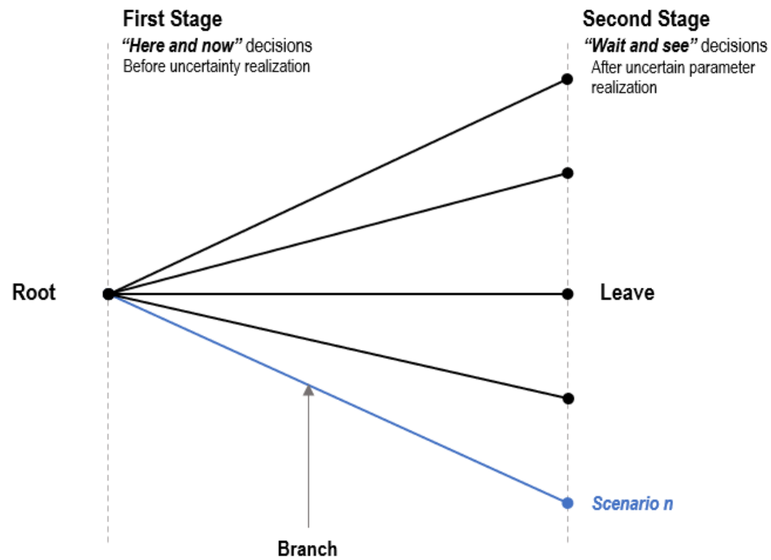
Subject to:

$$Ax = b \quad [3]$$
$$T(w)x + W(\omega)y(\omega) = h(\omega) \;, \forall \; \omega \in \Omega$$
$$x \in X$$
$$x \in X, y(\omega) \in Y, \forall \omega \in \Omega$$

In this approximation, every scenario has a probability of occurrence represented by the $\pi(\omega)$ component, meaning some scenarios are more likely to happen than others, and the optimization model integrates them based on their importance to make the best decision. Conejo et al. (2010) state that a scenario tree can easily represent a reduced TSP model (Figure 1) where the root indicates the time point for making the first-stage decision, each branch represents a distinct realization of the uncertain parameter, and the leaves denote the second-stage decisions made per scenario, depending on the first-stage decision and parameter realization.

Figure 1

TSP scenario tree representation



4.2. Scenario generation

Scenario generation strategies aim to build representative scenarios representing different parameter paths (represented by a scenario tree) in stochastic optimization models or decision-making approaches under uncertainty. There are multiples strategies used for scenario generation, and every strategy relies on the characteristics of the specific problem, available data, and the precision required (Conejo et al., 2010; Jitka Dupačová et al., 2000; Jamal et al., 2018). In Stochastic optimization models' uncertain parameters are usually modeled or represented by random variables, that is, a variable that takes on different values based on random events. The main strategies for modeling data under uncertainty through random variables are:

- Path-base methods: based on time-series modeling or econometric models that represent data behavior, this strategy uses past data to build a model that allows copying data patterns and using them to generate new paths called fan.
- Moment matching: represents a strategy focused on the method of moments to determine the real moments of the parameter to generate scenarios using a specific probability distribution function (Xu et al., 2012).
- Internal sampling: is a strategy that uses an iterative data sample selection (properly and strictly chosen) from the historical data as representations of the parameter (Høyland & Wallace, 2001).
- Scenario reduction: these methods relate scenario fan reduction based on a specific metric that commonly represents the dissimilitude between scenarios. It looks to provide a final set that is close enough to the real scenarios set, considering the statistical characteristics of the original fan (J. Dupačová et al., 2003; Jitka Dupačová et al., 2000)

It is worth remarking that the scenario generation process usually states two steps: (i) determining the first set of scenarios generated through the selected strategy and (ii) a scenario selection focusing on preserving most of the inherent statistical characteristics of the original scenario set.

4.3. Path-base methods

Path-based methods focus on fitting models based on evolving random variables (stochastic processes) over time. These methods aim to construct an appropriate model replicating the historical behavior using a simulation process, where each simulation represents an equiprobable and distinct scenario.

4.3.1. Autoregressive Integrated Moving Average (ARIMA). It is a time series forecasting model that analyzes and predicts data with a trend. It represents the combination of three components: autoregression (AR), representing the historical lag dependency relationship (autocorrelation); integration (I) which relates a differentiation factor determining how many times must be pulled out the mean from the series; and moving average (MA) establishing the relationship between observations and lagged residual errors (Cryer & Chan, 2008; Gujarati & Porter, 2013) The series data needs to meet the wide-sense stationary condition to fit an ARIMA model. Such conditions establish that the series mean and variance remain invariant in time ($E(Y_t) = \mu$ and $Var(Y_t) = \sigma^2$), and the covariance depends only on the realizations of the random variables or the distance between lags ($Cov(Y_t) = E[(Y_t - \mu)(Y_{t+k} - \mu)] = \gamma_t$). An ARIMA model usually has the following structure:

$$y_t = \theta + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \beta_0 u_t + \beta_1 u_{t-1} + _{\dots} + \varphi_n y_{t-n} + \beta_n u_{t-n} \quad [4]$$

Where y_t is the variable in time t, θ regression constant, $\varphi_1, \varphi_2, \varphi_n$ are autoregressive parameters associated with every significant lag: $y_{t-1}, y_{t-2}, y_{t-n}$, and $\beta_0, \beta_1, \beta_n$ moving average parameters related to the lagged errors u_t, u_{t-1}, u_{t-n}.

4.3.2. Backpropagation Artificial Neural Networks (ANN). It is a machine learning technique based on the structure and functioning of the human brain, with interconnected nodes, or neurons, that process and transmit information to generate desired outputs (Graupe, 2013). It allows modeling data and fit a non-parametric model that describes the series behavior by identifying patterns in data. Backpropagation is the type of training strategy used in the ANN learning process, representing a training process where the error is propagated backward through the network after the first epoch to adjust the weights of the connections between neurons. This adjustment usually follows the gradient descent optimization based on the partial derivatives of the error concerning the weights, supporting a rapid convergence of the model by reducing the error in the direction of the derivatives (see Appendix A for a deeper review). The ANN can provide a good strategy for scenario generation, integrating a random variable that follows the series residuals pdf in every step simulation (Vagropoulos et al., 2016) as follows:

$$
\begin{gathered}
y_{t+1} = BpANN_t + u_t, \\
y_{t+2} = BpANN_{t+1} + u_t, \\
. \\
. \\
. \\
y_{t+n} = BpANN_{t+n-1} + u_t \\
y_t = \theta + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \beta_0 u_t + \beta_1 u_{t-1} + \ldots + \varphi_n y_{t-n} + \beta_n u_{t-n}
\end{gathered} \quad [5]
$$

Where y_{t+n} is the future value of the uncertain parameter, $BpANN_{t+n-1}$ represents the ANN output in time t, and u_t is a random variable generated by the residual's series pdf.

4.4. Moment Matching

This method represents a strategy based on calculating a data series' real moments (e.g., mean, variance, skewness, and kurtosis) to match its probability distribution with some theoretical distribution. Then, the main goal of the moment matching is to capture the series behavior using the method of moments to find a suitable distribution that closely approximates or matches the moments of a target distribution to generate uncertain parameter scenarios (Xu et al., 2012). The method of moments is a statistical technique used to estimate the parameters of a probability distribution by equating population moments with sample moments.

4.5. Scenario reduction strategies

Scenario reduction strategies, also known as scenario aggregation or scenario selection, are techniques used to reduce the number of scenarios in stochastic optimization problems while preserving key characteristics of the original set of scenarios. This strategy aims to balance computational efficiency and accurately represent uncertainty, and it is mainly related to the second step of the scenario generation process. There are two classical scenario reduction paths, forward selection and backward reduction. The forward selection strategy is typically applied to reduce trees with few scenarios, presenting a better approach for scenario selection than the backward strategy in problems with small trees. This strategy focuses on selecting a set of scenarios that preserve the main characteristics of the original set, selecting one by one, and adding them to the new set (Conejo et al., 2010). Nevertheless, the greater the number of scenarios considered, the backward strategy states a better and faster scenario reduction strategy, which makes it a good technique for scenario reduction based on large decision trees (Heitsch & Römisch, 2003). The backward reduction starts with a scenario set as large as the original and subtracts scenarios that, when removed, will remain a scenario set close to the original set (J. Dupačová et al., 2003).

4.6. Scenario reduction using the Kantorovich distance

The backward scenario reduction technique typically uses the Kantorovich probabilistic distance to determine a final set of scenarios similar to the original (Conejo et al., 2010). This strategy compares the entire set of original scenarios through a distance function (e.g., Euclidean), selecting the most similar pair of scenarios based on their probability distribution similarity. The technique drops the scenario that allows for preserving the characteristics of the original set, that is, the scenario that has the least impact on the set of scenarios or is easily replaced by another. Eq [6] shows the Kantorovich distance between a pair of distributions (or scenarios):

$$D_K(P,Q) = \sum_{\omega \in \Omega \setminus \Omega_S} \pi_\omega \times \min_{\omega' \in \Omega_S} v(\omega, \omega') \quad [6]$$

Where ω and ω' represents the scenario from P and Q probability distributions, π_ω is the probability of every scenario, and $v(\omega, \omega')$ is the cost function or norm between ω and ω', which is nonnegative, continuous, and symmetric. See Appendix A for further details.

4.7. Quality metrics

The Expected Value of Perfect Information and the Value of stochastic solution are two metrics that allow for determining the importance and appraising the interest in applying stochastic models that integrate the analysis and modeling of uncertainty into an optimization framework.

4.7.1. Expected Value of Perfect Information (EVPI). EVPI measures the maximum value a decision-maker could gain by having complete and perfect information about the future before deciding. It quantifies the potential benefit of eliminating all uncertainties and making decisions based on perfect knowledge. According to Conejo et al. (2010) and Birge & Louveaux (2011)**,** EVPI represents how much the decision-maker is willing to pay to get a better parameter forecast, which means trying by external means (e.g., meteorological experts) to vanish the uncertainty present in parameter estimation. The difference between the optimal value of the Two-Stage Stochastic model (z^{S*}) and the TSP with non-anticipativity constraints or relaxed TSP (z^{P*}) allows obtaining the EVPI metric:

$$EVPI_{min} = z^{\mathrm{S}*} - z^{\mathrm{P}*} \quad [7]$$

A relaxed TSP model represents a model where the first-stage decision can be made after uncertainty realization, assuming that decision-makers can postpone their initial choices until they have perfect information. This relaxation allows evaluation of the maximum potential benefit that could be achieved by eliminating all uncertainties and making decisions based on perfect knowledge.

4.7.2. Value of stochastic solution (VSS). In a deterministic approach, decision-makers base their decisions on fixed or single-point parameter estimation, assuming they know how the future will unfold in a known and predictable manner. On the other hand, in a stochastic approach, decision-makers consider multiple possible outcomes and their associated probabilities, considering the inherent uncertainty and variability of the system or problem under study. VSS metric is a metric used to assess the benefit, which means that a positive VSS indicates that using a stochastic approach or probabilistic approach in decision-making yields higher expected values compared to a deterministic approach. It implies that considering variability and uncertainty can lead to better decision outcomes, improved risk management, and enhanced resource management. Then, VSS provides insights into the worth of developing further studies, experiments, or data collection efforts to reduce uncertainty and enhance decision-making in stochastic base decision modeling. The difference between the optimal value of a TSP deterministic version (z^{D*}) and the TSP optimal value (z^{S*}) allows performing the metric estimation according to Birge & Louveaux (2011):

$$VSS_{min} = z^{D*} - z^{S*} \tag{8}$$

The z^{D*} component states for the solution of a TSP using a fixed parameter configuration in the first-stage decision based on the value estimated through the multiplication of every scenario by its occurrence probability.

4.8. Water balance in crops

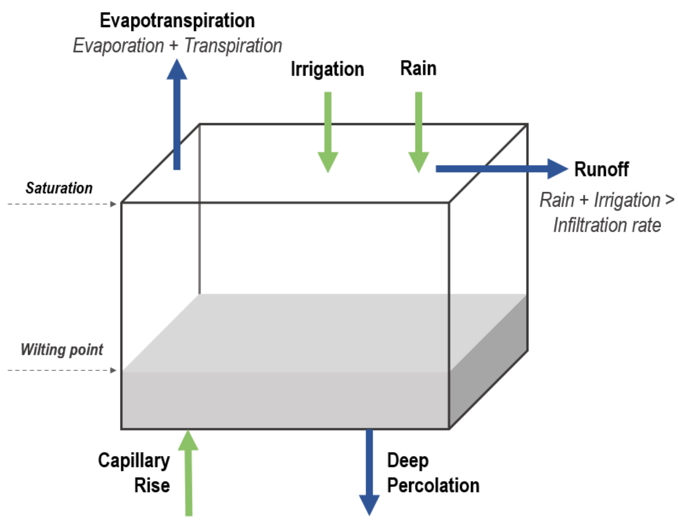
The water balance represents the equilibrium between the water losses and the water gaining in the plant root zone. In this relationship, the water inflow needs to be at least higher than the water outflow to support the minimum water requirements, considering the water loss rate (i.e., the amount of water plants can extract) and water gaining rates based on external factors and management decisions. Various inflow and outflow parameters are related to the crop conditions to determine the water balance in the root zone Figure 2. Precipitation (P), Capillary Rise (CR), and Irrigation (I) are the typical inflow parameters and represent the main ways considered in the water-gaining process. Precipitation and capillary rise factors represent climatic and soil conditions, respectively; precipitation refers to water obtained by water falling in the form of rain from the atmosphere to the Earth's surface, while the capillary rise is the movement of water from deeper zones (i.e., water table or saturated zones below the soil surface) drawn upward through tiny spaces of the soil matrix (Allen et al., 2006). Irrigation is the only water inflow factor affecting the water balance under the farmer's management.

On the other hand, there are the outflow parameters to complete the water balance components. Namely, Runoff (RO), Deep Percolation (DP), and Evapotranspiration (ET_c) represent the main forms of water-losses in the plant root zone. Runoff occurs when the frequency of irrigation and rainfall exceeds the soil's water absorption capacity, resulting in water loss due to surpassing the water holding capacity (Critchley & Siegert, 1991). Deep percolation describes the downward movement of water from the root zone to the water table when the soil becomes saturated (Bethune et al., 2008), representing the opposite of capillary rise (i.e., root zone water-gaining process considering non-saturated soils). Evapotranspiration encompasses the combined water loss through plant transpiration and soil evaporation (Allen et al., 2006). That relation supports estimating the soil's water depletion (i.e., water losses) and allows for determining the water required to restore moisture conditions regarding decision-makers considerations, such as water irrigation policies, management strategies, and technological irrigation systems. Figure 2 shows a simple relationship of the water balance process, where saturation is the maximum amount of water the soil can retain, and the wilting point is where the crop suffers from water stress. Eq [9] relates the relationship between water inflows and outflows in the root zone water balance (Appendix B presents a better development and parameters definition involved in the water balance process):

$$I + P + CR = DP + RO + ET_c \tag{9}$$

Figure 2

Water balance in the root zone. Adapted from Allen et al. (2006)



4.9. Yield response to water stress

The response of crop yield to water stress is a critical factor in agricultural productivity and the overall success of farming systems. The extent of yield reduction under water stress depends on several factors, including the crop species, growth stage, severity, and duration of the stress. Steduto et al. (2012) developed an equation that allows estimating the yield response considering water limitations in the root zone:

$$\left(1 - \frac{Y_a}{Y_m}\right) = k_y \left(1 - \frac{ET_a}{ET_c}\right) \tag{10}$$

Where Y_a is the actual crop yield or crop yield under water stress conditions $[kg/ha]$, Y_m is the maximum crop yield $[kg/ha]$, ET_a represents the actual evapotranspiration $[mm/day]$, ET_c the total maximum crop evapotranspiration $[mm/day]$, and k_y is a dimensionless parameter that represents the crop yield response factor and relates the amount of yield reduction regarding the evapotranspiration reduction in water stress conditions. See Appendix C and D for an in-depth understanding.

4.10. Evapotranspiration

Evapotranspiration combines the processes of water evaporation from the soil surface and water transpiration from plants. Solar radiation and heat primarily supply the energy that converts water at the soil surface into water vapor during evaporation. On the other hand, plants absorb water through their roots and release it into the atmosphere through small openings on their leaves called stomata during transpiration. Transpiration plays a vital role in plant cooling, nutrient uptake, and the transport of water and minerals from the roots to the leaves, supporting the generation of crop biomass. The Penman-Monteith (Allen et al., 2006) formula is the most accepted way of estimating evapotranspiration:

$$ET_o = \frac{0.408 * \Delta(R_n - G) + \gamma * \frac{900}{T + 273} u_2 * (e_s - e_a)}{\Delta + \gamma * (1 + 0.34 * u_2)} \qquad [11]$$

Where, ET_o is the reference evapotranspiration $[mm/day]$, T the air temperature at two meters height $[°C]$, Δ slope vapor pressure curve $[KPa/°C]$, R_n is the net radiation at the crop surface $[MJ/m^2day]$, G represents the soil heat flux density $[MJ/m^2day]$, γ the psychrometric constant $[KPa/°C]$, u_2 the wind speed at two meters height $[m/s]$, e_s the saturation vapor pressure $[KPa]$, and e_a actual vapor pressure $[KPa]$. See Appendix E and F for in-depth knowledge.

5. Methodology

This study is objectivist since it addresses a theoretical study that models a real case situation that exists and is real Lewis-Beck et al. (2004), implying a quantitative methodology. This research applies a six-stage framework to build optimization models for agricultural water management supporting water decision-making in cocoa crops. (Figure 3). The first two stages cover the literature review process and focus on analyzing the available scientific studies oriented to water resources management in the agricultural sector based on an adaptation of the Kaur et al. method (2021). While the remaining stages adapt Hillier & Lieberman's framework to formulate and solve optimization models (2010) (meeting the second, third, and fourth objectives).

Figure 3

Operations research studies methodological steps. An adaptation from (Hillier & Lieberman, 2010)

Stage one: Literature review protocol (First objective)

- The review process definition.

- The search equations construction.
- The final retrieved articles using the PRISMA statement and the Backward Snowball methodology.

Stage two: Literature review report (Second objective)

- The analysis and synthesis of guiding questions defined.
- The definition of emerging questions, future research and construction of methodological framework.

Stage three: Define the problem of interest and gather relevant data (Second objective)

This stage uses the literature review stages to support developing a well-defined statement of the study problem related to agricultural crop water management, thus defining the relevant data.

- Setting the study problem assumptions related to the operational costs, cultivar conditions and management aspects.
- Defining the study problem variables and parameters based on the water balance equation.
- Collecting data of study problem parameters using the NASA-POWER, and FAO databases.
- Univariate time-series analysis (i.e., correlation, stationary analysis, outliers' treatment, and Box-Cox data transformation).
- Time-series and pdf fitting process.
- Data trajectories generation (i.e., simulation) through the ARIMA and ANN-Bp models, and the Beta pdf function.
- Scenario reduction strategy implementation through the Kantorovich distance function.

Stage four: Formulate a mathematical model to represent the problem (Third objective)

It includes formulating the problem and constructing the mathematical model (stochastic and deterministic) that represents the essence of decision-making related to water resources management in uncertain cocoa crop situations using the water balance equation in the root zone.

- The objective function construction (stochastic model.) regarding the w scenarios.
- The objective function construction (deterministic model.)
- The mathematical programming constraints construction (stochastic model.) regarding the w scenarios and the assumptions in the case study.
- The mathematical programming constraints construction (deterministic model.) regarding the assumptions in the case study.

Stage five: Develop a computer-based procedure for deriving solutions to the problem from the model (Third objective)

The procedure contains the mathematical programming model solution proposed by:

- Select an algorithm technique for large-scale linear programming problems decomposition for the stochastic model.
- Apply the decomposition technique to solve the stochastic model if applicable.
- Apply an exact method to solve the deterministic model.

Stage six: Test and compare the model (Fourth objective)

This stage aims to test the stochastic optimization model built using two metrics that show the model's relevance and applicability in cocoa crop water management. The stage also compares models' performance concerning the equivalent deterministic model that uses the expected value of the uncertainty parameters.

- Determine the Expected Value of Perfect Information (Raiffa, 1968) and the Value of the Stochastic Solution quality metrics (Birge, 1982).

- Compare the models' performance using simulated climate scenarios.

6. Literature Review

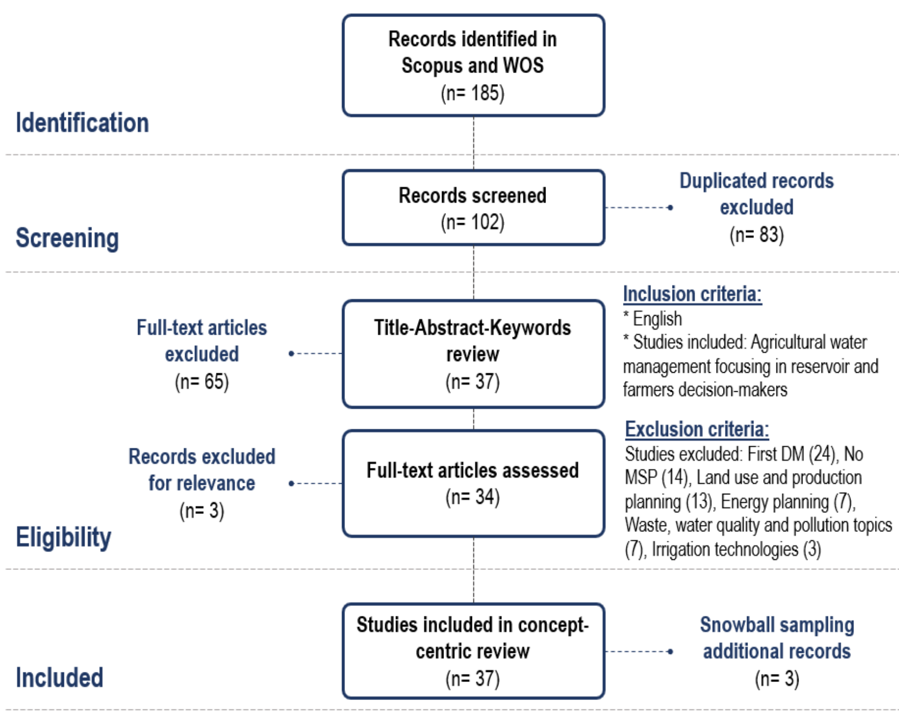
This study developed a systematic literature review on water resources management (allocation and irrigation) focused on second-level (i.e., basin manager) and third-level (farmers) decision-makers to provide a methodological framework for applying uncertainty optimization techniques on agricultural water management. This work applies the practical five-step framework proposed by (Arksey and O'Malley, 2005) for reviewing the scoping study to map the available literature, research gaps, and theories on the topic supported by the PRISMA statement (Liberati et al., 2009) and the Backward Snowball Sampling Methodology (Irshad et al., 2018). Such methodology guarantees a better-supported and detailed scoping and SLR study covering a more significant number of results through a retrospective review. The answer to the three following guiding questions supports the review, enabling the collection and establishment of essential points associated with various factors involved in Multistage stochastic programming modeling strategies:

- **Q1:** Which are the main crops analyzed in selected studies?
- **Q2:** What are the primary sources of uncertainty decision-makers face and the most suitable techniques for uncertain parameter modeling?
- **Q3:** What main modeling strategies are related to MSP and the most common algorithms or solution methods?

This research built a three-layer search query that focused on applying Multistage Stochastic Programming (MSP) to water management in agriculture (Figure 4). The first three layers relate MSP components (first layer: *stochast**, second layer: *two-stage* OR *multi-stage* OR *multistage*, third layer: *programm** OR *optim**) while the fourth and fifth allows filtering studies to water management activities (*water* OR *irrigation* OR *allocate**) and agriculture domain (*crop* OR *cultivation* OR *agricult**), respectively. This work applied the three layers into titles, keywords, and abstracts to acquire the articles. This study only uses Scopus and WOS databases considering the rigorous review and publication process of articles and scientific documents to support the retrieval of proper and pertinent studies. (Bakkalbasi et al., 2006; Falagas et al., 2008; Harzing & Alakangas, 2016). The PRISMA and Snowball hybrid methodology allows for retrieving 37 studies spread over 14 years (2005-2021), concentrating most of the studies in the 2016-2021 period (62%). About 89% of the studies are applied case studies, with most studies related to China (87%). A low proportion of hypothetical case studies consider the research field performance.

Figure 4

SLR flow chart based on the PRISMA statement (Liberati et al., 2009)



6.1. Answer to the first guiding question: Which are the main crops analyzed in selected studies?

Agricultural water allocation is complex due to various uncertainties and factors impacting water system performance and agricultural productivity (Hou et al., 2016; Kang et al., 2017; Samian et al., 2015). Decision-makers must consider farming systems, crop types, and their sensitivity to water scarcity to support agricultural water productivity, reduce poverty, and meet food demand (Dai & Li, 2013; J. Zhao et al., 2017). The studies focus primarily on annual crops like wheat, corn, and rice, which are important for regions facing water scarcity and high demand (Ji et al., 2020; Xiaoyun Li et al., 2016), with 70% of the crops studied. These crops are crucial in achieving global food security (Grote et al., 2021; Shiferaw et al., 2013). Jointly allocating water resources to annual and perennial crops is also a problem addressed in some studies establishing the importance of supporting every crop condition in agroforestry systems. However, the SLR results show a lack of study awareness of permanent crops such as fruits and berries (only four studies). Additionally, although the study focuses on the second and third levels of decision-makers, the results indicate a limited interest in the farmer's perspective and irrigation scheduling using MSP for water allocation.

6.2. Answer to the second guiding question: What are the primary sources of uncertainty decision-makers face and the most suitable techniques for uncertain parameter modeling?

This study explores the sources of uncertainty in the agricultural water allocation problem and identifies four primary factors: socioeconomic, hydrological, climatic, and productive conditions. Hydrological parameters related to the water cycle and availability of water bodies are the most common source of uncertainty. Socioeconomic parameters, including market conditions, political regulations, and social context, are the second source of uncertainty, where economic benefits and penalties play a significant role in this aspect. On the other hand, even though productive parameters and climatic conditions are important sources in modeling and managing water resources have not received adequate attention lately, which still presents a research gap. Nowadays, dealing with such uncertainties requires various strategies depending on data availability, quality, and reliability (Fu, Li, Li, et al., 2018; Y. Wang et al., 2020). When sufficient data is available, researchers commonly employ strategies based on Random Parameters (RP) using stochastic processes to represent uncertainty through parameter distribution (Marques et al., 2010). Alternatively, the Interval Parameter (IP) strategy is more suitable when data is lacking, but establishing parameter bounds is reliable (Xiaoyan Li et al., 2014). On the other hand, Fuzzy Parameters (FPa) modeling based on membership functions is a good strategy when uncertainties involve ambiguity and vagueness (S. Guo et al., 2019). Nevertheless, some situations demand more than one data availability, quality, and uncertainty strategy.

6.3. Answer to the third guiding question: What main modeling strategies are related to MSP and the most common algorithms or solution methods?

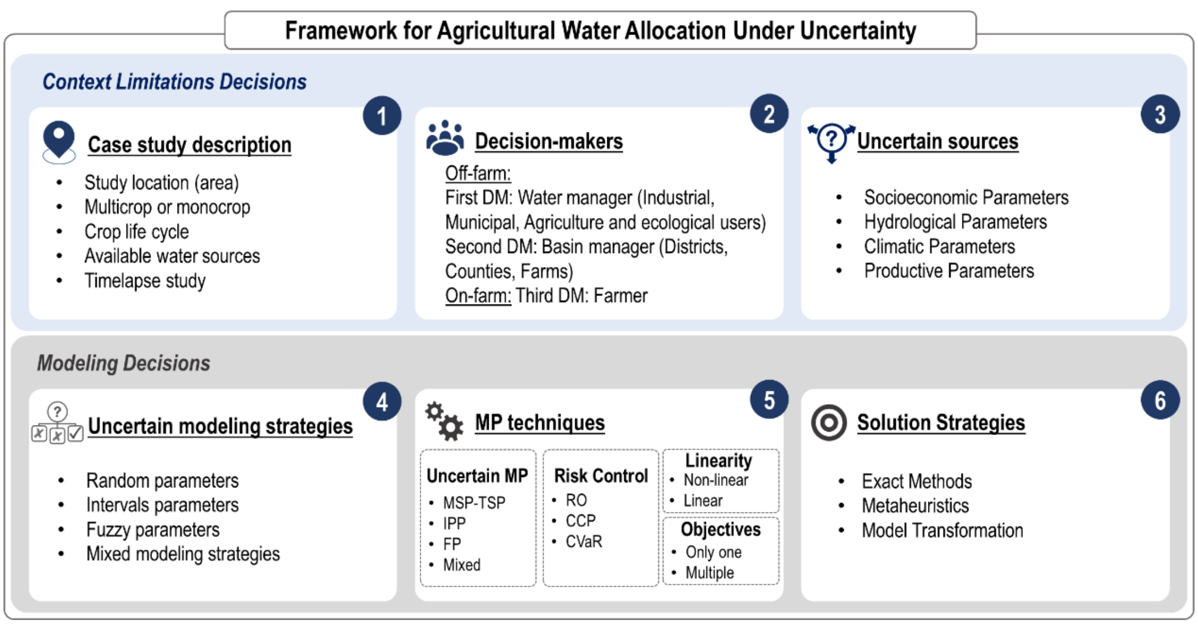
This study analyzed three main programming approaches that researchers use in agricultural water allocation problems under uncertainty: Multistage Stochastic Programming, Interval Parameter Programming, and Fuzzy Programming. MSP is a flexible decision-making framework that incorporates parameter uncertainty over time (C. Li & Grossmann, 2021), and TSP (a variant of MSP using only two-stage decisions) is the strategy most commonly implemented. However, MSP is less prevalent due to its mathematical complexity and computational cost. On the other hand, IPP, which represents uncertain parameters as intervals, is more widely used than FP, considering studies face more data limitations than ambiguous situations. Combining these programming approaches allows researchers to address multiple forms of uncertainty simultaneously. Including risk control strategies such as Chance-Constraint Programming (CCP), Fuzzy-CCP, Conditional Value at Risk (CVaR), and Robust Optimization (RO) in the modeling process allow for providing better decision schemes (Fu, Li, Cui, et al., 2018; Sahinidis, 2004; Youzhi et al., 2021; W. J. Zhang et al., 2021). Such strategies help tackle resource planning issues and quantify and reduce economic risk from uncertainties. Several studies also include Multiobjective and Non-linear Programming (e.g., Quadratic Programming, Fractional Programming) to handle problems with conflicting objectives and non-linear relationships (Marques et al., 2010; F. Zhang et al., 2019). Depending on the modeling strategies implemented, there are three main solution techniques metaheuristics, exact methods, and model transformation. Model transformation with hybrid strategies is the most commonly used method for solving these mathematical problems: MSP-IPP and MSP-FP, among others.

6.4. Framework

This study proposes a framework based on the results acquired in the literature review. The proposed framework (Figure 5) presents a 6-step methodology exposed in two stages (i.e., context limitation decisions and modeling decisions), supporting the agricultural water allocation modeling process under uncertain conditions. This framework seeks to reduce efforts in identifying critical factors associated with multiple conditions of mathematical modeling under uncertainty, providing an overview of the main strategies used in the study field. Proposals of this type allow leading future research based on current and pertinent information associated with the existing difficulties in the problem-solving process while proposing a helpful path to face the agricultural water system's complexities. The review establishes the current research landscape, allows building research framework for future studies, and provides the following research critical topics that will contribute to understanding and facing difficulties in the agricultural water allocation domain: (i) future studies focused on the third-level decision-maker to support regional water management from the lowest decision level; (ii) studies oriented to include different modeling strategies under uncertainty (e.g., stochastic programming) and (iii) a bigger spectrum of perennial crops considering their importance in the agricultural economy. Appendix G has additional detailed information on the review process, considering it is the article developed in this activity. In addition, regarding the knowledge gathered in the SLR, Marquez et al. (2022), and Talero et al. (2023) present preliminary results of water management in whole levels decisions and a complementary analysis addressing a cocoa crop optimization production considering selling price uncertainty (i.e., another type of uncertainty identified in the literature review).

Figure 5

Framework for developing water allocation optimization models under uncertainty



7. Case study

7.1. Problem definition

A decision-maker (DM) has a cocoa farm at Longitude: -73.4723, Latitude: 6.8773 with an altitude of 658 m.a.s.l. in San Vicente del Chucurí in Santander, Colombia. The DM faces adequately managing the place's available water resources due to irregular precipitation periods and random weather patterns annually. Figure 6 presents a brief overview of the conditions associated with this study to support crop yield. The decisions related to the optimization models consist of two moments. At each moment, the DM must determine the amount of water he must reserve or contract to meet the crop's water needs. In this sense, the DM observes the current precipitation behavior and determines the feasibility of securing water resources now (first period) or waiting for the horizon to advance (second period) concerning every scenario realization. The decision models support annual water planning split into 26 decision periods comprised of 14 days each, which means this study iteratively solves 26 optimization models by strategy for annual water planning. This case study presents the following assumptions considering those mentioned above:

- The planning period answers 14 days (2 weeks), corresponding to the number of days in which the crop reaches approximately the critical level of humidity (45.92) with a mean evapotranspiration $ET_o = 3.5$.
- This study establishes fixed costs through the whole horizon plan.
- The two main water sources' associated cost is different, viewing that the first source relates to water resources collected from rainfall, and the second contracts with external entities to supply the resource in scarcity conditions. The study stablishes a 10% cost increase using water from the second source.
- The DM knows that determining the amount of water to use by each source at the first moment relates to a lower cost, considering that planning the future allows for getting water properly. If the DM waits for the uncertainty realization, the associated costs are higher, and therefore there is a penalty for the delayed decision affecting the farmer's benefit (Liu

et al., 2017). The second stage cost increased by 4.5% from the second source regarding the first stage cost decisions.

- This study ignores sloping plantations and only addresses models on flat landscapes.
- This study addresses the cocoa Clon ICS 95 to support the case study (Cerón Salazar et al., 2020).
- The soil structure remains constant through the whole horizon plan.
- This study addresses a crop monoculture system since it is the main cultivar system in San Vicente del Chucurí, with 98.8% of cocoa farms (DANE, 2016b).
- The optimization models consist of supporting the water balance per plant, which means that each model represents management decisions focused on a single plant.
- This study considers only healthy cocoa crop plantations older than four, as the early fourth years represent unproductive ages (De Almeida & Valle, 2008; Djuideu et al., 2021).
- This study addresses a cocoa crop with two annual harvest periods (January-June and July-December), considering that every harvest takes between 165 and 181 days (Romero Vergel et al., 2022).
- Water availability based on rainwater reserves corresponds to 20%.

Figure 6

Main decision models conditions

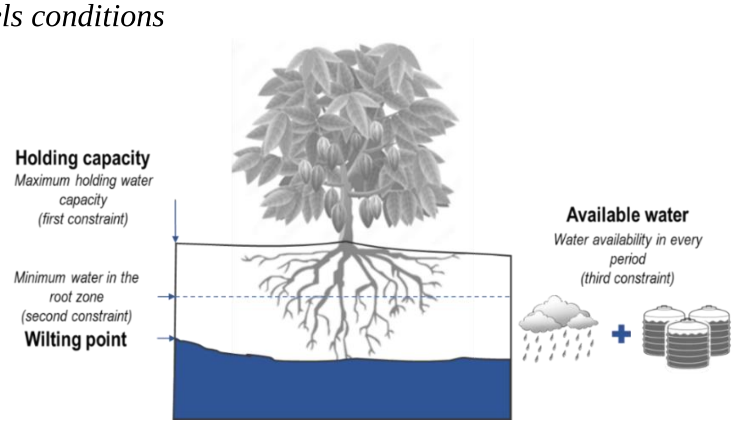
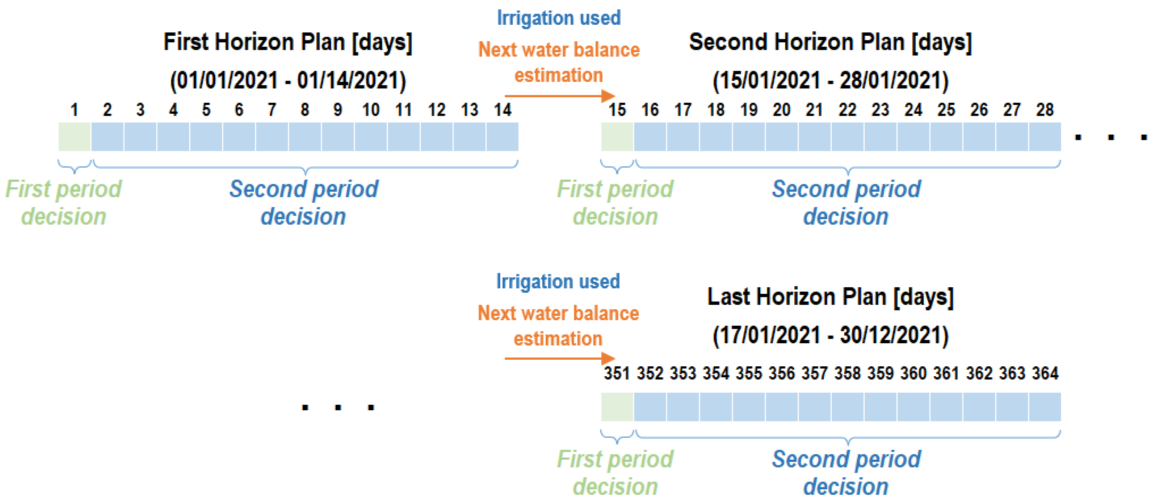


Figure 7

Decision modeling representation



7.2. Data collection and parameter estimation

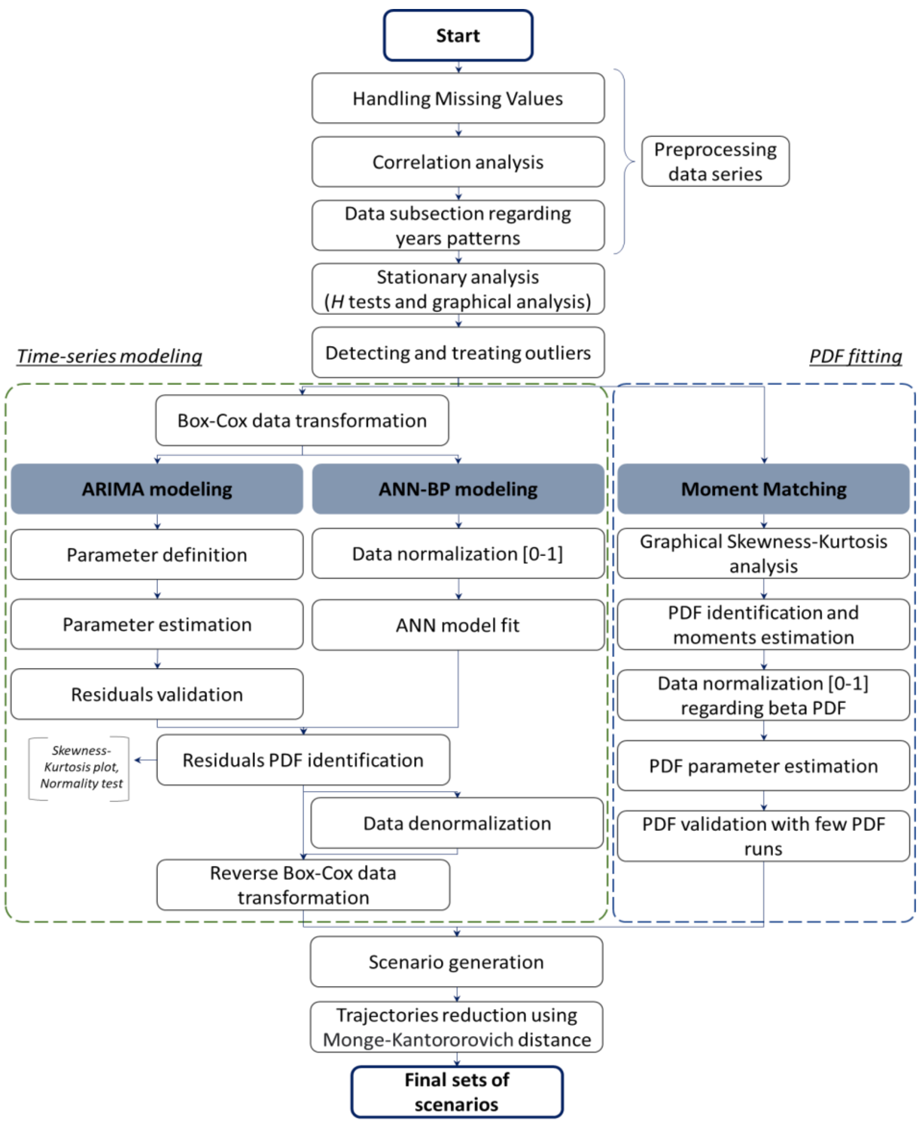
Through Eq [12], this study first establishes the parameters associated with supporting the water balance in the plant's root zone. Of the six parameters related to the water balance equation, two are negligible (CR_t y DP_t) considering the distance from the water table to the ground surface (Somers & McKenzie, 2020), which interferes with a water supply in the root zone due to porous transport or loss by saturation. The I_t and RO_t factors in the water balance represent the required irrigation and runoff produced after reaching soil holding capacity. These two parameters can be stated as managerial decisions in the optimization model as the need for irrigation (I_t) and drainage needs (D_t) in the field, respectively. The parameter P_t is the precipitation (uncertain parameter) every period, and ET_t^c is the main source of water extraction from the soil. Consequently, this study generates a statistical-descriptive analysis of the precipitation parameter to build future scenarios using multiple sources of information (Appendix H).

$$I_t + P_t + CR_t = DP_t + RO_t + ET_t^c \qquad [12]$$

For this purpose, this study collects climatic data stored in the NASA databases (NASA, 2021) in the 1981-2020 time window (Appendix I) which support an analysis of the precipitation behavior and subsequent estimation of the parameter in constructing optimization models under uncertainty. Figure 8 summarizes the tasks applied in the entire process of generating scenarios related to the modeling of precipitation from models based on time series and pdf fits. This study presents a detailed and more in-depth analysis of the results in Appendix A.

Figure 8

Scenario generation and reduction methodology



In the climatic parameters historical series, there are only missing data in the Radiation (RS) series, which has missing data corresponding to 3 years (Figure 9). Considering this lack of data, a listwise treatment is generated in the other parameters to perform a multivariate data analysis considering only the reported values. This study develops a simple correlation analysis to determine possible correlation between the parameter series. It is relevant to work with autoregressive univariate modeling processes considering the little correlation between parameters to explain the precipitation behavior, where relative humidity has the highest correlation (0.4). In the univariate modeling process, it is important to work with data series subsets that correspond to the most recent period of the parameter (2016-2020) since including prior periods can affect the model's fit and shift them from the current behaviors. Such a decision derives from the distribution shown in Figure 10, which presents four main periods with different behaviors: first (1984-2000), second (2001-2009), third (2010-2015), and fourth (2016-2020) being the most reasonable to work with last.

Figure 9

Radiation (RS) parameter series

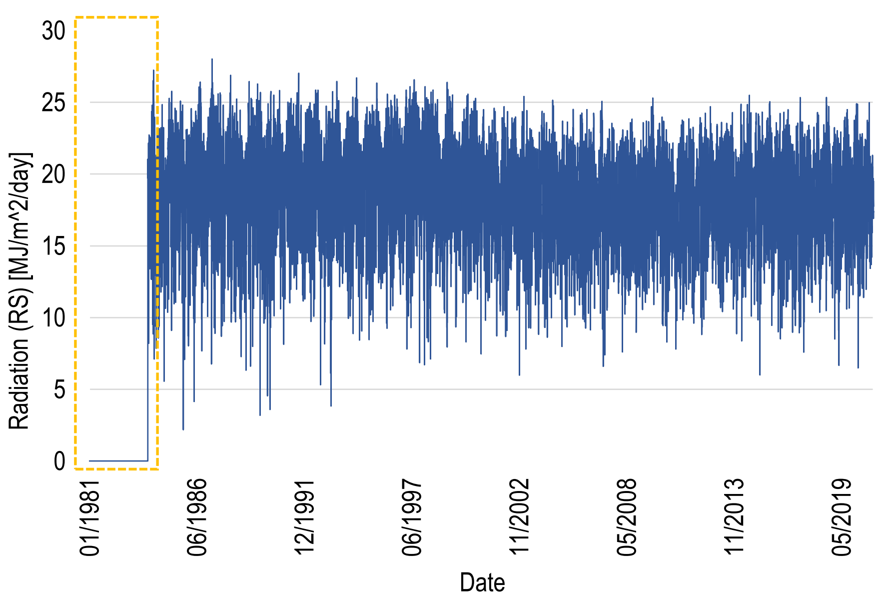
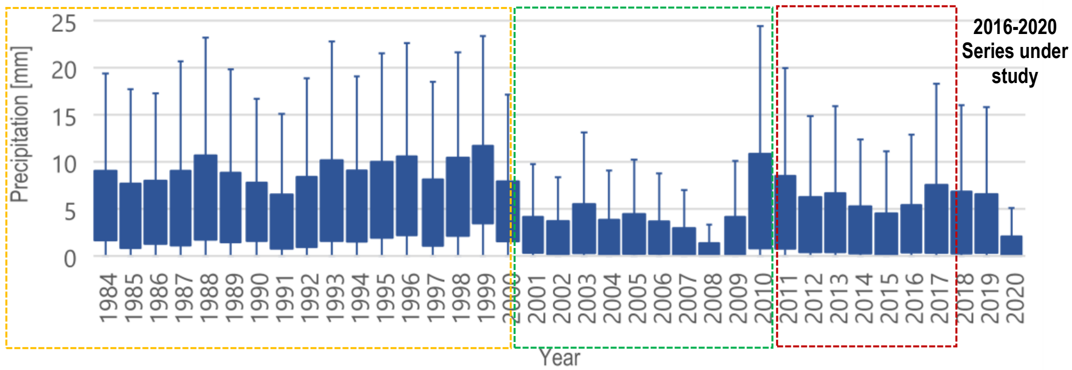


Figure 10

Precipitation year box-plot distribution



The main condition the series must meet to fit univariate autoregressive models is that the series must be stationary, meaning the mean and variance invariant behavior. This study applies the Augmented Dickey-Fuller (ADF) hypothesis tests with trend and drift, ADF without trend, ADF without trend and drift, and the Phillips-Perron (PP) test to validate the stationarity conditions (Dickey & Fuller, 1979; Phillips & Perron, 1986). For this purpose, the study performs seven partitions (with 261 data) of the historical series with the objective that the analysis guarantees stationarity in the complete series and its corresponding subsections. Table 1 presents the $p-values$ of the contrast of the hypotheses where only the first and fourth range of data presents a non-stationary behavior according to the ADF test ($p > 0.05$). However, considering the rest of the tests' contrast with $p - values \leq 0.05$ there is enough evidence of stationarity conditions, then, this study defines the precipitation series as stationary.

Table 1

$p - values$ of precipitation series stationary tests

Data subsets		ADF	ADF no trend	ADF no trend and no drift	PP
First range	(1-261)	0.2554	0	0	0.01
Second range	(262-522)	0	0	0	0.01
Third range	(523-783)	0	0	0	0.01
Forth range	(784-1044)	0	0	0	0.01
Fifth range	(1045-1305)	0.4552	0	0	0.01
Sixth range	(1306-1566)	0.002	0	0	0.01
Seventh range	(1567-1827)	0.0046	0	0	0.01

Considering the stationarity condition of the series, it is possible to fit autoregressive models such as ARMA (i.e., no integration required). However, before fitting a model, a data preprocessing is first performed on the series to fix outliers and reduce the risk of fitting inappropriate models. In this case, the preprocessed data correspond to records that exceed the maximum and minimum boundaries in the whisker box plot, with the lower whisker: 0 [mm] and the upper whisker: 13.5125 [mm]. This study uses an imputation based on the average of the seven before and after historical data (k-Nearest Neighbor (Fix & Hodges, 1989)) to the outlier to produce a replacement driven by recent historical behavior. Additionally, a Box-Cox transformation based on the fourth root of the series allows for reducing the series' variance and supports the model fitting properly (Box & Cox, 1964). Figure 11 shows the adjusted series' behavior from the outliers' imputation and the Box-Cox transformation.

Figure 11

Precipitation distribution before (a) and after (b) data transformation

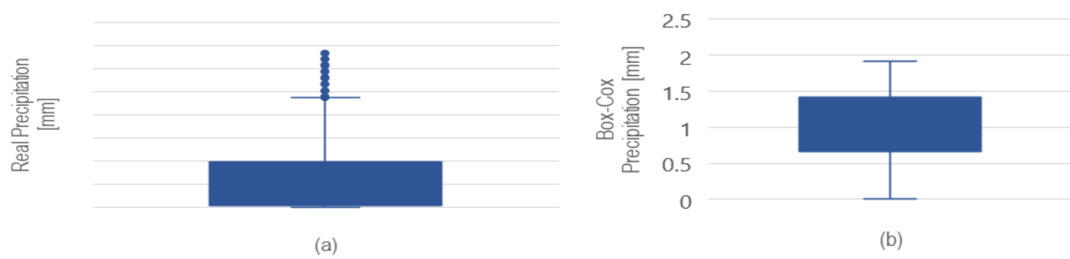
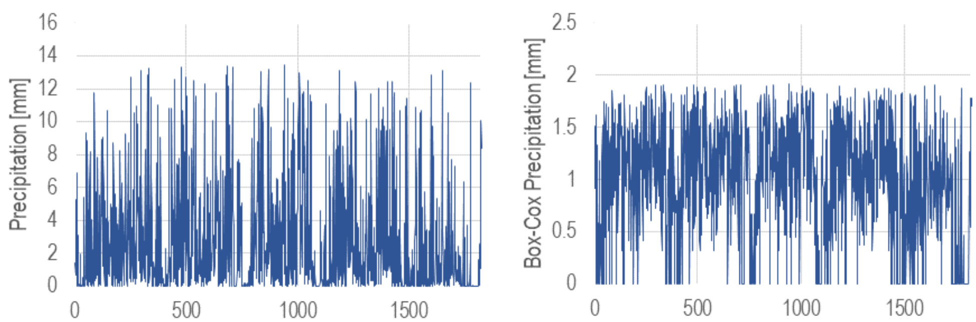


Figure 12

Precipitation and Box-Cox precipitation



7.2.1. ARIMA model. This study builds or fits different models based on time series to reproduce the historical behavior of precipitation and support the scenario generation. The first model built is an ARIMA model developed using the Box-Jenkins four-step methodology (Geurts et al., 1977): (i) selection of p (autoregressive component), d (integration component), q (moving-average component); (ii) ARIMA parameter estimation; (iii) white-noise residuals behavior (i.e., $IIDN(0, \sigma^2)$); (iv) forecast. In the ARIMA model, the d component is zero, considering the series is stationary, so there is no differentiation process. This study applies an exhaustive method based on a ten-step iterative process to determine the best combination of components varying p, d, q in the range (0-9) to obtain a proper available model based on the Akaike criterion (i.e., quality measure relative) regarding lower AIC values indicates a better fit (Weng et al., 2019). From this iterative process, the best model option is an $ARIMA(6,0,3)$, which has the best-fit value of $AIC = 1933.254$. The model proposed from the adjustment process is the following, in which all the coefficients are significant in the explanation of the series (Table 2):

$$\begin{aligned} y_t = {} & 2.1743 y_{t-1} - 2.3710 y_{t-2} + 1.7724 y_{t-3} - 0.7637 y_{t-4} + 0.3327 y_{t-5} \\ & - 0.1451 y_{t-6} - 1.6177 u_{t-1} + 1.3778 u_{t-2} - 0.7143 u_{t-3} \end{aligned} \tag{13}$$

Table 2

$p - values$ of precipitation series stationary tests

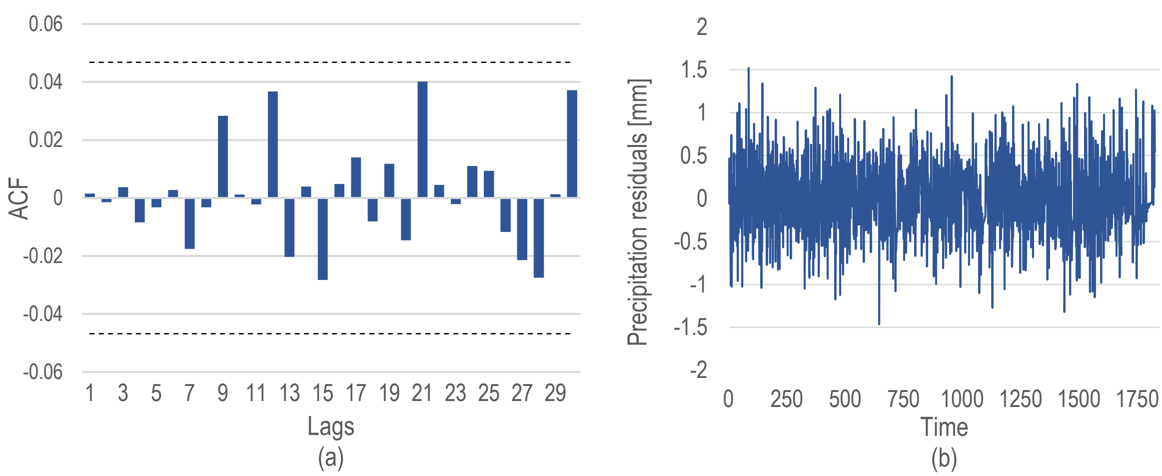
Coefficients	Estimate	Std. Error	Z value	$\Pr(> \lvert z \rvert)$	Significance
Ar1	2.1743	0.1264	17.1913	2.2e-16	***
Ar2	-2.3710	0.2428	-9.7629	2.2e-16	***
Ar3	1.7724	0.1941	9.1297	2.2e-16	***
Ar4	-0.7637	0.0976	-7.8239	5.120e-15	***
Ar5	0.3327	0.0579	5.7441	9.242e-09	***
Ar6	-0.1451	0.0260	-5.5730	2.503e-08	***
Ma1	-1.6177	0.1266	-12.7776	2.2e-16	***
Ma2	1.3778	0.1780	7.7389	1.003e-14	***
Ma3	-0.7143	0.0891	-8.0161	1.091e-15	***

****: represents that under 99.999% confidence interval ARIMA coefficients are statistically significant*

A white noise behavior in the residuals allows for validating the relevance of the generated model, considering that the model could capture all the autocorrelation behavior of the series. For this purpose, the LJung-Box hypothesis test and the graphical analysis support the validation of a white noise behavior in the model's residuals. The LJung-Box test allows us to confirm the non-existence of correlation (null hypothesis (Ho)) in the lags of the time series where the model presents a p-value=0.993, which shows that the residuals are not correlated. The graphical analysis of the residuals in Figure 5 makes it possible to demonstrate the result obtained by the statistical test where the values of the autoregressive function do not exceed the significance bands (Hyndman & Athanasopoulos, 2014). The residuals are white noise with behavior defined by $IIDN(0, \sigma \,\hat{}\, 2)$. Therefore, this study establishes that the model appropriately captures the series' historical behavior and fulfills the right conditions. Thus, the $ARIMA(6,0,3)$ model is relevant to support the generation of precipitation scenarios.

Figure 13

Residuals ACF and series behavior



7.2.2. ANN model. A model based on Artificial Neural Networks with backpropagation learning supports this study's second scenario generation strategy. Considering that it is an ANN model based on lags, the modeling must identify the significant lags for building the model. The partial autocorrelation function (PACF) allows defining the series lags with correlation. The model works only with the first significant lags (regarding significance bands), considering that the series has no long-term memory. Figure 14 (a) shows the series' most important lags: Lag 1, Lag 3, Lag 5, and Lag 6. After identifying the major lags, this study restructures the data producing a new matrix. The first column represents the independent value (value to predict), followed by the variables collected based on the lags order. A transformation $[0-1]$ suits the data to the range supported by the sigmoid logistic activation function. The training process follows the structure shown in Figure 14 (b), assuming a roll-window prediction (Ahumada & Cornejo, 2016) scheme where the value predicted in the previous stage (Orange) enters as the first lag for predicting the following period dropping the last lag in the process. Table 3 presents the hyperparameters list used to fit the model.

Figure 14

Main lags in ANN-bp training (a) and training process (b)

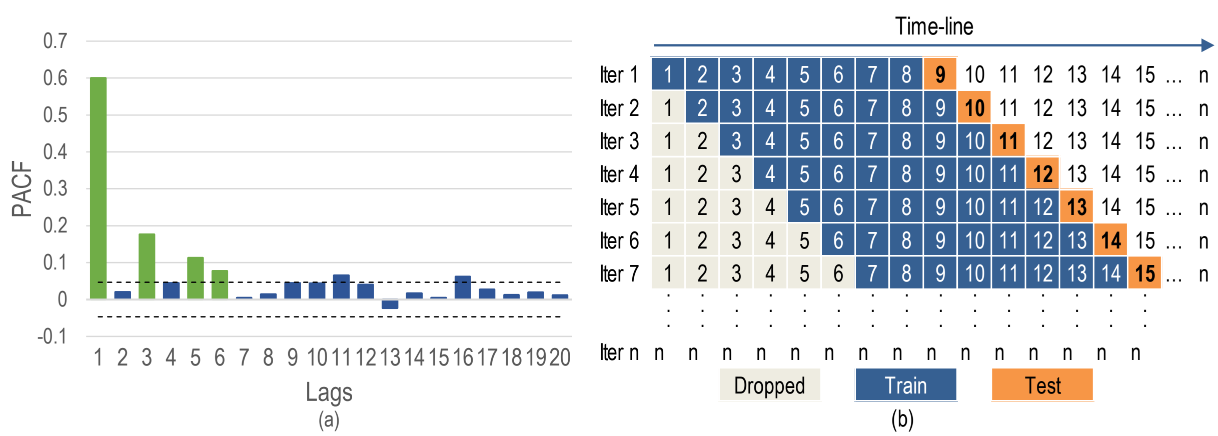


Table 3

ANN-bp hyperparameter list

Hyperparameter	Value/Description
Formula	$y_t = f(significant\ lags)$
Data	Precipitation data scaled and transformed.
Threshold	0.01
Output type	Numeric output
Learning rate	0.005
Activation function	Logistic Sigmoid
Hidden layers	2
Hidden neurons	5 (in each layer)
Learning algorithm	Backpropagation
Maximum number of steps	100.000

7.2.3. BETA Probability Distribution Fit. This study fits a Probability Distribution Function (pdf) to the series providing the compendium of strategies developed for generating scenarios. The BETA pdf selection derives from the evidence that according to the series kurtosis, skewness, and using a thousand resamples of the previously imputed data series (procedure applied in ARIMA and ANN models), data fall in the area of the pdf distribution (a grey area in Figure 15) according to Cullen and Frey graph (Bailer, 2001). A data transformation into the [0,1] range allows for adjusting the pdf parameters considering the distribution's range of the BETA function. This study shows that the Beta pdf with shape 1: 0.3785048 and shape 2: 1.5828210 is a good option to reflect the data distribution regarding 5 simulations of the fitted distribution and the real series (Table 4). This study ran 5 data simulations based on the fitted pdf to compare with the original series moments to support the BETA pdf selection. Table 4 presents the results reached in the simulations, which imply that the pdf is suitable to represent the distribution of the historical data:

Figure 15

Cullen and Frey series pdf estimation retrieved from Colab

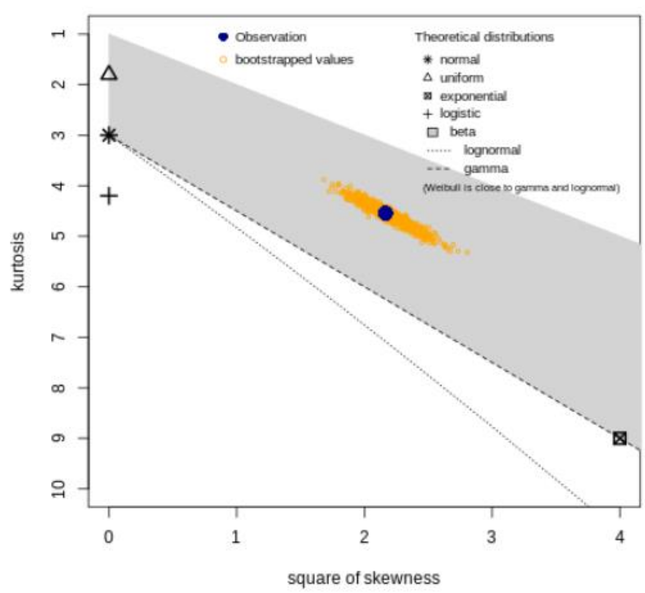


Table 4

BETA simulations moments comparison

Moments	Series	Second Simulation	Third Simulation	Fourth Simulation	Fifth Simulation
First moment	0.1929	0.1922	0.1954	0.1868	0.1909
Second moment	0.0526	0.0505	0.053	0.0528	0.0534
Third moment	1.4703	1.305	1.4273	1.3719	1.4205
Fourth moment	4.5340	4.1092	4.0433	3.9667	4.1680

7.2.4. Scenario generation. This study generates a thousand scenarios fan as a means to represent a large number of equiprobable future paths using the previously constructed models (see Appendix J). A simulation allows for generating scenarios based on a random variable that follows a pdf using the ARIMA and ANN models fitting and training errors. The Cullen and Frey graph (Figure 16) shows that a normal-type pdf could support the residual distribution of both models considering the proximity of the real-series and the thousand resampling's (blue dot and yellow dots) with a normal theoretical pdf (asterisk). The Kolmogorov-Smirnov (1951) hypothesis test supports and guarantees this assumption by comparing the distribution of the residuals of both models concerning a normal theoretical pdf obtaining p-values greater than 0.05, and accepting the null hypothesis (Ho: there is no significant difference between both distributions). Therefore, this study uses a normally distributed random variable to develop the 1000 scenarios fan using the first two strategies. The fitted normal pdf for each model's residuals and the $p-value$ obtained in the hypothesis test are:

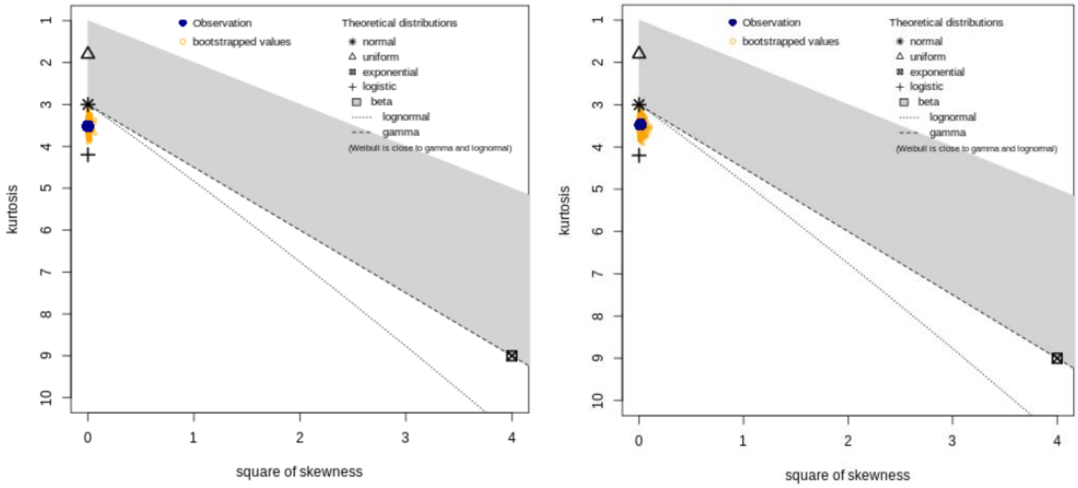
Table 5

Residuals pdf fit and KS test result

Moments	ARIMA residuals	ANN residuals
Mean	0.007849776	-4.181425e-06
Variance	0.408044721	2.053083e-01
KS test $p-value$	0.2252	0.4286

Figure 16

Cullen and Frey residuals pdf definition retrieved from Colab



7.2.5. Scenario Reduction. This technique allows building a scenario set with the most relevant scenarios in the generated fan. The preceding pursues identifying scenarios that will enable an adequate representation of the possible future paths of precipitation, reducing the subsequent computational cost. The Euclidean distance function (1951) allows for scenario reduction considering the similarity (distance) between pairs of scenarios (considering historical data), providing a merge of the most identical. The reduction process in each iteration determines the most similar scenarios and drops the scenario that generally produces the lowest cost regarding the rest of the set. Subsequently, the remaining scenario inherits the probability of the removed scenario. This process continues until the required number of scenarios are obtained, fulfilling the merging tolerance representing the maximum possible scenarios merging in each iteration (see Appendix A for more detailed information on the reduction process). Furthermore, a data aggregation processes every 14 days (based on the models' decision periods) allowed for reducing the scenarios and decreasing the computational cost, considering the aggregated method only runs 26 comparisons (i.e., 364/14) per pair of scenarios instead of 364 comparisons in each iteration (Appendix K and Appendix L relates the aggregated data and the notebook built). Figure 17 presents the original scenarios fans, and Figure 18 shows the reduced fan (5 scenarios) of the ARIMA, ANN models, and BETA pdf, respectively. It is important to state that this study uses 1, 3, 5, and 10 scenarios for model building by each strategy.

Figure 17

Scenarios of the original fan

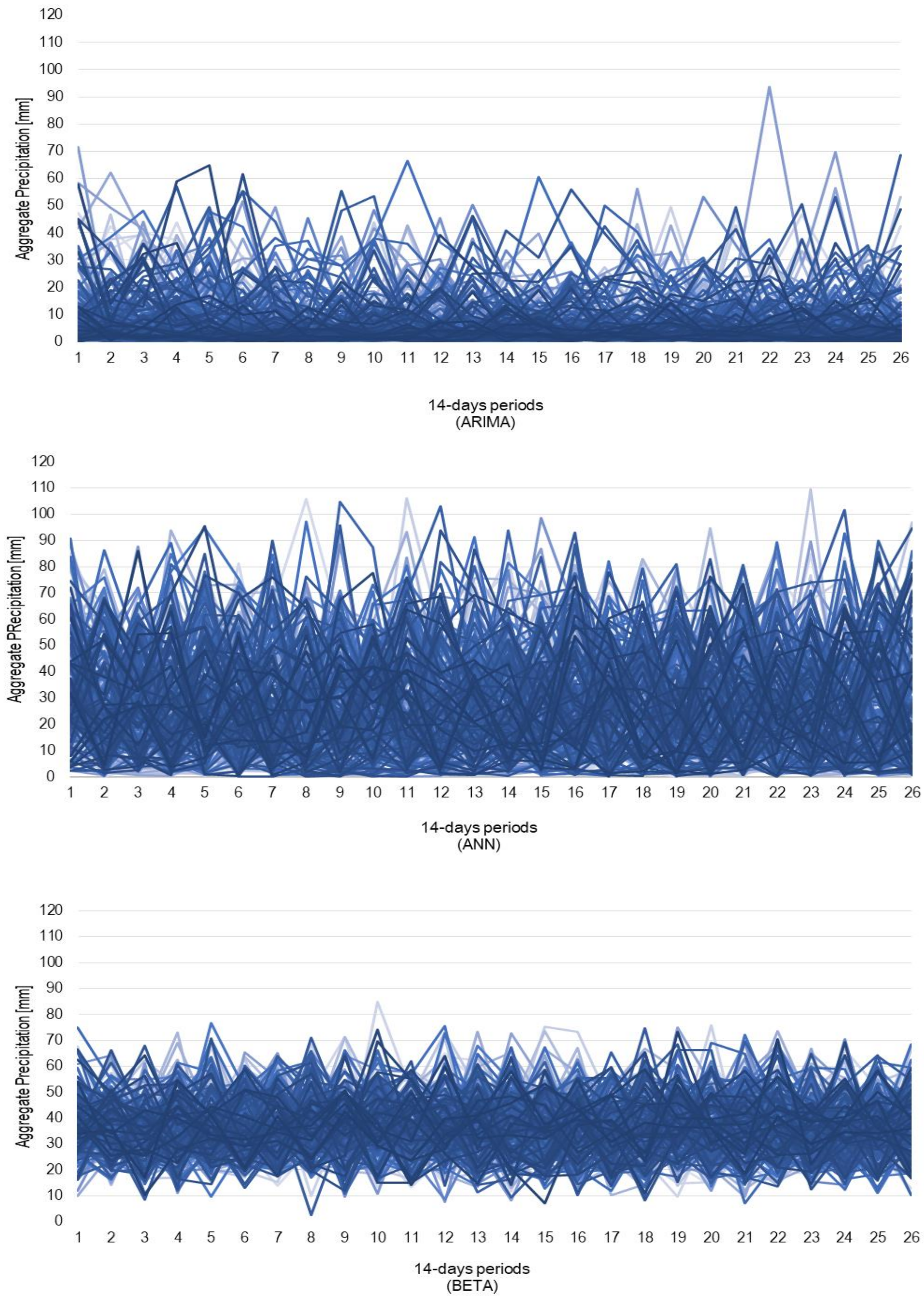
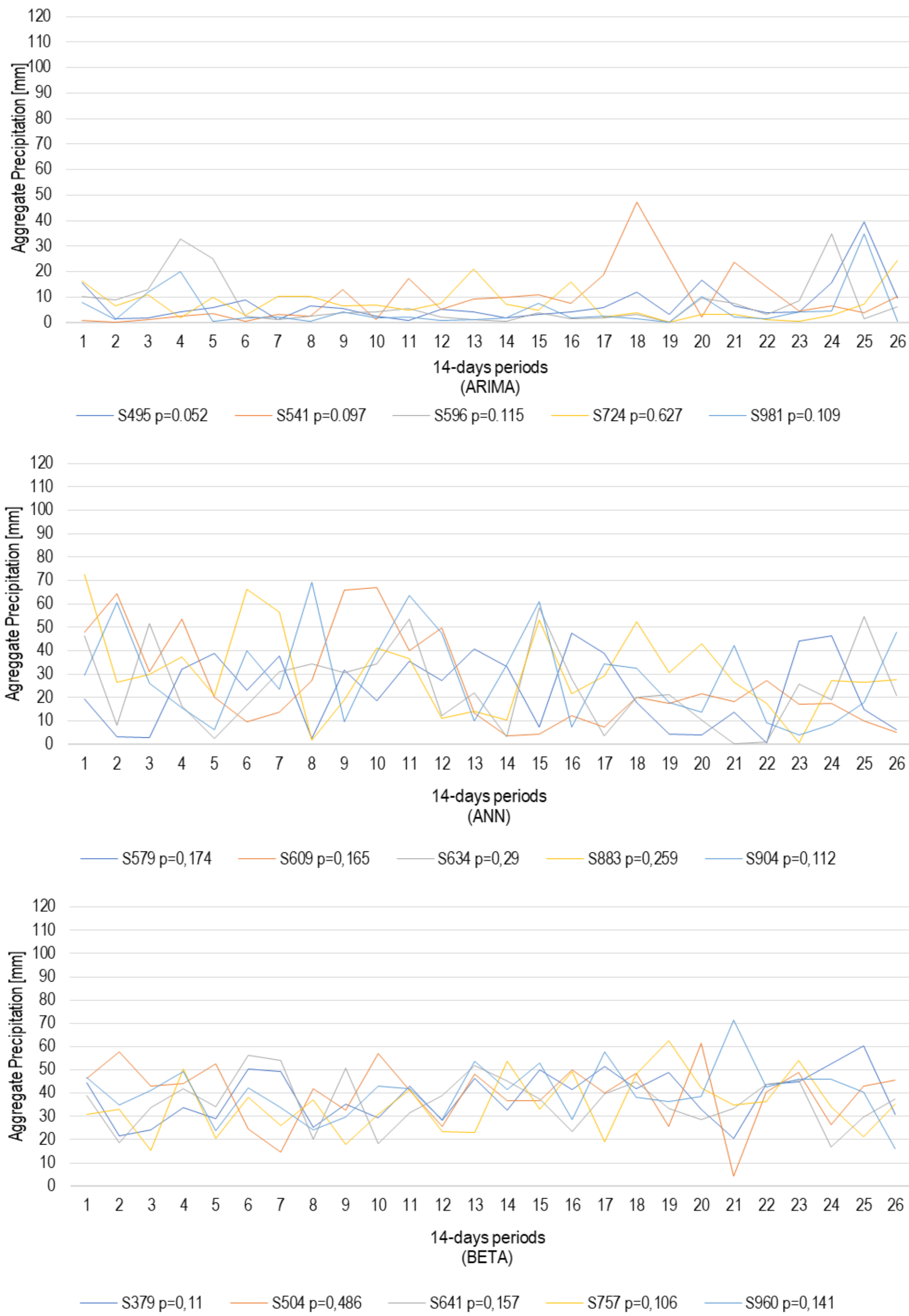


Figure 18

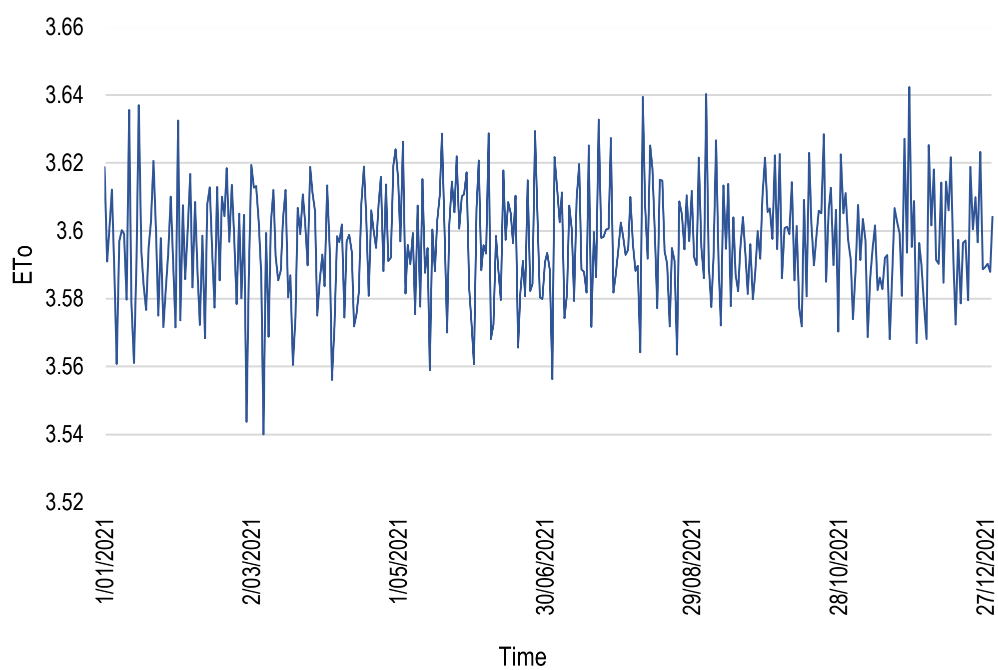
5-scenarios reduction



7.2.6. Evapotranspiration parameter definition. This step relates to the evapotranspiration parameter estimation considering its critical contribution to the water balance equation for supporting the cocoa crop model-building process. This study first estimates the reference evapotranspiration values for the five years (2016-2020) before the study period (2021) based on the Penman-Monteith formula to build possible future evapotranspiration values (Allen et al., 2006). Secondly, it fits a pdf to evapotranspiration values according to the Cullen and Frey plot and validates the selected and fitted pdf through the Kolmogorov-Smirnov hypothesis test. This study defines a normal pdf with $\mu = 3.5962403$ and $\sigma = 0.5430954$, which enables the generation of 1000 ETo simulations. The simulations' average supports producing a single path line representing the future evapotranspiration trend. Figure 11 allows us to observe the behavior of the scenario considered in constructing the models.

Figure 19

ETo



7.3. Mathematical model formulation

7.3.1. Maximum holding water capacity. The DM must only provide irrigation until the water in the root zone does not exceed the holding capacity. Such an assumption means water is only released while the soil water content is under the maximum amount of water the soil can retain. If the water balance surpasses the holding capacity, it is necessary to drain the water. This decision represents the cost of overusing the available water resource. This study starts by defining that water released and water drained must always be lower than the holding capacity in the water balance process (Q. Li & Hu, 2020), following the next equation:

$$(H - ACCD_{rt-1}) + I_{i_t} * IE - D_t + P_t - ET_{ct} - 0.2 * ET_t^o \leq H \;\; \forall\, t \tag{14}$$

The expression $H - ACCD_{rt-1}$ relates the water reaming in the root zone in time t, where H represents the water holding capacity estimated by the Total Available Water (TAW) parameter, which is the maximum amount of water present in the plant root zone (see Appendix D Sheet TAW-Kc) and $ACCD_{rt-1}$ the amount of water depleted in the root zone in the last period. I_{i_t} and D_t irrigating and draining decision variables, IE the irrigation efficiency of drip irrigation (i.e., 0.9 according to Yang et al. (2023)) P_t the estimated precipitation, ET_{ct} the evapotranspiration value in the planning horizon estimated by the ET_o simulation process. The expression $0.2 * ETo_t$ represents an estimated amount of water evaporated daily regarding rainfall and irrigation processes (Allen et al., 2006). This study establishes the first constraint by rearranging the Eq [15]:

$$I_{i_t} * IE - D_t \leq -P_t + ET_t^c + 0.2 * ET_t^o + ACCD_{rt-1}$$

Considering:

$$D_{rt} = -P_t + ET_t^c + 0.2 * ET_t^o \tag{15}$$

Then:

$$I_{i_t} * IE - D_t \leq D_{rt} + ACCD_{rt-1}$$

Where D_{rt} represents the amount of water depleted in the time t.

7.3.2. Minimum water available in the root zone. The DM comprises the minimum amount of water needed in the plant root zone to prevent yield detriment in the crop (Q. Li & Hu, 2020). Then decision-maker must provide the resource once water depletion reaches this threshold. Through the Steduto et al. (2012) equation, the model can represent the cocoa crop yield losses regarding water stress levels:

$$\left(1 - \frac{Y_t^a}{Y_t^m}\right) = k_y \left(1 - \frac{ET_t^a}{ET_t^c}\right) \quad [16]$$

where Y_t^a is the actual crop yield or crop yield under water stress conditions $[kg/ha]$, Y_t^m is the maximum crop yield $[Kg/ha]$, ET_t^a represents the actual evapotranspiration $[mm/day]$, ET_t^c the total maximum crop evapotranspiration $[mm/day]$, and k_y is a dimensionless parameter that represents the crop yield reduction regarding water stress. This study allows a 1% yield reduction as the maximum regarding the absence of water resources. Therefore, the expression is as follows:

$$(1 - 0.99) = k_y \left(1 - \frac{ET_t^a}{ET_t^c}\right)$$

Or:

$$0.01 = k_y \left(1 - \frac{ET_t^a}{ET_t^c}\right)$$

This expression represents that DM only endures yield reduction up to a maximum of 1%, which means the water soil content in any period must be enough to reach a 99% yield. Consequently, the equation now is:

$$0.01 \geq k_y \left(1 - \frac{ET_t^a}{ET_t^c}\right)$$

Where ET_t^a relies on:

$$ET_t^a = k_s k_c ET_t^o$$

Where k_s describes the effect of water stress on crop transpiration, k_c the crop evapotranspiration coefficient and ET_t^o the grass reference evapotranspiration. Now regarding ET_t^a the expression is as follows:

$$0.01 \geq k_y \left(1 - \frac{k_s k_c ET_t^o}{ET_t^c}\right)$$

$$\frac{0.01}{k_y} \geq \left(1 - \frac{k_s k_c ET_t^o}{ET_t^c}\right)$$

Where $ET_c = k_c ET_o$, so:

$$\frac{0.01}{k_y} \geq \left(1 - \frac{k_s \cancel{ET_t^c}}{\cancel{ET_t^c}}\right)$$

Therefore:

$$\frac{0.01}{k_y} \geq (1 - k_s) \quad [17]$$

Now, considering k_s relies on:

$$k_s = \frac{TAW - Dr}{TAW - RAW} \quad [18]$$

Where RAW represents the Readily Available Water and represents the amount of water crop can deplete by evapotranspiration way just before suffering water stress and yield reduction. Now the expression is:

$$\frac{0.01}{k_y} \geq \left(1 - \frac{TAW - Dr}{TAW - RAW}\right) \quad [19]$$

Where Dr the water depleted. Through the last expression, this study can determine the maximum amount of water depleted allowed in the root zone or the minimum amount of water that needs to exist in the root zone to guarantee the expected yield. In this study, $TAW = 140$, $RAW = 42$ regarding $RAW = cd * TAW$ where cd is 0.3 for the cocoa crop (see Appendix D). This work implements a cocoa crop yield response factor of 0.25 based on a fit generated from a simulation analysis using the Simple model (C. Zhao et al., 2019). Such an analysis allows for determining the value of k_y that produces a reduction in cocoa yield considering evapotranspiration using an irrigation lack yield estimation and a fully irrigated yield estimation in 2020 (see Appendix D for further details). Then the equation is as follows:

$$\frac{0.01}{0.25} \geq \left(1 - \frac{140 - Dr}{140 - 42}\right)$$

Now isolating Dr this study determines that the amount of water to support that yield reduction at maximum needs to be:

$$Dr = 45.92$$

Then the minimum amount of water that always needs to be in the root zone is:

$$WP = TAW - Dr \quad [20]$$

$$WP = 140 - 45.92 = 94.08$$

Where WP is the wilting point represents the minimum amount of water in the root zone to prevent water stress and yield reduction. Nevertheless, this value means a moment in which if the irrigation planning fails, the crop will suffer water stress quickly, producing a loss yield; therefore, we define a middle point that allows irrigation actions without incurring risks due to irrigation failures as follows:

$$WP = TAW - \frac{Dr}{2} = 140 - \frac{45.92}{2} = 117.04$$

Then, the second constraint relates to the following expression:

$$WP = H - Dr$$

Now, regarding Dr is the water depleted every day, following the next expression:

$$Dr_t = ACCDr_{t-1} - P_t - I_{it} * IE + ET_t^c + 0.2 * ET_t^o \quad [21]$$

The expression is as follows:

$$WP = H - Dr_t$$

$$WP = H - ACCDr_{t-1} + P_t + I_{it} * IE - ET_t^c - 0.2 * ET_t^o$$

Rearranging the equation:

$$I_{it} * IE \geq WP - H + ACCDr_{t-1} - P_t + ET_t^c + 0.2 * ET_t^o$$

Considering:

$$D_{rt} = -P_t + ET_t^c + 0.2 * ET_t^o$$

Then the second constraint is:

$$I_{it} * IE \geq WP - H + ACCDr_{t-1} + D_{rt} \quad [22]$$

7.3.3. Water availability in every period. The water availability constraint describes a maximum irrigation water use restriction (Fu, Li, Cui, et al., 2018; Q. Li & Hu, 2020) related to two sources to water the cocoa crop. The first relates to the amount of water DM can recollect from rainfall, and the second refers to a contract with another supplier to provide the water needs whenever the farmer requires it. The first one relates to no cost since it derives from rainfall capture. The second refers to the cost of acquiring the resource elsewhere (e.g., manager basin, river, government contract). In this sense, the water supported by the second alternative is unlimited, while the rainfall water stored water is limited in every period. This study establishes that the farmer can only store around 20% of the rainfall in every period, then:

$$WA_t = P_t * 0.2 \qquad [23]$$

Where WA_t is the water availability in every period regarding rainfall periods. Therefore, the constraint that relates to the usable water from the first source is as follows:

$$\sum_{t=1}^{2} I_{1t} \leq WA_t \qquad [24]$$

7.3.4. Nonnegative variables. The final model assumptions regarding decision variables relate to their nonnegative properties in the model.

$$I_{it}, D_t \geq 0 \qquad [25]$$

7.3.5. Objective function. The decision model seeks to produce an optimal water management scheme considering the costs associated with irrigating or draining the crop. In this sense, the objective function of a generic model relates to the following equation:

$$\min z = \sum_{t=1}^{2} \sum_{i=1}^{2} C_i * I_{it} + \sum_{t=1}^{2} DC * D_t \qquad [26]$$

Where C_i is the cost of water from the i source, I_{it} the irrigation water required from the i source in time t, DC the draining cost and D_t the amount of water drained in time t. Then, the deterministic and stochastic models in this study are:

7.3.6. Cocoa crop water management Deterministic Model.

Indexes

i the different types of water sources available to supply the resource {1,2}.

t the unit time where every decision is made {1,2}.

Parameters

C_{ij} cost of i source of water [$ $Currency\ Units - CU$].

DC draining cost [$ $Currency\ Units - CU$].

H the soil water holding capacity [mm].

IE the irrigation efficiency of the drip water irrigation system [%].

Dr_2 water depletion at the end of the second period [mm].

$ACCDr_2$ the cumulated water depletion in the root zone at the end of the first period in the second period [mm].

WP the minimum required amount of water in the root zone [mm].

WA_2 represents the amount of water the DM can store from the first source at the end of the second period [mm].

Variables

I_{it} the amount of water required to support productivity.

D_t is the draining needs once the water content in the root zone is higher than H.

Objective function

$$\min z \rightarrow \sum_{t=1}^{2}\sum_{i=1}^{2} C_{it} * I_{it} + \sum_{t=1}^{2} DC * D_t \qquad [27]$$

Subject to

$$\sum_{t=1}^{2}\sum_{i=1}^{2} I_{it} * IE - \sum_{t=1}^{2} D_t \leq Dr_2 + ACCDr_2 \qquad [28]$$

$$\sum_{t=1}^{2}\sum_{i=1}^{2} I_{it} * IE \geq WP - H + Dr_2 + ACCDr_2$$

$$\sum_{t=1}^{2} I_{1t} \leq WA_2$$

$$I_{it}\ , D_t \geq 0 \quad \forall i$$

Therefore, the model disclosed in equations 27 and 28 presents a two-period decision scheme to minimize irrigation operating costs. The objective function resembles the irrigating and draining costs. On the other hand, the constraints relate to requirements associated with supporting the minimum water crop conditions and the cultivar system capacities. The first constraint relates to the maximum limit of soil water retention conditioned by the irrigation system; the second constraint is the minimum level needed to avoid a yield loss in the crop; the third constraint is the available water capacities, and the last constraint is non-negativity conditions of the decision variables. The decisions relate mainly to the second period considering that the conditions are established for the end of the planning horizon (i.e., the end of the second period).

7.3.7. Cocoa crop water management Two-stage Stochastic Model.

Indexes

i the different types of water sources available to supply the recourse $\{1,2\}$.

t the unit time where every decision is made $\{1,2\}$.

w every scenario related $\{3\ scen: 1-3; 5\ scen\ model: 1-5; 10\ scen\ model: 1-10\}$.

Parameters

C_{i1} cost of every source in the first period [\$ CU].

C_{i2w} cost of every source in the second period regarding the scenario w [\$ CU].

DC draining cost [\$ CU].

p_w is the probability of each scenario w considered [%].

H the soil water holding capacity [mm].

IE the irrigation efficiency of the drip water irrigation system [%].

Dr_{2w} water depletion at the end of the second period considering every scenario w [mm].

$ACCDr_{2w}$ the cumulated water depletion in the root zone at the end of the first period in the second period considering every scenario w $[mm]$.

WP the minimum required amount of water in the root zone $[mm]$.

WA_{2w} represents the amount of water the DM can store from the first source at the end of the second period regarding every scenario w $[mm]$.

Variables

I_{i1} the amount of water required from any source to support productivity in the first period.

I_{i2w} the amount of water required from any source to support productivity in the second period regarding every scenario w.

D_1 a variable that represents the cost production of extracting water once the water content in the root zone is higher than H in the first period.

D_{2w} a variable that represents the cost production of extracting water once the water content in the root zone is higher than H in the second period and regarding every scenario w.

Objective function

$$\min z \rightarrow \sum_{i=1}^{2} C_{i1} * I_{i1} + DC * D_1 + \sum_{w=1}^{w} \sum_{i=1}^{2} p_w * (C_{i2w} * I_{i2w} + DC * D_{2w}) \qquad [29]$$

Subject to

[30]

$$\sum_{i=1}^{2} I_{i_1} * IE + \sum_{i=1}^{2} I_{i_{2w}} * IE - D_1 - D_{2w} \leq Dr_{2w} + ACCDr_{2w} \ \forall \ w$$

$$\sum_{i=1}^{2} I_{i_1} * IE + \sum_{i=1}^{2} I_{i2w} * IE \geq WP - H + Dr_{2w} + ACCDr_{2w} \ \forall \ w$$

$$I_{11} + I_{12w} \leq WA_{2w} \quad \forall w$$

$$I_{i1}, D_1, I_{i2w}, D_{2w} \geq 0 \quad \forall w, \forall i$$

$$\sum_{w=1}^{w} p_w = 1$$

In this sense, the model exposed in equations 29 and 30 presents a two-stage decision scheme for irrigation cost minimization under uncertain conditions. In this model, the objective function seeks to reduce the average costs associated with irrigation by considering multiple uncertain scenarios. The objective function relates to two parts: the first describes the irrigation and draining costs in the first stage, and the second relates the irrigation and draining costs weighted by an occurrence probability (P) regarding each precipitation scenario. Hence, this model's constraints have the same essence as the previous model but connect a constraint for each scenario. Then, if the model integrates three uncertain scenarios, each constraint relates three constraints of the same type: one constraint for each scenario considered. Therefore, the model must determine the irrigating and draining decisions in both stages, simultaneously considering all the scenarios of the second period and thus supporting a robust decision in the face of different future scenarios. The decisions relate mainly to the second period considering that the conditions are established for the end of the planning horizon (i.e., the end of the second period).

7.4. Computer-based solution procedure

Programming the models in the Google Colab environment using Python's Pyomo library allows this study to solve the formulated cocoa decision models (see Appendix M for every model result). Considering that the formulated models integrate reduced precipitation fans (i.e., 1, 3, 5, and 10 instead of 1000), solution strategies to solve large-scale problems are irrelevant, regarding that the complex model (i.e., any strategy of 10 scenarios) takes around 0.18 seconds to provide a solution. Such a situation is supported considering that the mathematical models are linear with continuous variables and the model constraints have a matrix full of zeros since each constraint relates a maximum of two variables (i.e., corresponding to an identity matrix). The CPLEX solver integrated into the Colab environment using the Pyomo library allows for obtaining computationally efficient and optimal solutions for the developed models. This study builds an online repository that hosts all programmed models to support further analysis of decision processes, the importance of integrating parameter modeling strategies under uncertainty, and the appropriate use of these strategies (see Appendix N). Appendix O relates the data used for the optimization models.

7.5. Optimization model results and models comparison

This study addresses the optimization of water resources using deterministic and Two-stage Stochastic optimization models. Figure 20 presents all the models built in this research (six deterministic and nine stochastic-based models).

Figure 20

Study optimization models

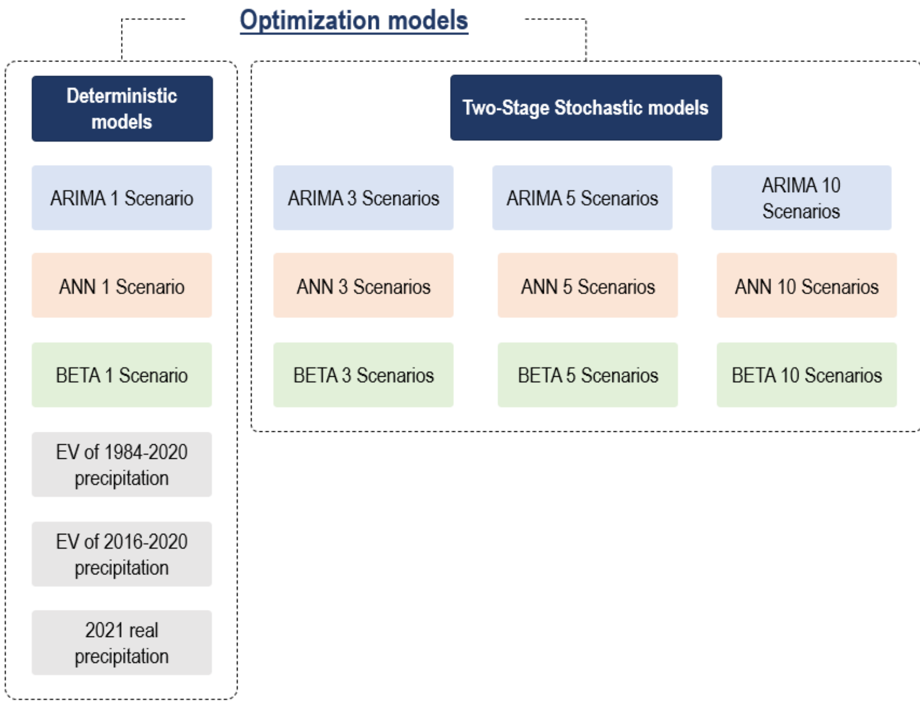


Figure 21 presents the monetary cost of the deterministic decision models (orange bars) and stochastic models (blue bars) contrasted against a deterministic optimization model that uses the real value of precipitation during the study period, which reflects the optimal and real decision of the decision (green bar) allows this study to establish that the decision models that support a reduction in costs are predominantly the deterministic model that uses the expected precipitation value for the 2016-2020 period, followed by decision models based on the BETA pdf. However, although these strategies provide an optimal decision process in economic terms, they reflect a non-veridical reality compared to the real precipitation model (data in Appendix P). This situation derives from the deterministic method (2016-2020 average precipitation), and the models based on BETA tend to overestimate rainfall, which leads to making wrong decisions considering that, in reality, the water supply regarding rainfall periods is less. Under such a premise, the deterministic models that use the expected value and the BETA strategies represent counterproductive approaches to replicate the parameter behavior, promoting incorrect decisions derived from the crop water balance.

Figure 21

Irrigation cost of every strategy

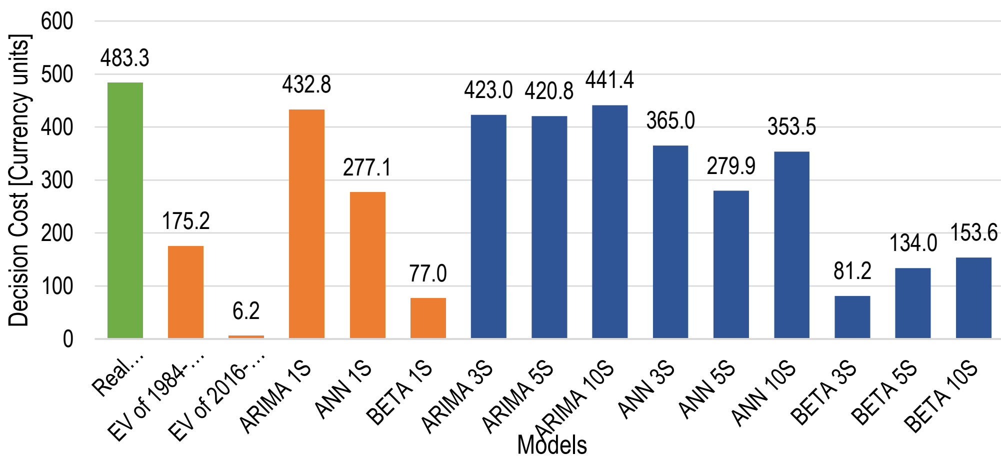
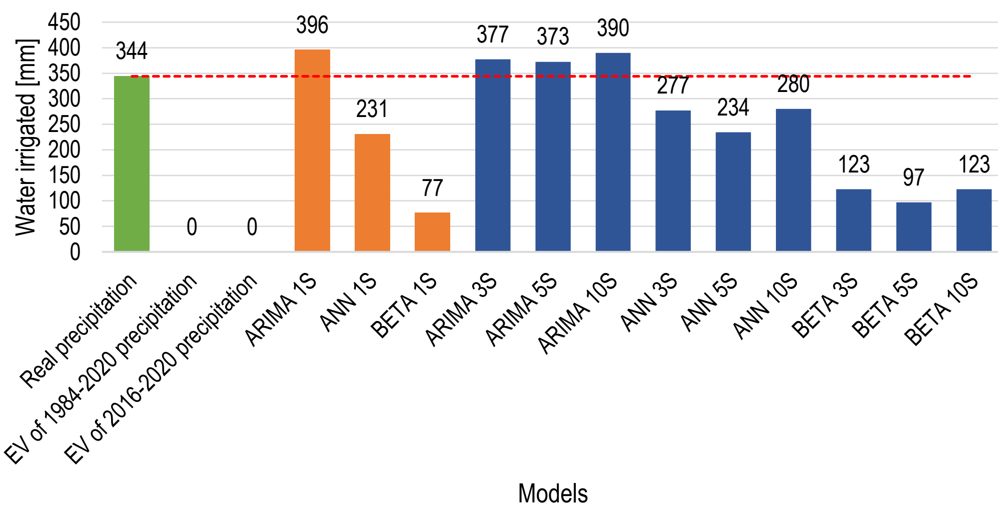


Figure 22 supports the statement above by displaying that the water volume determined for annual planning is zero in the expected value deterministic strategies followed by BETA models. In this way, it is clear that using average precipitation values does not represent an acceptable procedure to support the water management process. Additionally, strategies based on probability function fits, such as the BETA pdf, are equally unsuitable alternatives to support the decision process. Of the models built, the best techniques derive from autoregressive or machine learning models. Of the TSP models based on ARIMA parameter modeling, the model that allows the lowest cost and provides a water management scheme similar to the real one is the ARIMA with five scenarios with a cost of $420.8\ [cu]$ and a use of $373\ [mm]$. On the other hand, the best ANN-based TSP strategy is not clear, considering that the 5-scenario ANN model has the lowest cost $(279.9\ [cu])$ while the 10-scenario ANN model is the best water management strategy $(280\ [mm])$. Considering that the 10-scenario ANN model is closer to the real plan in irrigation terms, this study determines the TSP model based on 10-scenario ANN as the optimal one within the compendium of ANN strategies.

Figure 22

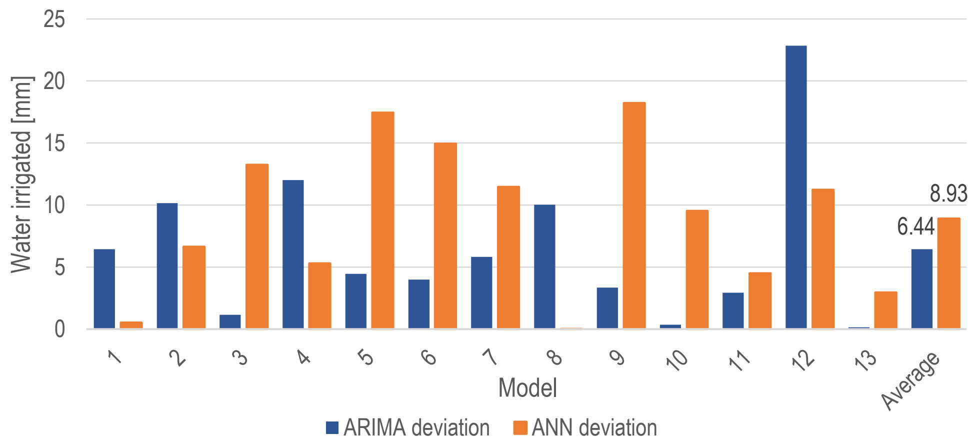
Used irrigation water



This study considers that the TSP strategy based on the ARIMA 5-scenarios model represents the best option to solve water management problems under conditions of uncertainty in the cocoa crops case study. Although the ARIMA-based TSP model presents a higher irrigation cost than the ANN-based TSP, it supports a better water allocation scheme considering that the amount of planned water resource represents the smallest deviation in most periods and the year. Similarly, both strategies present optimization schemes that reduce irrigation costs concerning the real scheme. Figure 23 presents both models' absolute deviations regarding the real irrigation scheme. It is important to state that the higher value in the irrigation cost of the real strategy in Figure 21 concerning the TSP-ARIMA 5 scenarios model (from now on) derives from a drainage process in the last period due to atypical rainfall.

Figure 23

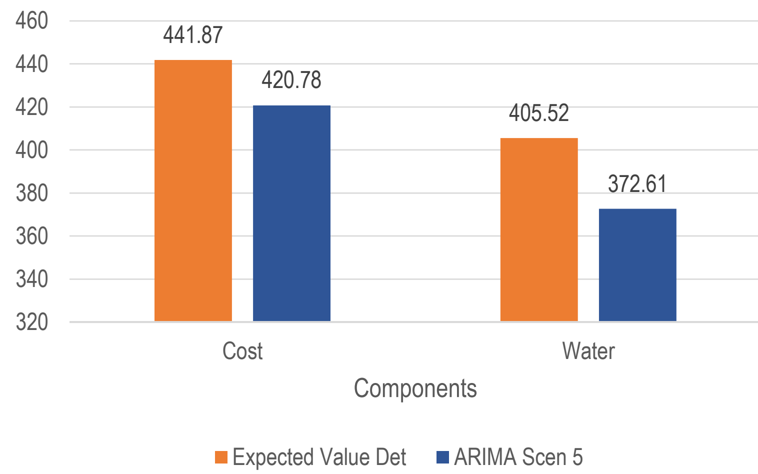
TSP ARIMA 5 scenarios and TSP ANN 10 scenarios used irrigation water



7.5.1. ARIMA 5 scenarios expected value deterministic model and TSP-ARIMA 5 scenarios comparison. This study develops a last deterministic model built from the TSP-ARIMA 5 scenarios model to expand the analysis of the performance of models under uncertainty against deterministic models. The optimization model takes the precipitation values as the expected value of the five scenarios generated from the ARIMA model and the reduction of scenarios. Figure 24 compares the models based on the irrigation cost in the annual planning and the water resource optimization. Figure 24 allows this study to show a superior performance of the TSP-ARIMA 5 scenarios model even against its deterministic counterpart (i.e., the expected value of the TSP-ARIMA 5 scenarios model). This comparison supports stating that using proper parameter modeling techniques and optimization techniques under uncertainty, such as TSP, are relevant alternatives to support cocoa crops water management (i.e., per plant), reducing the associated risk and incorrect decisions.

Figure 24

Cost and used irrigation water of Expected Value deterministic and TSP ARIMA 5 scenarios model

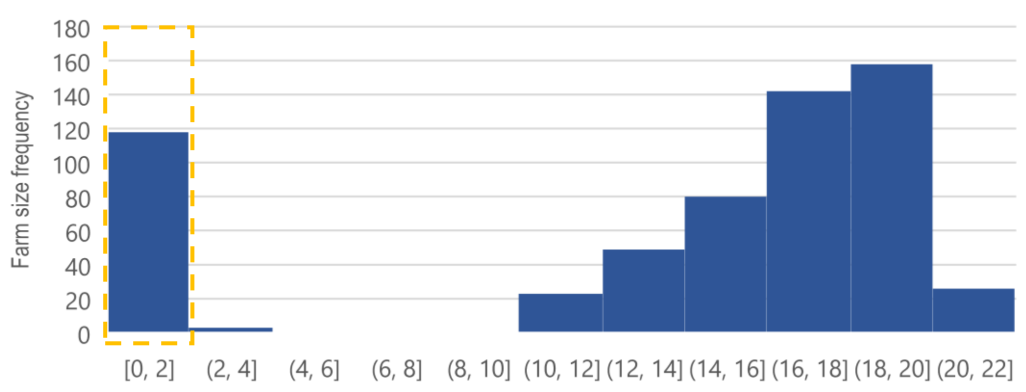


The previous results represent the cost reduction and water irrigation use per cocoa plant in one year. Regarding the size of the cocoa plantations in the San Vicente del Chucurí region and the distance from plant to plant, this study can determine approximately how much water resources the models allow to save annually. Note that in a typical cocoa plantation, the trees are approximately 3 m apart (PNUD, 2014), and around 1111 plants fit in one hectare ($10000\ m^2/9m^2$ $-$per plant (assuming square plantation)) if the land use is appropriate. According to the 2014 national agricultural census (DANE, 2016a) in San Vicente, by 2013, the largest small farmers had farms of 2 hectares maximum (Figure 25) (Appendix Q). Considering that the study focuses mainly on the small farmers and the number of possible cultivable plants in one hectare, this study determines that the water saved in a $2\ ha$ farm per year is $73.326\ liters$ (Eq [31]), which represents a significant result in the process of water management in agriculture.

$$Water\ saving = plants\ in\ 2\ ha * water\ saved\ per\ plant$$
$$Water\ saving = 2222 * 33 = 73.326\ liters\ annually \quad [31]$$

Figure 25

Cocoa farm size distribution in San Vicente del Chucurí, Santander



7.5.2. Quality metrics performance. Additionally, two metrics widely discussed in the literature allow determining the use of TSP strategies' significance in solving problems under uncertain conditions. The EVPI and VSS metrics establish what a DM will pay for real process information and the value of developing optimization models compared to deterministic models. Each metric is applied to the first iteration of the TSP-ARIMA model of 5 (see Appendix R), regarding the model is the same applied repeatedly (only changing the water balance in the rest of the iterations). The EVPI metric in this case study shows that a decision maker would be willing to pay $6.42\ [cu]$ or 1.53% of the annual planning cost per plant to improve the decision considering that $EVPI = 0.24\ [cu]$. The value of EVPI supports stating that concerning the TSP-ARIMA 5 scenarios model, paying for more accurate information does not significantly improve the decision. On the other hand, the VSS metric exposes the value of implementing TSP strategies in water resource optimization, considering that the annual cost reduction concerning a deterministic approach based on the parameters defined in the same model is 15.81% considering that in the first planning period the $VSS = 2.56\ [cu]$. The metric value supports and guarantees a better performance of TSP models over deterministic ones in the cocoa case study.

8. Discussion

In Colombia, cocoa has gained significant importance as one of the key crops, with the government making efforts to boost national production since it supports as a substitute alternative for illicit crops (Fedecacao, 2021; Minagricultura, 2016). However, cocoa crops face difficulties in ensuring optimal conditions affecting crop yield (Carr & Lockwood, 2011; Chapman et al., 2021a; Cilas & Bastide, 2020; Finagro, 2018; Plazas et al., 2017). These challenges primarily stem from weather factors, especially precipitation, which restricts water availability and impacts irrigation, eventually affecting biomass production and cocoa crop yield (Fedecacao, 2018; P. Guo et al., 2010). Thus, it is essential to develop and apply strategies that enable effective water resource management under climate variability. The MSP and TSP strategies represent flexible optimization schemes that generate a decision process composed of multiple stages where uncertainty is present in future periods of the planning horizon. Such strategies denote appropriate alternatives to support water management problems in cocoa crops, considering the technological advance and the adoption of data collection and processing technologies. Regarding such situations, this study's main objective was to develop TSP models supporting decision-making focused on optimizing water resources in cocoa crops while providing an updated panorama of the cocoa crop challenges regarding uncertain conditions and climate change.

The SLR allows for defining research gaps about optimal schemes in cocoa crop water management under climatic uncertainty regarding the following findings: (i) the significant studies focused on the farmer or final decision-maker support decisions based on water requirements simulation, with a lack of studies based on agricultural optimization strategies. (ii) there is an inequality favoring research supporting the first and second decision levels (i.e., government and basin manager) concerning the last decision level (i.e., farmer), which represents a counterproductive approach considering that farmers are the most significant proportion of water resource demands among the industrial, municipal, and ecological sectors (iii) there is a predominance of studies focused on annual crops relating cereals primarily with few studies focused on perennial crops. However, the SLR establishes that supporting water management processes under multiple types and forms of uncertainty with uncertain optimization models at the first and second decision levels represents an acceptable strategy to face water management problems under uncertainty in several crops.

The main alternatives developed established using the MSP and TSP model, Interval Programming, Fuzzy Programming, and various hybrid strategies derived from integrating these techniques to answer the problem's uncertainty adequately. Each technique has specific characteristics allowing more suitable than others for uncertain situations. Interval programming is a strategy that allows appropriate optimization schemes in situations of a lack of data associated with the uncertain parameter (Xiaoyan Li et al., 2014). Fuzzy programming is appropriate for facing uncertainties in the form of ambiguity or vagueness, requiring experts' knowledge (S. Guo et al., 2019). On the other hand, strategies based on stochastic techniques (i.e., MSP or TSP) represent relevant alternatives when the uncertainty derives from random processes, and there is sufficient available data to model the variability of the parameter (Marques et al., 2010).

Among every uncertain optimization modeling strategy, TSP represents a well-structured and flexible technique to support cocoa crop water management representing a lower cost decision modeling compared to MSP (Liu et al., 2017). Then, this research builds different deterministic and TSP models supported by different fitted models based on time series to provide a contrast about using both strategies. Such contrasts between the results of the models allow for highlighting the significance of modeling under uncertainty and the impact on cocoa crop water decision-making. This research builds five deterministic models using approaches based on one path and nine TSP models that integrate precipitation uncertainty using scenarios in model formulation and solution. Optimization results allow concluding that, in general, the strategies that integrate uncertainty in the decision model promote adequate water management considering precipitation variability. Of the multiple modeling strategies applied, the ARIMA model presented the best alternatives to support appropriate water management in its different versions (i.e., 1, 3, 5, and 10) scenarios. Although ANN-based strategies allow a better approach decision than traditional strategies, they also present significant deviations from the real scenario. On the other hand, optimization strategies based on BETA pdf settings are just not good. This situation results from generating possible precipitation values independently using the pdf instead of strategies that reflect the data's historical behavior.

In this sense, the results of the optimization model under uncertainty TSP-ARIMA 5 scenarios model (i.e., the best TSP model) compared to the deterministic version using the expected value of the five scenarios (equivalent expected model) allow this study to establish that the stochastic model performs better in managing water resources under uncertain conditions. Such a statement answers the study hypothesis that a stochastic optimization model integrated with a good parameter modeling technique allows a water optimization scheme under conditions of climatic uncertainty superior to the equivalent expected model. Additionally, the EVPI metric demonstrates that, based on the TSP-ARIMA 5 scenarios model, paying for more accurate information on precipitation does not substantially improve the current decision, which makes it possible to confirm that the model built adequately responds to the problem. On the other hand, the VSS metric shows that the optimization model under uncertainty represents a proper strategy to improve water management in cocoa cultivation by reducing annual costs by approximately 15% compared to the equivalent expected model.

This study condensed a significant effort into supporting parameter definition associated with the optimization models. Promoting relevant parameter estimation strategies means that decisions are appropriate and similar to reality when formulating and solving different mathematical models. Nevertheless, establishing the proper related parameters in the cocoa crop context denotes a difficult task based on the little information available in the literature. Therefore, although the optimization models represent straightforward decision schemes, the previous work to define and estimate the model factors entails a significant effort as it allows defining parameters not available in the literature, such as S_{water}or the $k_y$factor. Then, this study defines a promising first approach to establishing the relationship between water stress and crop yield, considering the existing literature does not fully define such a relationship (Zuidema et al., 2005). Following the Simple study (C. Zhao et al., 2019), a crop S_{water} factor of 0.65 represents a reasonable approximation of biomass sensitivity to water stress conditions since it falls into a specific range of other known crop species ($0.4 - 2.5$). Additionally, a crop yield response factor to water stress conditions (k_y) of 0.25 reflects a proper assessment since the cocoa crop is a perennial crop with some resilience to water stress conditions but is sensitive to long-term drought periods reflecting a low but existing yield detriment depending on the cocoa genotype (Lahive et al., 2019; Steduto et al., 2012; Zuidema et al., 2005).

Considering the little research on optimization schemes for the allocation and optimal use of resources in cocoa crops (Tosto et al., 2023), the results of this research establish a first effort in modeling cocoa crop water management under uncertainty. The optimization models focus on supporting the minimum conditions required by the cocoa, allowing the plant not to reach yield detriment. Such conditions derive from keeping soil moisture within limits that allow the proper growth and plant production while including system constraints. Regarding the almost null research associated with cocoa crops, the proposed models are supported primarily on parameters defined by the FAO (Allen et al., 2006; Steduto et al., 2012) and related studies (León-Moreno et al., 2019; Romero Vergel et al., 2022), deriving modeling decisions from similar studies (Fu, Li, Cui, et al., 2018; Jamal et al., 2018; Q. Li & Hu, 2020) Therefore, this study allows formulation and solving optimization models that represent the crop water problem supported by accurate and reliable data exhibited in several studies and different exponents of the research field.

This study reveals that the selection and application of the scenario generation techniques are equally crucial to the optimization model selection. The modeling technique, coupled with the scenario reduction strategy, provides a reasonable estimate of the actual future behavior of the uncertain parameter through a proper scenario fan. In this sense, the findings of this study support the hypothesis that stochastic optimization models integrated with parameter modeling techniques perform better in water resource management; however, it is essential to note that most results drive a robust based on actual data into a theoretical approach. Such a situation states that the parameter estimation process for the optimization models requires significant effort due to their significance in model solutions and the limited information available in the literature about cocoa crops. Then, defining parameters properly still represents significant limitations requiring proper addressing when developing optimization models to reflect real-world decisions. The optimization models might provide wrong results if they relate wrong or improper assumptions (e.g., invariant irrigation cost, stable soil conditions, crop age, cocoa type, and healthy conditions) which can severely impact the optimal solution and decision risk reduction in cocoa crop management considering that such parameters fluctuate over time, leading to more complex decision scenarios. Therefore, establishing adequate parameters defines an essential and critical point for the problem solution, which can lead to inappropriate decisions and produce inaccurate decision schemes.

Since this study only relies on data-driven optimization models and historical data, there are still approaches that may provide better results regarding additional types of uncertainty than randomness (Fu, Li, Cui, et al., 2018; S. Guo et al., 2019; Y. P. Li & Huang, 2011; Youzhi et al., 2021; W. J. Zhang et al., 2021). Hybrid strategies can provide suitable alternatives to model uncertain conditions in cocoa crop water by integrating randomness processes, ambiguous situations in establishing parameters behavior, and not representing constant behavior in parameters while providing risk-aware decision-making. Therefore, future studies incorporating advanced modeling techniques to improve data collection methods, integrating a more comprehensive range of system components variability and the multiple forms of uncertainty during the planning period can enhance this study's findings' accuracy, robustness, and applicability. In this sense, such studies should integrate simulation strategies (i.e., AquaCrop (Y. Wang & Guo, 2021), DSSAT (Linker, 2021), CropWat 8 (Gabr & Fattouh, 2021), Case-2 (Zuidema et al., 2005) among others) considering the plant's physiological and phenological characteristics while considering system uncertainty derived from climatic scenarios and market and population irregular behaviors in the market, among others. Additionally, this research established a first approach to cocoa crop water management under uncertain conditions based on parameter definitions related to cocoa crop yield and water conditions. However, future studies in situ will make it possible to validate these estimates and adjust the value considering a parameter definition process when analyzing the daily variation. Research with this approach would allow better-defined and robust parameters to be integrated into decision schemes under uncertainty, presenting more complete decision processes.

9. Conclusions

Planning the available natural resources in the agricultural sector represents one of the main axes to support agricultural development and meet global food needs by reacting to the multiple conditions demanded by climate change. In Colombia, one of the crops with the most significant importance in recent years is cocoa, considering the effort made by the government to boost national production and the position of the crop internationally. At the national level, Santander is the leading cocoa producer, with a total production of 40% of the national production. However, this crop presents difficulties with guaranteeing the minimum conditions required to avoid affecting yield. This situation derives from the effect of weather factors such as precipitation, which restricts water availability and conditions irrigation, which is a factor that directly impacts biomass production and, therefore, plant yield. Hence, strategies that allow facing water resources management under climate variability are currently relevant to reduce the decision-maker risk, manage the resource properly, and influence the economy of those involved.

This study supports establishing three primary research gaps related to the research field based on the systematic literature review. 1. The largest proportion of studies associated with government and reservoir managers (first decision levels) regarding they face the highest risk in the water allocation process. 2. There is a global hurry to develop studies that support decisions on specifically annual crops over perennial ones, considering the risk of the crop and the significance of this crop for food security. 3. Most studies primarily focus on strategies based on production process simulation to support DM, wasting the benefits of optimization techniques to support optimal resource management. Therefore, this study developed a framework to provide a methodological structure supporting future research on optimal agricultural water resources. The framework describes six detailed steps relating to the main aspects of answering water management problems in multiple forms of uncertainty. However, the SLR clearly defines a lack of studies based on optimization strategies that allow proper water resource management and the minimum conditions for cocoa crops.

Therefore, supporting the appropriate use of water resources in cocoa crops through optimization schemes under uncertainty represents a pertinent and essential issue to guarantee crop yield while preserving currently available resources. However, building appropriate optimization models entails a demanding parameter estimation process and complex modeling strategies selection to reproduce the behavior of the uncertain parameter. The parameters definition related to the cocoa crop presents challenges, considering that despite the extensive research associated with the crop, the studies still dismiss the impact of water stress on the crop. Most studies establish a relationship between the lack of water resources and crop yield, but such a relationship is still unknown. This research results allow for defining this relationship from a simulation process considering the productivity of a type of cocoa clone (ICS 95) and the lack of water resources. The simulation process allows defining a yield factor $k_y = 0.25$ and biomass growth $S_{water} = 0.65$ regarding the crop water stress. Such estimations agree with the literature presenting a moderate relationship between crop yield and drought conditions.

Then, based on an appropriate definition of parameters related to cocoa crop conditions, different optimization strategies support crop water conditions, properly manage resources, and guarantee crop yield. Such strategies must allow the integration of uncertainty in the decision process to constitute robust and appropriate strategies that support decisions at all decision levels and reduce the associated risk. Hence, this study developed multiple TSP models to compare the uncertain strategies performance concerning deterministic models in cocoa crop water management, addressing a case study focused on the San Vicente del Chucuri area in Santander, Colombia. Additionally, the proposed models present a decision scheme for using water resources supported by an exhaustive and detailed definition of parameters to lay the foundations for future optimization models derived from this study. The uncertain models developed in this study derive from integrating the precipitation uncertainty into a TSP framework using scenario representations built from time series and pdf-based models. On the other hand, deterministic model approaches derive from approximations based on the expected value and the complete reduction of the scenarios.

The optimization models present a scheme that covers 14 days horizon decision, considering that this period relates to a time window where the crop begins to suffer stress due to lack of soil moisture. The TSP optimization model (TSP-ARIMA 5 scenarios) allows an adequate representation of reality, considering that the use of water resources per plant presents a deviation of 29 [mm] concerning the real scenario, producing a better result than the developed deterministic versions. In this sense, the TSP models developed represent adequate alternatives to respond to the water management problem, producing superior decision schemes considering the model results. The superior performance of the TSP strategies over the deterministic models derives from considering multiple possible realizations of the uncertain parameter, allowing a decision with more information and robustness regarding several scenarios and reducing the error incurred when defining a crisp value. Therefore, this study's results establish that data modeling techniques based on time series for scenario generation and using a technique to reduce the fan integrated with the application of TSP models provide appropriate water management schemes under climatic uncertainty in the case of cocoa study performing better results than deterministic approaches.

Consequently, this study's results contribute to the research panorama on managing water resources in conditions of climatic uncertainty in cocoa crops. The study lists significant gaps in agricultural water management in uncertain conditions. It also provides a well-structured TSP model-building methodological approach based on available scientific literature related to cocoa modeling, offering a synthetic and precise decision scheme on the major problem-associated factors. Regarding the research results, it is pertinent to establish that in the cocoa case study, an optimization model based on TSP combined with proper scenario generation and reduction techniques enables better water resource management than a traditional strategy providing a good base for answering the hypothesis stated in this research.

References

Agbenyo, W., Jiang, Y., Ding, Z., Titriku, J. K., & Ntim-Amo, G. (2022). Impact of Climate Change on Cocoa Production in Africa: An Approach of Cross-sectional ARDL. *International Journal of Environmental Research*, *16*(5), 91. https://doi.org/10.1007/s41742-022-00471-0

Agronet. (2020, November 13). *UPRA presentó soluciones alternativas de acceso al agua para mejorar la productividad de la agricultura familiar*. Agronet. https://www.agronet.gov.co/Noticias/Paginas/UPRA-presentó-soluciones-alternativas-de-acceso-al-agua-para-mejorar-la-productividad-de-la-agricultura-familiar.aspx

Ahumada, H., & Cornejo, M. (2016). Forecasting food prices: The case of corn, soybeans and wheat. *International Journal of Forecasting*, *32*(3), 838–848. https://doi.org/10.1016/j.ijforecast.2016.01.002

Ali, M. H., & Talukder, M. S. U. (2008). Increasing water productivity in crop production-A synthesis. *Agricultural Water Management*, *95*(11), 1201–1213. https://doi.org/10.1016/j.agwat.2008.06.008

Allen, R., Pereira, L., Raes, D., & Smith, M. (2006). Crop evapotranspiration - Guidelines for computing crop water requirements. In FAO (Ed.), *Evapotranspiración del cultivo Guías para la determinación de los requerimientos de agua de los cultivos. ESTUDIO FAO RIEGO Y DRENAJE 56.*

Amfo, B., Ali, E. B., & Atinga, D. (2021). Climate change, soil water conservation, and productivity: Evidence from cocoa farmers in Ghana. *Agricultural Systems*, *191*, 103172. https://doi.org/https://doi.org/10.1016/j.agsy.2021.103172

Arksey, H., & O’Malley, L. (2005). Scoping studies: Towards a methodological framework.

International Journal of Social Research Methodology: Theory and Practice, *8*(1), 19–32. https://doi.org/10.1080/1364557032000119616

Bailer, A. (2001). Probabilistic Techniques in Exposure Assessment. A Handbook for Dealing with Variability and Uncertainty in Models and Inputs. A. C. Cullen and H. C. Frey. *Statistics in Medicine*, *20*. https://doi.org/10.1002/sim.958

Bakkalbasi, N., Bauer, K., Glover, J., & Wang, L. (2006). Three options for citation tracking: Google Scholar, Scopus and Web of Science. *Biomedical Digital Libraries*, *3*(2003), 1–8. https://doi.org/10.1186/1742-5581-3-7

Beale, E. M. L. (1955). On Minimizing a Convex Function Subject to Linear Inequalities. *Journal of the Royal Statistical Society: Series B (Methodological)*, *17*(2), 173–184. https://doi.org/10.1111/j.2517-6161.1955.tb00191.x

Bellman, R., & Zadeh, L. (1970). Decision-Making in a Fuzzy Environment. *Management Science*, *17*(4). https://doi.org/10.1142/9789812819789_0004

Bethune, M. G., Selle, B., & Wang, Q. J. (2008). Understanding and predicting deep percolation under surface irrigation. *Water Resources Research*, *44*(12). https://doi.org/https://doi.org/10.1029/2007WR006380

Birge, J. R. (1982). The value of the stochastic solution in stochastic linear programs with fixed recourse. *Mathematical Programming*, *24*(1), 314–325. https://doi.org/10.1007/BF01585113

Birge, J. R., & Louveaux, F. (2011). *Introduction to Stochastic Programming*. Springer New York. https://doi.org/10.1007/978-1-4614-0237-4

Bomdzele, E., & Molua, E. L. (2023). Assessment of the impact of climate and non-climatic parameters on cocoa production: a contextual analysis for Cameroon. *Frontiers in Climate*, *5*. https://doi.org/10.3389/fclim.2023.1069514

Box, G. E. P., & Cox, D. R. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, *26*(2), 211–252. http://www.jstor.org/stable/2984418

CAF. (2020). Iniciativa Latinoamericana del Cacao. In *Scioteca* (Vol. 8). http://scioteca.caf.com/handle/123456789/1530

Caracol Radio. (2016). *50% de la producción del cacao en Antioquia se daña por el fenómeno de 'El Niño' | Medellín | Caracol Radio*. https://caracol.com.co/emisora/2016/02/02/medellin/1454418858_518366.html

Caracol Radio. (2020). *Ola de calor en el Huila afecta 2.000 hectáreas de cacao*. https://caracol.com.co/emisora/2020/10/25/neiva/1603644889_799813.html

Caracol Radio. (2022). *Agricultura tropical ¿El cambio climático está impactando la producción de cacao? : ¿El cambio climático está impactando la producción de cacao?* Caracol Radio. https://caracol.com.co/programa/2022/01/29/al_campo/1643415363_140373.html

Carr, M. K. V., & Lockwood, G. (2011). The water relations and irrigation requirements of cocoa (Theobroma cacao L.): A review. *Experimental Agriculture*, *47*(4), 653–676. https://doi.org/10.1017/S0014479711000421

Cerón Salazar, I. X., García Muñoz, M. C., Rodríguez Polanco, L., Cubillos Varela, A., Polanco Díaz, E., López Hernández, M. del P., Parra Alferes, E. B., Moreno Turriago, J. M., & Quiñones Méndez, F. (2020). Estrategias tecnológicas para el manejo del cultivo y el beneficio del cacao. In *Estrategias tecnológicas para el manejo del cultivo y el beneficio del cacao*. https://doi.org/10.21930/agrosavia.institutional.7403572

Ch, B., & F, C. (2021). Cambio Climático y Cacao en Colombia. *Encuesta Mensual de Servicios*, *4*(23), 2021.

https://www.inei.gob.pe/media/MenuRecursivo/boletines/boletin_servicios_9.pdf

Chapman, R., Cock, J., Samson, M., Janetski, N., Janetski, K., Gusyana, D., Dutta, S., & Oberthür, T. (2021a). Crop response to El Niño-Southern Oscillation related weather variation to help farmers manage their crops. *Scientific Reports, 11*(1), 1–8. https://doi.org/10.1038/s41598-021-87520-4

Chapman, R., Cock, J., Samson, M., Janetski, N., Janetski, K., Gusyana, D., Dutta, S., & Oberthür, T. (2021b). Crop response to El Niño-Southern Oscillation related weather variation to help farmers manage their crops. *Scientific Reports*, *11*(1), 1–8. https://doi.org/10.1038/s41598-021-87520-4

Cilas, C., & Bastide, P. (2020). Challenges to Cocoa Production in the Face of Climate Change and the Spread of Pests and Diseases. *Agronomy*, *10*(9), 1–8. https://doi.org/10.3390/agronomy10091232

Conejo, A. J., Carrión, M., & Morales, J. M. (2010). *Decision Making Under Uncertainty in Electricity Markets* (Vol. 153). Springer US. https://doi.org/10.1007/978-1-4419-7421-1

Critchley, Wi., & Siegert, K. (1991). Water harvesting. In *FAO*. https://www.fao.org/3/u3160e/u3160e00.htm

Cryer, J. D., & Chan, K.-S. (2008). *Time Series Analysis with applications in R*. Springer New York. https://doi.org/10.1007/978-0-387-75959-3

Dai, Z. Y., & Li, Y. P. (2013). A multistage irrigation water allocation model for agricultural land-use planning under uncertainty. *Agricultural Water Management*, *129*, 69–79. https://doi.org/10.1016/j.agwat.2013.07.013

DANE. (2016a). *COLOMBIA - Tercer Censo Nacional Agropecuario - 2014 -3er CNA*. http://microdatos.dane.gov.co/index.php/catalog/513/get_microdata

DANE. (2016b). *Tercer Censo Nacional Agropecuario - 2014 -3er CNA*. https://microdatos.dane.gov.co/index.php/catalog/513/get_microdata

Dantzig, G. B. (1955). Linear Programming under Uncertainty. *Management Science*, *1*(3–4), 197–206. https://doi.org/10.1287/mnsc.1.3-4.197

De Almeida, A. A., & Valle, R. R. (2008). Ecophysiology of the cacao tree. *Brazilian Journal of Plant Physiology*, *19*(December), 425–448.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *Journal of the American Statistical Association*, *74*(366), 427–431. http://www.jstor.org/stable/2286348

Djuideu, C. T. L., Bisseleua, H. D. B., Kekeunou, S., & Ambele, F. C. (2021). Rehabilitation practices in cocoa agroforestry systems mitigate outbreaks of termites and support cocoa tree development and yield. *Agriculture, Ecosystems and Environment*, *311*(October 2020). https://doi.org/10.1016/j.agee.2021.107324

Dos Santos, I. C., De Almeida, A. A. F., Anhert, D., Da Conceição, A. S., Pirovani, C. P., Pires, J. L., Valle, R. R., & Baligar, V. C. (2014). Molecular, physiological and biochemical responses of theobroma cacao L. genotypes to soil water deficit. *PLoS ONE*, *9*(12), 1–31. https://doi.org/10.1371/journal.pone.0115746

Dupačová, J., Gröwe-Kuska, N., & Römisch, W. (2003). Scenario reduction in stochastic programming. *Mathematical Programming*, *95*(3), 493–511. https://doi.org/10.1007/s10107-002-0331-0

Dupačová, Jitka, Consigli, G., & Wallace, S. W. (2000). Scenarios for Multistage Stochastic Programs. *Annals of Operations Research*, *100*(1), 25–53. https://doi.org/10.1023/A:1019206915174

EL TIEMPO. (2021). *Chocolate podría extinguirse en menos de 30 años por cambio climático - Medio Ambiente - Vida - ELTIEMPO.COM*. EL TIEMPO. https://www.eltiempo.com/vida/medio-ambiente/chocolate-podria-extinguirse-en-menos-de-30-anos-por-cambio-climatico-618216

Falagas, M. E., Pitsouni, E. I., Malietzis, G. A., & Pappas, G. (2008). Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses. *The FASEB Journal*, *22*(2), 338–342. https://doi.org/10.1096/fj.07-9492lsf

FAO. (2017). Water for Sustainable Food and Agriculture. In *A report produced for the G20 Presidency of Germany*. www.fao.org/publications

Fedecacao. (2018). *El agua, clave para el desarrollo del cacao*. Fedecacao. http://www.fedecacao.com.co/portal/index.php/es/2015-04-23-20-00-33/585-el-agua-clave-para-el-desarrollo-del-cacao

Fedecacao. (2020, February 25). *Así se comportó la producción de cacao por departamentos en el 2020*. Fedecacao. https://www.fedecacao.com.co/portal/index.php/es/2015-04-23-20-00-33/1385-asi-se-comporto-la-produccion-de-cacao-por-departamentos-en-el-2020

Fedecacao. (2021, April 5). *Gobierno invirtió más de $11 mil millones para sustituir cultivos ilícitos por cacao*. Fedecacao. https://www.fedecacao.com.co/portal/index.php/es/2015-04-23-20-00-33/1409-gobierno-invirtio-mas-de-11-mil-millones-para-sustituir-cultivos-ilicitos-por-cacao

Finagro. (2018). Inteligencia de mercado: Cacao. In *Finagro: Vol. II*. https://www.finagro.com.co/sites/default/files/node/basic-page/files/ficha_cacao_version_ii.pdf

Finagro. (2020). Ficha de inteligencia: Cacao. In *Finagro*.

https://www.finagro.com.co/sites/default/files/node/basic-page/files/ficha_tabaco_version_ii.pdf

Fix, E., & Hodges, J. L. (1989). Discriminatory Analysis. Nonparametric Discrimination: Consistency Properties. *International Statistical Review / Revue Internationale de Statistique*, *57*(3), 238–247. https://doi.org/10.2307/1403797

Fu, Q., Li, L., Li, M., Li, T., Liu, D., Hou, R., & Zhou, Z. (2018). An interval parameter conditional value-at-risk two-stage stochastic programming model for sustainable regional water allocation under different representative concentration pathways scenarios. *Journal of Hydrology*, *564*(April), 115–124. https://doi.org/10.1016/j.jhydrol.2018.07.008

Fu, Q., Li, T., Cui, S., Liu, D., & Lu, X. (2018). Agricultural Multi-Water Source Allocation Model Based on Interval Two-Stage Stochastic Robust Programming under Uncertainty. *Water Resources Management*, *32*(4), 1261–1274. https://doi.org/10.1007/s11269-017-1868-2

Gabr, M. E., & Fattouh, E. M. (2021). Assessment of irrigation management practices using FAO-CROPWAT 8, case studies: Tina Plain and East South El-Kantara, Sinai, Egypt. *Ain Shams Engineering Journal*, *12*(2), 1623–1636. https://doi.org/https://doi.org/10.1016/j.asej.2020.09.017

Gateau-Rey, L., Tanner, E. V. J., Rapidel, B., Marelli, J. P., & Royaert, S. (2018). Climate change could threaten cocoa production: Effects of 2015-16 El Niño-related drought on cocoa agroforests in Bahia, Brazil. *PLoS ONE*, *13*(7), 1–17. https://doi.org/10.1371/journal.pone.0200454

Geurts, M., Box, G. E. P., & Jenkins, G. M. (1977). Time Series Analysis: Forecasting and Control. *Journal of Marketing Research*. https://doi.org/10.2307/3150485

Graupe, D. (2013). *Principles of artificial neural networks* (3rd ed.). World Scientific.

Grote, U., Fasse, A., Nguyen, T. T., & Erenstein, O. (2021). Food Security and the Dynamics of Wheat and Maize Value Chains in Africa and Asia. *Frontiers in Sustainable Food Systems*, *4*(February), 1–17. https://doi.org/10.3389/fsufs.2020.617009

Gujarati, D. N., & Porter, D. C. (2013). Econometría. In *Mc Graw Hill* (Vol. 53, Issue 9). https://doi.org/10.1017/CBO9781107415324.004

Guo, P., Huang, G. H., & Li, Y. P. (2010). Inexact fuzzy-stochastic programming for water resources management under multiple uncertainties. *Environmental Modeling and Assessment*, *15*(2), 111–124. https://doi.org/10.1007/s10666-009-9194-6

Guo, S., Zhang, F., Zhang, C., Wang, Y., & Guo, P. (2019). An improved intuitionistic fuzzy interval two-stage stochastic programming for resources planning management integrating recourse penalty from resources scarcity and surplus. *Journal of Cleaner Production*, *234*, 185–199. https://doi.org/10.1016/j.jclepro.2019.06.183

Harzing, A. W., & Alakangas, S. (2016). Google Scholar, Scopus and the Web of Science: a longitudinal and cross-disciplinary comparison. *Scientometrics*, *106*(2), 787–804. https://doi.org/10.1007/s11192-015-1798-9

Heitsch, H., & Römisch, W. (2003). Scenario Reduction Algorithms in Stochastic Programming. *Computational Optimization and Applications*, *24*(2), 187–206. https://doi.org/10.1023/A:1021805924152

Hillier, F. S., & Lieberman, G. J. (2010). *Introducción a la investigación de Operaciones*. Mc Graw Hill.

Hou, J., Fan, X., & Liu, R. (2016). Optimal spatial allocation of irrigation water under uncertainty using the bilayer nested optimisation algorithm and geospatial technology. *International Journal of Geographical Information Science*, *30*(12), 2462–2485.

https://doi.org/10.1080/13658816.2016.1181264

Høyland, K., & Wallace, S. W. (2001). Generating Scenario Trees for Multistage Decision Problems. *Management Science*, *47*(2), 295–307. https://doi.org/10.1287/mnsc.47.2.295.9834

Huang, G. H., & Loucks, D. P. (2000). An inexact two-stage stochastic programming model for water resources management under uncertainty. *Civil Engineering and Environmental Systems*, *17*(2), 95–118. https://doi.org/10.1080/02630250008970277

Hyndman, R. J., & Athanasopoulos, G. (2014). Forecasting: Principles and Practice. *OTexts, july 1994*, 46–51. https://doi.org/10.1017/CBO9781107415324.004

Inter Press Service. (2020). *La Niña Weather Phenomenon Could Endanger Colombia's Food Security*. http://www.ipsnews.net/2020/12/la-nina-weather-phenomenon-endanger-colombias-food-security/

Irshad, M., Petersen, K., & Poulding, S. (2018). A systematic literature review of software requirements reuse approaches. *Information and Software Technology*, *93*(September 2017), 223–245. https://doi.org/10.1016/j.infsof.2017.09.009

Jamal, A., Linker, R., & Housh, M. (2018). Comparison of Various Stochastic Approaches for Irrigation Scheduling Using Seasonal Climate Forecasts. *Journal of Water Resources Planning and Management*, *144*(7), 04018028. https://doi.org/10.1061/(asce)wr.1943-5452.0000951

Ji, L., Zhang, B., Huang, G., & Lu, Y. (2020). Multi-stage stochastic fuzzy random programming for food-water-energy nexus management under uncertainties. *Resources, Conservation and Recycling*, *155*(September 2019), 104665. https://doi.org/10.1016/j.resconrec.2019.104665

Jr., F. J. M. (1951). The Kolmogorov-Smirnov Test for Goodness of Fit. *Journal of the American*

Statistical Association, *46*(253), 68–78. https://doi.org/10.1080/01621459.1951.10500769

Kang, S., Hao, X., Du, T., Tong, L., Su, X., Lu, H., Li, X., Huo, Z., Li, S., & Ding, R. (2017). Improving agricultural water productivity to ensure food security in China under changing environment: From research to practice. *Agricultural Water Management*, *179*, 5–17. https://doi.org/10.1016/j.agwat.2016.05.007

Kang, S., Su, X., Tong, L., Shi, P., Yang, X., Abe, Y., Du, T., Shen, Q., & Zhang, J. (2004). The impacts of human activities on the water-land environment of the Shiyang River basin, an arid region in northwest China. *Hydrological Sciences Journal*, *49*(3), 413–427. https://doi.org/10.1623/hysj.49.3.413.54347

Kaur, P., Dhir, A., Tandon, A., Alzeiby, E. A., & Abohassan, A. A. (2021). A systematic literature review on cyberstalking. An analysis of past achievements and future promises. *Technological Forecasting and Social Change*, *163*(September 2020), 120426. https://doi.org/10.1016/j.techfore.2020.120426

La opinión. (2015). *Sequía afecta la siembra y cosecha de cacao, café y arroz*. https://www.laopinion.com.co/economia/sequia-afecta-la-siembra-y-cosecha-de-cacao-cafe-y-arroz

Läderach, P., Martinez-Valle, A., Schroth, G., & Castro, N. (2013). Predicting the future climatic suitability for cocoa farming of the world's leading producer countries, Ghana and Côte d'Ivoire. *Climatic Change*, *119*(3), 841–854. https://doi.org/10.1007/s10584-013-0774-8

Lahive, F., Hadley, P., & Daymond, A. J. (2019). The physiological responses of cacao to the environment and the implications for climate change resilience. A review. *Agronomy for Sustainable Development*, *39*(1). https://doi.org/10.1007/s13593-018-0552-0

León-Moreno, C., Rojas-Molina, J., & Castilla-Campos, C. (2019). Physicochemical characteristics of cacao (Theobroma cacao L.) soils in colombia: Are they adequate to improve productivity? *Agronomia Colombiana*, *37*(1), 52–62. https://doi.org/10.15446/agron.colomb.v37n1.70545

Lewis-Beck, M. S., Bryman, A., & Liao, T. F. (2004). *The SAGE encyclopedia of qualitative research methods Vol.3*. 907–908.

Li, C., & Grossmann, I. E. (2021). A Review of Stochastic Programming Methods for Optimization of Process Systems Under Uncertainty. *Frontiers in Chemical Engineering*, *2*(January), 1–20. https://doi.org/10.3389/fceng.2020.622241

Li, Q., & Hu, G. (2020). Multistage stochastic programming modeling for farmland irrigation management under uncertainty. *PLoS ONE*, *15*(6), 1–21. https://doi.org/10.1371/journal.pone.0233723

Li, Xiaoyan, Lu, H., He, L., & Shi, B. (2014). An inexact stochastic optimization model for agricultural irrigation management with a case study in China. *Stochastic Environmental Research and Risk Assessment*, *28*(2), 281–295. https://doi.org/10.1007/s00477-013-0748-4

Li, Xiaoyun, Liu, N., You, L., Ke, X., Liu, H., Huang, M., & Waddington, S. R. (2016). Patterns of cereal yield growth across China from 1980 to 2010 and their implications for food production and food security. *PLoS ONE*, *11*(7), 1–18. https://doi.org/10.1371/journal.pone.0159061

Li, Y. P., & Huang, G. H. (2011). Planning agricultural water resources system associated with fuzzy and random features. *Journal of the American Water Resources Association*, *47*(4), 841–860. https://doi.org/10.1111/j.1752-1688.2011.00558.x

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Journal of Clinical Epidemiology*, *62*(10), e1–e34. https://doi.org/10.1016/j.jclinepi.2009.06.006

Linker, R. (2021). Stochastic model-based optimization of irrigation scheduling. *Agricultural Water Management*, *243*(August 2020), 106480. https://doi.org/10.1016/j.agwat.2020.106480

Liu, J., Li, Y. P., Huang, G. H., Zhuang, X. W., & Fu, H. Y. (2017). Assessment of uncertainty effects on crop planning and irrigation water supply using a Monte Carlo simulation based dual-interval stochastic programming method. *Journal of Cleaner Production*, *149*, 945–967. https://doi.org/10.1016/j.jclepro.2017.02.100

Marques, G. F., Lund, J. R., & Howitt, R. E. (2010). Modeling Conjunctive Use Operations and Farm Decisions with Two-Stage Stochastic Quadratic Programming. *Journal of Water Resources Planning and Management*, *136*(3), 386–394. https://doi.org/10.1061/(asce)wr.1943-5452.0000045

Marquez, J., Talero-Sarmiento, L. H., & Lamos, H. (2022). *Multistage Stochastic Programming to Support Water Allocation Decision-Making Process in Agriculture: A Literature Review*. 26. https://doi.org/10.3390/iocag2022-12307

Minagricultura. (2016, August 26). *Cacao es ejemplo de paz y de sustitución de importaciones*. Minagricultura. https://www.minagricultura.gov.co/noticias/Paginas/Cacao-es-ejemplo-de-paz-y-de-sustitución-de-importaciones.aspx

Minagricultura. (2020, November 19). *Minagricultura anuncia plan de renovación de diez mil hectáreas de cacao en 2021 para mejorar la productividad*. Minagricultura. https://www.minagricultura.gov.co/noticias/Paginas/Minagricultura-anuncia-plan-de-renovación-de-diez-mil-hectáreas-de-cacao-en-2021-para-mejorar-la-productividad.aspx

Muhammad, Y. S., & Pflug, G. C. (2014). Stochastic vs deterministic programming in water management: the value of flexibility. *Annals of Operations Research*, *223*(1), 309–328. https://doi.org/10.1007/s10479-013-1455-8

Naranjo-Merino, C. A., Ortíz-Rodriguez, O. O., & Villamizar-G, R. A. (2017). Assessing green and blue water footprints in the supply chain of cocoa production: A case study in the Northeast of Colombia. *Sustainability (Switzerland)*, *10*(1). https://doi.org/10.3390/su10010038

NASA. (2021). *POWER | Data Access Viewer*. https://power.larc.nasa.gov/data-access-viewer/

Noticias Canal TRO. (2019). *Sequía afecta cultivos de cacao y aguacate en San Vicente de Chucurí*. https://noticias.canaltro.com/sequia-afecta-cultivos-de-cacao-y-aguacate-en-san-vicente-de-chucuri/

Organización Internacional del trabajo. (2023). *Oportunidades, retos y necesidades para promover trabajo decente en la cadena de suministro del cacao en Colombia*.

Pannell, D. J., Malcolm, B., & Kingwell, R. S. (2000). Are we risking too much? Perspectives on risk in farm modelling. *Agricultural Economics*, *23*(1), 69–78. https://doi.org/10.1016/S0169-5150(00)00058-X

Phillips, P., & Perron, P. (1986). Testing for a Unit Root in Time Series Regression. *Cowles Foundation, Yale University, Cowles Foundation Discussion Papers*, *75*. https://doi.org/10.1093/biomet/75.2.335

Plazas, J. E., López, I. D., & Corrales, J. C. (2017). A tool for classification of cacao production in Colombia based on multiple classifier systems. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in*

Bioinformatics), *10405*(857), 60–69. https://doi.org/10.1007/978-3-319-62395-5_5

PNUD. (2014). *Guía para el cultivo del cacao*. https://issuu.com/pnudcol/docs/cartilla-final

Portafolio. (2009). *Cae producción mundial de cacao; Fenómeno de "El Niño" le pasó la cuenta de cobro a los cultivadores del grano*. https://www.portafolio.co/economia/finanzas/cae-produccion-mundial-cacao-fenomeno-nino-le-paso-cuenta-cobro-cultivadores-grano-465622

Raiffa, H. (1968). *Decision Analysis: Introductory Lectures on Choices Under Uncertainty*.

Robers, P. D., & Ben-Israel, A. (1969). *Interval Programming: New approach to linear programming with applications to chemical engineering problems*. 5. https://doi.org/10.1021/i260032a011

Romero Vergel, A. P., Camargo Rodriguez, A. V., Ramirez, O. D., Arenas Velilla, P. A., & Gallego, A. M. (2022). A Crop Modelling Strategy to Improve Cacao Quality and Productivity. *Plants*, *11*(2). https://doi.org/10.3390/plants11020157

Sahinidis, N. V. (2004). Optimization under uncertainty: state-of-the-art and opportunities. *Computers & Chemical Engineering*, *28*(6), 971–983. https://doi.org/https://doi.org/10.1016/j.compchemeng.2003.09.017

Samian, M., Mahdei, K. N., Saadi, H., & Movahedi, R. (2015). Identifying factors affecting optimal management of agricultural water. *Journal of the Saudi Society of Agricultural Sciences*, *14*(1), 11–18. https://doi.org/10.1016/j.jssas.2014.01.001

Schwendenmann, L., Veldkamp, E., Moser, G., Hölscher, D., Köhler, M., Clough, Y., Anas, I., Djajakirana, G., Erasmi, S., Hertel, D., Leitner, D., Leuschner, C., Michalzik, B., Propastin, P., Tjoa, A., Tscharntke, T., & Van Straaten, O. (2010). Effects of an experimental drought on the functioning of a cacao agroforestry system, Sulawesi, Indonesia. *Global Change*

Biology, *16*(5), 1515–1530. https://doi.org/https://doi.org/10.1111/j.1365-2486.2009.02034.x

Semana. (2014). *Sequía prende alarmas de cacaoteros*. https://www.semana.com/afectacion-del-fenomeno-del-nino-produccion-cacao/199018/

Shiferaw, B., Smale, M., Braun, H. J., Duveiller, E., Reynolds, M., & Muricho, G. (2013). Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. *Food Security*, *5*(3), 291–317. https://doi.org/10.1007/s12571-013-0263-y

Somers, L. D., & McKenzie, J. M. (2020). A review of groundwater in high mountain environments. *WIREs Water*, *7*(6), e1475. https://doi.org/https://doi.org/10.1002/wat2.1475

Steduto, P., Hsiao, T. C., Fereres, E., & Raes, D. (2012). *Crop yield response to water*. FAO. https://www.fao.org/3/i2800e/i2800e00.htm

Talero, L., Marquez, J. D., & Lamos, H. (2023). A roadmap to solve two-stage stochastic problems implementing scenario reduction for agricultural production planning under uncertainty A ROADMAP TO SOLVE TWO-STAGE STOCHASTIC PROBLEMS IMPLEMENTING SCENARIO REDUCTION FOR AGRICULTURAL PRODUCTION PLANNING U. *IX Congreso Internacional Industria y Organizaciones*, *May*.

Tosto, A., Morales, A., Rahn, E., Evers, J. B., Zuidema, P. A., & Anten, N. P. R. (2023). Simulating cocoa production: A review of modelling approaches and gaps. *Agricultural Systems*, *206*, 103614. https://doi.org/https://doi.org/10.1016/j.agsy.2023.103614

Unión Europea. (2017). *El fenómeno de El Niño y la agricultura latinoamericana*. http://www.ncdc.noaa.gov/ersst/

United Nations. (2015). Transforming our world: the 2030 Agenda for Sustainable Development

Preamble. In *United Nations*. https://doi.org/10.1163/157180910X12665776638740

Vagropoulos, S. I., Kardakos, E. G., Simoglou, C. K., Bakirtzis, A. G., & Catalão, J. P. S. (2016). ANN-based scenario generation methodology for stochastic variables of electric power systems. *Electric Power Systems Research*, *134*, 9–18. https://doi.org/https://doi.org/10.1016/j.epsr.2015.12.020

Wang, Y., & Guo, P. (2021). Irrigation water resources optimization with consideration of the regional agro-hydrological process of crop growth and multiple uncertainties. *Agricultural Water Management*, *245*, 106630. https://doi.org/10.1016/j.agwat.2020.106630

Wang, Y., Li, Z., Guo, S., Zhang, F., & Guo, P. (2020). A risk-based fuzzy boundary interval two-stage stochastic water resources management programming approach under uncertainty. *Journal of Hydrology*, *582*(August 2019), 124553. https://doi.org/10.1016/j.jhydrol.2020.124553

Wang, Y. Y., Huang, G. H., & Wang, S. (2017). CVaR-based factorial stochastic optimization of water resources systems with correlated uncertainties. *Stochastic Environmental Research and Risk Assessment*, *31*(6), 1543–1553. https://doi.org/10.1007/s00477-016-1276-9

Weng, Y., Wang, X., Hua, J., Wang, H., Kang, M., & Wang, F. Y. (2019). Forecasting Horticultural Products Price Using ARIMA Model and Neural Network Based on a Large-Scale Data Set Collected by Web Crawler. *IEEE Transactions on Computational Social Systems*, *6*(3), 547–553. https://doi.org/10.1109/TCSS.2019.2914499

Xu, D., Chen, Z., & Yang, L. (2012). Scenario tree generation approaches using K-means and LP moment matching methods. *Journal of Computational and Applied Mathematics*, *236*(17), 4561–4579. https://doi.org/10.1016/j.cam.2012.05.020

Yang, P., Wu, L., Cheng, M., Fan, J., Li, S., Wang, H., & Qian, L. (2023). Review on Drip

Irrigation: Impact on Crop Yield, Quality, and Water Productivity in China. *Water*, *15*(9). https://doi.org/10.3390/w15091733

Youzhi, W., Alexander, F., & Ping, G. (2021). A model integrating the system dynamic model with the risk based two-stage stochastic robust programming model for agricultural-ecological water resources management. *Stochastic Environmental Research and Risk Assessment*, *8*. https://doi.org/10.1007/s00477-021-01972-8

Zhang, F., Guo, P., Engel, B. A., Guo, S., Zhang, C., & Tang, Y. (2019). Planning seasonal irrigation water allocation based on an interval multiobjective multi-stage stochastic programming approach. *Agricultural Water Management*, *223*(17), 105692. https://doi.org/10.1016/j.agwat.2019.105692

Zhang, W. J., Tan, Q., & Zhang, T. Y. (2021). A risk-averse stochastic quadratic model with recourse for supporting irrigation water management in uncertain and nonlinear environments. *Agricultural Water Management*, *244*(March 2020), 106431. https://doi.org/10.1016/j.agwat.2020.106431

Zhao, C., Liu, B., Xiao, L., Hoogenboom, G., Boote, K. J., Kassie, B. T., Pavan, W., Shelia, V., Kim, K. S., Hernandez-Ochoa, I. M., Wallach, D., Porter, C. H., Stockle, C. O., Zhu, Y., & Asseng, S. (2019). A SIMPLE crop model. *European Journal of Agronomy*, *104*(January), 97–106. https://doi.org/10.1016/j.eja.2019.01.009

Zhao, J., Li, M., Guo, P., Zhang, C., & Tan, Q. (2017). Agricultural water productivity oriented water resources allocation based on the coordination of multiple factors. *Water (Switzerland)*, *9*(7). https://doi.org/10.3390/w9070490

Zuidema, P. A., Leffelaar, P. A., Gerritsma, W., Mommer, L., & Anten, N. P. R. (2005). A physiological production model for cocoa (Theobroma cacao): model presentation, validation and application. *Agricultural Systems*, *84*(2), 195–225. https://doi.org/https://doi.org/10.1016/j.agsy.2004.06.015