



End of studies project

Topic

Pricing Of An Index-based Flood Insurance Product Using Satellite Data

Case Of Cereal Production Of The Wilaya of El Tarf -Algeria-

Presented and defended by:

CHIKHI Loubna

Supervised by:

Mr. OUESLATI Ghazi

Student belonging to:

Ministry of Finance – Algeria -

ACKNOWLEDGMENTS

In the name of Allah, the Most Gracious, the Most Merciful.

First and foremost, I extend my deepest gratitude to my supervisor, Mr. OUESLATI Ghasi, whose support and encouragement guided me in choosing and pursuing this research. His mentorship provided invaluable insights and direction throughout this journey, helping to shape this paper into its final form. For this, I am profoundly grateful.

I would also like to express my sincere thanks to Mr. AMARI Djamel, Director of Agricultural Insurance, Mrs. KEBAB Fatiha, Director of State Funds, and Mr. IMOUDECHE Farouk, Director of the Actuarial and Statistics Unit at CNMA. Their generous time, invaluable information, and support were instrumental in advancing my research. I am also grateful to the entire team at CNMA's training cell for their consistent encouragement during my internship. Their collective efforts have significantly contributed to the success of this work.

Furthermore, I wish to express my sincere appreciation to the staff of IFID, particularly Mr. Khaled ZOUARI, for his exceptional dedication to our professional development. The comprehensive training and unwavering support I received at IFID have greatly contributed to the successful completion of this thesis.

Lastly, I would like to thank all those who have contributed to this work, whether directly or indirectly. Your efforts and support have been greatly appreciated.

DEDICATION

To my mother, whose unwavering belief in me has been my greatest strength.

To my father, whose constant support has guided me throughout this journey.

To my brother, Ramzy, and my sisters, Amel, Maroua, and Nadine, for always having my back.

To my cherished nieces and nephew, who bring joy to my life.

I dedicate this work to all of you, with deep love and gratitude.

SUMMARY

ACKNOWLEDGMENTS

DEDICATION

LIST OF ABBREVIATIONS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

GENERAL INTRODUCTION..... 1

**CHAPTER ONE: THEORICAL FRAMEWORK: THE INSURANCE AGAINST
AGRICULTURAL FLOOD RISK..... 4**

INTRODUCTION 4

SECTION 1: UNDERSTANDING FOOD RISKS 5

SECTION 2: INSURANCE AGAINST FLOOD RISK IN AGRICULTURE 10

SECTION 3: MODELING EXTREME EVENTS 17

COCLUSION..... 26

**CHAPTER TWO: EMPIRICAL ANALYSIS: CONSTRUCTION AND PRICING OF
AGRICULTURAL FLOOD INDEX-BASED INSURANCE 28**

INTRODUCTION 28

SECTION 1: ASSESSMENT OF THE SITUATION AND PRODUCT CONSTRUCTION 29

SECTION 2: MODELING FLOOD VULNERABILITY 42

SECTION 3: MODELING OF EXTREME PRECIPITATION AND PRODUCT PRICING 56

CONCLUSION..... 72

GENERAL CONCLUSION..... 74

BIBLIOGRAPHY

APPENDICES

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AAL	Average Annual Loss	GDP	Gross Domestic Product
AIC	Akaike Information Criterion	GEV	Generalized extreme value
ANRH	Agence nationale des ressources hydrauliques	GLM	Generalized Linear Model
ASAL	L'Agence spatiale algérienne	GPD	Generalized Pareto Distribution
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer	Ha	Hectare (10 000 m ²)
CDF	Cumulative Distribution Function	INSID	Institut National des Sols, de l'Irrigation et Drainage
CERES	Clouds and the Earth's Radiant Energy System	ITGC	Institut Technique des Grandes Cultures
CNA	Conseil National des Assurances	JAXA	Japanese Aerospace Exploration Agency
CNMA	Caisse Nationale de mutualité Agricole	MADR	Ministère de l'agriculture et du développement rural
DACA	Dispositif d'Assurance des Calamités Agricoles	MENA	Middle East North Africa
DEM	Digital Elevation Model	MERRA	Modern-Era Retrospective analysis for Research and Applications
ESA	European Space Agency	MPCI	Multi-peril Crop Insurance
ESRI	Environmental Systems Research Institute	NASA	National Aeronautics and Space Administration
EVT	Extreme Value Theory	ONM	Office national de la météorologie
FAO	Food and Agriculture Organization	POT	peak-over-threshold
FGCA	Fonds de Garantie des Calamités Agricoles	TWI	Topographic wetness index
GDAP	Gross Domestic Agricultural Production	VIF	Variance Inflation Factor

LIST OF FIGURES

Figure 1: Diagram of different risk categories.....	5
Figure 2: GEV density with $\mu=0$, $\sigma=1$ for different values of the shape parameter ξ	22
Figure 3: Agriculture, value added (% of GDP) from 2000 to 2022	30
Figure 4: Flood frequency in Algeria from 1980 to 2022.....	32
Figure 5: Geographic situation of the Wilaya of El Tarf	35
Figure 6: The variations in monthly precipitation	37
Figure 7: Histogram of daily precipitation.....	37
Figure 8: Precipitation cumulative (mm) for a flood event	39
Figure 9: Map showing flooded and non-flooded locations in agriculture areas	43
Figure 10: Distance to river's map of the Wilaya of El Tarf	46
Figure 11: Elevation's map of the Wilaya of El Tarf	46
Figure 12: Slope's map of the Wilaya of El Tarf	46
Figure 13: TWI's map of the Wilaya of El Tarf	46
Figure 14: Soil types' map of the Wilaya of El Tarf	46
Figure 15: Soil's moisture map of the Wilaya of Tarf.....	46
Figure 16: Methodological Flowchart for Flood Vulnerability	47
Figure 17: Destruction function of the Wilaya of El Tarf	55
Figure 18: Illustration of dependency in two days cumulative precipitation.	57
Figure 19: Mean Excess Function.....	60
Figure 20: Hill Plot	61
Figure 21: Goodness of Fit of the Data to the Gumbel Distribution.....	64
Figure 22: Convergence of the sample mean to the theoretical mean.....	67
Figure 23: Sensitivity of the pure premium to the choice of the attachment point.....	70

LIST OF TABLES:

Table 1: The difference between deterministic and probabilistic models	18
Table 2 : Historical flood events' detection with 100 mm threshold.....	40
Table 3: Permibility characteristics of the soil in the Wilaya of El Tarf	45
Table 4: Spearman Rank Correlation Matrix	49
Table 5: Variance Inflation Factors.....	51
Table 6: Hosmer–Lemeshow test.....	53
Table 7: Confusion matrix for the logistic regression.....	53
Table 8: KPSS test results	58
Table 9: The Phillips-Perron test results	59
Table 10: Results of GPD estimation	62
Table 11: Results of GEV estimation.....	63
Table 12: Anderson-Darling goodness of fit statistic test.....	65
Table 13: The pure premium for a 20% destruction rate	69
Table 14: Sensitivity of the pure premium to the choice of the attachment point	69

ABSTRACT

Flooding in Algeria is the most frequent natural catastrophe and ranks second in terms of severity. The agricultural sector, which plays a significant role in Algeria's economy and is a cornerstone of national food security, is particularly vulnerable to the devastating risks posed by floods. Traditional indemnity-based insurance allows farmers to transfer the financial consequences of agricultural flood risk through the payment of a premium, and the insured is indemnified in the event of a claim, following an assessment of loss declarations. However, such insurance faces several constraints that can limit accessibility for farmers.

To address these challenges, we tried in this research to develop an index-based insurance product aimed at providing coverage to Algerian farmers against flood risk. The indemnity under this index insurance is triggered when a predefined threshold is exceeded, rather than relying on individual loss assessments. The goal in this research is to define the index and characterize the probability distribution to accurately price the insurance product.

Due to the absence of historical loss data, we employed a probabilistic model, starting with assessing the vulnerability of agricultural land to flooding. This model leverages a Generalized Linear Model (GLM) framework, incorporating meteorological and geographic data obtained from satellite sources. We constructed a damage function specific to our study area, linking precipitation levels to probable damage. Additionally, extreme value theory was applied to model extreme precipitation events, enabling us to determine the actuarial cost of coverage.

Keywords: Index-based flood insurance, Extreme precipitation, Extreme value theory, Satellite data, Generalized Linear Model.

RÉSUMÉ

Les inondations en Algérie représentent la catastrophe naturelle la plus fréquente et se classent au deuxième rang en termes de gravité. Le secteur agricole, qui occupe une place significative dans l'économie algérienne et constitue un pilier de la sécurité alimentaire nationale, est particulièrement vulnérable aux risques dévastateurs posés par les inondations. L'assurance traditionnelle permet aux agriculteurs de transférer les conséquences financières des risques d'inondation agricole via le paiement d'une prime, l'assuré est indemnisé en cas de sinistre, après une évaluation des pertes. Cependant, ce type d'assurance fait face à plusieurs contraintes qui peuvent limiter l'accessibilité pour les agriculteurs.

Pour relever ces défis, nous avons tenté dans cette recherche de développer un produit d'assurance indicielle visant à offrir une couverture aux agriculteurs algériens contre le risque d'inondation. L'indemnisation dans cette assurance indicielle est déclenchée lorsqu'un seuil prédéfini est dépassé, plutôt que de se baser sur des évaluations individuelles des pertes. L'objectif de cette recherche est de définir l'indice et de caractériser la distribution de probabilité afin de tarifier précisément le produit d'assurance.

En raison de l'absence de données historiques sur les pertes, nous avons utilisé un modèle probabiliste, en commençant par évaluer la vulnérabilité des terres agricoles aux inondations. Ce modèle s'appuie sur un cadre de Modèle Linéaire Généralisé (GLM), intégrant des données météorologiques et géographiques obtenues à partir de sources satellitaires. Nous avons construit une fonction de dommage spécifique à notre zone d'étude, reliant les niveaux de précipitation aux dommages probables. De plus, la théorie des valeurs extrêmes a été appliquée pour modéliser les événements de précipitations extrêmes, nous permettant ainsi de déterminer le coût actuariel de la couverture.

Mots-clés : Assurance indicielle contre les inondations, Précipitations extrêmes, Théorie des valeurs extrêmes, Données satellitaires, Modèle linéaire généralisé.

GENERAL INTRODUCTION

GENERAL INTRODUCTION

Flooding is one of the most devastating and frequent natural disasters globally, causing extensive damage and loss of life. According to the Swiss Re Institute, floods have accounted for over a third of all-natural catastrophe-related fatalities since 2011, and they continue to pose a significant threat to communities worldwide. Between 2012 and 2022, cumulative insured losses from flood events exceeded USD 88 billion, underscoring the growing financial impact of these disasters. This increase is closely tied to climate change, which has intensified both the frequency and severity of extreme weather events. As global temperatures rise, changes in precipitation patterns and increased atmospheric moisture have led to more intense rainfall and, consequently, more severe floods (Amarnath, Taron, Alahacoon, & Ghosh, 2023).

The Middle East and North Africa (MENA) region, recognized by the World Bank as one of the most climate-vulnerable regions globally, faces significant risks due to climate change. Within this region, Algeria, part of the Mediterranean basin—a known "hot spot" for climate change—has seen a marked increase in both the frequency and intensity of floods. These climatic shifts have profound implications, particularly for Algeria's agricultural sector, which is a critical pillar of the nation's economy.

Agriculture in Algeria, contributing 11.6% to the national GDP and providing employment to nearly 23% of the population, is dominated by smallholder farmers. These farmers, who rely heavily on their crops for both sustenance and income, are especially vulnerable to the growing threat of floods. The sector's susceptibility is compounded by the lack of adequate risk transfer mechanisms, making it one of the most vulnerable sectors to climate-induced flood risks (Sahnoune, Belhamel, Zelmat, & Kerbach, 2013).

Despite its economic importance, agricultural flood risk in Algeria is considered as an uninsurable risk under conventional insurance frameworks. The existing safety net, provided by the "The Agricultural Disaster Guarantee Fund" (FGCA), is constrained by limited resources, and the process of indemnification is often slow. This highlights the urgent need for innovative and

more effective risk transfer tools that can provide timely, transparent, and affordable protection, particularly for smallholder farmers.

Index-based flood insurance has emerged as a potential solution to these challenges. By linking payouts to a predefined index, such as rainfall levels, rather than actual loss assessments, this approach offers the possibility of faster and more objective compensation with relatively lower premiums. However, the successful implementation of index-based insurance in Algeria faces significant obstacles, chief among them the absence of comprehensive historical data required for accurate risk modeling and product pricing.

That leads us to our research question:

“How can we transfer agricultural flood risk via an index-based solution in the absence of historical damage data in Algeria?”

To answer the main question Our attention has focused on the following questions:

- What is flood risk, and what are its parameters? What makes the agricultural sector vulnerable to this risk?
- What are the challenges faced by indemnity-based insurance in covering flood risk in agriculture, and how can index-based insurance provide a solution to these challenges?
- What is the appropriate index for flood insurance, and what pricing model should be proposed in light of the absence of historical data?

This study addresses the challenge of transferring flood risk in agricultural landscapes through the development of index-based insurance, particularly where historical damage data, crucial for traditional insurance modeling, is absent. Our research will focus on insurance of cereal production, as it holds a strategic position within both the national food system and the economy. It will also be limited to the Wilaya of El Tarf that was chosen as our study area.

Previous academic researches on flood index insurance have predominantly utilized simple linear equation to establish a relationship between precipitation and resultant damages. However, our study seeks to advance beyond this methodology by developing a method inspired by probabilistic models that integrates both vulnerability and damage functions, specifically tailored to the characteristics of the study area. By leveraging satellite data, the research aims to construct

a vulnerability function that more accurately reflects the potential impact of specific precipitation events on agricultural land.

While satellite data has been effectively employed in academic researches that aim to develop drought index insurance, to the best of our knowledge, its application in the context of flood index insurance remains unexplored. This research endeavors to fill this gap by developing a more precise and context-specific index, thereby enhancing the effectiveness of flood risk management in agriculture.

This dissertation explores the insurability of agricultural flood risk using index-based insurance through a comprehensive theoretical framework and an empirical analysis.

The first chapter delves into the theoretical aspects, beginning with a conceptual framework of natural catastrophes, followed by an in-depth examination of flood risk. It then explores the insurability of flood risk in agriculture, discussing traditional indemnity insurance and introducing the principles of index-based flood insurance. The chapter concludes with a discussion on modeling extreme events, emphasizing extreme value theory and Monte Carlo simulation.

The second chapter transitions into an empirical analysis, focusing on the construction and pricing of agricultural flood index-based insurance. It begins with an assessment of the situation in Algeria, presenting the study area and data before detailing the construction of the index. The chapter then models flood vulnerability in the selected area. Finally, it addresses the modeling of extreme precipitation and the pricing of the index-based flood insurance product.

CHAPTER ONE: THEORICAL FRAMEWORK: THE INSURANCE AGAINST AGRICULTURAL FLOOD RISK

CHAPTER ONE: THEORICAL FRAMEWORK: THE INSURANCE AGAINST AGRICULTURAL FLOOD RISK

INTRODUCTION

Flooding poses a significant threat to global populations, impacting more individuals than any other risk. Approximately 29% of the world's population is vulnerable to flood risks (Rentschler & Salhab, 2020).

Furthermore, floods have a more extensive impact on agricultural land than on urban areas and other sectors. This is particularly significant for agricultures, where smallholder farmers often represent the primary production units. Notably, between 1960 and 2015, flood-related damages were responsible for 60% of the total agricultural production losses. The exposure to floods and resultant damages is on the rise, driven by increasingly severe rainfall and storms attributed to climate change (Thomas, Tellman, Osgood, DeVries, & Akm , 2023).

As natural catastrophes become more frequent and severe, the insurability of agricultural flood risk has become a pressing issue. This chapter provides a theoretical framework for understanding the intricacies of flood risk in agriculture and explores the various approaches to mitigating this risk through insurance mechanisms while focusing on the different approaches and modeling methods for extreme precipitation events that leads to floods.

SECTION 1: UNDERSTANDING FOOD RISKS

This section lays the foundational understanding of flood risks within the broader context of natural catastrophes. we will exploring the conceptual framework that underpins the study of natural disasters, emphasizing the unique characteristics of flood events. By examining the factors that contribute to the vulnerability of the agriculture to floods.

1.1 Conceptual Framework of natural catastrophes

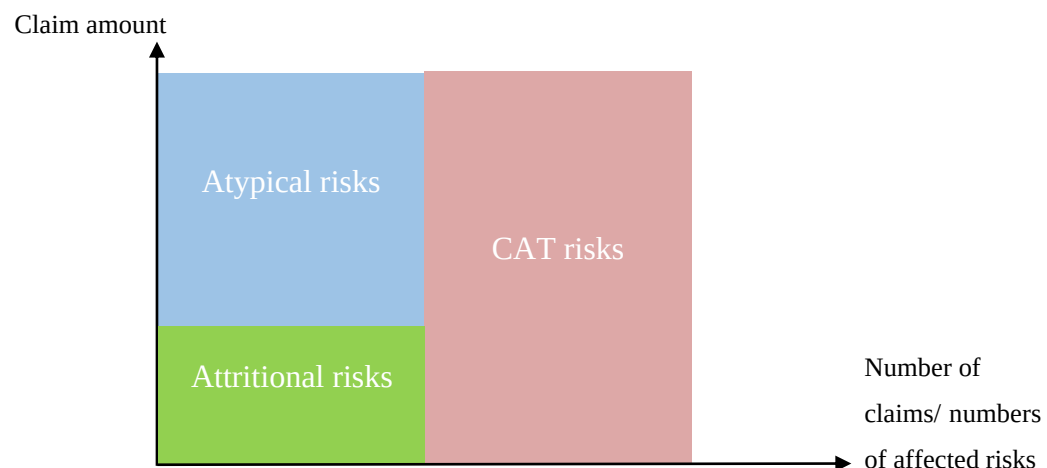
As a general definition, risk is the possibility of damage or a financial loss resulting from exposure to a hazard. It comprises the probability of occurrence of a feared event and the severity of its consequences.

1.1.1 The Different Types of Risks

Risks are usually divided into three categories:

- Attritional risks: which represent frequent individual risks with low damage value.
- Atypical risks: which represent highly impactful risks (e.g., the fire of a building).
- CAT risks: Catastrophic risk refers to a scenario in which a significant number of individuals or goods, face the potential for substantial losses due to the unfolding of a perilous event.it is a collective risk caused by a rare event qualified as a catastrophe of significant magnitude (Vasudeva , 2009).

Figure 1: Diagram of different risk categories



Source: (Vasudeva , 2009)

Catastrophic risks typically exhibit two primary characteristics. Firstly, many Catastrophes, demonstrate what is commonly referred to as a “heavy-tailed” distribution. the term “heavy-tailed” generally suggests that the likelihood of an event decreasing slowly relative to its severity. Consequently, the most extreme event observed thus far could surpass the second most severe event by several multiples.

Secondly, catastrophic risks entail spatially correlated losses, implying that numerous properties and activities in close proximity are impacted simultaneously. With the escalation of an event's scale, the number affected in the vicinity also increases, further reinforcing the correlation (Kousky, 2011).

1.1.2 Definition of Natural catastrophes

In general, natural catastrophes refer to sudden and violent events of natural origin, often resulting in loss of life, significant material damage, and serious socio-economic consequences. These events, typically unpredictable, can have a devastating impact. In some cases, the impact is primarily material, especially in wealthier nations, while in others, particularly in poorer countries, the toll is primarily human. In emerging economies, both material and human losses can be significant, as high exposure values often exist without adequate precautionary measures (Hoeppe, 2016) .

1.2 Flood risk

Flood originates from the occurrence of heavy and sustained rainfall over measurable time intervals, typically in hours, half-days or accumulative days (Tabet Helal, Baghli, Bensaoula, Mami, & Ghellaï, 2012). According to the National Insurance Council of Algeria, flooding is a temporary inundation by water of land that is not normally submerged, regardless of its origin.¹ Floods manifest due to the swift buildup and subsequent discharge of runoff water from higher to lower areas, triggered by intense rainfall. The flow peaks rapidly and then recede nearly as rapidly (Yashon & Ryutaro , 2014).

¹ <https://cna.dz/faq/assurance-contre-les-effets-des-catastrophes-naturelles-cat-nat/>

1.2.1 Flood types

There are several types of flooding. In this part, we will mention the most common ones, which are river flood, flash flood, and storm surge.

- **River flood:** River floods ensue when a river breaches its banks, leading to the flooding of nearby regions. The occurrence of extreme river floods is commonly triggered by intense and/or extended periods of rainfall. Throughout the twentieth century, river floods were responsible for approximately 7 million fatalities. Additionally, the direct global average annual loss (AAL) attributed to river floods is estimated at USD 104 billion as of 2015 (Merz, Blöschl, Vorogushyn, & Dottori, 2021).
- **Flash flood:** Flash floods occur when there is a rapid increase in water levels within a short period, usually within hours of heavy rainfall. They represent one of the most prevalent and profoundly destructive natural disasters, causing substantial harm to infrastructure, socio-economic stability, and, tragically, loss of life. Annually, over 5000 individuals worldwide lose their lives due to flash floods, a figure approximately four times higher than fatalities from other types of flood events. The exceptionally destructive nature of flash floods often stems from the sheer volume of intense rainfall over a brief timeframe, leading to significant surface runoff (Pham, Avand, Janizadeh, & Phong, 2020).
- **Storm surge:** refers to an anomalous elevation in water level along coastal regions, surpassing the usual astronomical tide. It arises from the combined forces of strong winds, heavy rainfall, waves, and reduced atmospheric pressure during severe storms. Notably perilous, storm surge has the potential to inundate extensive coastal areas. With climate change, there is an anticipated rise in the frequency of storm surge occurrences, particularly in the western and central regions of the Mediterranean Sea (Androulidakis, Kombiadou, Makris, & Baltikas, 2015) .

1.2.2 Flood risk parameters

Assessing flood risk relies on various factors, including the period of return (frequency of occurrence), duration of the flood, and the velocity of flooding (speed of floodwater flow).

- **Return period:** The return period represents the anticipated (average) interval, typically measured in years, between surpassing a specific extreme threshold. Traditionally, it serves as a measure of how often an event is expected to occur (Messner, Penning-Rowsell, Green, & meyer, 2007). In a given year, a 10-year flood has a 10 percent chance of occurring, a 50-year event has a 2 percent chance, a 100-year event has a 1 percent chance. Although rare, it is feasible to experience two 100-year or even 500-year floods within close proximity, such as within years or months of each other (Deraman, Abd Mutalib, & Mukhtar, 2017).
- **Duration:** Extended submersion of plants can lead to their demise. It does not only impact crops in the field but also stored and harvested crops, livestock, rural infrastructure, machinery, soil, and perennial plants. Typically, the duration is measured in terms of the number of days submerged. It's noteworthy that even during the most severe floods, the duration is unlikely to surpass three days (Antolini, Tate, Dalzell, & Young, 2019).
- **Flow velocity:** high-velocity water can pull down trees or wash away farm land. (Vozinaki, Karatzas, Sibetheros, & Varouchakis, 2015) The flow velocity is conditioned by the slope of the bed and its roughness, which is a characteristic of the surface state of the soil.

1.2.3 Flood vulnerability

Flooding is primarily driven by intense episodes of rainfall, but its impact is significantly influenced by various geographical, topographical, and landform factors that exacerbate an area's vulnerability. Agriculture is particularly vulnerable to flooding due to a combination of these factors, many of which are inherently linked to the geographical and environmental characteristics of agricultural regions.

Agricultural lands are often strategically located in proximity to rivers, lakes, and other water bodies to ensure reliable access to irrigation. Additionally, these lands are frequently situated in low-lying areas and floodplains because of their fertile soils, which are regularly enriched sediments deposited by rivers. However, these same factors heighten flooding, particularly during periods of heavy rainfall.

In this part, we will delve into the various factors that contribute to flood vulnerability, as identified in the literature, to provide a comprehensive understanding of the risks faced by agricultural regions.

- **Elevation:** Elevation is a critical topographic factor that affects the direction and velocity of water flow during flooding events. Lower elevations areas are more prone to inundation as water flows from higher to lower terrain. Consequently, floodplains, which represent the lowest topographic regions within a catchment, are especially susceptible to flood conditions (Schwarz & Kuleshov, 2022).
- **Slope:** Flooding is a physiographic phenomenon intricately linked to slope, with water accumulating in lower-slope regions (Megahed, Abdo.M, Abdelrahman.A.E, Scopa, & Hegazy.N, 2023). Areas characterized by gentle slopes, experience frequent flooding because the water discharge is impeded by the lack of gradient.
- **Soil types:** Different soil types exhibit varying characteristics in response to rainfall. Infiltration—the process through which water penetrates the soil—varies across soil types (Basri, Syakur, & Azmeri, 2022). For instance, soils with a high clay content tend to have small pore spaces that rapidly become saturated, resulting in increased surface runoff rather than allowing water to infiltrate leading to floods (Rawls, W, Ahuja,L, Brakensiek,D, & Shirmohammadi, 1993).
- **Soil moisture:** There is an inverse relationship between initial soil moisture content and infiltration rate (Goel & Garg, 2019). This means that as soil moisture content rises, the rate of infiltration decreases, which in turn increases the likelihood of flooding.
- **Topographic wetness index (TWI):** TWI is a quantitative measure used to predict areas of potential water accumulation based on terrain characteristics. By integrating slope and elevation data, TWI assesses how topography influences the accumulation and flow of rainfall. TWI serves as a critical tool in hydrological modeling and flood vulnerability studies by forecasting regions prone to excessive water accumulation during rainfall events.
- **Distance to rivers:** Distance to the river is a key factor influencing flood vulnerability, with proximity to the river significantly increasing the likelihood of flooding. Areas closer to a river are more susceptible to inundation due to the river's capacity to overflow its banks during heavy rainfall events (Buddha & Binu, 2023).

SECTION 2: INSURANCE AGAINST FLOOD RISK IN AGRICULTURE

As we have seen in the previous section, flooding is considered one of the most disastrous phenomena, particularly for activities such as agriculture, which is highly vulnerable to this risk, it impacts not just smallholder farmers, who constitute the primary food suppliers in many developing nations, but also affects various stakeholders throughout the local economy who rely on agricultural outcomes. (Carter, De Janvry, Sadoulet, & Sarris, 2017).

Protection against this catastrophe is generally provided through indemnity-based traditional insurance. Farmers in developing countries typically do not have access to traditional insurance against the risk of poor harvests. Insurance based on meteorological indices represents a recent innovation potentially capable of mitigating the consequences of adverse weather events for crops, such as flooding. In this section, we will discuss the limitations of traditional insurance for agricultural protection against the risk of flooding and the innovation in index-based insurance.

2.1 Insurability of flood risk

It's important for us to define the insurability of flood risk in agricultural insurance framework. In fact, as per Swiss Re, floods stand out as a peril that challenges the fundamental principles of insurability more than any other (Swiss-re, 2012). An insurance scheme's viability and longevity hinge upon the existence of insurable risks. according to (Schmit, 1986) , for a risk to be perfectly insurable it has to satisfy some key criteria.

- **Large Number of Homogeneous Exposure Units:** This requirement ensures that insurers can predict losses accurately based on the law of large numbers. When a sufficient number of similar exposure units exist within an agriculture's portfolio, insurers can reliably anticipate both the frequency and severity of losses.
- **Accidental and Unintentional Loss:** Insurance typically covers losses that are accidental and unintentional. Deliberate actions leading to loss, such as intentional flooding of fields, are not covered, as this would undermine the fundamental principle of random occurrence upon which insurance calculations are based.
- **Determinable and Measurable Loss:** Losses covered by insurance must be definite, measurable, and significant enough to cause economic hardship. This criterion ensures that

insurers can accurately assess and quantify the extent of the loss, determining factors such as the place, and the amount.

- **No Catastrophic Loss:** Ideally, insurance avoids catastrophic losses where a significant proportion of exposure units incur losses simultaneously.
- **Calculable Chance of Loss:** Insurers must be able to calculate both the frequency and severity of future losses with reasonable accuracy. This calculation informs the setting of premiums, ensuring they are adequate to cover all claims and expenses while still yielding a profit.
- **Economically Feasible Premium:** Premiums for flood insurance must be economically feasible for farmers, representing a reasonable cost relative to the value of the coverage provided. Insurers aim to set premiums at a level where they cover potential losses while remaining attractive to customers.

In accordance with these criteria, flood risk is often not considered an “ideal” insurable risk. This is due to several factors.

Firstly, flood risk exhibits significant variability in terms of its occurrence frequency and severity across different regions and return periods. This variability makes it challenging to establish a large number of homogeneous exposure units, which are essential for accurate risk assessment and premium calculation.

Additionally, flood risk is a highly correlated risks, leading to challenges in treating damages as independent and identically distributed random variables. This unpredictability makes it difficult to estimate the precise amount of damage that may occur in any given year (Kast & Lapied, 2012).

Moreover, flood risk is characterized by its catastrophic nature, posing a fundamental challenge to the insurance principle of risk spreading. When entire regions are affected by flooding simultaneously, insurers may struggle to cover the resulting losses adequately, leading to financial strain and potential insolvency.

Furthermore, the high cost of insuring against flood risk presents a barrier to many farmers, particularly in emerging economies where spare resources for insurance premiums are limited. The prohibitive cost of flood insurance further exacerbates the problem of low affordability,

constraining the levels of insurance coverage available to vulnerable farmers (Lamond & Penning-Rowsell, 2014).

In another hand, quantifying risk is a delicate process, but advancements in understanding natural disasters and the development of their modeling techniques enable a more accurate assessment of risk. Also, Reinsurance can be useful in various ways to address these challenges. Firstly, it helps to avoid excessively large losses due to extreme claims. Consequently, its primary goal is to stabilize and make the insurer's results more predictable. Furthermore, it allows for risk pooling across geographic areas by implementing geographic diversification to reduce risk dependency. Flood risk is therefore practically insurable.

2.2 Addressing agricultural flood risk with traditional indemnity insurance

Traditional agricultural insurances, which are based on the assessment and evaluation of crop damage, have existed since the 1930s starting from the United States. They have recently expanded more specifically in the form of multi-peril insurances. However, these insurances have only seen widespread development where they are heavily subsidized (Leblois & Quirion, 2013).

2.2.1 Generalities about Flood insurance in agriculture

Indemnity based insurance is the classical form of coverage against agricultural risks, in exchange for an insurance premium paid. The insured is indemnified in the event of a claim, following an assessment of loss declarations. The agricultural risks covered mainly include climatic risks, and health risks (livestock mortality due to diseases and preventive culling orders), in addition to traditional coverage for buildings and equipment on the farm against fire, theft, etc.

Agricultural insurance against flooding is generally provided as an extension within Multi-peril Crop Insurance (MPCI). In such insurance products, the combined impact of different perils is reflected in reduced crop yields, which form the basis for insurance payouts, it covers the loss of quantity caused to plants, following overflow from excessive rainwater. As of now, flood insurance remains unavailable as an independent insurance offering within the agricultural sector worldwide.

2.2.2 Challenges in insuring agricultural risk with a traditional indemnity insurance

The application of traditional agricultural insurance presents several inherent structural challenges, which complicates its application in covering flood risks.

I. Challenges related to information asymmetry

The insurer and the farmer are not equal in information possession; for instance, the farmer has an idea about the quality of the land and its vulnerability and exposure to flood risk. Thus, they have an advantage of being ahead of the insurer.

Adverse selection is a direct consequence of information asymmetry, where one party (the insured or the insurer) has more information than the other. The insured knows whether they are a good or bad risk, which is not the case for the insurer; farmers whose risks exceed the premiums are more willing to subscribe to insurance. According to (Mcmahon, 2019) adverse selection is a particularly significant issue in flood insurance, as it primarily attracts those who are most at risk of experiencing floods. This results in a situation where only individuals who are likely to suffer flood damage are interested in purchasing flood insurance. Once again, the insurer is condemned to suffer losses, leading to the failure of the insurance market (Carter, de Janvry, Sadoulet, & Sarris, 2015).

Moral hazard is also a consequence of information asymmetry, where the farmer or the insured tends to limit their effort since they will be compensated if the harvest is poor.

II. Challenges related to the protection gap

The global challenge of the protection gap, representing the variance between losses covered by insurance and those left uninsured, persists across the globe. This gap signifies a diminished resilience among individuals, businesses, communities, and even entire nations compared to what could be achieved if the gap were eliminated. Over the past decade, only around 30% of losses stemming from natural disasters have been mitigated through insurance coverage. Particularly in developing nations, the proportion of economic losses not covered by insurance frequently surpasses 90% (Clyde&co, 2018). The protection gap is no different in the agriculture sector, In the aftermath of significant weather-related calamities like floods, as much as 90% of agricultural losses remain uninsured (Schanz, 2018).

The effect of the uninsured risk can be devastating for an economy; The Bank of International Settlements calculates that in developing markets, the worst natural catastrophes can permanently reduce a country's GDP by almost 2% (Garaix, 2020) .

Affordability emerges as a prominent factor contributing to underinsurance, particularly among farmers with lower incomes. Notably, researchers have identified a considerable price sensitivity of flood insurance demand, especially within developing nations (Holzheu & Turner, 2017). A portion of the elevated cost associated with indemnity-based insurance can be attributed to the significant expense of expert estimation and the cost of information asymmetry, reflected in insurance premiums (Leblois & Quirion, 2013).

The gap also reveals deficiencies in the current practices of risk underwriting, and insurance contract design. The relatively limited transparency in risk underwriting and pricing schemes, along with delays in claims settlements and sometimes disagreements in loss appraisal lead to lower-than-anticipated demand for insurance and thus a larger protection gap (Lin & Kwon, 2020).

2.3 Principles of Index based flood insurance

Index-based weather insurance is a significant institutional advancement that has the potential to transform the availability of formal insurance for many farmers and associated individuals. It has been implemented in a pilot or experimental capacity in numerous nations, either at the individual or institutional level.

2.3.1 Definition of index-based insurance

Index-based insurance operates by utilizing triggers based on weather indicators, rather than relying on direct estimates of losses at the individual farm level. All policyholders in a specific geographic area determined by the weather station are entitled to receive payments based on an agreement and measurement as determined by the designated weather station (Amarnath, Taron, Alahacoon, & Ghosh, 2023) .

Index insurance is characterized by a standardized format, eliminating the need for specialized professionals to verify or validate losses, resulting in relatively cheap administrative expenses. This mitigates the need for loss adjustment by both the insurer and the insured, so reducing risks, fostering trust, and preventing modifications. Due to the decrease in administrative

expenses, insurance premiums may also be reduced, resulting in more inexpensive cover for end users (Annibale, Martina, Monteiro, & Galasso, 2021). This insurance effectively mitigates moral hazard and adverse selection issues that might otherwise skew behavior in insurance markets. The reason for this is because payouts are set by external information that is not influenced by either hidden traits (adverse selection) or subsequent choices (moral hazard) (Carter, De Janvry, Sadoulet, & Sarris, 2017).

2.3.2 Types of indices

An index must be objectively and easily measurable, publicly verifiable, and not susceptible to manipulation by both the insurer and the insured party. The decision about the threshold and the insured level can be delegated to the provider and the client, with the cost being the actuarially reasonable price of the associated risk plus a competitive loading (Carter, De Janvry, Sadoulet, & Sarris, 2017). Indices used for the implementation of parametric insurance can vary depending on the type of risk covered. For example:

- **Yield estimation:** This is a direct index that involves estimating yields based on agronomic models, which will then be used to construct aggregations to set trigger thresholds for indemnities (Kouli, 2018).
- **Cumulative rainfall:** for extreme rainfall, the commonly used index is related to the accumulation of precipitation over several days. In this context, an extreme rainfall event is defined as the continuous period during which the cumulative precipitation exceeds a certain threshold. During periods of heavy rainfall, data collected from nearby meteorological stations is compared to specific criteria. When the amount of rainfall surpasses the specified level, the farmers who have joined the program are eligible to receive a payment (Amarnath, Taron, Alahacoon, & Ghosh, 2023).
- **Drought indices:** This type of index quantifies drought based on rainfall and degree-days (temperatures) on a daily time step.
- **Vegetation indices: Normalized Difference Vegetation Index (NDVI):** derived from satellite imagery of vegetative cover, measures the physiological activity of plants. It is used in agriculture to assess vegetation health and quantity through remote sensing analyses (Kouli, 2018).

2.3.3 Barrier for the application of index insurance

Despite the solutions presented by index-based insurance, this approach faces many barriers.

- **Basic risk:** Insurance policies are designed to cover specific risks, and providers must analyze the probability distribution of these risks to price the insurance product accurately. Agricultural index insurance follows the same principle, where the risk, represented by the probability distribution of a relevant index, should align with the probability distribution of the crop yield (Carter, De Janvry, Sadoulet, & Sarris, 2017). The difference between the two probability distributions lies in the basis risk, suggesting that the risk experienced by an individual household may vary from the risk assessed or indicated by the index. In another way the correlation between the net transfer from index insurance and the consumer's net loss is assumed to be imperfect. Therefore, by purchasing index insurance it is possible for a consumer to experience a loss yet receive no net compensation from the index insurance product, or to have no loss but receive a positive net income (Clarke D. , 2016).
- **Liquidity:** Premiums must be paid before the planting season, frequently with significant intervals between seasons. In order to maintain sufficient funds from one harvest to the following planting season, it is crucial to have savings tools and access to credit. Nevertheless, farmers in several developing nations sometimes have difficulties in accessing these financial instruments owing to the lack of widespread financial inclusion. Research done in Malawi reveals that the financial situation of farmers, who face rainfall as the main risk, affects their interest in index insurance (Giné & Yang, 2009).
- **Learning difficulties:** The introduction of index insurance may be challenging to understand, especially for those who are not accustomed with the notion, particularly in developing nations. Farmers often expect to be reimbursed by the provider if no adverse events occur over the year. Research conducted by (Giné & Yang, 2009) indicates a notable positive association between years of schooling and the desire for index insurance among farmers. Furthermore, index insurance presents an added layer of complexity due to basis risk.

SECTION 3: MODELING EXTREME EVENTS

Flood insurance plays a crucial role in enhancing resilience, yet pricing it poses significant challenges for actuaries. A major difficulty lies in the fact that floods are infrequent events, with historical claims data often being sparse or unreliable. As a result, actuaries must depend on a variety of models and probability distributions to assess flood risk. These tools are essential for developing insurance products that accurately reflect the risks posed by flooding, despite the inherent uncertainties.

1.1 Extreme events modeling approaches

To accurately assess and mitigate the impact of extreme events, two main modeling approaches are employed: deterministic and probabilistic modeling.

3.1.1 Deterministic Modeling

Deterministic models offer a straightforward approach for modeling losses caused by natural disasters. This method is based on analyzing historical data of natural events to estimate the frequency and severity of flood risk. Specific statistical distributions, such as Poisson or Negative Binomial for frequency and Log-Normal, Weibull, or Pareto for severity, are applied to assess these risks. By examining historical events, the parameters of these distributions can be estimated. This modeling approach utilizes significant past events, adjusted to current insured values using an "As-If" scenario. One major challenge with the As-If analysis of natural risks is accounting for changes in the vulnerability and the exposure of insured assets.

Geographic variables play a crucial role in loss modeling—as we mentioned lands near a river or catchment area are more prone to flooding than those located further away.

Another limitation of this method is the lack of extensive historical flood data, many modeling issues ultimately stem from data rarity and quality issues. As (Uy, 2019) pointed out, especially in private insurance market that lacks sufficient historical loss data, further complicating modeling efforts. In the realm of agricultural insurance, the task is compounded. Specifically, the availability of precise and current data stands as an important factor for the efficacy of agricultural insurance.

3.1.2 Probabilistic Modeling

In contrast to the deterministic approach, which relies on a relatively short historical record, probabilistic modeling simulates possible events. This approach generates a catalog of catastrophic loss events, each associated with a probability of occurrence and an expected loss amount. It consists of three main components:

- **Hazard Assessment:** The hazard module aims to simulate fictitious but realistic events in order to measure the potential impact of more extreme events than those recorded in the historical data for all possible perils. Initially, the module relies on the historical record of past losses before varying certain parameters, such as precipitation intensity.
- **Vulnerability Analysis:** the vulnerability module aims to determine the average loss ratio based on the intensity of the hazard. The average loss ratio is defined as the total amount of the loss relative to the total value of all insured objects in the considered area. This module allows for the creation of a vulnerability curve (destruction function) that represent the relationship between hazard intensity and the average loss ratio.
- **Loss Estimation:** The damage model aims to calculate the amount of losses based on the hazard and vulnerability.

Each of these modeling approaches has its advantages and disadvantages, summarized as follows:

Table 1: The difference between deterministic and probabilistic models

Model	Advantages	Disadvantages
Deterministic Model	- Quick to develop - Based on historical loss data	- Limited historical data - Past events may not accurately predict future outcomes
Probabilistic Model	- Incorporates vulnerability - Adjusts for changes in exposure	- Requires detailed and precise information

Source: (Idir, 2021)

3.2 Extreme value theory

Floods are extreme events characterized by their rarity and high impact, making them a critical focus for effective risk management strategies. Classical statistical approaches, which typically center on the behavior of random variables around their mean values, often fall short in adequately capturing these rare and severe events. These traditional methods primarily address average conditions and typical outcomes, which may not sufficiently account for the extreme values that drive the most significant flood events. In contrast, extreme value theory is specifically used to model and predict the occurrence and magnitude of such rare, high-impact events, which are situated in the tails of the distribution.

3.2.1 Fundamental theories

Classical extreme value statistics consists of two fundamental approaches: the block maxima (BM) method and the peak-over-threshold (POT) approach.

I. Extreme value distributions

Let $X_1, X_2, \dots, X_n$ be a sample of n independent and identically distributed random variables with a common cumulative distribution function F . In "classical" statistics, the focus is primarily on the random variable $S_n = X_1 + X_2 + \dots + X_n$. The Weak Law of Large Numbers ensures that if X_1 is integrable, then $\frac{S_n}{n} \xrightarrow{p} E(X_1)$. Furthermore, if X_1 is square-integrable, the Central Limit Theorem guarantees that:

$$\frac{S_n - E(S_n)}{\sqrt{\text{Var}(S_n)}} = \frac{S_n - n\mu}{\sqrt{n} \sigma} \xrightarrow[n \rightarrow \infty]{loi} N(0,1).$$

With $\mu = E(X_1)$ and $\sigma = \sqrt{\text{Var}(X_1)}$

The sum S_n thus provides results regarding the central part of the distribution. In our case, we are interested in the tail of the distribution, and it is therefore natural to study the behavior of the sample maximum.

$$M_n = \text{Max}\{X_1, X_2, \dots, X_n\} \quad (3.1)$$

In practical applications, the X_i can represent measurements of a process taken at regular intervals, with M_n denoting the maximum value observed over n time periods. Theoretically, the distribution of M_n can be determined using the cumulative distribution function F as follows:

$$P\{M_n \leq z\} = P\{X_1 \leq z, \dots, X_n \leq z\} \quad (3.2)$$

$$= P\{X_1 \leq z\} \times \dots \times P\{X_n \leq z\} \quad (3.3)$$

$$= \{F(z)\}^n \quad (3.4)$$

However, the function F is typically not known. Even if F can be estimated, predicting the behavior of M_n as n becomes very large is challenging. Typically, classic distributions such as the exponential, gamma, and log-normal distributions are used for modeling the X_i . The issue with these standard distributions is that they often significantly underestimate the importance of the distribution tails, as these tails represent only a minimal portion of the data. Standard calibrations focus on capturing the central tendency, providing an accurate "average" depiction. In our case, we aim for accuracy in the "high quantile" range. We aim for a limiting behavior that converges to a non-degenerate distribution. To achieve this, it is necessary to standardize the sample's extreme values. The following theorem helps to address this issue when M_n does not degenerate:

Theorem 3.1: (theorem of Fisher–Tippett)

If there exists a sequence of pairs of real numbers (a_n, b_n) such that $a_n > 0$ and:

$\lim_{n \rightarrow +\infty} P\left(\frac{M_n - b_n}{a_n} \leq x\right) = G_\xi(x)$ for a non-degenerate distribution $G_\xi(x)$, then $G_\xi(x)$ is a Fréchet distribution, a Gumbel distribution, or a Weibull distribution. With :

- **Type I:** Gumbel: $G(x) = \exp\left\{-\exp\left[-\left(\frac{x-b}{a}\right)\right]\right\} \quad -\infty < x < \infty \quad (3.5)$

- **Type II :** Fréchet : $G(x) = \begin{cases} 0 & \text{if } x \leq b \\ \exp\left\{\left(\frac{x-b}{a}\right)^{-\alpha}\right\} & \text{else} \end{cases} \quad (3.6)$

- **Type III:** Weibull: $G(x) = \begin{cases} \exp\left\{-\left[\frac{x-b}{a}\right]^{-\alpha}\right\} & \text{if } x < b \\ 1 & \text{else} \end{cases} \quad (3.7)$

According to the Fisher-Tippett theorem, if the number of extreme observations is sufficiently large, the distribution of these extreme values converges to the Generalized Extreme Value (GEV) distribution. The Gumbel, Weibull, and Fréchet distributions represent the three types of domains of attraction within extreme value theory. The most general form of the limiting distribution $G(x)$ for extreme values, as established by Gnedenko (1943), Von Mises (1954), and Jenkinson (1955), indicate that the behavior of the tails of most common distributions can therefore be explained by a single distribution, defined below, where a single parameter ξ controls the heaviness of the tails. It is defined by:

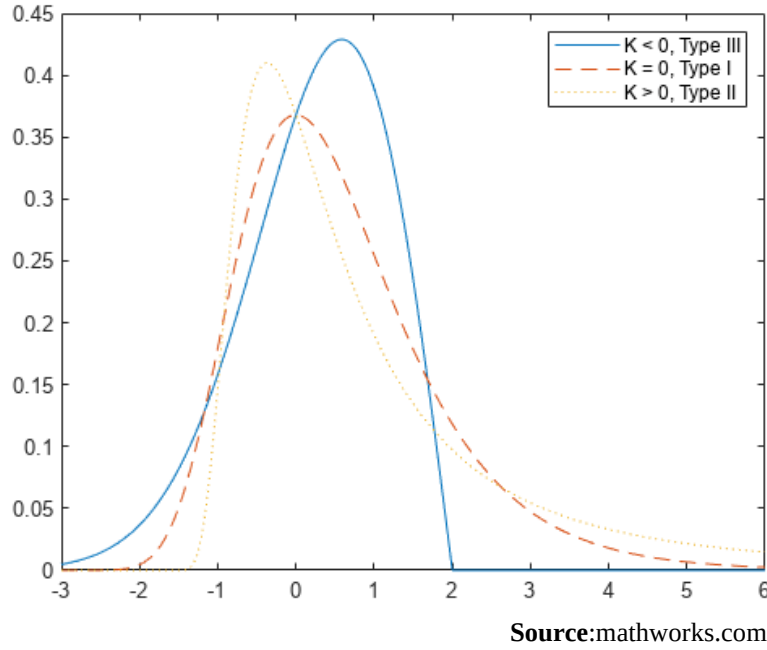
Definition: The Generalized Extreme Value (GEV) distribution $G(x)$ is defined by the cumulative distribution function:

$$G_{(\xi, \mu, \beta)}(x) = \begin{cases} \exp \left[- \left(1 + \xi \frac{x - \mu}{\beta} \right)^{-\frac{1}{\xi}} \right] & \text{if } \xi \neq 0 \\ \exp \left[- \exp \left(- \frac{x - \mu}{\beta} \right) \right] & \text{if } \xi = 0 \end{cases} \quad (3.8)$$

This distribution has three parameters:

- μ (location parameter), which indicates the average size of the extremes in the parent distribution.
- β (scale parameter), which characterizes the dispersion.
- ξ (shape or tail index), which provides an indication of the shape (or the heaviness of the tails) of the limiting distribution. The smaller the tail parameter ξ , the heavier the considered distribution's tails.

Figure 2:GEV density with $\mu=0$, $\sigma=1$ for different values of the shape parameter ξ .



II. Peak over threshold

Suppose $X_1, X_2, \dots, X_n$ are iid random variables with distribution $F(x)$ and a relatively high threshold u above which the excesses are determined. Consider $Y_1, Y_2, \dots, Y_n$ the positive excesses defined by the expression $Y_i = X_i - u$ for $i=1, 2, \dots, n$. Thus, we can define the cumulative distribution function of the excesses above the threshold u , denoted by $F_u(y)$, as follows:

$$F_u(y) = P(X - u \leq y \mid X > u)$$

$$F_u(y) = \frac{F(y + u) - F(u)}{1 - F(u)} = \frac{F(x) - F(u)}{1 - F(u)} \quad (3.9)$$

The result from (3.8) allows us to write:

$$[F(x)]^n \approx \exp\left[-\left(1 + \xi \frac{x - \mu}{\beta}\right)^{-\frac{1}{\xi}}\right]$$

$$\log(F(x)) \approx -\frac{1}{n} \left(\left(1 + \xi \frac{x - \mu}{\beta}\right)^{-\frac{1}{\xi}} \right)$$

$$1 - F(x) \approx -\frac{1}{n} \left(\left(1 + \xi \frac{x - \mu}{\beta} \right)^{-1/\xi} \right)$$

We can then deduce an approximation of $F_u(y)$ from Equation (3.9), we get I final:

$$F_u(y) \approx 1 - \left(1 + \xi \frac{y}{\phi} \right)^{-\frac{1}{\xi}}$$

where $\phi = \beta + \xi(u - \mu)$. This final expression is the characteristic function of the Generalized Pareto Distribution (GPD).

Proposition: If a random variable X follows a GPD (ξ, β) distribution, then for any threshold $u > 0$, the conditional random variable $(X - u | X > u)$ follows a GPD $(\xi, \beta + \xi u)$ distribution.

This formula provides the probability of excess losses given that the values of X are beyond the threshold u . The distribution of X is unknown and can follow any type of distribution (Normal, Lognormal, Student's t , or Beta). Estimating the distribution of extreme values from the parent function F poses a major challenge, as the latter is not known. To overcome this limitation, researchers (Fisher-Tippett (1928)) have focused on the asymptotic behavior of the random variable y . The goal is to determine the distribution towards which the extreme values converge as n approaches infinity, in order to replace F with this distribution for large values of n , the theorem of Balkema-Pickands-de Haan have characterized the distribution of extreme values without being conditioned by the parent distribution. This theorem states that for a relatively high threshold, the distribution $F_u(y)$ converges to a Generalized Pareto Distribution (GPD), with the cumulative

$$\text{distribution function } G_{\{\xi, \beta\}}(y) \text{ given by: } G_{\{\xi, \beta\}}(y) = \begin{cases} 1 - \left(1 + \xi \frac{y}{\beta} \right)^{-\frac{1}{\xi}} & \text{if } \xi \neq 0 \\ 1 - \exp\left(-\frac{y}{\beta}\right) & \text{if } \xi = 0 \end{cases}$$

This distribution has only two parameters to estimate: a dispersion parameter β and a shape or tail index parameter ξ :

- If $\xi < 0$, the tails of the distribution are bounded.
- If $\xi = 0$, the tails decay exponentially.
- If $\xi > 0$, the tails of the distribution are heavy.

3.2.2 Threshold Selection

The application of the Peak Over Threshold (POT) method entails some analytical challenges, particularly in selecting a suitable threshold level for extreme events. Determining this threshold involves balancing variance and bias. A higher threshold value ensures more reliable tail asymptotic for estimating the distribution and minimizes bias. However, this comes at the cost of increased variance due to the reduced sample size available for parameter estimation. In other words, a threshold that is too low leads to a poor fit between the distribution function and the GPD, resulting in biased estimators. Conversely, a threshold that is too high results in a small sample size of exceedances, leading to volatile estimators.

The mean excess: Suppose $X_1, X_2, \dots, X_n$ are independent and identically distributed random variables with cumulative distribution function F , and introduce the order statistic $X_1 \geq X_2 \geq X_3 \geq \dots \geq X_n$. The mean excess function is defined as the sum of the excesses exceeding a certain high threshold divided by the number of observations that exceed this threshold. The empirical estimator of the mean excess function $e_n(u)$ is defined by the following expression:

$$e_n(u) = \frac{1}{N_u} \sum (X_i - u) \quad u > 0$$

The choice of threshold is based on the behavior of the mean excess curve. Indeed, if the Mean Excess plot appears to be linear above a certain positive value of u , this indicates that the excesses beyond this threshold follow a GPD. In practice, this involves identifying values of u from which $e_n(u)$ is approximately linear. Graphically, this is indicated by a change in the slope of the curve.

3.3 Monte Carlo Simulation

The asymptotic theory of extreme values operates under the premise that the sample size tends towards infinity. However, practical applications often involve finite data sets, resulting in deviations from this theoretical assumption. This limitation complicates the accurate prediction of the true distribution of extreme value samples. As a result, evaluating the prediction errors that arise from relying on asymptotic distribution assumptions becomes inherently challenging (Siyuan , Wenhua, & Arvid , 2023) .Monte Carlo simulation mitigates these limitations by generating extreme value samples from the estimated distribution.

we will focus on the Monte Carlo simulation technique through the inversion of the cumulative distribution function (CDF). This method is particularly favored when the inversion of the CDF is feasible. The generalized inverse of the cumulative distribution function F of a real-valued random variable X is given by:

$$F^{-1}(y) = \inf\{x \mid F(x) \geq y\}$$

Lemma: If U is a random variable uniformly distributed on $[0,1]$ and F is the CDF of the random variable X , then $F^{-1}(U)$ follows the same distribution as X .

Thus, to simulate n i.i.d. realizations from a distribution with CDF F , one needs to generate n independent realizations of a uniformly distributed random variable on $[0,1]$, and then apply the inverse of the CDF to each of these values.

COCLUSION

The global agricultural sector is increasingly exposed to the escalating risks posed by floods, a phenomenon exacerbated by climate change and environmental degradation. In all various types of floods, the primary driver remains extreme precipitation within a limited timeframe. However, the extent of damage to agricultural land is significantly influenced by the inherent characteristics of the land itself. Crop lands are particularly vulnerable to flooding due to their strategic location near rivers and water bodies, which facilitates irrigation, and their placement in low-lying areas, often chosen for their fertile soils.

Traditional flood insurance in agriculture, typically provided through indemnity-based multi-peril contracts, faces numerous challenges. These challenges include information asymmetry, high administrative costs, and prolonged indemnity periods, all of which contribute to the low penetration of such insurance, leaving many farmers, especially smallholders, vulnerable to financial distress.

To mitigate these challenges, index-based insurance has emerged as an alternative. This form of insurance is characterized by its transparency and reduced administrative costs, as it eliminates the need for individual loss assessments. Index-based insurance operates on predefined triggers based on weather indices; in the context of floods, if precipitation levels exceed a predetermined threshold within a specific timeframe, as measured by meteorological stations or satellite data, farmers covered under the scheme receive compensation. This approach offers a viable solution for the agricultural sector, enabling it to manage climatic risks such as floods more effectively.

The development of index-based insurance hinges on the creation of an index that accurately reflects the potential damage from extreme precipitation events. However, in the absence of historical data, deterministic models fall short of providing a reliable basis for this index. In the subsequent chapter, we will examine the flood risk context in Algeria and proceed to develop an index-based insurance model tailored to a specific study area. This model will aim to construct an insurance scheme that accurately reflects the vulnerabilities and potential damages specific to the region, thereby offering a realistic and effective flood risk management tool for agriculture.

**CHAPTER TWO: EMPIRICAL ANALYSIS:
CONSTRUCTION AND PRICING OF AGRICULTURAL
FLOOD INDEX-BASED INSURANCE**

CHAPTER TWO: EMPIRICAL ANALYSIS: CONSTRUCTION AND PRICING OF AGRICULTURAL FLOOD INDEX-BASED INSURANCE

INTRODUCTION

Building on the theoretical framework of flood insurance and the potential of index-based solutions, this chapter turns to empirical analysis. The challenge in designing an effective index insurance product lies in its ability to predict damages resulting from the exceedance of precipitation thresholds, especially in the absence of pre-existing damage assessments.

To develop an index that aligns with the realities of the region, we begin by assessing the current situation in Algeria, with a particular focus on our study area. This assessment serves as the foundation for constructing the index product.

As highlighted in the theoretical discussion, traditional deterministic modeling methods are often inadequate for insurance pricing when data are non-available or when dealing with geographically variable risks like floods, where damage can vary significantly across different areas. In response to these challenges, we will try to approach the challenge with a probabilistic model, the second section of this chapter will examine the potential for modeling vulnerability functions using satellite data to create a more precise and regionally specific destruction function. The final section will then focus on modeling extreme precipitation events and integrating them with our destruction function to establish the pricing framework for the proposed insurance scheme.

SECTION 1: ASSESSMENT OF THE SITUATION AND PRODUCT CONSTRUCTION

In this section, we will provide a comprehensive analysis of Algeria's agricultural landscape and its crucial role in the economy and food security, with a focus on the important of cereal production. We will examine precipitation patterns, assess flood risk, and discuss the coverage of this risk within the agricultural sector. We will focus on our specific study area, detailing its characteristics to accurately construct an index-based insurance model, a critical step in any index insurance project.

1.1 State of Affairs and Situation in Algeria

1.1.1 The Agricultural Landscape of Algeria

Agricultural development is recognized as a key driver of economic growth, resulting in poverty reduction and improved food security (Bellout, Vas, & Penfound, 2020).

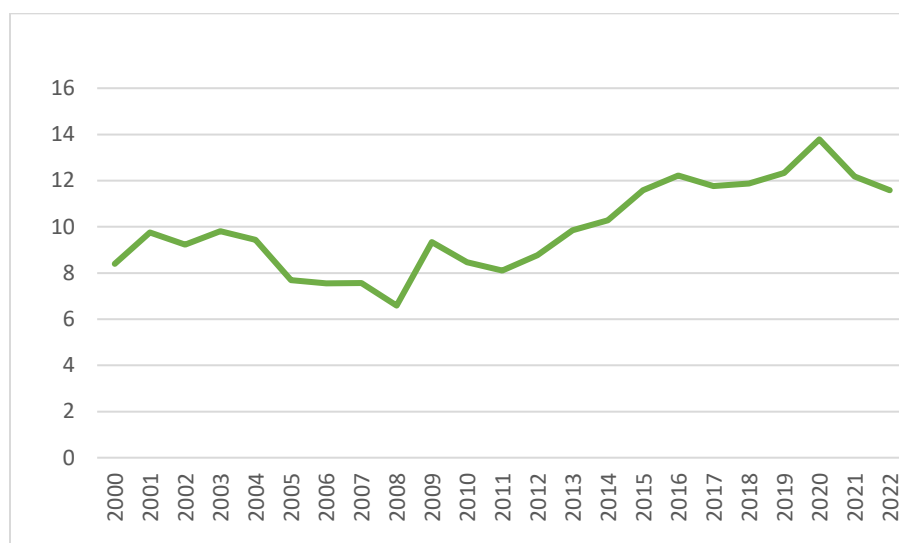
I. Agricultural land in Algeria

Although it is the largest country in Africa, with its 238 million hectares, Algeria has one of the smallest arable land areas in the region. According to statistics from the World Bank, Algeria ranks 204th globally and 4th in the Arab Maghreb region with an agricultural land percentage of 17.34% in 2022, with only 41.316 million ha and a percentage of 4% of arable land.

II. Agriculture in the Algerian economy

From 2000 to 2022, the agricultural sector in Algeria has experienced significant growth in its contribution to the GDP. The Gross Domestic Agricultural Production (GDAP) as a percentage of GDP has shown an increase from 8.4% in 2000 to 11.6% in 2022, according to available world bank's data. This upward trend reflects advancements in both vegetable and animal production, underlining the sector's resilience and importance within the national economy.

Figure 3: Agriculture, value added (% of GDP) from 2000 to 2022



Source: Established by the author using world banks' data

Regarding employment in the agricultural sector, according to the world bank, this sector employed 22.68% of the active population in 2022, with 2.7 million workers.

III. Cereal production in Algeria

In Algeria, cereal crops hold a strategic position within both the national food security and the economy. Cereal production occupies approximately 80% of the country's arable land. In 2022, the total area sown with cereals amounted to 3,342,552 hectares. Of this, 58% (1,937,985 hectares) was harvested, yielding an average of 14.3 quintals per hectare, according to data from the Ministry of Agriculture and Rural Development (MADR). Nearly 60% of the agricultural workforce are engaged in cereal farming.

Cereal consumption in Algeria is estimated at approximately 205 kg per capita per year. Cereals and their derivatives form the backbone of the Algerian food system, providing over 60% of caloric intake and 75 to 80% of the protein intake in the diet. Algeria is the second-largest wheat consumer in Africa, following Egypt. These figures underscore the importance of developing a strategic plan for the enhancement of cereal production and the critical need to protect this production against various climatic risks, such as floods.

1.1.2 Rainfall profile of Algeria

The variability of rainfall patterns in Algeria is largely influenced by its diverse climate, which is shaped by three distinct climatic zones:

- **Mediterranean Climate (North Algeria):** The northern coastal region of Algeria experiences a Mediterranean climate, characterized by mild, wet winters and hot, dry summers. Notably, January and December exhibit the highest precipitation levels, averaging 80mm and 90mm respectively, followed by a gradual decline leading into the drier summer season. Even during the driest month, July, the region still receives an average of 5mm of rainfall. The yearly cumulative precipitation for the Mediterranean climate stands at 600 mm. the rainfall pattern is influenced by Mediterranean cyclones and seasonal winds.
- **Semi-Arid Climate (Central Algeria):** Moving southward, the climate becomes semi-arid. This zone acts as a transition between the wetter Mediterranean coast and the arid Sahara Desert. The yearly cumulative precipitation for the Semi-arid climate stands at 390mm.
- **Arid Climate (Southern Algeria/Sahara Desert):** The vast southern region of Algeria falls under the arid climate of the Sahara Desert. Rainfall is extremely scarce and highly unpredictable, often occurring as isolated and irregular events. The yearly cumulative precipitation for the Desert climate stands at 155mm.

The precipitation patterns in Algeria exhibit notable variability, with significant fluctuations and irregularities in rainfall amounts and distribution over time. This variability is particularly evident in the Mediterranean climate zone, which is known for its occurrence of synoptic patterns marked by intense precipitation events (Lionello, Bhend , & Buzzi, 2006).

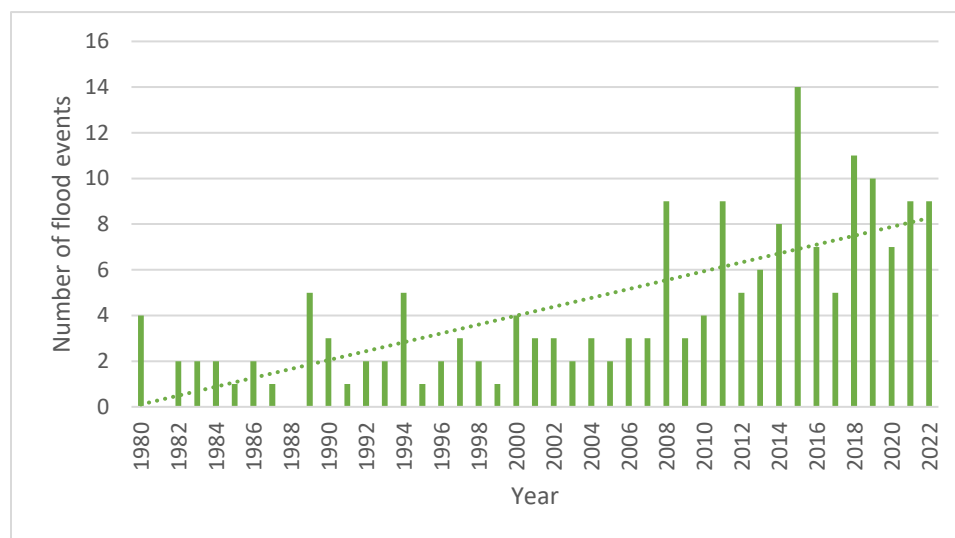
According to a study by (Bessaklia, Serrano-Notivoli, & Nekkache Ghenim, 2021), the analysis of extreme precipitation reveals that a significant portion of rainfall along coastal areas results from high-intensity events. The study indicates that approximately 20% to 25% of the total precipitation occurs during these extreme events.

1.1.3 Flood risk in Algeria

In Algeria, floods pose a significant risk, ranking second in severity after earthquakes and first in frequency according to the United Nations. This classification underscores the considerable tangible and intangible losses associated with flood occurrences. Over the past two decades, Algeria has witnessed numerous catastrophic flood events, resulting in hundreds of fatalities, thousands of injuries, and property damage exceeding billions of dollars. The majority of these flood disasters have been attributed to intense storm events (Bouamrane, Derdous, Dahri, Tachi, & Boutebba, 2020).

I. frequency analysis of flood events

Figure 4: Flood frequency in Algeria from 1980 to 2022



Source: Hafnaoui, et al., 2023

The analysis of flood events in Algeria spanning from 1980 to 2022 unveils a dynamic landscape of occurrences over the decades. During the initial years, from 1980 to 1988, the incidence of flood events remained relatively sporadic, with some years registering minimal to no occurrences. However, a noticeable shift occurred from 1989 onwards, marked by a discernible uptick in flood frequency. Subsequently, from 2015 onward, Algeria experienced a sustained period of heightened flood activity, with several years recording double-digit occurrences. The greatest number of floods was documented in 2015, totaling 14 occurrences, followed closely by 2018 and 2019, which experienced 11 and 10 floods, respectively.

II. Agricultural damaged caused by floods in Algeria

In Algeria, as in many developing countries, the challenge of accurately assessing flood severity is compounded by a lack of comprehensive historical data, particularly in the agricultural sector. The available records from the Ministry of Agriculture and Rural Development are limited to only four significant events.

1. Ghardaia's floods of October 2008, which resulted in agricultural losses estimated at 302 million Algerian dinars.
2. Oum El Bouaghi's floods of August 2013, with agricultural losses estimated at 517 million Algerian dinars.
3. Khenchela's floods of October 2013, where total damages were estimated at 837 million dinars, including a cereal loss valued at 14.25 million Algerian dinars.
4. El Tarf's floods of February 2012, where total damages were assessed at 2.326 billion dinars, with cereal losses amounting to 470 million Algerian dinars, representing 34% of the wilaya's cereal production in 2012.

These figures underscore the significant damage caused by floods in the agricultural sector. At the same time, the scarcity of data highlights the challenges in developing accurate risk models and insurance schemes to mitigate this risk. The limited information available emphasizes the urgent need for more systematic data collection and analysis to better understand flood impacts and enhance the resilience of agricultural practices in Algeria.

1.1.4 Coverage Against Flood Risk in Algeria

In general, according to the 2023 report by the CNA, agricultural insurance accounts for only 1.4% of the revenue in the Algerian insurance market. Seven companies offer agricultural insurance, with CNMA holding a 75% market share. However, the penetration of this insurance remains very low, with only 10% of farmers being insured.

Cereal crop insurance is provided under a multi-risk agricultural contract, with fire and hail as the basic coverage. Flood risk and climatic risks in general are still considered uninsurable in

Algeria, with CNMA being an exception by offering flood insurance starting in 2020, though on a limited scale.

Till this day, flood risk in agriculture is classified as an agricultural calamity. According to Executive Decree No. 12-251 of June 5, 2012, an agricultural calamity is defined as an uninsurable, exceptional damage caused by abnormal variations in the intensity of a natural agent, where the usual preventive or curative measures in agriculture have proven insufficient or ineffective.

The coverage of damages resulting from agricultural calamities, such as floods, is provided by the Agricultural Calamity Guarantee Fund (FGCA), which is managed by the CNMA. Compensation procedures are only initiated after the disaster area is declared by an inter-ministerial decree involving the ministries responsible for finance, interior, and agriculture, and are limited to the resources available for the fund. Moreover, the fund is marked by the administrative burden in handling indemnification files.

The fund does not cover damages unless they exceed 10% of the total gross product of the affected Wilaya, and the coverage is limited to 80% of the damage.

“DACA” a system for agricultural disaster insurance has been the subject of several studies, starting in 1988 and continuing to the present day. The proposed product aims to cover farmers against risks related to climate change, including forest fires, floods, and droughts. However, to date, the DACA system has not yet been implemented.

The agricultural insurance system in Algeria remains inadequate, offering limited coverage for floods through the FGCA. The significant gaps in coverage and numerous limitations highlight the pressing need for a new approach to ensure that farmers are effectively protected against flood risks.

1.2 Presentation of the study area and the Data

For our study on flood index insurance, we selected the Wilaya of El Tarf as the study area, based on two key considerations. First, the agricultural sector in El Tarf is notably vulnerable to flooding. As previously discussed, the estimation by the FGCA for damages caused by the February 2012 flood amounted to 2.326 billion Algerian dinars. This sum represents the highest damage assessment among the four major floods impacting the agricultural sector, underscoring the significant risks faced by farmers in the region. Second, the flood events recorded in this area have been documented through satellite imagery from various sources. This documentation provides valuable data that supports the development of the vulnerability function, which will be further discussed in the subsequent sections of this dissertation.

1.2.1 Presentation of the study area

The wilaya of El Tarf is located in the extreme northeast of Algeria. It is bounded to the north by the Mediterranean Sea, to the east by Tunisia (Governorate of Jendouba), to the south by the wilaya of Souk Ahras, to the southwest by the wilaya of Guelma, and to the west by the wilaya

Figure 5: Geographic situation of the Wilaya of El Tarf



Source : decoupageadministrativealgerie.com

of Annaba. The region has a significant hydrological network and numerous bodies of water, including two watersheds and several wetlands composed of five lakes and two dams.

The area has experienced several floods, one of the most severe being on February 22-23, 2012, resulting in three fatalities and more than one-third of the land being completely flooded.

Although the wilaya possesses other economic potentials such as fishing and tourism, agriculture remains the primary economic sector, employing more than a third of the active population and serving as a key driver of job creation and wealth generation. The wilaya, with its cereal production vocation, yielded 529,480 quintals in 2021.

1.2.2 Precipitation Data

We have at our disposal daily precipitation data in millimeters for the Wilaya of El Tarf spanning the past 38 years, from 1985 to 2023. These real-time satellite data are collected from NASA's CERES and MERRA-2 datasets, derived from NASA's Global Precipitation Measurement (GPM). This satellite was launched in cooperation between the United States' NASA and Japan's JAXA to provide detailed spatiotemporal precipitation data across all continents. This product offers the best estimation of precipitation by combining data from multiple sensors aboard several satellites.

It is important to note that the station used covers a grid cell that spans the entire Wilaya of El Tarf. Therefore, we put forward the hypothesis that the precipitation is homogeneous across the entire Wilaya of El Tarf.

- The average daily precipitation is approximately 3.86 mm. The 90th percentile is 11.26 mm. This distribution is highly concentrated around lower values, suggesting that while heavy precipitation events are less frequent, they are significantly more intense when they do occur.
- Approximately 33.16% of the days in the dataset have zero precipitation over the study period, reflecting a pattern of relatively rare but significant precipitation events.
- Notably, January, and December experience the highest cumulative precipitation levels, with averages of 96 mm, and, 100 mm respectively. This is followed by a gradual decrease

as the region transitions into the drier summer months. With an average of 6.5 mm of rainfall during July. The annual cumulative precipitation for the area is approximately 650 mm, which is higher than the average annual precipitation of the Mediterranean climate.

- Significant precipitation events were recorded in 2012 and 2023, with the highest single-day precipitation of 118.97 mm on February 22-23rd, 2012, and 120 mm on May 21st, 2023. These dates are associated with two of the most devastating floods that impacted the region.

Figure 7: Histogram of daily precipitation

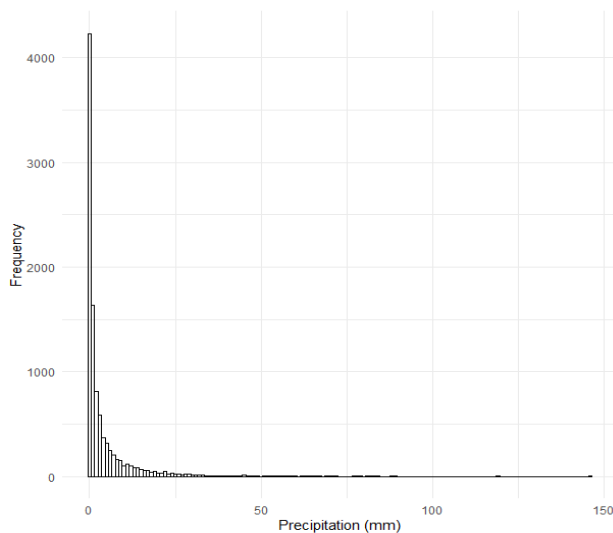
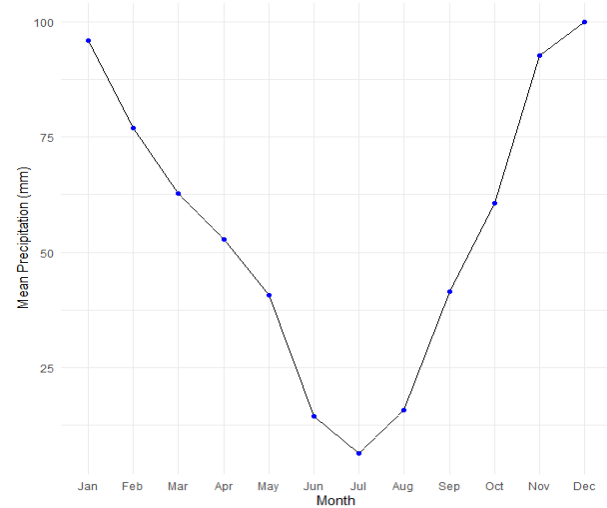


Figure 6: The variations in monthly precipitation



Source: Established by the author using R software

1.2.3 Vulnerability Data:

As previously mentioned, floods are meteorological hazards caused by extreme precipitation over a short period of time. the severity of a flood event is particularly influenced by geography and environmental characteristics, such as the presence of a watercourse or a catchment area in the region. These factors can exacerbate the impact of heavy rainfall, leading to severe flooding events.

Due to absence of historical data that allows us to model the vulnerability and thus the function explaining the damage occurred, in our study, we aim to model the vulnerability to flood risk by integrating geographical variables. By using a Generalized Linear Model (GLM), we can statistically analyze how these factors contribute to the likelihood and severity of flooding.

Flood triggers were identified through an extensive review of existing literature and discussions with researchers from the ITGC (Technical Institute of Large Crops), the National Institute of Soils, Irrigation, and Drainage (INSID) and the National Agency for Water Resources (ANRH). We will leverage data from multiple satellite sources available of internet that are designed specially to conduct comprehensive geological and meteorological analyses.

The goal of this modeling is to explain the rate of destruction based on the meteorological variable(precipitation), and the geological variables.

- The land use map was obtained from the Environmental Systems Research Institute (ESRI), derived from ESA Sentinel-2 imagery at a 10m resolution.
- Digital Elevation Model (DEM) images were sourced from the ASTER Global Digital Elevation Model, a collaborative effort between NASA and Japan's Ministry of Economy, Trade, and Industry (METI), providing a 30m resolution. This data, available on NASA's Earth Data Search, was used to extract slope, elevation, and TWI maps.
- The digital soil map was obtained from the Food and Agriculture Organization (FAO)
- The soil moisture data was sourced from NASA's MERRA-2 Reanalysis.

1.2.4 Exposure data:

For an index-based insurance designed to cover flood risk for cereal production. Data on agricultural land and cereal production were collected from the Ministry of Agriculture. This data will be applied to the damage rate of our study area to accurately assess the value of the damage.

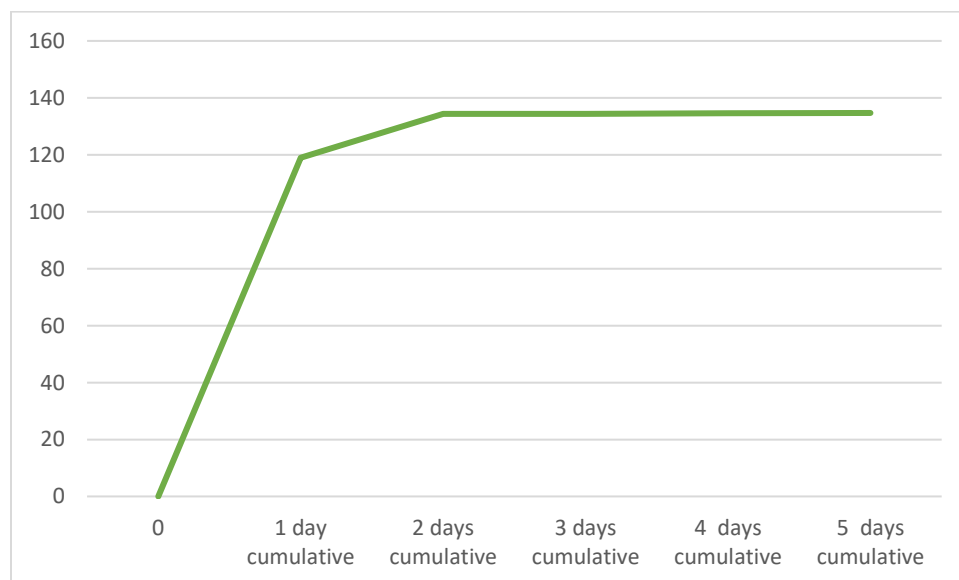
1.3 The construction of the Index:

The construction of the index is the most critical step in the implementation of index-based insurance. Given the absence of claim adjustment post-disaster, it is crucial for the index to be highly correlated with actual incurred losses. Thus, the index must be clearly identifiable, external, objective, easily measurable.

As previously discussed in the literature review, within the framework of flood index-based insurance, continuous rainfall measurements serve as a reliable proxy for assessing the occurrence and severity of flood events. Following an analysis of multiple events and consultations with an expert from the ANRH, who confirmed that extreme precipitation events in Algeria generally do not exceed two days, we have concluded that a two-day cumulative precipitation measure provides a more accurate representation of flood occurrences.

For example, in analyzing the major floods that affected our study area in February 2012, we observed that after the second day, the intensity of precipitation progressively decreases.

Figure 8: Precipitation cumulative (mm) for a flood event



Source: Established by the author

After determining the necessary number of days for setting the threshold, it is important to establish the exceedance threshold for precipitation over two days to trigger the coverage.

To achieve this, we conducted an analysis of historical flood events that affected our study area. The summary of this analysis is presented in the following table.

Table 2:Historical flood events' detection with 100 mm threshold

Flood events	01 Day	Detection	02 Days cumulative	Detection
21-mai-2023	120.71	Yes	154.74	Yes
23-june-2023	51.13	No	100.44	Yes
26-mai-2019	30.2	No	81.5	No
17-Oct-2018	58.37	No	116.74	Yes
10-feb-2017	83.1	No	106.9	Yes
28-Nov-2015	44.19	No	114.72	Yes
23-24-feb-2012	118.97	Yes	134.34	Yes

Source: Established by the author

The 100 mm threshold enables the detection of 5 out of the 6 major flood events in our study area. The analysis in the accompanying table further corroborates that cumulative precipitation over two days is a more effective measure than daily precipitation for identifying flood occurrences. Based on these analyses, we conclude that a cumulative precipitation of 100 mm over two days should be used as the trigger index for the insurance guarantee. This threshold corresponds to a return period of two years.

To build the index the following steps are going to be follows:

- The cumulative rainfall over two days is calculated on a daily basis From September to July, which coincides with the coverage period of our index insurance. It should be noted that the cereals, which are the subject of our insurance coverage, are planted in September and harvested by July at the latest.
- An event is triggered when the cumulative precipitation over two days exceeds 100 mm and ends when it falls below this threshold. For each detected event, only the peak of the cumulative two-day rainfall is used to calculate the loss rate.
- Linking the rainfall to a loss rate, commonly referred to as the vulnerability rate. The modeling of vulnerability rates will be addressed in detail in our next section.
- Construct a damage curve, indicating the damage caused by floods for each level of precipitation in the area.
- Apply the damage rate to our exposure, which is the cereal production in the Wilaya under study.

SECTION 2: MODELING FLOOD VULNERABILITY

Flood vulnerability in a region is influenced by various geo-environmental factors. A specific precipitation can lead to total destruction in one area while causing minimal impact in another. The contributing factors include the low capacity of watercourses or river networks, catchment area characteristics, and topography. Floods arise from a combination of hydro-meteorological and physiographic factors.

This section aims to link meteorological factors, specifically cumulative precipitation, with geo-environmental factors to develop a vulnerability function for our study area. This will enable us to derive a destruction function that reflects the area's specific characteristics.

Based on the extent of our research, the use of satellite imagery for vulnerability assessment has been widely applied in fields like hydrology, geology, and environmental science. However, we have not identified an academic study that specifically models these factors to derive a destruction function for the pricing of an index-based flood insurance. Previous research, such as by (Ritling & Nguyen, 2014) and (Idir, 2021), utilized a linear function to express the relationship between precipitation and damage rates. Our study aims to expand upon this by incorporating geo-environmental factors alongside meteorological data to develop a comprehensive destruction function.

2.1 Selection of the dataset

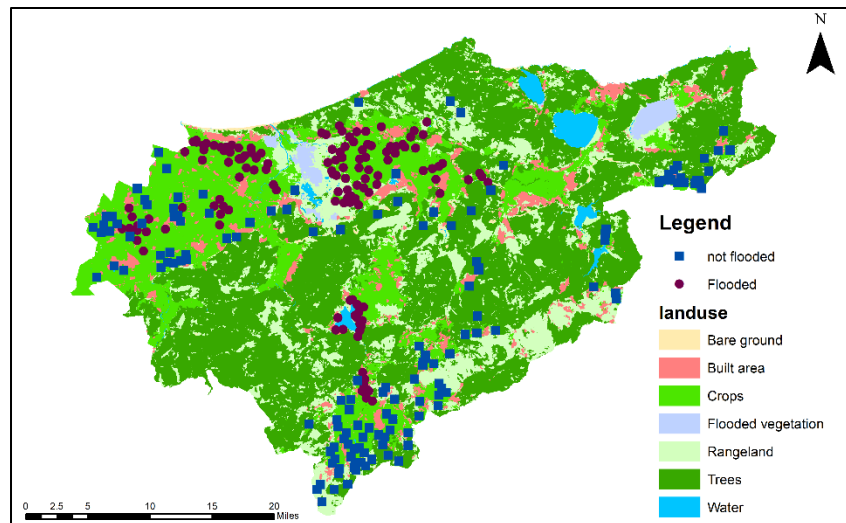
To assess flood vulnerability accurately, it is crucial to select the appropriate factors, as a controlling factor might be highly significant in one area but not in another (Megahed, Abdo, Abdelrahman, Scopa, & Hegazy, 2023). The selection of vulnerability indicators involved consulting prior literature and seeking expert opinions from institutions such as the INCID, ANRH, and ITGC.

2.1.1 Response Variable

Our Response variable is the occurrence of flooding in an area. To model this variable, we will identify flood locations based on historical flood records. This step involves using satellite images of significant flood events, sourced from the Algerian Space Agency (ASAL) and the International Charter Space and Major Disasters, using RADARSAT-2 data, a trademark of the Canadian Space Agency. The satellite images used are listed in the appendices. By analyzing past events, we can estimate potential locations for future extreme events and identify their conditioning factors. This is the first major step in assessing flood vulnerability.

The flood inventory map shows the locations of flooded areas in the agricultural zones studied. We identified 144 flood points and 144 non-flood points. Flood and non-flood points were assigned values of 1 and 0, respectively.

Figure 9: Map showing flooded and non-flooded locations in agriculture areas



Source: Established by the author using ArcMap Software

2.1.2 Predictor variables

- **Precipitation:** Precipitation is a critical predictor factor for flooding, as it directly influences the likelihood and severity of flood events (Morante-Carballo, Montalván-Burbano, Arias-Hidalgo , & Domínguez-Granda , 2022). To establish the vulnerability function and subsequently the damage rate function, we linked the detected flood event with the corresponding two-day cumulative precipitation data related to the events, allowing us to quantify the relationship between the amount of rainfall and the resulting flood and predict the extent of damage that might occur under different precipitation conditions in the future.
- **Distance to rivers:** El Tarf is characterized by a network of rivers and wadis that significantly influence the area's vulnerability to floods. The hydrological dynamics of these rivers, coupled with the region's topography, create conditions conducive to flooding, particularly during heavy rainfall events. (Figure 10). The map shows that 48% of crop land is located within a distance of less than one kilometer from rivers.
- **Elevation:** Low-lying areas are more prone to flooding (Pujari, S & Wayal, A, 2023). The elevation map was created using a Digital Elevation Model (DEM). The study area exhibits complex topographical features, with the upstream region characterized by very high altitudes ranging from 450 meters to over 1367 meters. In contrast, the coastal area, which is primarily urban and agricultural, lies in lowland areas with elevations varying from 8 meters below sea level to 100 meters above (figure 11). Crop land in the Wilaya of El Tarf is predominantly located in low-lying areas, with 75% of it situated on land with an elevation ranging from -8 to only 70 meters above sea level.
- **Slope:** As previously mentioned, slope is a key physiographic feature that significantly influences flood vulnerability. Analysis of the slope map reveals that the gradient in the region ranges from 0 to 55 degrees. 65% of the crop land is located in areas with zero slope, making them prone to flow accumulation during extreme rainfall events (Figure 12).
- **TWI:** The Topographic Wetness Index (TWI) map was generated from a digital elevation model (DEM). TWI calculates the potential for water to accumulate at any given point, considering the influence of gravity. In this context, lower TWI values indicate regions with limited water accumulation and runoff, while higher values suggest areas where water is more likely to gather. The analysis of the TWI map reveals that TWI values, range from -8 to 13.3 (Figure 13). 50% of the agricultural land in our study area is located in zones with a positive

TWI, reaching up to 12.8. These zones are likely to experience greater water retention, which, in the event of heavy rainfall, could lead to significant surface runoff and flooding.

- **Soil type:** According to the FAO classification, the soil in our study area is categorized into 4 types. The table below summarizes the permeability properties of each type:

Table 3:Permibility characteristics of the soil in the Wilaya of El Tarf

Code	Soil	Area%	permeability	Code
Lg	Luvisols, podzolic	2.3%	Moderate ²	1
BK	Cambisols	52.8%	Moderate ³	1
Vc/Vp	Vertisols	44.2%	Low ⁴	2
Wr	Planosols	0.8%	Low ⁵	2

Source: Established by the author

According to Figure 14, the soil in the northern part of the study area is characterized by low permeability, whereas the southern part exhibits moderate permeability. For crop land, the soil of 70% of this later have a low permeability.

- **Soil moisture:** Figure 15 shows a gradient of soil moisture content across the region, with higher moisture levels typically found in the northeastern part and lower levels in the western part. This distribution is influenced by factors such as precipitation patterns, proximity to the coast, and the type of soil, which in this area tends to be less permeable.

This primary analysis gives us an idea on the high vulnerability of crop land to floods in our study area.

² <https://indico.ictp.it/event/a0124/material/1/6.pdf>, retrieved :12 7,2024.

³ <https://museum.isric.org/monoliths/reference-soil-netherlands-07>, retrieved: 12 7 .2024.

⁴ <https://doi.org/10.1016/B978-0-12-809239-2.00008-5>, retrieved: 12 7 .2024.

⁵ <https://www.isric.org/>, retrieved: 12 7 .2024.

Figure 10: Distance to Rivers' map of the Wilaya of El Tarf

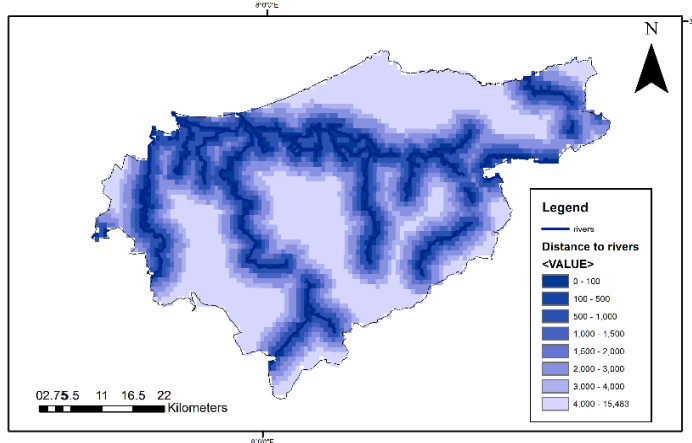


Figure 11: Elevation's map of the Wilaya of El Tarf

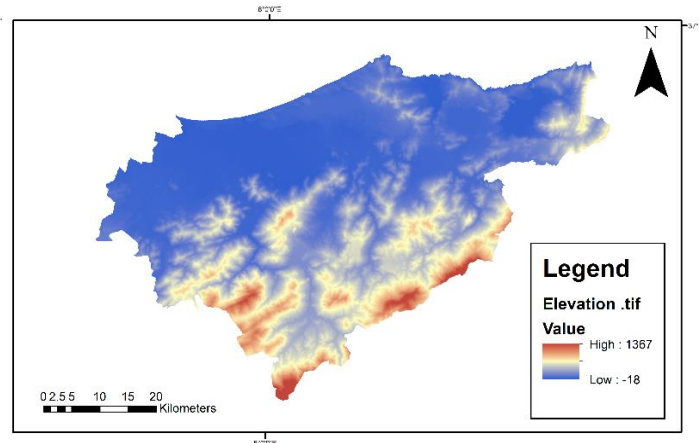


Figure 12: Slope's map of the Wilaya of El Tarf

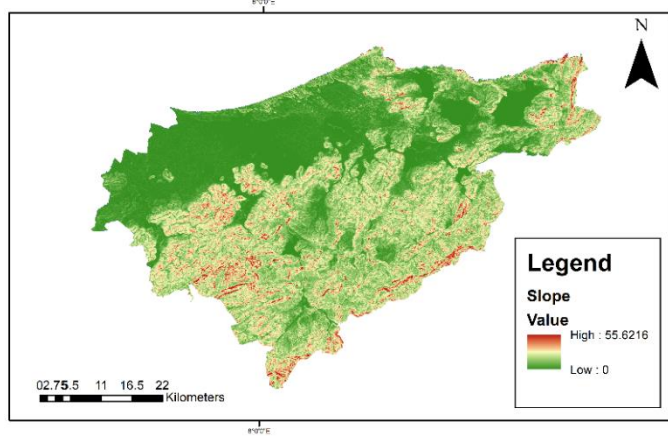


Figure 13: TWI's map of the Wilaya of El Tarf

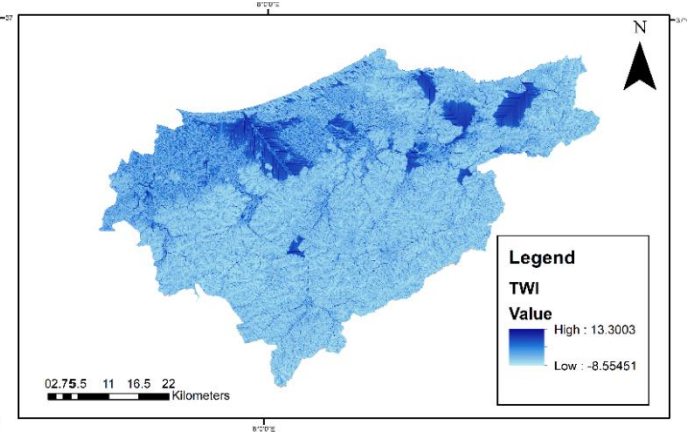


Figure 14: Soil Types' map of the Wilaya of El Tarf

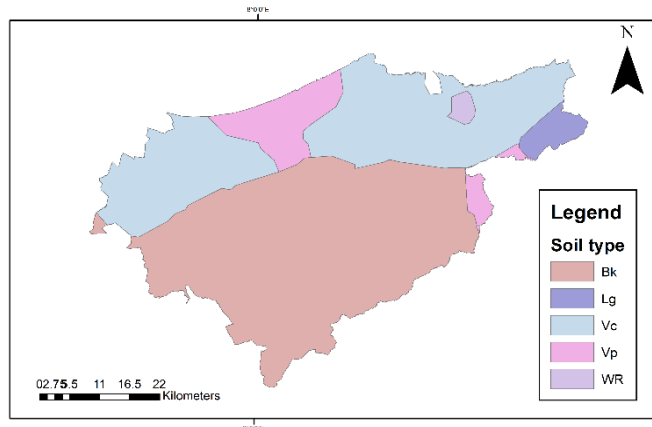
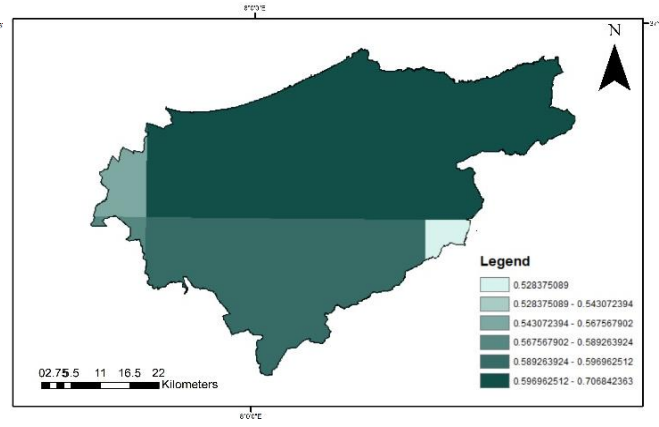


Figure 15: Soil Moisture's map of the Wilaya of El Tarf



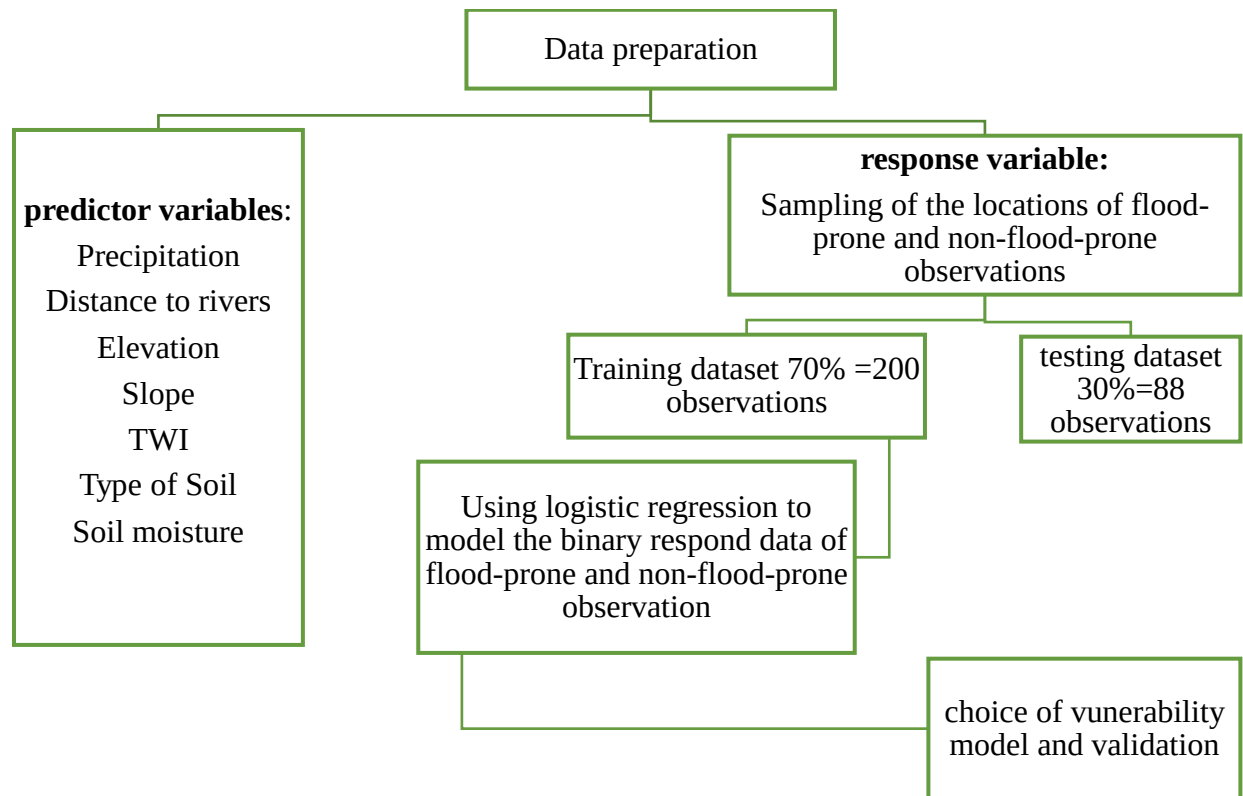
Source: Established by the author using ArcMap Software

2.2 Methodology and results

In this part of the study, the methodology is summarized in Figure 16. The first step in developing an appropriate GLM model is to select the training and testing datasets. A total of 288 flood-prone and non-flood-prone observations were selected and randomly split into two datasets in a 70/30 ratio. Since the risk of flood is treated as a binary classification task in GLM models, flood observations were assigned a value of 1, while non-flood-prone observations were given a value of 0. Finally, the values of the flood influence factors were extracted from these observations and organized into a file, which was processed using R software.

A brief overview of the GLM Model is provided in the appendices.

Figure 11: Methodological Flowchart for Flood Vulnerability



Source: Established by the author

2.2.1 Estimation of Logistic Regression Coefficients

For the estimation we used a Stepwise selection which is a method used in statistical modeling to select a subset of predictor variables that provides the best fit for the model. It Combines both forward selection and backward elimination. For our model we chose to start with a full model that contains all variable:

```
Call:
glm(formula = RD ~ soil+ precipitation + twi + distance + moisture + elevat
ion + slope , family = binomial(), data = detection_data)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.65409	-0.20758	-0.00637	0.05641	2.73009

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-9.2349102	5.7627806	-1.603	0.109043
soil	-0.6802383	0.9064224	-0.750	0.452975
precipitation	0.0061285	0.0025418	2.411	0.015889 *
moisture	24.9213302	11.4877797	2.169	0.030054 *
distance	-0.0006934	0.0003516	-1.972	0.048590 *
twi	1.2838917	0.3814987	3.365	0.000764 ***
elevation	-0.0104107	0.0343587	-0.303	0.761596
slope	0.1139497	0.1484255	0.768	0.442651

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 277.179 on 199 degrees of freedom

Residual deviance: 64.768 on 192 degrees of freedom

AIC: 80.768

Number of Fisher Scoring iterations: 8

```
> mcfadden_r2 <- pr2(model)["McFadden"]
fitting null model for pseudo-r2
McFadden's R-squared: 0.766333
```

Significant predictors (indicated by p-values less than 0.05) include precipitation, soil moisture, distance to rivers, and TWI. These variables have noteworthy impacts on the model, while others like soil, elevation, and slope not statistically significant.

I. Spearman Rank Correlation

The Spearman rank correlation coefficient measures the strength and direction of the monotonic relationship between two variables. It is a non-parametric measure, meaning it does not assume a specific distribution of the data. Unlike the Pearson correlation, which measures linear relationships, Spearman's rank correlation assesses the monotonicity of the relationship, it does not require the data to meet specific parametric assumptions (like normality), making it suitable for the non-linear and potentially non-normal nature of logistic regression predictions.

Table 4: Spearman Rank Correlation Matrix

	Precipitation	Distance-rivers	TWI	Soil-moisture	Soil-type	Elevation	Slope
Precipitation	1.0000	-0.0804	0.3896	0.3180	0.4050	-0.6426	-0.4565
Distance-rivers	-0.0804	1.0000	-0.3679	-0.1661	-0.2833	0.3249	0.3666
TWI	0.3896	-0.3679	1.0000	0.4644	0.4769	-0.4883	-0.8763
Soil-moisture	0.31806	-0.1663	0.4644	1.0000	0.7024	-0.5193	-0.4514
Soil-type	0.40504	-0.2833	0.4769	0.7024	1.0000	-0.6760	-0.6012
Elevation	-0.6426	0.3249	-0.4883	-0.5193	-0.6760	1.0000	0.7384
Slope	-0.4565	0.3666	-0.8763	-0.4514	-0.6012	0.7384	1.0000

Source: Established by the author using R software

According to the correlation matrix, the variables soil type, elevation, and slope show significant correlations both among themselves and with other variables. Notably, there is a strong correlation between TWI and slope, as well as between slope and elevation. It's important to note that these three variables are derived from the same model (DEM).

A notable correlation also exists between soil type and soil moisture. As soil moisture increases, the permeability tends to decrease (with soil types classified by their permeability from least to most permeable).

Additionally, it can be observed that the variables TWI, precipitation, soil moisture, and distance to rivers exhibit low correlations with each other.

II. Variance Inflation Factor (VIF)

In order to confirm the existence or not of multicollinearity we use VIF. Collinearity occurs when predictor variables in a regression model are highly correlated, which increases the standard errors of the regression coefficients, leading to wider confidence intervals and less significant p-values. The VIF quantifies this effect (Miles, 2005).

$$VIF_{(k)} = \frac{1}{1-R_{(k)}^2}$$

Where $R_{(k)}^2$ represents the coefficient of determination from a regression of all other predictors on the predictor k. When there is only one predictor in the model, or when predictors are uncorrelated, the term $\frac{1}{1-R_{(k)}^2}$ simplifies to 1. If the VIF exceeds 1, it suggests that the predictors may be experiencing some level of correlation. Values ranging from 5 to 10 indicate a high degree of correlation that could be problematic, while a VIF above 10 implies significant multicollinearity, potentially leading to unreliable estimates of the regression coefficients. Such cases warrant careful attention and corrective measures (Akinwande, Dikko, & Samson, 2015).

Table 5: Variance Inflation Factors

moisture	TWI	distance	precipitation	slope	elevation	soil
2.226642	2.332311	1.421402	4.147528	2.621955	2.508062	2.1759

moisture	TWI	distance	precipitation
2.089712	1.121755	1.183385	1.115594

Source: Established by the author using R software

After eliminating the correlated variables (slope, elevation, and soil type), we observe a notable reduction in the VIF values, bringing them closer to 1. This indicates that the remaining four variables exhibit a low level of correlation.

After addressing multicollinearity in the model and applying a stepwise approach to include and exclude variables, The selected model is the one that has the highest McFadden's R-squared and the lowest AIC.

```
Call:
glm(formula = RD ~ precipitation + twi + distance + moisture,
    family = binomial(), data = detection_data)
```

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.36707  -0.19361  -0.01985   0.02968   2.61506
```

```
Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -2.317e+01  5.244e+00  -4.419  9.91e-06 ***
precipitation  3.771e-02  1.227e-02   3.074  0.00211 **
twi           1.167e+00  2.424e-01   4.815  1.47e-06 ***
distance     -9.593e-04  3.165e-04  -3.030  0.00244 **
moisture      3.888e+01  8.308e+00   4.680  2.87e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

(Dispersion parameter for binomial family taken to be 1)

```
Null deviance: 277.179  on 199  degrees of freedom
Residual deviance:  64.848  on 195  degrees of freedom
AIC: 74.848
```

```
Observations      200
McFadden's R^2    0.808
```

The p-values indicate that all predictors are statistically significant in the model. TWI, precipitation, and soil moisture have positive coefficients in the model, indicating that an increase in these variables raises the likelihood of flooding. This finding aligns with the literature. Conversely, the distance to rivers has a negative coefficient, meaning that as the distance decreases (closer proximity to rivers), the likelihood of flooding increase.

Furthermore, McFadden's R^2 is 80.8%, indicating that the chosen variables explain 80% of the variability in the likelihood of flooding in the area. This high R^2 value suggests a strong explanatory power of the model for predicting flood events based on these factors.

2.2.2 Model evaluation

I. Hosmer–Lemeshow test

The Hosmer-Lemeshow (HL) test is a widely utilized global goodness-of-fit (GOF) measure designed to evaluate the overall fit of a logistic regression model. This test statistic is based on the Pearson chi-squared statistic, which compares the actual event counts to those expected (Surjanovic & Loughin, 2023). First, the data is sorted based on the predicted probabilities, and then divided into g groups. The Hosmer-Lemeshow test statistic is then computed using the following formula:

$$G_{HL}^2 = \sum_{j=1}^g \frac{(O_j - E_j)^2}{E_j \left(1 - \frac{E_j}{n_j}\right)} \sim \chi_{HL}^2$$

- χ^2 represents the chi-squared distribution.
- n_j is the number of observations in the j^{th} group.
- O_j is the observed number of cases in the j^{th} group.
- E_j is the expected number of cases in the j^{th} group.
 - Null Hypothesis (H_0): There is no significant discrepancy between the observed and expected outcomes.
 - Alternative Hypothesis (H_1): The model's predictions do not adequately reflect the actual outcomes.

Table 6: Hosmer–Lemeshow test

X-squared (Chi-Squared Statistic)	Degrees of Freedom (df)	p-value:
2.1078	8	0.9775

Source: Established by the author using R software

p-value=0.9775 > 0.05: We fail to reject the null hypothesis. There is no significant difference between the observed and expected frequencies. This means the model fits the data well according to the Hosmer-Leeshawn test.

II. Confusion matrix

A confusion matrix is a fundamental tool for evaluating the performance of a classification model, including logistic regression models. It is a table that summarizes the model's predictions against the actual outcomes and helps in understanding how well the model distinguishes between different classes. Our testing dataset contain 88 observations, the matrix of confusion for our model is presented in the table:

Table 7: Confusion matrix for the logistic regression

	false	true
false	44	3
true	4	37

Source: Established by the author

Out of 41 observed flooding cases, the model correctly identified 37 as True Positives (TP), while misclassifying 4 cases as False Negatives (FN). For the 47 observations where flooding did not occur, the model correctly predicted 44 as True Negatives (TN), with 3 cases being False Positives (FP).

- **Accuracy:** The proportion of correct predictions is 92.045%.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} = 92.045\%$$

- **Precision (Positive Predictive Value):** The proportion of positive predictions that are actually correct is 92.5%.

$$\text{Precision} = \frac{TP}{TP+FP} = 92.5\%$$

- **Recall (Sensitivity or True Positive Rate):** The proportion of actual positives that are correctly identified by the model.

$$\text{Recall} = \frac{TP}{TP+FN} = 90.24\%$$

- **Specificity (True Negative Rate):** Specificity measures the proportion of actual negative cases that the model correctly identified.

$$\text{Specificity} = \frac{TN}{TN+FP} = 93.62\%$$

These metrics suggest the model is well-calibrated and capable of effectively distinguishing between flooding and non-flooding cases, making it a reliable tool for predicting such events.

The vulnerability function can be written as:

$$\pi = p(x = 1) = \frac{1}{(1 + e^{-(23.17 + 0.0377\text{precipitation} + 1.167\text{TWI} - 0.0009593\text{distance} + 38.88\text{moisture})})}$$

2.3 Determination of destruction rate function for the Wilaya of El Tarf

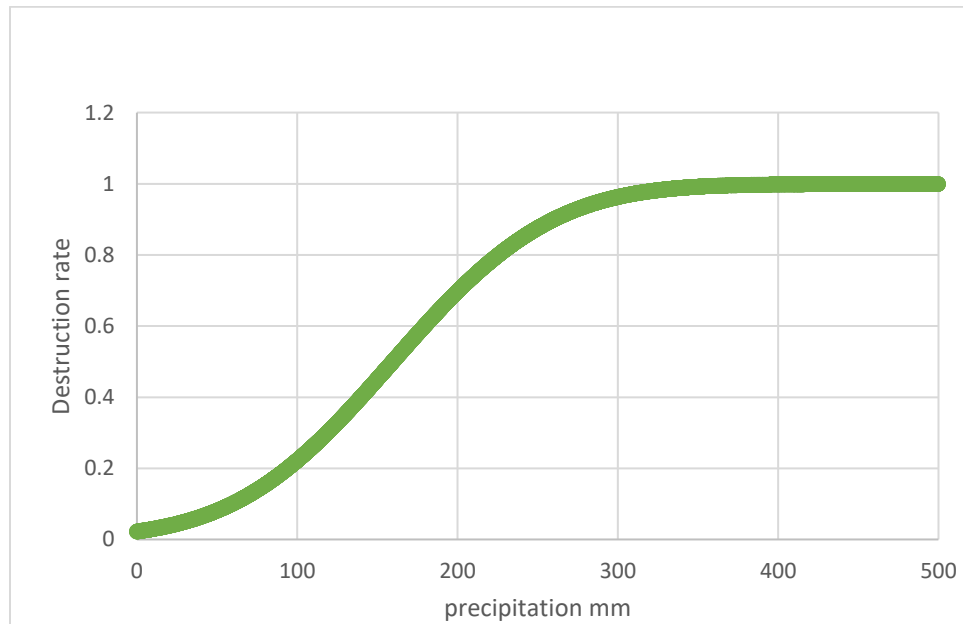
The primary goal of determining the vulnerability function is to develop a damage function that connects our index with a damage rate that is specific to the region under study. By establishing this relationship, we aim to reflect the impact of an extreme precipitation event in terms of potential damage.

To estimate the damage function for our study area, we assumed that cereal production is uniform across the entire crop area. We then randomly selected 20,000 locations within this area (See appendix 2). By taking the average damage rate from these randomly chosen locations, we can represent the average damage rate for the entire study area.

The only concrete evidence we have is from reports by the FGCA, which indicate that the damage rate for cereal crops during the flood event on February 22-23, 2012, was 34.34%. Our model estimated the damage rate for the same event at 35.12%, demonstrating a detection accuracy

of 97.77%. While this single data point cannot fully confirm or refute the model's accuracy, it is valuable. Throughout this research, we aim to leverage all available information, given the lack of historical damage data.

Figure 12: Destruction function of the wilaya of El Tarf



Source: Established by the author

The figure indicates that the crop area approaches near-total destruction (99.99%) when the two-day cumulative precipitation reaches 500 mm, thereby confirming the high vulnerability of agricultural land in El Tarf to flood risk.

In addition of the destruction function, we have to define the payment function which is designed to determine compensation based on the severity of the catastrophe. The lower limit, commonly referred to as the attachment point, is set such that any losses with an index below this threshold are not covered. For our analysis, we have set this index at 100 mm, a threshold that captures extreme and catastrophic events, resulting in a 20% damage rate. We can adjust this threshold to offer different levels of coverage, starting from as low as a 15% or 10% damage rate.

SECTION 3: MODELING OF EXTREME PRECIPITATION AND PRODUCT PRICING

As discussed in the first chapter, extreme value theory (EVT) is crucial for accurately modeling the tail of a distribution, allowing for a precise representation of rare and extreme events—something that traditional statistics, which primarily focuses on the distribution's mean, cannot achieve. This approach is particularly relevant to our index insurance product, where we aim to model instances of two-day cumulative precipitation exceeding a specific threshold. EVT enables a more accurate risk assessment, allowing insurers to set appropriate coverage levels and design policies that reflect the true likelihood and potential impact of extreme flooding events.

3.1 Modeling the extreme precipitation

Understanding the behavior of extreme precipitation events is crucial for effective hazard assessment and insurance modeling, after modeling the vulnerability we will explore the application of EVT in modeling the exceedances of a two-day cumulative precipitation threshold, in order to determine the pure premiums.

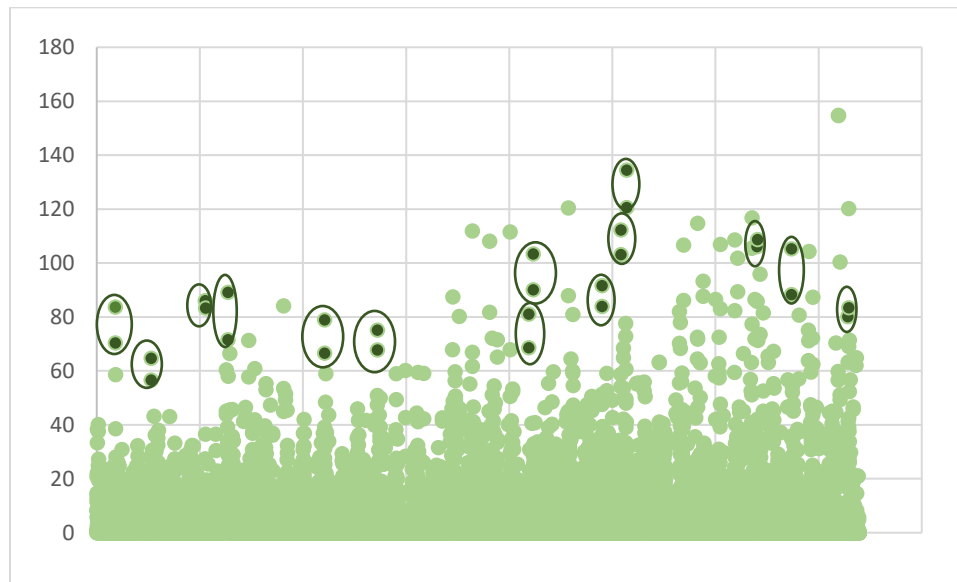
3.1.1 Addressing the dependency

Ensuring the independence of data series is a critical challenge in applying extreme value theory. In the context of meteorological data, heavy rainfall on one day can significantly influence the likelihood of subsequent rainfall events. This dependence is particularly relevant when analyzing extreme precipitation events, as the occurrence of one heavy rainfall event can set up favorable conditions for further rainfall (Barlow, 2019).

In our case, it is essential to calculate the aggregation of two-day rainfall. While the assumption of independence between daily precipitation events might seem acceptable, the same cannot be said for aggregated precipitation events.

The following graph illustrates some dependent observations of subsequent events using two-day cumulative precipitation.

Figure 13: Illustration of dependency in two days cumulative precipitation.



Source: Established by the author

To work with EVT, it is common to assume that extremes are independent when they exceed a sufficiently high threshold u . This essentially means assuming that extreme events are sufficiently spaced out in time to be considered independent. This can be achieved through the decluttering technique which is the most commonly used approach (Coles S. , 2001):

1. **Identify Clusters:** Define the cluster which is a continuous period where a threshold is exceeded. This definition aligns with the concept of a flood event, as the event insured by the index product is triggered when the threshold is surpassed and ends when the precipitation falls below this threshold. To ensure independence of events, we can introduce a constraint on the minimum temporal spacing between clusters. After analyzing the exceedances, we found that a temporal spacing of three days allows us to consider the events as independent.
2. **Maximum Exceedance:** Identify the maximum exceedance within each cluster. The index will be calculated based on the maximum value of the exceedance.
3. **Probability Distribution Adjustment:** Fit the probability distribution to the maxima of the clusters.

To confirm the result of the decluttering we use The Wald-Wolfowitz Runs Test which is a non-parametric test for independent samples.

- Null Hypothesis (H₀): There are no significant patterns or clustering
- Alternative Hypothesis (H₁): There are significant patterns or clustering in the data.

statistic	p-value
1.1548	0.2482

The p-value is 0.2482, which is greater than the common significance levels of 0.05. We fail to reject the null hypothesis. There are no significant patterns or clustering in the Data.

3.1.2 Addressing the stationarity of extreme events

Extreme value analysis traditionally relies on the assumption of stationarity, which implies that statistical properties of the data remain constant over time. However, recent research has questioned this assumption due to the rise in extreme weather events. This shift has the potential to lead to an underestimation of flood quantiles and an increased risk of undervaluing insurance premiums.

To inspect the stationarity, first we are going to use Kwiatkowski–Phillips–Schmidt–Shin (KPSS) tests which test the null hypothesis that an observable time series is stationary around a deterministic trend.

- Null Hypothesis (H₀): The time series is level or trend stationary.
- Alternative Hypothesis (H₁): The time series is not stationary.

The results of the test are as it follows:

Table 8: KPSS test results

KPSS Level Statistic	Truncation Lag Parameter	p-value:
0.043138	3	>0.1

Source: Established by the author using R software

The p-value of 0.1 indicates that the test does not reject the null hypothesis of trend stationarity at the 10% significance level. In other words, there is no significant evidence against the null hypothesis, the series can be considered trend stationary.

To inspect furthermore the stationarity, we will use The Phillips-Perron (PP) Unit Root Test which is a statistical test used to determine whether a time series has a stochastic trend.

- Null Hypothesis (H0) the time series has a unit root, implying that the series is non-stationary
- Alternative Hypothesis (H1) the time series does not have a unit root, implying that the series is stationary

The results of the test are as it follows:

Table 9: The Phillips-Perron test results

Dickey-Fuller Z(alpha):	Truncation lag parameter:	p-value	alternative hypothesis
-61.26	3	0.01	stationary

Source: Established by the author using R software

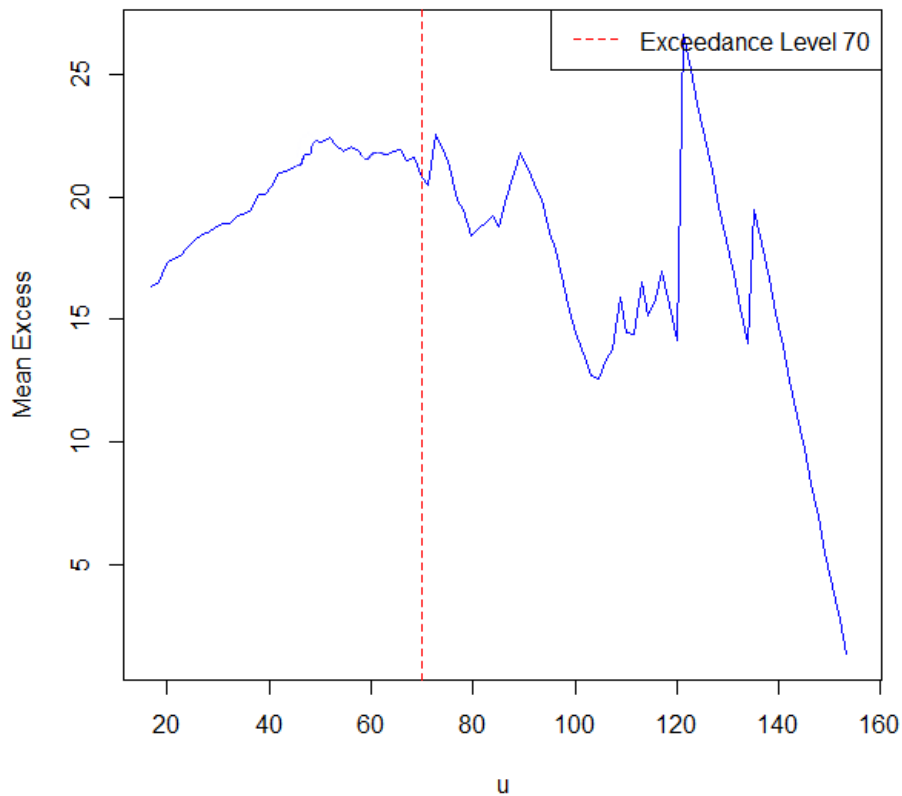
P-value < 0.05 we reject the null hypothesis of a unit root. there is statistically significant evidence to suggest that the time series is stationary.

In the absence of a determinist and stochastic trend we can say that statistically our data series is stationary.

3.1.3 threshold determination

We need to determine a threshold to model our distribution. The selection of this threshold is among the analytical complexities of the POT method. This threshold must be sufficiently high to ensure that the approximation by the distribution is valid and to guarantee the economic reliability of the model. However, it should not be too high to ensure that there are enough observations above the level u to estimate the parameters. To determine this threshold, we will use the mean excess plot as a graphical method for threshold determination, utilizing the “ismev” package in R. Below is the "Mean Excess Plot".

Figure 14: Mean Excess Function

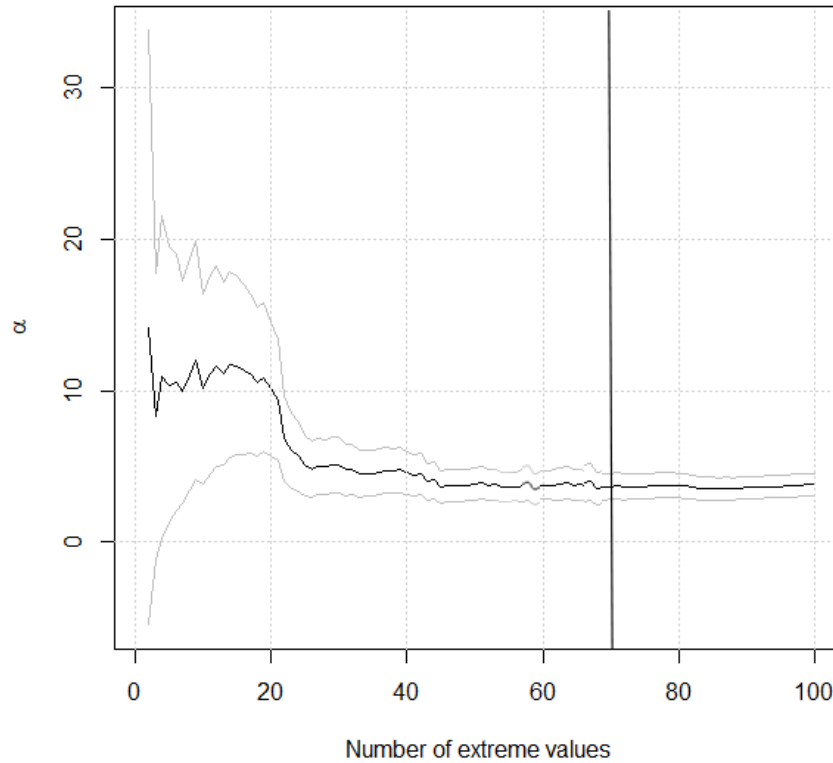


Source: Established by the author using R software

The threshold selection is made from the point where the curve $e(u)$ appears to have a linear behavior. According to the figure, the curve is almost linear when the threshold u is around 70 mm.

To confirm that this threshold represents the beginning of extreme values, we also use the Hill plot. The choice of the threshold is made by observing the stability of the tail index value. The number of exceedances corresponding to the threshold, and is determined by observing the stability region on the graph.

Figure 15: Hill Plot



Source: Established by the author using R software

The number of observations to retain corresponds to the start of the stability region, which is indeed around 70 mm. Therefore, we retain this threshold for our extreme event modeling.

3.1.4 Index modeling

An index is defined as the exceedance of the threshold for aggregated rainfall over 2 days. Therefore, our primary approach will be to model these threshold exceedances using the POT method, which employs the Generalized Pareto Distribution (GPD). We will be studying the distribution of $(X|X > 70 \text{ mm})$ which is the threshold used for modeling and then derive $(X|X > 100 \text{ mm})$ that will be used for pricing. Below is a summary table of the estimated parameters.

Table 10: Results of GPD estimation

Threshold	Scale parameter β	Shape parameter ξ
70	29.3632080	-0.2701485

Source: Established by the author using R software

The shape parameter in this case is $-0.2701 < 0$, For $\xi < 0$, the GPD has a finite upper bound. This means there is a maximum value, beyond which the probability of exceeding this value is zero. This upper bound is calculated as follows:

$$X_{cap} = u - \frac{\beta}{\xi}$$

According to the current estimations, the maximum value of the two-day aggregated rainfall is projected to be 178.69 mm. However, this hypothesis is contradicted by historical flood events extending beyond our study period. For instance, on December 29, 1984, exceptional rainfall exceeding 250 mm over four days and 195 mm in a single day occurred, leading to severe flooding (Lahlah, 2004). It is also crucial to note that El Tarf has a Mediterranean climate, and similar regions in Algeria with the same climate have experienced even higher precipitation levels. For example, from March 28 to 31, 1974, Tizi Ouzou (approximately 360 km from El Tarf) recorded 688 mm over four days and 381 mm in a single day (Lahlah, 2004). Consequently, capping the maximum precipitation at 178.69 mm may underestimate the potential risk, leading to insufficient insurance coverage. While the GPD model may fit statistically, it does not adequately reflect historical records and extreme events. Therefore, we will explore the Generalized Extreme Value (GEV) distribution as an alternative modeling approach.

The results of the modeling are as follows:

Table 11: Results of GEV estimation

	Location	scale	shape
Estimate	83.5282	12.6840	0.1781
Standard errors	2.2687	1.8612	0.1881

Source: Established by the author using R software

The initial analysis suggests that the two-day cumulative precipitation may follow a Fréchet distribution. To validate this hypothesis, we need to test whether the shape parameter ξ is significantly different from zero. We compute the student's t-statistic for ξ as follows:

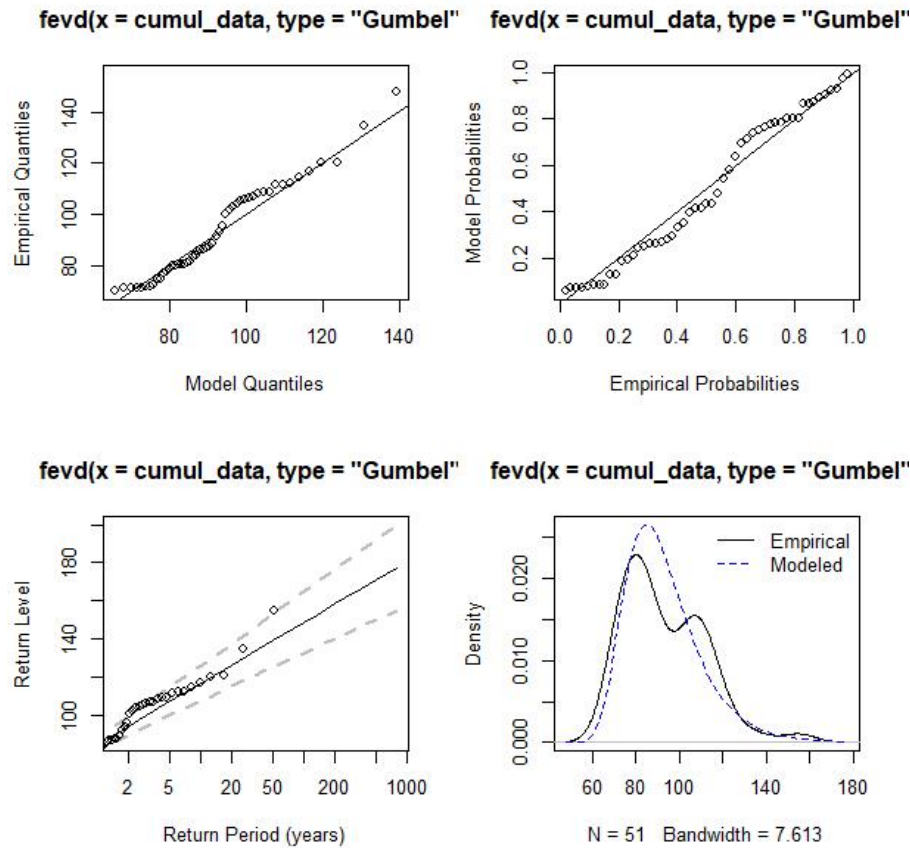
$$T - Statistic = \frac{SE(\hat{\xi})}{\hat{\xi}} = \frac{0.1781}{0.1881} = 0.9468$$

$$P\text{-value} = 0.34846$$

- Null Hypothesis (H0): The shape parameter ξ is equal to zero.
- Alternative Hypothesis (H1): The shape parameter ξ is not equal to zero.

For a student's t-test at the 5% significance level, we fail to reject the null hypothesis that ξ equals zero. This result implies that the underlying Generalized Extreme Value (GEV) distribution can be identified as a Gumbel distribution, which is characterized by a lighter tail compared to the Fréchet distribution.

Figure 16: Goodness of Fit of the Data to the Gumbel Distribution



Source: Established by the author using R software

The Q-Q plot compares the empirical quantiles of the data with the theoretical quantiles of the fitted Gumbel distribution. While The P-P plot compares the empirical probabilities with the model probabilities of the fitted Gumbel distribution. In both Q-Q plot and P-P plot the points lie close to the diagonal line, indicating that the empirical probability and quantiles closely match the theoretical one. the diagnostic plots suggest that the Gumbel distribution is a good fit for the 2 days cumulative precipitation data.

To further confirm the goodness of fit of the Gumbel distribution to the two-day cumulative precipitation data, we employed the statistical test of the Anderson-Darling.

The Anderson-Darling test is a statistical methodology utilized to assess whether a sample of data conforms to a specified probability distribution. An extension of the Kolmogorov- Smirnov (KS) test, the Anderson-Darling test places greater emphasis on the tails of the distribution, thereby

providing a more sensitive measure of goodness-of-fit in the extremities. This feature renders the Anderson-Darling test particularly advantageous for applications in extreme value theory. According to (Abidin. N.Z, 2012) it is considered as the most powerful Goodness Of Fit test for Gumbel distribution.

$$AD = - \sum \frac{2i - 1}{n} \{ \log[F(x_i)] + \log[1 - f(x_{n+1} - i)] \} - n$$

1. Null Hypothesis (H0): The sample data follows Gumbel distribution.
2. Alternative Hypothesis (H1): The sample data does not follow the Gumbel distribution.

Table 12: Anderson-Darling goodness of fit statistic test

Test Statistic:	0.965
P-Value:	0.376

Source: Established by the author

The p-value of 0.376 is greater than the significance level of 0.05. Therefore, we do not reject the null hypothesis, the Gumbel distribution is a reasonable fit for the two-day cumulative precipitation data according to the Anderson-Darling test.

3.2 The pricing of the index-based flood insurance

Having modeled both the hazard and vulnerability, our next step is to integrate these components to estimate potential damages and accurately calculate the pure premium.

3.2.1 principle of the Pure Premium's calculation

Mathematically, the pure premium is defined as the expected value of the indemnities that will be paid out. The main challenge of the pure premium is to estimate the expected losses, in other words, how much the insurer expects to pay:

$$\text{Pure Premium} = E[S]$$

Where S is the random variable to be estimated: “total claims expense for the period”. The collective approach provides a simple theoretical framework for determining the distribution of the random variable S . This approach involves considering the portfolio as a whole. The total loss load at the portfolio level is then expressed as:

$$S = \begin{cases} \sum_{i=1}^N X_i & \text{if } N > 0 \\ 0 & \text{if } N = 0 \end{cases}$$

Where N is the random variable representing the number of claims reported over a given period and X_i is the random variable representing the cost of the i^{th} claim ($i=1, 2, \dots, N$).

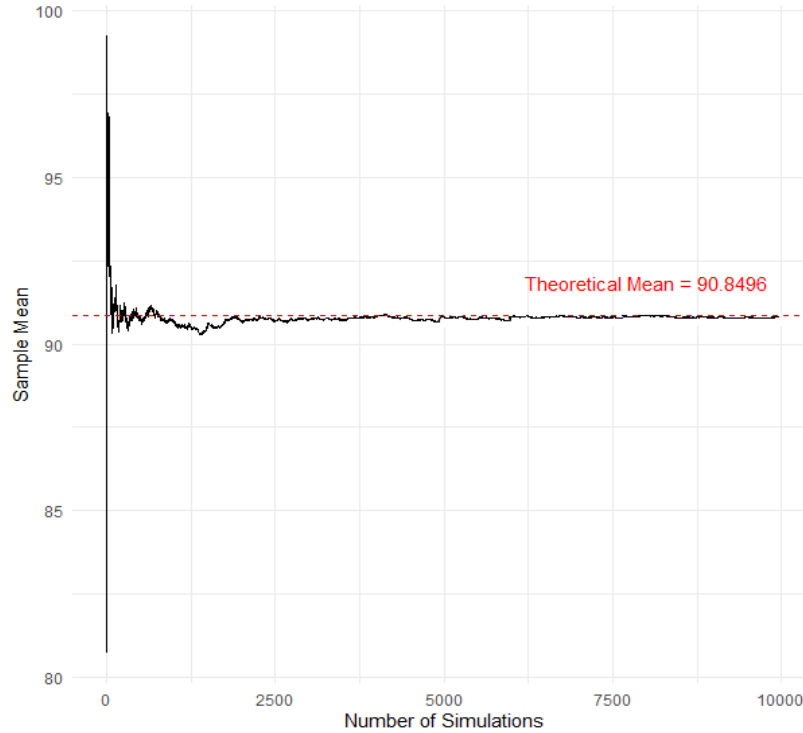
To facilitate the calculation of the pure premium, we assume that the occurrence of events is independent of the associated payments ($\forall i, N \perp X_i$). Furthermore, we posit that the payments for each event are independent and identically distributed (i.i.d.). Under these conditions, the pure premium can be expressed as:

$$\begin{aligned} \text{Primium} &= E \left[\sum_{n=1}^N \text{paiement}_i \right] \\ &= E \left\{ \left[E \left(\sum_{i=1}^n (\text{paiement}_i \mid N) \right) \right] \right\} \\ &= E(N \times \text{paiement}) \\ &= E(N) \times E(\text{paiement}) \\ &= E(N) \times E[f(X \mid X > 100)] \end{aligned}$$

Where X denotes the cumulative precipitation over 2 days, and $E(N)$ represents the annual frequency of flood events. In the previous section, the distribution of the random variable X was estimated. However, to stabilize the pricing model and achieve accurate estimates of $E(X)$, we will utilize the Law of Large Numbers through a Monte Carlo simulation.

The following figure illustrates the convergence of the Gumbel distribution with estimated parameters towards the mean as a function of the number of simulations. It is observed that stability improves as the number of simulations increases.

Figure 17: Convergence of the sample mean to the theoretical mean



Source: Established by the author using R software

We will focus on the Monte Carlo simulation technique through the inversion of the cumulative distribution function (CDF). For a Gumbel distribution random variable, the distribution function is given by:

$$G(x) = e^{-e^{-(\beta x - \mu)}}$$

The inverse distribution function is expressed as:

$$G^{-1}(y) = \mu - \beta \ln(-\ln y)$$

We can generate random variables that follow the Gumbel distribution. For a uniformly distributed random variable v on $[0,1]$ the Gumbel-distributed random variable X can be expressed as:

$$X = \mu - \beta \ln(-\ln(v))$$

This formula enables the generation of random variables following the Gumbel distribution. The modeling of extreme precipitation events, where X denotes the aggregated precipitation over 2 days.

3.2.2 pure premium calculation for the Wilaya of El Tarf

The pure premium required for the region is equal to the expected value of the indemnities paid out. We have at our disposal a damage function that links the index to the losses, thereby estimating the probable amounts to be paid for each simulated level of precipitation.

The distribution of $(X|X>100)$ will be estimated using the parameters of the already modeled Gumbel distribution, with the modeling threshold selected at $u=70$. To obtain a stable model, we will simulate the variable $(X|X>100)$ 100,000 times. The conditional CDF of X given $X>100$ is:

$$F_{X|X>100}(x) = \frac{S(x)}{S(100)} = \frac{1 - \exp(-\exp(-\frac{x-\mu}{\beta}))}{1 - \exp(-\exp(-\frac{100-\mu}{\beta}))}$$

Since our dataset for the Wilaya of El Tarf is stationary indicating the absence of a significant trend in our study zone, $E(N)$ is estimated by the annual average of historical events in the considered Wilaya. The pure premium for a province is calculated using the formula:

$$S = E(N) \times E[(X | X > 100)]$$

We have defined the attachment point of the index at a damage rate of 20%, which allows for the coverage of events we would seek to insure. The coverage limit is set at the total cereal production for the Wilaya of El Tarf, according to the most recent figures from the Ministry of Agriculture available.

The table below summarizes the modeling results:

Table 13: The pure premium for a 20% destruction rate

Destruction rate	Coverage (M DZD)	Premium (M DZD)	Premium rate	Premium per Ha (DZD/Ha)	
20%	3,061.083	191.605	6.26%	Durum wheat	10003.1499
				Soft wheat	10640.9943
				Barley	3670.35898

Source: Established by the author

The high pure premium rate of 6.26% reflects the flood risk in the wilaya of El Tarf. Indeed, the wilaya of El Tarf receives an average of 650 mm of annual precipitation, whereas the average for the Mediterranean climate is 600 mm. Additionally, as we observed in the vulnerability study, most crop land is located in catchment areas and in close proximity to rivers.

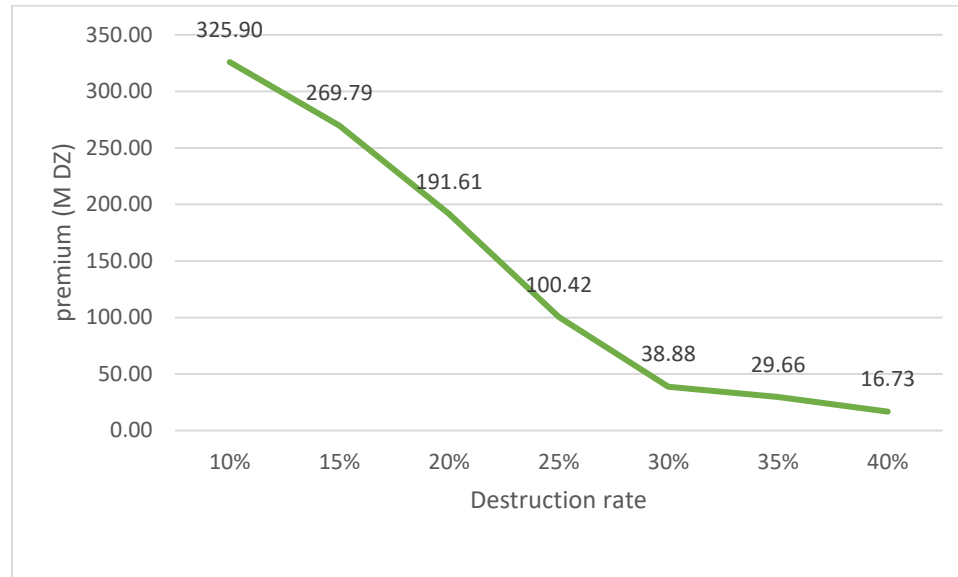
We can offer other coverage options for farmers by adjusting the attachment to align with the risk profile in the area. Increasing the attachment lowers the rate, and vice versa, lowering it increases the pure premium.

Table 14: Sensitivity of the pure premium to the choice of the attachment point

Destruction rate	Premium (DZD)	Premium rate
10%	325,903,871.05	10.65%
15%	269,790,349.05	8.81%
20%	191,605,700.14	6.26%
25%	100,421,351.58	3.28%
30%	38,878,600.26	1.27%
35%	29,655,321.37	0.97%
40%	16,734,099.44	0.55%

Source: Established by the author

Figure 18: Sensitivity of the pure premium to the choice of the attachment point



Source: Established by the author

The table and graph illustrate the impact of increasing the attachment point on the reduction of the premium. While there is an opportunity to pay a lower premium, this comes at the cost of covering fewer events. The reduction in the premium is particularly significant for damage rates exceeding 30%, reflecting the light tail characteristic of the Gumbel distribution utilized in the index modeling.

3.3 Necessity for financing mechanisms and a private public partnership

To ensure the entire winter cereal production in the Wilaya of El Tarf, the required annual premium for coverage spanning from September to July is 191.605 million Algerian dinars, corresponding to a premium rate of 6.26% for a 20% destruction rate. If the coverage is extended to include a 10% destruction rate, the annual premium increases to 325.903 million Algerian dinars, with a premium rate of 10.65%. This premium level may be prohibitive for smallholder farmers. As highlighted in the first chapter, index insurance is recognized as a potentially effective mechanism for providing affordable protection against catastrophic, correlated risks, particularly for low-income households in developing countries.

Adopting an index-based insurance approach allows for a reduction in the adjustment costs embedded in the premium, which, based on reports from the MADR, account for 3% of the total damages. Moreover, the financial strain on farmers could be alleviated by implementing effective financing mechanisms.

The most successful index insurance programs, such as the Modified National Agricultural Insurance Scheme (mNAIS) in India and the Index-Based Livestock Insurance Program (IBLIP) in Mongolia, have typically been established through public-private partnerships, with significant involvement and contributions from the public sector (Clarke & Wren-Lewis, 2013). It is widely observed that governmental participation in agricultural insurance schemes is prevalent, with many governments offering formal or informal safety nets to farmers in the event of catastrophic events.

According to a communication from the CNA during the international seminar on agricultural insurance, the Algerian government allocates, on average, one billion DA annually for indemnification purposes. However, the compensation distributed is contingent upon the fund's capacity and may not suffice to cover the totality of incurred damages.

Government intervention, particularly through the provision of premium subsidies for the index-based flood insurance, serves to mitigate the financial burden on insured farmers and provides the government with the strategic advantage of predetermining budget allocations for agricultural disaster relief.

Moreover, the establishment of index insurance in Algeria may require partnerships with various public authorities and institutions, such as the ONM, the ANRH, and ASAL. The data necessary for constructing an index can be costly to obtain from the public authorities that own it. Additionally, the creation of public-private partnerships (PPP) with statistical institutes and meteorological stations can partially address the issue of the cost and improve the precision of the index which will allow the reduction of the basic risk.

CONCLUSION

Through spatial modeling of vulnerability using GLM and modeling of extreme precipitation, we constructed a destruction function and determined the premium required for transferring agricultural flood risk through an index-based solution. To achieve this, we made two key assumptions: first, the homogeneity of precipitation within the study area, and second, the homogeneity of cereal production across the crop lands.

These assumptions can be addressed through potential partnerships with the ANRH and the ONM for precipitation data, as both agencies have numerous stations across each region. Additionally, collaboration with agricultural chambers that maintain detailed records of production by individual farmers could provide more accurate data.

While our empirical analysis successfully validated our models statistically, the lack of historical data limited our ability to compare results with multiple events. Therefore, it is essential to consider these limitations and uncertainties when using these results for decision-making purposes.

GENERAL CONCLUSION

GENERAL CONCLUSION

Algeria's climate is marked by significant variability, with precipitation patterns exhibiting considerable fluctuations and irregularities. This variability is especially pronounced in the Mediterranean climate zone, known for intense precipitation events that account for approximately 20% of total annual rainfall. With climate change, Algeria is anticipated to be highly vulnerable to extreme weather events, particularly affecting the agriculture sector, which is crucial to the national economy and the food security.

Currently, agricultural flood risks in Algeria are considered non-insurable. The existing system, managed by the FGCA, provides limited coverage and faces substantial management challenges. In addition, traditional insurance struggle with issues such as information asymmetry, high costs, and prolonged indemnity periods, which can disadvantage smallholder farmers.

Index-based insurance represents an effective alternative, characterized by an expedited indemnity process and reduced administrative expenses. In this model, payouts are initiated when cumulative precipitation, as recorded by meteorological stations or satellite data, surpasses a specified threshold. In our study, we set this threshold at 100 mm for the cumulative precipitation over two days.

The study was limited to cereal production in the Wilaya of El Tarf and employed a probabilistic model to simulate potential losses due to agricultural flood risk. By constructing an index based on extreme precipitation data, supplemented with satellite imagery and data to assess regional vulnerability, we propose an index-based insurance solution that links extreme precipitation events to cereal crop damage.

However, several limitations must be addressed. The absence of historical damage data necessitated the use of probabilistic models and simulations, which, while insightful, may not fully capture the complexities of flood-related damages. Additionally, reliance on satellite data introduces potential accuracy issues and discrepancies with ground-level conditions. Despite the statistical significance of the models, the lack of data limits the ability to correlate simulations with real-world outcomes.

To enhance the effectiveness of index-based flood insurance in agriculture, the following areas require attention:

- **Data Collection and Integration:** Establishing a comprehensive database on agricultural flood risk and statistics is crucial for rational actuarial calculations and premium setting.
- **Public-Private Partnerships:** Strengthening collaborations between government agencies, private insurers, and international organizations can support the development of sustainable and accessible insurance schemes for smallholder farmers that minimize the basic risk.

While this research lays a groundwork for index-based flood insurance, ongoing refinement and adaptation are necessary to fully leverage its potential. The recommendations outlined here offer a roadmap for creating a more resilient agricultural sector amidst an increasingly volatile climate. index-based insurance can become a viable tool for managing agricultural flood risk in Algeria, ultimately contributing to the broader goal of climate resilience in the region.

BIBLIOGRAPHY

1. Abidin. N.Z, M. ., (2012). The Goodness-Of-Fit Test For Gumbel Distrubution: A Comparative Study. *Matematika*, 18, 35-48.
2. Akinwande, M., Dikko, H., & Samson, A. (2015). Variance Inflation Factor: As a Condition for the Inclusion of Suppressor Variable(s) in Regression Analysis. *Open Journal of Statistics*, 15(7), 754-767. doi: 10.4236/ojs.2015.57075.
3. Allianz:re. (2023, march 9). *Agriculture Insurance in Asia:Protection gap and embracing technology*. Retrieved 4 2024, 21, from https://www.allianzre.com/content/dam/onemarketing/azre/allianzre/documents/Agriculture_Insurance_in_Asia.pdf.
4. Amarnath, G., Singh, M., & Avinandan, T. (2021). *Scaling up Index-based Flood Insurance (IBFI) for Agricultural Resilience and flood-proofing Livelihoods in Developing Countries*. Colombo: International Water Management Institute (IWMI). doi: <https://doi.org/10.5337/2021.213>.
5. Amarnath, G., Taron, A., Alahacoon, N., & Ghosh, S. (2023, December 1). Bundled Climate-smart Agricultural Solutions for Smallholder Farmers in Sri Lanka. *Frontiers in Sustainable Food Systems*, 7. doi:<https://doi.org/10.3389/fsufs.2023.1145147>.
6. Androulidakis, Y., Kombiadou, K., Makris, C., & Baltikas, V. (2015). Storm Surges in the Mediterranean Sea: Variability and Trends under Future Climatic Conditions. *Dynamics of Atmospheres and Oceans*, 71, 56-82.
7. Annibale, V., Martina, M., Monteiro, R., & Galasso, C. (2021). Precipitation-based Extreme Event Detection for Flood Index Insurance. *International Journal of Disaster Risk Reduction*, 55, 102-108. doi:10.1016/j.ijdr.2021.102108.
8. Antolini, F., Tate, E., Dalzell, B., & Young, N. (2019). Flood Risk Reduction from Agricultural Best Management Practices. *JAWRA Journal of the American Water Resources Association*, 56(1), 161-179. doi:10.1111/1752-1688.12812.
9. Barlow, M. G. (2019). North American Extreme Precipitation Events And Related Large-scale Meteorological Patterns: A Review Of Statistical Methods, Dynamics, Modeling, And Trends. *Clim Dyn*, 53, 6835–6875. doi:<https://doi.org/10.1007/s00382-019-04958-z>
10. Basri, h., Syakur, H., & Azmeri, A. (2022). Floods And Their Problems: Land Uses And Soil. In I. Publishing (Ed.), *IOP Conference Series: Earth and Invironvemental Science*, 951. doi:10.1088/1755-1315/951/1/012111.
11. Bayeux, A., & Truy, A. (2014). Modelisation d'un systeme d'alert pour le risque d'inodation . ENSA paristech.

12. Bellout, A., Vas, E., & Penfound, E. (2020). Rethinking Agricultural Land Use in Algiers: a Spatial Analysis of the Eastern Mitidja Plain. *Habitat International*, 104, 102-239. doi:10.1016/j.habitat.2020.102239.
13. Bessaklia, H., Serrano-Notivoli, R., & Nekkache Ghenim, A. (2021). Extreme Precipitation Trends In Northeast Algeria Using A Highresolution Gridded Daily Dataset. *International Journal of Climatology*.
14. Bessaoud, O., Pellissier, J., Rolland, P., & Khechimi, W. (2019, Mai 28). Rapport de synthèse sur l'agriculture en Algérie. *CIHEAM-IAMM*, p.
15. Bouamrane, A., Derdous, O., Dahri, N., Tachi, S.-E., & Boutebba, K. (2020). A Comparison of the Analytical Hierarchy Process and the Fuzzy Logic Approach for Flood Susceptibility Mapping in a semi-arid Ungauged Basin (Biskra basin: Algeria). *International Journal of River Basin Management*, 1-11. doi:10.1080/15715124.2020.1830786.
16. Boutaghane, H., Boulmaiz, T., Lameche, E., Lefkir, A., Hasbaia, M., & Abdelbaki, C. (2022). Flood Analysis and Mitigation Strategies in Algeria. *Natural Disaster Science and Mitigation Engineering: DPRI reports*, 95-118. doi:10.1007/978-981-16-2904-4_3.
17. Brémond, P., Grelot, F., & Agenais, A. (2013). Review Article: Economic Evaluation of Flood Damage to Agriculture – Review and Analysis of Existing Methods. *Natural Hazards and Earth System Sciences*, 13(10), 2493-2512. doi:10.5194/nhess-13-2493-2013
18. Brunner, M., Slater, L., Tallaksen, L., & Clark, M. (2021). Challenges in Modeling and Predicting Floods and droughts: a Review. *WIREs Water*, 8(3). doi:10.1002/wat2.1520.
19. Buddha, S., & Binu, D. (2023). Flood Hazard Mapping Using a Multi-Criteria Decision Analysis Approach Over the Indrawati River Basin. *Interdisciplinary Journal of Atmospheric and Hydrospheric Sciences*, 2(2), 25–42. doi:https://doi.org/10.3126/jalawaayu.v3i1.52058.
20. Carter, M., de Janvry, A., Sadoulet, E., & Sarris, A. (2015). Assurance climatique indicielle pour les pays en développement : examen des faits et propositions visant à augmenter le taux de souscription. *Revue d'économie du développement*, 23(1), 5-57. doi:10.3917/edd.291.0005.
21. Carter, M., De Janvry, A., Sadoulet, E., & Sarris, A. (2017). Index Insurance for Developing Country Agriculture: a Reassessment. *Annual Review of Resource Economics*, 9, 421-438.
22. Clarke, D. (2016). A Theory of Rational Demand for Index Insurance. *American Economic Journal: Microeconomics*, 8(1), 283-306. doi:10.1257/mic.20140103.

23. Clarke, D., & Wren-Lewis, L. (2013). The Role of Government in Index Insurance for Individuals. *FRENDI policy brief*, 70. doi: <http://dx.doi.org/10.2139/ssrn.2659210>.
24. Clyde&co. (2018). *Parametric Insurance: Closing the Protection Gap - Legal Considerations*. Retrieved 4 20, 2024, from https://www.unisdr.org/preventionweb/files/56544_clydecoresilienceparametricinsuranc.pdf.
25. Coles, S. (2001). Extremes of Dependent Sequences. Dans Springer (Éd.), *An Introduction to Statistical Modeling of Extreme Values* (pp. 340-363). London: Springer Series in Statistics. doi: https://doi.org/10.1007/978-1-4471-3675-0_5.
26. Deraman, W., Abd Mutalib, N., & Mukhtar, N. (2017, 8 1). Determination of Return Period for Flood Frequency Analysis Using Normal and Related Distributions. (I. P. Ltd, Ed.) *Journal of Physics: Conference Series*, 890(1). doi:10.1088/1742-6596/890/1/012162
27. Garaix, J.-C. (2020). Closing The Protection Gap With Parametrics. *AGRICULTURE & WEATHER*, PUP1006-11-19. Retrieved 4 24, 2024 from <https://www.libertymutualre.com/static/2020-11/Closing-the-gap.pdf>.
28. Giné, X., & Yang, D. (2009). Insurance, Credit, And Technology Adoption: Field Experimental Evidence from Malawi. *Journal of Development Economics*, 89(1). doi:10.1016/j.jdevco.2008.09.007.
29. Goel, s., & Garg, A. (2019). Infiltration—A Critical Review. *Sustainable Engineering. Lecture Notes in Civil Engineering*, (pp. 111–120). Singapore: Springer. doi: https://doi.org/10.1007/978-981-13-6717-5_11.
30. Hafnaoui, M., Boultif, M., & Dabanlı, I. (2023, December). Floods In Algeria: Analyzes And Statistics. *Larhyss Journal*, 1112-3680(56), 351-369.
31. Hoeppe, P. (2016). Trends In Weather Related Disasters – Consequences For Insurers And Society. *Weather and Climate Extremes*, 11, 70-79. doi:10.1016/j.wace.2015.10.002.
32. Holzheu, T., & Turner, G. (2017). The Natural Catastrophe Protection Gap: Measurement, Root Causes and Ways of Addressing Underinsurance for Extreme Events. *The Geneva Papers on Risk and Insurance - Issues and Practice*, 43(1), 37-71.
33. Idir, S. (2021, Juin). Etude d'un produit d'assurance paramétrique contre le risque d'inondation au Maroc. *PFE N° 19*. INSTITUT NATIONAL DE STATISTIQUE ET D'ECONOMIE APPLIQUEE, Maroc.
34. Kast, R., & Lapied, A. (2012). *Economics and Finance of Risk and of the Future*. (1, Ed.) John Wiley & Sons Ltd. doi:DOI:10.1002/9781119205951.

35. Kouli, t. (2018). étude relative à la constryction et au parametrage du modèle actuariel nécessaire à la conception et à l'exploitation d'une assurance paramétrique,mémoire pour l'optention du diplome du mater actuariat et l'admission à l'institut des actuaires ,paris doph. groupe Finactu.
36. Kousky, C. (2011, 6 7). Managing Natural Catastrophe Risk: State Insurance Programs in the United States. *Review of Environmental Economics and Policy*, 5(1), 153-171. doi:10.1093/reep/req020.
37. Lahlah, S. (2004). Les inondations en Algerie:Actes des journees techniques/Risques naturels:inondation,prevision,protection. Batna: Le CRSTRA.
38. Lamond, J., & Penning-Rowsell, E. (2014). The Robustness of Flood Insurance Regimes Given Changing Risk Resulting from Climate Change. *Climate Risk Management*, 2, 1-10. doi:10.1016/j.crm.2014.03.001.
39. Leblois, A., & Quirion, P. (2013). *Les assurances agricoles basées sur des indices météorologiques : bilan des connaissances et agenda de recherche*. Paris: Peter Lang. Récupéré sur <https://enpc.hal.science/hal-00801267>.
40. Lin , X., & Kwon, W. (2020). Application of Parametric Insurance in Principle-compliant and Innovative Ways. *Risk Management and Insurance Review*, 23(2), 121-150. doi:10.1111/rmir.12146.
41. Lionello, P., Bhend , J., & Buzzi, A. (2006). *Cyclones In The Mediterranean Region: Climatology And Effects On The Environment*. Developments in Earth and Environmental Sciences.
42. Maref, N. (2019). Démarche méthodologique d'adaptation d'un système de prévision aux risques d'inondation en Algérie, cas du bassin versant de l'Oued Mekerra (N.W Algérien),THÈSE Pour l'obtention du diplôme de Doctorat en Hydraulique. Tlemcen: UNIVERSITE ABOU BEKR BELKAID TLEMCEN.
43. McMahan, L. (2019, 1 7). *Modeling Flood Risks: Opportunities and Challenges*. Retrieved 8 14, 2024, from actuarial review: <https://ar.casact.org/modeling-flood-risks-opportunities-and-challenges/>.
44. Mdejerab, A., & Henia, L. (2014, Mai 12). VARIATIONS SPATIO-TEMPORELLES DE LA SECHERESSE. *Courrier du Savoir ;N°11,,* pp. 71-79.
45. Megahed, H., Abdo.M, A., Abdelrahman.A.E, M., Scopa, A., & Hegazy.N, M. (2023). Frequency Ratio Model as Tools for Flood Susceptibility Mapping in Urbanized Areas: A Case Study from Egypt. *Appl. Sci*, 13(9445). doi: <https://doi.org/10.3390/app13169445>

46. Merz, B., Blöschl, G., Vorogushyn, S., & Dottori, F. (2021, 8 10). Causes, Impacts and Patterns of Disastrous River Floods. *Nature Reviews Earth & Environment*, 2, pages592–609. doi:<https://doi.org/10.1038/s43017-021->
47. Messner, F., Penning-Rowsell, E., Green, c., & meyer, V. (2007). *Evaluating Flood Damages: Guidance And Recommendations On Principles And Methods*. FLOOD Site Project Report. Retrieved from https://www.floodsite.net/html/partner_area/project_docs/T09_06_01_Flood_damage_guidelines_D9_1_v2_2_p44.pdf.
48. Miles, J. (2005). Tolerance and Variance Inflation Factor. In B. S. Howell (Ed.), *Encyclopedia of Statistics in Behavioral Science* (Vol. Volume 4, pp. pp. 2055–2056). John Wiley & Sons, Ltd, Chichester. Retrieved from ISBN-13: 978-0-470-86080-9.
49. Mohand, B. (2019). CLIMATOLOGIE. Setif , algerie: Université Ferhat Abbas, Sétif 1 ,IAST-UFAS.
50. Morante-Carballo, F., Montalván-Burbano, N., Arias-Hidalgo , M., & Domínguez-Granda , L. (2022). Flood Models: An Exploratory Analysis and Research Trends. *Water*, 14(16)(2488). doi: <https://doi.org/10.3390/w14162488>.
51. Pham, B., Avand, M., Janizadeh, S., & Phong, T. (2020, 3 2). GIS Based Hybrid Computational Approaches for Flash Flood Susceptibility Assessment. *Water*, 12(3), 683. doi:10.3390/w12030683.
52. Pujari, S, S., & Wayal, A, S. (2023). Mitigation of Flooding in Low-Lying Area with Development of Storage–Discharge-Frequency Curve. *J. Inst. Eng. India Ser*, 104, 643–652. doi:<https://doi.org/10.1007/s40030-023-00731-y>.
53. Rawls, W, J., Ahuja,L, R., Brakensiek,D, L., & Shirmohammadi, A. (1993). *Handbook Of Hydrology: Infiltration And Soil Water Movement*. New York: McGraw-Hill.
54. Rentschler, J., & Salhab, M. (2020, 11 12). *1.47 billion people face flood risk worldwide: for over a third, it could be devastating*. Retrieved 8 17, 2024, from Wold bank: <https://blogs.worldbank.org/en/climatechange/147-billion-people-face-flood-risk-worldwide-over-third-it-could-be-devastating>.
55. Ritling, F., & Nguyen, C. (2014, jun). Etude d’un produit d’assurance paramétrique contre le risque de pluie torrentielle en Jamaïque. *Mémoire d’actuariat*. ’ENSAE Paristech,filière Actuariat, France.
56. Sahnoune, F., Belhamel, M., Zemat, M., & Kerbach, R. (2013). Climate Change in Algeria: Vulnerability and Strategy of Mitigation and Adaptation. *Energy Procedia*, 1286-1294.

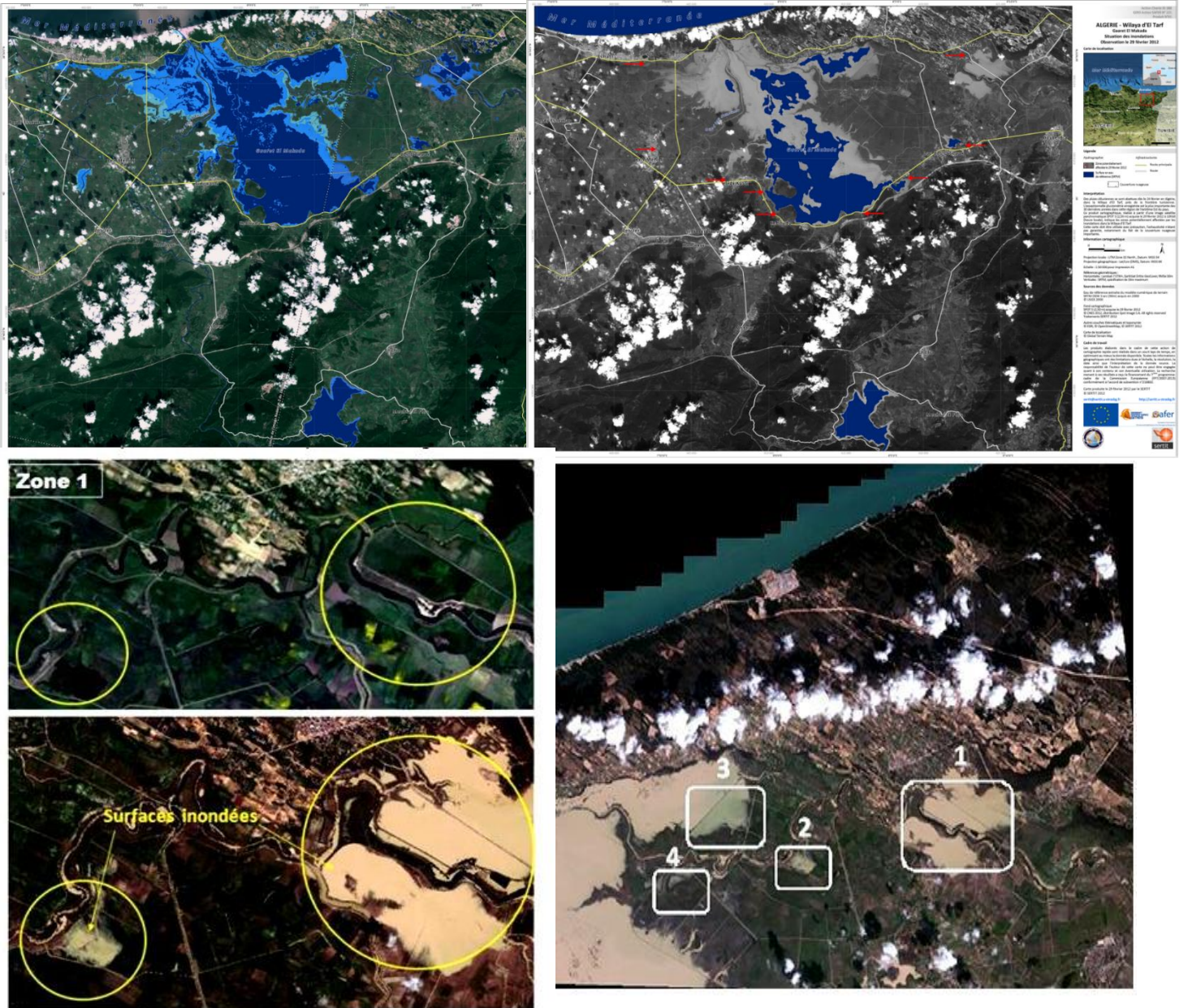
57. Sarkar, D., & Monda, P. (2020). Flood vulnerability mapping using frequency ratio (FR) model: : a case study on Kulik river basin, Indo-Bangladesh Barind region. (Springer, Ed.) *Applied Water Science*, 10(17). doi:<https://doi.org/10.1007/s13201-019-1102-x>.
58. Schanz, K.-U. (2018). *Understanding and Addressing Global Insurance Protection Gaps*. Zurich: The Geneva Association (The International Association for the Study of Insurance Economics). Retrieved from https://www.genevaassociation.org/sites/default/files/research-topics-document-type/pdf_public/understanding_and_addressing_global_insurance_protection_gaps.pdf.
59. Schmit, J. (1986). A New View of the Requisites of Insurability. *The Journal of Risk and Insurance*, 53(2), 320. doi:10.2307/252380.
60. Schwarz, I., & Kuleshov, Y. (2022,, 4). Flood Vulnerability Assessment and Mapping: A Case Study for Australia's Hawkesbury-Nepean Catchment. (J. Choi, Ed.) *Remote Sens*, 14(4894). doi:<https://doi.org/10.3390/>.
61. Siyuan , y., Wenhua, W., & Arvid , N. (2023). Extreme Value Prediction With Modified Enhanced Monte Carlo Method Based On Tail Index Correction,. *Journal of Sea Research*, 192. doi:<https://doi.org/10.1016/j.seares.2023.102354>.
62. Sumi, T., Kantoush, S., & Saber, M. (2022). *Wadi Flash Floods* (éd. 1). Springer Singapore.
63. Surjanovic, N., & Loughin, T. (2023). Improving The Hosmer-Lemeshow Goodness-Of-Fit Test In Large Models With Replicated Bernoulli Trials. *Journal of Applied Statistics*, 51(7), 1399–1411. doi:<https://doi.org/10.1080/02664763.2023.2272223>.
64. Swiss-re. (2012). Flood In Switzerland: An Underestimated Risk. Switzerland: Swiss Reinsurance Company Ltd. (Swiss. Retrieved from https://www.planat.ch/fileadmin/PLANAT/planat_pdf/alle_2012/2011-2015/Swiss_Re__Hg__2012_-_Floods_in_Switzerland.pdf.
65. Tabet Helal, M., Baghli, A., Bensaoula, F., Mami, E.-F., & Ghellaï, N. (2012). La crue centenaire de Bab EL Oued, Alger (Algérie). *Journées de l'hydraulique*, 34, 1-8. Récupéré sur https://www.persee.fr/doc/jhydr_0000-0001_2012_act_34_1_1266.
66. Thomas, M., Tellman, E., Osgood, D., DeVries, B., & Akm , S. (2023, 2 10). A Framework to Assess Remote Sensing Algorithms for Satellite-Based Flood Index Insurance. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 16, 2589 - 2604. doi:10.1109/JSTARS.2023.3244098.
67. Towler, E., Llewellyn, D., Prein, A., & Gilleland, E. (2020). Extreme-value analysis for the characterization of extremes in water resources: A generalized workflow and case study

on New Mexico monsoon precipitation,. *sciencedirect*, Volume 29.
doi:<https://doi.org/10.1016/j.wace.2020.100260>.

68. Uy, S. (2019, 1 7). *Modeling Flood Risks: Opportunities and Challenges*. Retrieved from Actuarial Review Magazine: <https://ar.casact.org/modeling-flood-risks-opportunities-and-challenges/>.
69. Vasudeva , S. (2009). Catastrophic Risk And Insurance. *Management Avenir*, n° 27(7), 225-240. doi:<https://doi.org/10.3917/mav.027.0225>.
70. Vozinaki, A.-E., Karatzas, G., Sibetheros, I., & Varouchakis, E. (2015). An Agricultural Flash Flood Loss Estimation methodology: the Case Study of the Koiliaris Basin (Greece), February 2003 Flood. *Natural Hazards*, 79(2), 899-920. doi:10.1007/s11069-015-1882-8
71. Yashon , O., & Ryutaro , T. (2014). Urban Flood Vulnerability and Risk Mapping Using Integrated. *Water*(6), 1515-1545.
72. Zainal abidin, n., Mohd Bakri, A., & Habshah, M. (2012). The Goodness-of-fit Test for Gumbel Distribution:A Comparative Study. *MATEMATIKA*, 28(1), 35–48.

APPENDICES

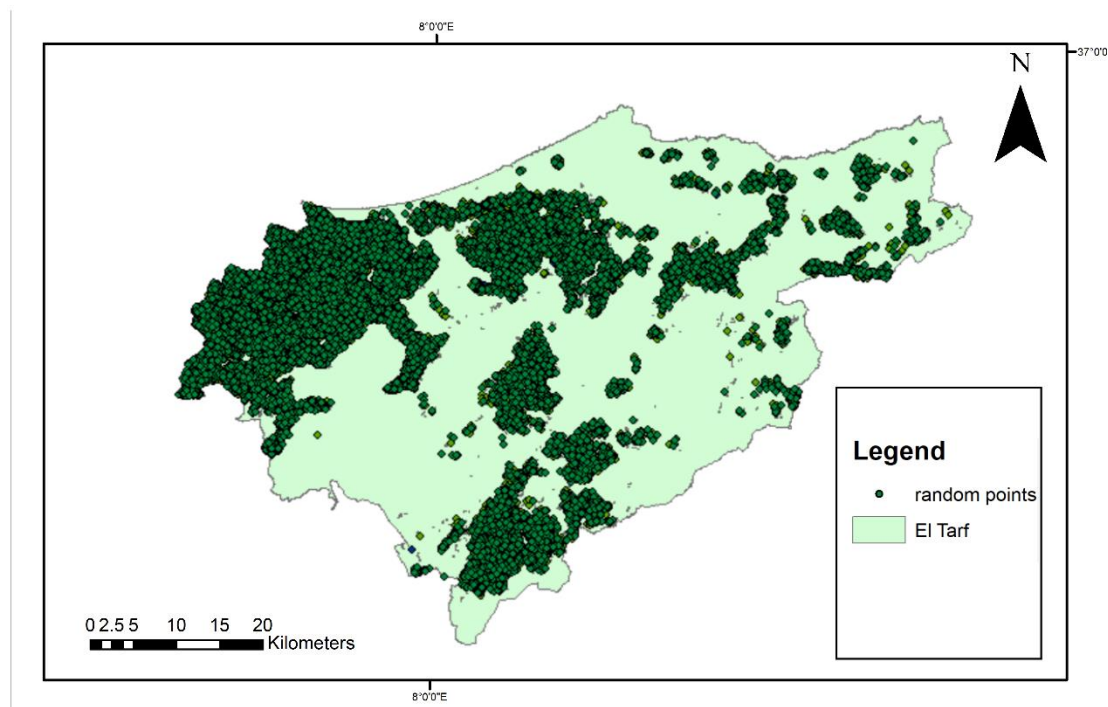
Appendix 1: Satellite images showing flood location is the Wilaya of El Tarf



Appendix 1: Satellite images showing flood location is the Wilaya of El Tar



Appendix 2: Generating 20 000 locations in crop area for the destruction function



Appendix 3: Overview of Logistic Regression:

Logistic regression is a widely used technique for modeling binary outcome data. In this context, the response variable is binary, usually represented as 1 or 0, where 1 typically denotes a success and 0 indicates a failure. In flood risk assessments, this can correspond to the presence or absence of flooding.

$$f(y_i; p) = p^{y_i}(1 - p)^{1-y_i}$$

where (y) represents the response variable being modeled, and (p) denotes the probability that (y) takes the value 1.

The Bernoulli probability density function (pdf) belongs to the exponential family of distributions. In 1972, Nelder and Wedderburn developed a unified algorithm for estimating models based on this family of distributions, known as Generalized Linear Models (GLM). This approach became a standard for estimating various types of regression models, including binary response models. The exponential distribution used in generalized linear models can be represented by the following form:

$$f(y_i; \theta_i, \varphi) = e^{\frac{(y_i\theta_i - b(\theta_i))}{\alpha(\varphi)}} + c(y_i; \varphi)$$

where θ_i represents the link function, $\alpha(\varphi)$ is the scale parameter, $b(\theta_i)$ is the cumulant function, and $c(y_i; \varphi)$ is the normalization term to ensure the probability sums to 1. For the Bernoulli distribution, which is a special case of the exponential family, the probability density function is:

$$f(y_i; \pi_i) = e^{y_i \ln\left(\frac{1-\pi_i}{\pi_i}\right) + \ln(1-\pi_i)}$$

Here, π_i is the probability of success, with the link function being $\ln\left(\frac{1-\pi_i}{\pi_i}\right)$ and the cumulant function being $\ln\left(\frac{1}{1-\pi_i}\right)$. For the Bernoulli distribution, the first and second derivatives of the cumulant are π_i and $\pi_i(1-\pi_i)$ respectively, representing the mean and variance functions.

Logistic regression offers a different approach compared to ordinary linear regression, particularly in how it models the response variable based on predictors. In general regression

analysis, the relationship between the response variable y and predictors $x_1, x_2, \dots, x_p$ is expressed as:

$$\hat{y}_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}$$

Here $\hat{y}_i$ represents the predicted value, which is also known as the linear predictor. Each term βx corresponds to a predictor x and its associated coefficient β . In linear regression, which utilizes the Gaussian (or normal) distribution, $\hat{y}_i$ is directly the output of the linear predictor. There is a straightforward linear relationship between the predicted values and the predictors in this model.

However, logistic regression differs significantly in this respect. the fitted values are derived through the link function, this function establishes the relationship between the predicted probability π and the linear predictor:

$$\beta x_i = \ln \left(\frac{1-\pi_i}{\pi_i} \right) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}$$

To derive the probability π from the linear predictor, the logistic model uses the inverse of the logit function, this function converts the linear predictor into a probability value.:

$$\pi = \frac{1}{1 + e^{-\beta x_i}}$$

TABLE OF CONTENTS

ACKNOWLEDGMENTS

DEDICATION

LIST OF ABBREVIATIONS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

GENERAL INTRODUCTION..... 1

CHAPTER ONE: THEORICAL FRAMEWORK: THE INSURANCE AGAINST AGRICULTURAL FLOOD RISK..... 4

INTRODUCTION 4

SECTION 1: UNDERSTANDING FOOD RISKS 5

1.1 Conceptual Framework of natural catastrophes 5

1.1.1 The Different Types of Risks..... 5

1.1.2 Definition of Natural catastrophes 6

1.2 Flood risk 6

1.2.1 Flood types..... 7

1.2.2 Flood risk parameters..... 8

1.2.3 Flood vulnerability 8

SECTION 2: INSURANCE AGAINST FLOOD RISK IN AGRICULTURE 10

2.1 Insurability of flood risk..... 10

2.2 Addressing agricultural flood risk with traditional indemnity insurance 12

2.2.1 Generalities about Flood insurance in agriculture 12

2.2.2 Challenges in insuring agricultural risk with a traditional indemnity insurance 13

2.3 Principles of Index based flood insurance 14

2.3.1 Definition of index-based insurance	14
2.3.2 Types of indices	15
2.3.3 Barrier for the application of index insurance	16
SECTION 3: MODELING EXTREME EVENTS	17
3.1 Extreme events modeling approaches	17
3.1.1 Deterministic Modeling	17
3.1.2 Probabilistic Modeling.....	18
3.2 Extreme value theory:	19
3.2.1 Fundamental theories	19
3.2.2 Threshold Selection	24
3.3 Monte Carlo Simulation.....	25
COCLUSION:.....	26
CHAPTER TWO: EMPIRICAL ANALYSIS: CONSTRUCTION AND PRICING OF AGRICULTTURAL FLOOD INDEX-BASED INSURANCE	28
INTRODUCTION	28
SECTION 1: ASSESSMENT OF THE SITUATION AND PRODUCT CONSTRUCTION	29
1.1 State of Affairs and Situation in Algeria	29
1.1.1 The Agricultural Landscape of Algeria	29
1.1.2 Rainfall profile of Algeria.....	31
1.1.3 Flood risk in Algeria	32
1.1.4 Coverage Against Flood Risk in Algeria	33
1.2 Presentation of the study area and the Data.....	35
1.2.1 Presentation of the study area	35
1.2.2 Precipitation Data.....	36
1.2.3 Vulnerability Data:.....	37

1.2.4 Exposure data.....	38
1.3 The construction of the Index	39
SECTION 2: MODELING FLOOD VULNERABILITY	42
2.1 Selection of the dataset	42
2.1.1 Response Variable.....	43
2.1.2 Predictor variables	44
2.2 Methodology and results	47
2.2.1 Estimation of Logistic Regression Coefficients	48
2.2.2 Model evaluation	52
2.3 Determination of destruction rate function for the Wilaya of El Tarf	54
SECTION3: MODELING OF EXTREME PRECIPITATION AND PRODUCT PRICING .	56
3.1 Modeling the extreme precipitation	56
3.1.1 Addressing the dependency.....	56
3.1.2 Addressing the stationarity of extreme events:.....	58
3.1.3 threshold determination	59
3.1.4 Index modeling.....	61
3.2 The pricing of the index-based flood insurance.....	65
3.2.1 principle of the Pure Premium’s calculation.....	65
3.2.2 pure premium calculation for the Wilaya of El Tarf.....	68
3.3 Necessity for financing mechanisms and a private public partnership.....	70
CONCLUSION.....	72
GENERAL CONCLUSION.....	74
BIBLIOGRAPHY	
APPENDICES	
TABLE OF CONTENTS	

