MILLET YIELD ESTIMATIONS IN SENEGAL: UNVEILING THE POWER OF REGIONAL WATER STRESS ANALYSIS AND ADVANCED PREDICTIVE MODELING

By

Enid Martha Banda

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Biosystems Engineering – Master of Science

2023

ABSTRACT

Crop yield is usually affected by impending weather and climate conditions, as well as human interventions like irrigation. Hence, the prompt detection of regions experiencing water scarcity can aid in implementing effective mitigation strategies. Our study utilized a data-driven approach to compute a Water-Demand-Index (WDI), which incorporates crucial first-order geophysical variables like ambient temperature, vegetation status, and soil moisture, to identify water-stressed fields in Senegal's agricultural regions during the millet planting, growing, and harvesting periods. We have also explored various scenarios for enhancing the accuracy of millet yield prediction by incorporating other drought indices, soil characteristics, and a bias correction factor. To optimize the hyperparameters of machine learning (ML) models, various techniques were utilized. Meanwhile, the performance of these ML models was evaluated using a nested cross-validation approach. The outcomes of the analysis demonstrate that the Random Forest Regressor model exhibits superior predictive performance. The outcomes of this study also indicate that integrating soil moisture-based indices generated from advanced satellite-based high-resolution soil moisture observations, accounting for individual phases of millet growth, and encompassing millet production regions at the department (administrative unit) level, can significantly enhance the overall predictive capacity of the model. The results imply that a holistic approach, encompassing diverse environmental factors and crop growth stages, could result in more precise and dependable millet yield predictions. Such refined yield predictions could aid in making informed agricultural planning and intervention decisions.

ACKNOWLEDGEMENTS

First and foremost, I would like to extend my deepest appreciation to Dr. Pouyan Nejadhashemi, my academic advisor, who has also served as my research advisor and mentor throughout this journey. Your guidance, expertise, and unwavering support have been invaluable in shaping the direction and quality of my work. Thank you for your genuine commitment to helping me succeed and for grooming me for excellence.

I am grateful to my committee, Dr. Narendra Das and Dr. Timothy Harrigan for their support in achieving this goal. Dr. Das' class that I took in my third semester is what equipped me with immense knowledge in remote sensing and has been very vital to my work. My sincere gratitude goes to you both for helping me with my work.

I also extend my gratitude to my lab mates, colleagues, and friends – Gurjeet Singh, Josue Kpodo, Ian Kropp, Anoh Kouao, Vahid Rafiei, Sebastian Suarez, Nicolas Hernandez, Hannah Ferriby, Anna Raschke, Kieron Moller, Hoda Razavi, Mervis Chikafa, Shashank Mohan, Talha Mohammad, Emmanuel Chima, and Edith Gondwe, whose collaboration and insightful discussions have greatly enriched my research experience and time at MSU. Your willingness to share knowledge and those lunch, ice cream, or coffee breaks have been pivotal in overcoming challenges and achieving my goals.

I want to express my sincere appreciation to my dear husband, Mr. Kondwani Ngwira, for being my biggest cheer leader, believing in me, and supporting me to pursue my dreams. Your unwavering and constant encouragement, patience, understanding, and unconditional support have been my pillars of strength and made this endeavour possible. I truly am grateful.

I would like to acknowledge the guidance and spiritual support provided by my pastor and life coach, Pastor Joel Gondwe, and his wife Mrs. Duwa Gondwe. Your wisdom, prayers, and words of encouragement have been a source of inspiration and motivation during my stay at MSU. Thank you for believing in me when I didn't believe in myself, and for inspiring me to dream. I am eternally grateful. In the same light, I am thankful for my friends and family God has blessed me with through you, who have been a source of joy, laughter, and great inspiration.

I extend my deepest gratitude to my parents, Mr. and Mrs. Banda, and Mrs. Nellie Don. Your support and belief in my abilities have been my driving force, and I am grateful to you for raising me well and sending me to school. I am forever indebted to you.

Last but certainly not least, I would like to thank the Foundation of a Smoke Free World for the financial support during my studies at MSU. This work was produced with the help of a grant from the Foundation of a Smoke Free World, Inc. Furthermore, this work has received partial support from by the United States Agency for International Development (USAID) Bureau for Resilience and Food Security/Center for Agriculture-led Growth under the Cooperative Agreement # AID-OAA-L-14-00006 as part of Feed the Future Innovation Lab for Collaborative Research on Sustainable Intensification (SIIL). The contents, selection, and presentation of facts, as well as any opinions expressed herein, are the sole responsibility of the author(s) and under no circumstance shall be regarded as reflecting the positions of the Foundation of a Smoke Free World, Inc or USAID.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AMSR-E	Advanced Microwave Scanning Radiometer for EOS
CHIRPS	Climate Hazard Group InfraRed Precipitation with Station
CIAT	International Centre for Tropical Agriculture
CIESIN	Center for International. Earth Science Information Network
CSA	Climate Smart Agriculture
CSSR	Climate Science Specialist Report
DAI	Drought Area Index
DSI	Drought Severity Index
EDI	Evaporation Deficit Index
ERS	Economic Research Service
FAO	Food and Agricultural Organization
GDD	Growing Degree Days
GDDI	Growing Degree Days Index
GDP	Gross Domestic Product
GIS	Geographic Information Systems
ICRISAT	International Crop Research Institute for the Semi-Arid Tropics
IPCC	International Panel on Climate Change
MSRRI	Multivariate Standardized Reliability and Resilience Index

NASA	National Aeronautics and Space Administration
NCAR	National Centre for Atmospheric Research
NDVI	Normalized Difference Vegetation Index
NRDC	Natural Resources Defense Council
OECD	Organization for Economic Cooperation and Development
PHDI	Palmer Hydrologic Drought Index
PSDI	Palmer Drought Severity Index
RDI	Reclamation Drought Index
SDI	Streamflow Drought Index
SEMED	Southern and Eastern Mediterranean
SMAP	Soil Moisture Active Passive
SMDI	Soil Moisture Deficit Index
SMI	Soil Moisture Index
SPEI	Standardized Precipitation-Evapotranspiration Index
SPI	Standardized Precipitation Index
SRI	Standardized Runoff Index
SWSI	Surface Water Supply Index
TCI	Temperature Condition Index

TRMM	Tropical Rainfall Measuring Mission
UN	United Nations
UNCCD	United Nations Convention to Combat Desertification
UNDP	United Nations Development Program
UNESCO	United Nations Educational, Scientific, and Cultural Organization
UNWA	United Nations Water Africa
US EPA	United States Environmental Protection Agency
USAID	United States Agency for International Development
USDA	United States Department of Agriculture
USGS	United States Geological Survey
VCI	Vegetation Conditioning Index
VHI	Vegetation Health Index
VI	Vegetation Index
WDI	Water Demand Index
WFP	World Food Program
WHO	World Health Organization
WSR	Water Storage Resilience Index

1.0 INTRODUCTION

Agriculture plays a vital role in ensuring food security globally. Food security may be defined as the availability, accessibility, and utilization of safe and nutritious food to meet the dietary needs and preferences of individuals (FAO, 2008; World Food Summit, 1996). According to the World Economic Forum, about 690 million people were undernourished in 2019, and the number is expected to rise to over 840 million by 2030 (World Economic Forum, 2020). Several factors contribute to food insecurity, but climate change plays a significant role. Rising temperatures, erratic weather patterns, and extreme weather events that result from the change in climate usually lead to the loss of crops and livestock, thus causing food insecurity, malnutrition, and various socioeconomic difficulties. Therefore, addressing food insecurity by transforming agriculture is critical to achieving the United Nations Sustainable Development Goal of zero hunger by 2030 (United Nations, 2020; FAO, 2018).

One of the African countries constantly affected by droughts is Senegal, located in West Africa (World Bank Group, 2021a). Agriculture accounts for a significant portion of Senegal's Gross Domestic Product (GDP) and employs a large proportion of its population (CIAT; BFS/USAID, 2016). The main crops that define the majority of Senegal's agricultural system are typically comprised of groundnuts, millet, sorghum, rice, maize, cowpea, cassava, cotton, mango, and vegetables, where millet, sorghum, rice, and maize are mostly grown by smallholder farmers, and groundnuts and cotton are the main cash crops (CIAT;BFS/USAID, 2016). Millet holds great importance in Senegal due to its higher resilience to drought compared to other staple crops. In fact, millet cultivation covers approximately one-third of the country's arable land (USDA, 2011).

Crop yield estimation is an essential component of agricultural management, as it helps farmers plan their planting and harvesting schedules, make informed decisions about inputs such as fertilizers and pesticides, and optimize their crop management practices, thereby improving productivity (Kropp et al., 2019). Traditionally, crop yield estimation has been done through ground-based surveys, which can be time-consuming and resource-intensive. For instance, ground-based data collection exercises take a long to complete, are heavily technical, and are difficult to execute at a larger scale. However, with advances in remote sensing technology, it has become possible to estimate crop yields from satellite and aerial imagery (Meroni et al., 2013; Panek & Gozdowski, 2020; Qiao et al., 2021). Remote sensing refers to the collection of information from an object or phenomenon using sensors mounted on platforms such as satellites. This data collection technique can provide much information about crops, including their growth stage, health, and yield potential. Therefore, by using remote sensing to estimate crop yield, farmers can make more informed decisions about crop management. For example, if remote sensing data indicates a crop is experiencing stress, farmers can take corrective actions such as adjusting irrigation or applying additional fertilizers. Similarly, if remote sensing data suggests a crop is nearing maturity, farmers can plan their harvesting schedule accordingly (de Castro et al., 2018). Although advancements in remote sensing have helped address different issues in the world, there are challenges associated with it, including the need for extensive computational resources and the construction of models for data analysis (Chadburn, 2020), which may not be readily accessible to smallholder farmers.

Remote sensing can be used to estimate crop yield in several ways. One approach uses vegetation indices, such as the Normalized Difference Vegetation Index (NDVI), which measures the difference between the reflectance of near-infrared and red light. NDVI is a widely used vegetation index that can provide valuable information about crop health and biomass (de Castro et al., 2018; Lumbierres et al., 2017). Multiple studies have employed remote sensing techniques to assess crop yield. Some studies have relied on a single variable (Panek & Gozdowski, 2020), while others have utilized a combination of variables and indices (Ines et al., 2013; Prasad et al., 2006a). In the context of Senegal, Fall et al. (2020) used precipitation-based and remote sensing-based indicators to identify the optimal combination of variables for yield prediction (Fall et al., 2021). Sarr & Sultan (2022) also tested several machine learning algorithms for yield predictability for different crops, including millet (Sarr & Sultan, 2022). However, the reliability of certain past studies is compromised by incomplete datasets caused by bad weather conditions or aggregated information to reduce the problem's complexity and computational time. For example, It can be quite challenging to obtain continuous evapotranspiration datasets in coastal areas due to prolonged cloudiness periods. Additionally, precipitation data is often provided in an aggregated format, such as annual or seasonal timeframes, which can make it difficult to calculate water stress for crops. Due to insufficient data on evapotranspiration water demand, it is currently unfeasible to accurately identify crucial periods of water stress in crops. Additionally, the prolonged drought period, which can harm crop growth, may go unnoticed due to precipitation data aggregation. Therefore, the proposed study tries to address these challenges using higher-resolution temporal and spatial remote sensing products. In addition, we are considering physiographical datasets to examine their relevance to improving the overall crop yield predictability model. Therefore the objective of the work is to

1) Address the challenge of missing and continued observation of remote sensing products by using the new generation of remote sensing products that can improve our understanding of the earth's water cycle, climate, and agricultural productivity.

2) Address the challenge of temporal aggregation by dividing the total growing season into distinct stages: planting, growing, and harvesting periods, and generate remote sensing products for each period separately.
3) Address the challenge of spatial aggregation of crop yield information at the department level by introducing a new bias correction index.

Through the integration of these novel strategies, we anticipate a substantial advancement in the timely prediction of millet yield, enabling us to forecast it several months in advance. This can assist policy makers with the information and tools they need to make better decisions and implement more effective strategies for ensuring food security in their countries. Ultimately, with the increasing availability and affordability of remote sensing data, techniques such as the one proposed here are poised to provide farmers with new opportunities to optimize their crop management practices and increase their productivity.

2.0 LITERATURE REVIEW

2.1 Food, Nutrition, and Water Security

2.1.1 Food and Nutrition Security

Food security can be defined as a state in which people have access to enough nutritious and safe food and are able to acquire it through socially accepted means (USDA ERS, 2022). On the contrary, food insecurity is a state in which people do not have access to enough safe and nutritious food. The United States Department of Agriculture (USDA) uses a scale to gauge food security, ranging from high to very low levels (USDA ERS, 2022). High food security indicates unhindered access to nutritious and healthy food without any difficulties. Marginal food security denotes some obstacles in accessing healthy and nutritious food, but it does not compromise the quality and quantity of the food consumed. Low food security implies a compromised state of food quality, variety, and desirability, while the quantity remains relatively unaffected. Lastly, very low food security signifies the inability to afford food for all family members, leading to significant disruptions in eating patterns.

About 2 billion people experience food insecurity globally, and about 820 million people are hit by hunger, of which about 514 million are in Asia, 256 million are from Africa, and about 42.5 million are from Latin America and the Caribbean (FAO et al., 2019). Another concerning issue is malnutrition, which ranges from being underweight to being obese, and according to the Food and Agriculture Organization of the United Nations 2019 report, about 4 million birth deaths are due to obesity worldwide (FAO et al., 2019). Africa and Asia are the most affected regions, with the most stunted and overweight children. This is also a significant issue in high-income countries, as regular access to healthy food is a challenge for about 8% of the people in Europe and Northern America (WHO, 2019).

One of the ways to deal with food and nutrition insecurity is by strengthening resilience through diversification of food supply, investing in food processing and preservation, and considering cross-sectoral approaches in dealing with food insecurity and malnutrition (World Bank Group, 2017). To achieve this, there is a need for collaboration from different sectors, including education, water and sanitation, agriculture, and health (FAO et al., 2019).

2.1.2 Water Security

The earth's surface measures about 510 million square kilometers. Of this surface area, 71% is covered by water, and the remaining 29% is the surface area of land. The 71% surface area covered by water comprises different forms, including water in the oceans, glaciers, lakes, rivers, groundwater, and the water that makes up soil moisture (Piani & Paris, 2021). Of all the water on the earth’s surface, 97.2% is the water in oceans, 2.15% is in Antarctica and Greenland glaciers, and the remaining 0.65% includes all surface and groundwaters (Sharp, 2017).

The combination of glaciers, surface, and groundwater accounts for 2.8% of the total water on the earth’s surface and is categorized as freshwater. Most of the water stored in glaciers is frozen, which means even less freshwater is available for human consumption and other uses, which is why water is becoming scarce with time (Smith & Southard, 2002). Meanwhile, accessible freshwater is not enough to meet human demands in many regions of the world. As a result, a large population experiences water shortage problems (Jackson et al., 2018). Currently, Southern and Eastern Mediterranean (SEMED) countries (Algeria, Libya, Cyprus, Morocco, Egypt, Palestine, Israel, Syria, Jordan, Lebanon, Turkey, and Tunisia) use more of their renewable water resources, with eight of the countries using over 50% of the water resources and two of the countries using over 100% of the renewable water resources (Baba et al., 2011).

It was estimated that the world's population will reach 9 billion by 2050, and with the growing population, there are increasing strains on water resources (Kurian, 2015). This means there will not be enough water to sustain agricultural production, industrial expansion, and human demand if we do not come up with an innovative solution to address the aforementioned challenges. By 2050, the world's need for water will increase by 55% due to increased agriculture, manufacturing, and electricity production, among other uses (OECD, 2012). Agriculture accounts for 70% of freshwater use worldwide (Smedley, 2017), and by 2050, the water demand for agriculture is expected to increase by 60% so that there is enough food production for the growing population (UN Water, 2019).

The world's population today uses six times more water than the world's population 100 years ago (UN WATER, 2020). Yet, the sources of water have remained the same. Regions that are currently experiencing water scarcity are likely going to have amplified water scarcity, and regions that are experiencing an abundance of water are likely going to experience water stress (UN Water, 2019). Some regions of the world that are currently experiencing water stress include Mumbai, Rio, Jakarta, and Nairobi. About 3 billion people worldwide do not afford basic handwashing services (Ki-moon & Verkooijen, 2021). In addition, approximately 4.2 billion people globally do not have safe sanitation facilities, and 2.2 billion do not have safe drinking water (UN WATER News, 2020).

The available water resources are not meeting current water demands, and local and federal governments need to start planning for a world that will not have enough water in years to come (Ki-moon &Verkooijen, 2021). Several solutions have been suggested to address this issue, such as water trading, water recycling, water desalination, and rainwater harvesting (Day, 2019). Water trading means exchanging water for money among people in different regions or countries (Chong

& Sunding, 2006). Water recycling means treating water regarded as wastewater into a usable and safe form for human use (UN-Environment, 2017). Water desalination means removing minerals and other inorganic substances from saline water so that it is safe for human use (WHO, 2011). This is the primary source of water for many countries, such as Israel, where the saline seawater is converted into potable water through desalination, and it is safe for human consumption. In this case, the desalinating seawater supplies over half of the country's drinking water and over 40% of its agricultural activities (Region, 2015). Lastly, rainwater harvesting means capturing or collecting rainwater from rooftops or any platform so that it is stored and used later. The amount captured through rainwater harvesting can be quite large. For example, 30% of the water demand in Singapore is met by rainwater harvesting (Smedley, 2017).

2.1.3 Water Security in Africa

Africa's water supply relies on both surface and groundwater sources. However, there has been a lack of comprehensive research and effective management of these vital water resources (Matondo et al., 2020). Africa is home to several of the world's largest lakes, which are primarily situated in the Great Lakes Region within the Great African Rift. The lakes within the Great Lakes Region of Africa play a crucial role in the continent's water supply, accounting for approximately 25% of the unfrozen freshwater available (Kiprop, 2020). Groundwater is also extensively utilized, with around 40% of the population relying on it as their primary source of drinking water (Rutten, 2012).

In terms of water consumption, agriculture dominates the water usage in Africa, accounting for approximately 85% of the fresh water utilized. Comparatively, industry utilizes only 5% of the water, while household activities account for the remaining 10% (AWF, 2021). Despite having sufficient freshwater resources to meet the continent's water requirements, the average daily water usage per person in Africa is considerably low, ranging from 30 to 40 liters. This falls significantly below the recognized daily water requirement of 100 liters per person (Howard et al., 2003).

The current water crisis in Africa is severe, with approximately 1 billion people lacking access to clean and safe water sources (Project, 2019). This situation is exacerbated by significant pressures on existing water resources throughout the continent, leading to the gradual depletion of major water bodies (UNWA, 2003). An alarming example is Lake Chad, once among Africa's largest lakes, which served as a vital water source for over 20 million people across four countries. However, it has significantly diminished in size and now occupies less than 10% of its original extent (FAO, 2017; Gao et al., 2011). The primary cause behind Lake Chad's shrinking is the excessive water withdrawal resulting from a rapid population increase over the past five decades (Jacobs, 2010).

The plight of Lake Chad is not unique, as other African lakes have also experienced substantial shrinkage. For instance, Lake Jipe and Lake Natron in Tanzania have diminished in size by approximately 60% and 65%, respectively, within the past few decades (Waigwa, 2007). These shrinking water bodies illustrate Africa's critical challenges in sustaining its water resources. The consequences of shrinking lakes are far-reaching. The reduction in water availability disrupts ecosystems, threatens biodiversity, and jeopardizes the livelihoods of communities reliant on these water sources for drinking, irrigation, and other economic activities (Gao et al., 2011; Jones & Fleck, 2020). To safeguard Africa's water security and ensure access to clean and safe water for all its inhabitants, managing population growth and investing in green infrastructures are necessary (Jacobs, 2010; Holtz & Golubski, 2021).

Africa requires a comprehensive approach to effectively manage surface and groundwater resources to conserve its precious water sources (UN Water, 2019). However, insufficient planning and inadequate management of water resources can have detrimental consequences, including the gradual disappearance of water bodies and detrimental impacts on the communities and ecosystems reliant on these resources (Plessis, 2019).

Proactive water resource management is crucial to prevent the depletion and degradation of water resources, which involves implementing sustainable practices that balance the needs of human populations, agriculture, industry, and the environment (The World Bank, 2022). In addition, ensuring collaboration and appropriate adaptation of different strategies for water resource management are necessary to ensure the equitable and efficient use of water while minimizing the negative impacts on ecosystems and vulnerable communities.

Efficient water allocation and distribution, supported by accurate data and scientific research, can help prevent the overuse and depletion of water resources, which requires robust monitoring systems to track water availability, consumption patterns, and ecological health. By understanding these dynamics of water resources, policymakers, and stakeholders can make informed decisions to regulate water usage and promote conservation efforts (Miller et al., 1997; OECD, 2015). Moreover, effective water resource management should prioritize protecting and restoring ecosystems that sustain water sources. This involves preserving wetlands, forests, and other natural habitats crucial in regulating water flow, enhancing water quality, and maintaining the overall ecological balance (US EPA, 2022). Integrated approaches that consider both the social and ecological aspects of water management are essential for long-term sustainability.

Additionally, raising awareness and promoting education about water conservation practices among the general public is crucial. Encouraging responsible water use at individual and community levels can significantly contribute to preserving water resources, and public campaigns, community engagement, and educational programs can empower individuals to adopt water-saving behaviors and contribute to sustainable water management efforts (Kelly et al., 2017; Nelson et al., 2021). By adopting holistic and proactive approaches to water resource management, Africa can safeguard its water resources for future generations, protect ecosystems, and ensure water security for all(UNWA, 2003). Collaboration among governments, local communities, researchers, and international organizations is vital to address the challenges and implement effective solutions that will lead to sustainable water management practices (UNESCO, 2019).

2.1.4 Food, Nutrition, and Water Security in Senegal

Senegal, located in West Africa, has been significantly impacted by natural disasters such as floods and drought, exacerbating the issue of food insecurity in the country (WFP, 2019). Approximately 50% of the population in Senegal lives below the poverty line, and the situation worsened in 2020 with a rise in cases of food insecurity and malnutrition due to low agricultural yields in the preceding years (Action Against Hunger, 2022).

To address the challenges posed by climate change, Senegal has recognized the importance of adopting climate-smart agriculture approaches alongside traditional farming methods (World Bank Group, 2021b). These approaches aim to mitigate the adverse effects of climate change on agriculture and enhance resilience in the face of climate-related challenges. The World Food Program (WFP) has been actively promoting the adoption of climate-smart agriculture in various parts of Senegal to support farmers in achieving food and nutrition security. Furthermore, preserving baobab trees has been a key focus for the WFP. These iconic trees contribute to reducing the concentration of carbon dioxide in the atmosphere, mitigating climate change effects, and producing baobab fruits that serve as valuable vitamin supplements, enhancing nutritional diversity in local diets (WFP, 2019).

By implementing climate-smart agriculture practices and protecting vital ecological resources like baobab trees, Senegal aims to build resilience in its agricultural sector, enhance food security, and improve the nutritional status of its population. Collaborative efforts among government agencies, international organizations, local communities, and farmers are crucial to implementing and scaling up these initiatives, ensuring their long-term sustainability and positive impact on Senegal's food system and environment (Ndoye, 2015; USDA, 2023).

2.2 Factors Affecting Food Security

2.2.1 Climate Change

Climate change is a phenomenon characterized by long-term alterations in weather patterns and environmental conditions (Muchuru & Nhamo, 2019). Over the past century, human activities and natural processes have led to a significant increase in carbon dioxide levels in the Earth's atmosphere, rising by 48% since 1850 (NASA, 2021). The consequences of climate change are far-reaching and impact various aspects of the planet. Sea-level rise is occurring as a result of melting ice caps and thermal expansion of the oceans, posing significant threats to coastal regions and island nations (National Geographic Society, 2019). Global temperatures are increasing, with the Earth's average temperature rising by approximately 1 degree Celsius in the past century, and projections indicate further warming in the coming decades (NASA, 2014). This temperature rise has numerous implications, including the shrinking of icebergs, the melting of glaciers, and alterations in precipitation patterns leading to more frequent and severe floods and droughts (CSSR, 2016).

Without effective measures to mitigate greenhouse gas emissions, the future projections paint a worrisome picture. Global temperatures may rise by up to 5 degrees Celsius by the end of the 21st century, with dire consequences for ecosystems, human livelihoods, and the planet's overall health (CSSR, 2016). Such significant temperature increases would exacerbate the risks of extreme weather events, further threaten biodiversity, and amplify the challenges associated with food security, water availability, and public health (US EPA, 2023).

Climate change has disrupted water cycles and weather patterns on a global scale, leading to unpredictable rainfall patterns characterized by either insufficient or excessive precipitation in different regions (Ki-moon and Patrick Verkooijen, 2021). This variability in rainfall has significant implications for water-related health effects linked to climate change. One such consequence is malnutrition, resulting from food scarcity caused by droughts or floods (UN WATER, 2020). Furthermore, extreme weather events like floods can lead to injuries, fatalities, and the spread of vector-borne diseases (Husain & Trak, 2018).

Agricultural production has been significantly impacted by climate change. However, implementing strategies such as cultivating hybrid crop varieties and establishing flood or drought-resistant infrastructure and systems can help communities adapt to the effects of climate change (Sarkodie & Strezov, 2019). These measures can enhance the resilience of agricultural systems and support food security in the face of changing climate conditions.

In Africa, the impact of climate change varies across different regions, with countries in the northern part generally facing lower risks compared to those in the south. However, certain areas within the continent, such as the Sahel, Somalia, parts of Ethiopia, and the southwest portion of the Arabian Peninsula, are particularly vulnerable due to limited adoption of adaptive strategies (UN WATER, 2020). As a result, these areas exhibit higher susceptibility to the adverse effects of climate change, compounded by challenges related to water resources and political interference. In countries frequently affected by conflict, political interference has disrupted the equitable distribution of water resources, exacerbating water inequality in the region. Furthermore, deteriorating water infrastructure, population displacement, and using water as a tool for diplomatic leverage pose additional challenges in these countries (UN WATER, 2020). These factors contribute to heightened vulnerability to the impacts of climate change and further exacerbate the water-related challenges these nations face.

Various solutions have been identified to address water inequality and enhance resilience to climate change in the region. Promoting research, innovation, and sustainable development through policy formulation and integration is essential. Through mitigation and adaptation strategies, this approach ensures the implementation of long-term solutions that can withstand the challenges posed by climate change (UN WATER News, 2020). By integrating climate change considerations into policy frameworks and fostering sustainable practices, countries in the region can work towards achieving water security, reducing vulnerability, and promoting equitable access to water resources.

2.2.2 Land Use and Land Cover Changes

Land use encompasses the various cultural and economic activities that shape the purpose and utilization of land. In contrast, land cover refers to the physical materials that cover the Earth's surface (MSU extension, 2013). Over time, profound changes in land use and land cover have occurred worldwide. For instance, forests have been cleared to make way for agriculture, urban development, and other human activities. Within the past century, these changes have been substantial, with approximately 75% of the Earth's land area experiencing alterations in land use. Meanwhile, over the past 60 years alone, about 32% of the planet has undergone changes in land use (Winkler et al., 2021).

These land-use transformations have wide-ranging consequences, with significant impacts on climate patterns, human health, ecological systems, watershed functions, and more (US EPA, 2021). Additionally, land cover change has been identified as a major environmental threat, playing a substantial role in the increase of carbon dioxide and other greenhouse gases in the atmosphere during the transition from the 20th to the 21st century (Arneth et al., 2014).

Another crucial aspect in the study of land use and land cover is land management change, which pertains to the human activities conducted on different land covers without altering them (Luyssaert et al., 2014). These activities may involve practices like irrigation, fertilizer application, and the implementation of specific cropping systems, among others. Over time, these practices can have environmental consequences, such as groundwater pollution resulting from nutrient leaching in the soil and the accumulation of salts, both of which can adversely affect agricultural productivity (Krasilnikov et al., 2022; Mostafazadeh-Fard et al., 2007; Pahalvi et al., 2021). As a result, the concepts of land use change, land cover change, and land management change contribute directly to food scarcity by impeding adequate farming land availability and exacerbating land degradation through unsustainable land management practices (Agidew & Singh, 2017).

Addressing these challenges requires a comprehensive approach considering sustainable land use and management practices. This includes promoting responsible land stewardship, implementing effective land management strategies, and adopting sustainable farming techniques. Proper land management can mitigate the negative environmental impacts associated with land use changes, help preserve soil fertility, prevent soil erosion, reduce water pollution, and promote long-term agricultural productivity (CIESIN, 2000; Findell et al., 2017). Moreover, integrating sustainable land management practices with land use planning and policy development is crucial for achieving resilient and sustainable agricultural systems (FAO, 2023). This entails supporting farmers with knowledge, resources, and incentives to adopt sustainable practices (El Fartassi et al., 2023), investing in research and innovation for improved land management techniques, and fostering collaboration among stakeholders to ensure the long-term viability of land resources (Karalliyadda et al., 2023).

By prioritizing sustainable land use, land cover management, and responsible land management practices, we can address the interconnected challenges of food security, environmental conservation, and land degradation, ultimately fostering a more sustainable and resilient future for generations to come.

2.2.3 Population Growth

The world's population is currently about 8 billion, with an average density of 62 people per square kilometer (Our World in Data, 2023) and is expected to increase to 9 billion by 2050 (USAID, 2022) and almost 10.9 billion by 2100 (Our World in Data, 2023). However, there is variability among countries and regions as some will continue to experience a population increase while some will experience a population decrease. In addition, studies have shown that by 2050, the need for food will rise to 100% from 70% due to the increasing population, and developing countries will need to double their food production to achieve food security (USDA, 2022).

Since rapid population growth often results in high demand for food, interventions like family planning may help slow down the food demand to become more manageable (Toolkits, 2012). Other strategies that will ensure food security include minimizing food waste (FAO, 2019; Parfitt et al., 2010), improved storage and transportation systems, and better consumer awareness initiatives (Parfitt et al., 2010). Pretty et al. (2018) also found that improving agricultural productivity through sustainable practices can increase food production without expanding agricultural land. This includes promoting agroecological approaches, precision agriculture, and efficient water and nutrient management techniques (Pretty et al., 2018). Investing in research and technology transfer to smallholder farmers can help improve yields and reduce the demand for additional agricultural land. Therefore, ensuring food security will eliminate hunger and promote economic stability (Wageningen University and Research, 2023).

2.3 Drought

Drought, a natural disaster, lacks a universal definition, as its interpretation varies across different disciplines. Nevertheless, drought can be characterized as a prolonged period of inadequate soil moisture, a state of dryness resulting in diminished water levels in streams, and a scarcity that adversely impacts the economy, agriculture, and water resources (Minucci, 2021). Unlike other natural disasters, drought has a widespread impact on global populations (NRDC, 2018), and its severity and frequency are escalating over time (UNCCD, 2019). Research examining the impacts of drought reveals its adverse effects on various sectors. These include reduced crop yields (Ray et al., 2018), negative repercussions on livestock production (Dzavo et al., 2019; Mare et al., 2018), compromised child nutrition (Cooper et al., 2019), and significant economic consequences (IPCC, 2019; Keyantash, 2002). In fact, from 2005 to 2015, approximately 80 percent of the economic losses incurred in developing nations were attributed to drought's impact on crops, livestock, and fisheries (NRDC, 2018). Moreover, according to the Atlas of Mortality and Economic Losses from Weather, Climate, and Water extremes, drought emerged as the deadliest natural disaster between 1970 and 2019, claiming over 600,000 lives (UN News, 2021).

2.3.1 Types of Drought

Drought can be classified into five distinct categories: meteorological drought, agricultural drought, hydrological drought, socioeconomic drought, and stream health drought (Esfahanian et al., 2017). Meteorological drought occurs when there is a prolonged period of reduced or no precipitation (Spinoni et al., 2019). Agricultural drought refers to insufficient soil moisture, negatively impacting crop growth (Dai et al., 2020). Hydrological drought is characterized by depleted water levels in lakes, streams, and groundwater sources (Keyantash, 2002). Socioeconomic drought describes a shortage of water supply, impacting society, the economy, and the environment (Meng et al., 2019). Stream health drought relates to low stream flow during the driest month, commonly known as index flow (Esfahanian et al., 2017). While each type of drought possesses its own definition, all are interconnected, originating from a deficit in precipitation. Meteorological drought precedes agricultural drought, which in turn influences stream health drought and socioeconomic drought (Xianfeng et al., 2016).

2.3.2 Drought Assessment Strategies

To effectively mitigate the impacts of extreme drought, it is crucial to accurately measure and assess its intensity and severity across different categories. Several indices have been developed for this purpose. The Palmer Drought Severity Index (PDSI) (Aiguo & NCAR, 2019), the Standardized Precipitation-Evapotranspiration Index (SPEI) (Stagge et al., 2014), and the Standardized Precipitation Index (SPI) (Palmer, 1965b) are commonly used indices for studying meteorological drought. For hydrological drought analysis, the Surface Water Supply Index (SWSI) (Shafer & Dezman, 1982), the Reclamation Drought Index (RDI) (Weghorst, 1996), the Streamflow Drought Index (SDI) (Aghelpour et al., 2020), and the Palmer Hydrologic Drought

Index (PHDI) (Palmer, 1965a) are widely employed. In the realm of agricultural drought assessment, commonly utilized indices include the Palmer Moisture Anomaly Index (Z-Index), the Soil Water Deficit Index (SWDI), the Evapotranspiration Deficit Index (EDI), and the Soil Moisture Deficit Index (SMDI) (Esfahanian et al., 2017). Evaluating stream health drought involves using the Current Drought Severity Model and the Future Drought Severity Model (Esfahanian et al., 2016). Lastly, for socioeconomic drought evaluation, the Multivariate Standardized Reliability and Resilience Index (MSRRI), the Inflow-demand Reliability Index (IDR), and the Water Storage Resilience Index (WSR) are commonly employed (Huang et al., 2016).

It's important to note that these indices not only aid in detecting and assessing drought severity but also play a crucial role in drought forecasting. Drought forecasting is essential as it enables governments and policymakers to proactively prepare for disasters, explore alternative resource options, and maintain economic sustainability (UN-Environment, 2017).

2.3.3 Drought in Senegal

Like other West African countries, Senegal is impacted by climate variability and change (Diatta et al., 2021). Research by Busby et al. (2014) highlights that Senegal and other African nations are particularly vulnerable to climate change effects, including increased drought occurrences, primarily due to rising heat waves in the region (Busby et al., 2014). Senegal has a history of being prone to droughts, with the African Sahel drought between 1968 and 1988 being one of the most severe and enduring droughts in history (Alahacoon & Edirisinghe, 2022). However, it is important to note that vulnerability to climate change varies among countries due to socioeconomic, environmental, and cultural factors (Alcamo et al., 2007). Developing countries, such as Senegal, often face greater exposure and have less capacity to adapt to climate events compared to developed nations (Sarkodie & Strezov, 2019).

To address the adverse effects of natural disasters, including drought and flood, countries in Africa have embraced initiatives such as the African Alliance for Climate Smart Agriculture. This alliance aims to enhance agricultural productivity and adaptability to climate change (World Bank Group, 2017). It promotes convergence and coordination of efforts toward common goals aligned with Climate Smart Agriculture (CSA), which involves supporting the planning, implementation, and evaluation of investment programs in agriculture, as well as mobilizing resources for these initiatives (Toure & Fane, 2015). Another approach is the development of crop models that utilize observed and simulated data to estimate crop yield under different weather and soil conditions. In a successful study, Faye et al. (2018) developed and calibrated a crop model to estimate peanut yield in Senegal by incorporating data on water stress and soil nutrient conditions. These strategies contribute to enhancing agricultural resilience and adaptation to climate change in Senegal.

2.4 Introduction to Remote Sensing

Remote sensing is a method of acquiring data that utilizes sensors mounted on various platforms, such as satellites, airplanes, or drones, to gather reflected or emitted radiation from the object or area of interest without direct contact (Lillesand et al., 2015). There are two primary categories of remote sensing: passive and active. Passive remote sensing relies on natural energy sources, particularly the sun, while active remote sensing utilizes artificial energy sources (GISGeography, 2022). Irrespective of the type of data acquisition, remote sensing finds numerous applications in natural resources studies, including the monitoring of land use and land cover changes, weather predictions (USGS, 2021), water resource monitoring, and natural disaster surveillance (Jyotsna, 2017).

2.4.1 Applications of Remote Sensing in Food Security

2.4.1.1 Drought Measurement through Remote Sensing

Traditional drought forecasting methods have limitations in terms of accuracy (Ren et al., 2008). However, the availability of remote sensing products with varying spatiotemporal resolutions enables large-scale drought monitoring and prediction (Choi et al., 2012; Sur et al., 2015; West et al., 2019). This is particularly crucial in regions with limited ground-level monitoring capabilities and insufficient financial resources. The ability to monitor and predict droughts through remote sensing can significantly enhance regional food security by serving as an early warning system (Kogan et al., 2015; Krishnamurthy R et al., 2020).

Drought indices and monitoring systems have been developed to measure and forecast the severity of droughts. These indices utilize numerical representations based on one or multiple physical variables to assess drought intensity on different scales (Hayes et al., 2012). The majority (around 90%) of remote sensing-based drought indices are used for monitoring agricultural drought, while approximately 10% are employed for meteorological and hydrological drought monitoring. More than 150 indices have been developed globally, and policymakers in different countries need to select the most suitable indices for their specific regions (Alahacoon & Edirisinghe, 2022).

To quantify agricultural drought, commonly used indices include the Normalized Difference Vegetation Index (NDVI), the Vegetation Health Index (VHI), and the Vegetation Condition Index (VCI) (Shahzaman et al., 2021). NDVI measures vegetation cover, with values near +1 indicating robust vegetation and lower values near -1 indicating sparse vegetation (J. Zhao et al., 2020). VHI assesses overall vegetation health and is calculated using two other indices: the Vegetation

Condition Index (VCI) and the Temperature Condition Index (TCI). VCI is employed for drought monitoring, with values below 50% indicating drought conditions and values below 35% indicating extreme drought (Dutta et al., 2015).

For quantifying hydrological drought, the Standardized Runoff Index (SRI), Surface Water Supply Index (SWSI), and Standardized Water Level Index (SWI) are commonly utilized drought indices (Zhu et al., 2018). SRI simulates surface runoff and considers various hydrologic processes that influence seasonal streamflow variations (Shukla & Wood, 2008). SWSI compares water availability across multiple basins to determine the extent and severity of drought(Garen, 1993; Zhu et al., 2018). SWI is another index used for monitoring hydrological drought and aids in assessing groundwater deficits, among other applications (Zhu et al., 2018). Although these were developed using ground-observed data, studies have also shown that they can be derived entirely from remotely sensed products. For instance, the SWSI has been developed using data obtained from Landsat 5, 7 and 8, Climate Hazard Group InfraRed Precipitation with Station (CHIRPS), and Sentinel-1 datasets that are entirely remotely sensed (Alahacoon & Edirisinghe, 2022). The Landsat series have been extensively used for different purposes, including drought monitoring, and has provided vegetation index data like NDVI and VCI (Ghaleb et al., 2015). The CHIRPS dataset was also used in the Feed the Future project to develop a remote sensing based SRI (Feed the Future, 2023).

Meteorological drought can be monitored using indices such as the Standardized Precipitation Index (SPI), the Palmer Drought Severity Index (PDSI), and the Drought Area Index (DAI) (Danandeh Mehr et al., 2022). SPI is particularly useful for assessing the frequency and impact of meteorological droughts, contributing to drought forecasting and climate studies (Dhawale & Paul, 2018). The CHIRPS dataset is a widely used remote sensing source for precipitation data and has been used to develop the SPI (Feed the Future, 2023), and The Tropical Rainfall Measuring Mission (TRMM) is another remote sensing-based precipitation data source that has been used to develop the SPI (Q. Zhao et al., 2018).

2.4.1.2 Yield Estimation

Several studies have investigated the relationship between remotely sensed data and crop yield estimation, employing two primary methods: [1] comparing remotely sensed information or parameters with available regional crop yield data, and [2] integrating remotely sensed data into existing crop models (Ren et al., 2008).

While some studies have used soil health and environmental factors data as a tool for yield estimation (Suruliandi et al., 2021), vegetation cover, as determined by vegetation indices, has also been a commonly utilized indicator for yield estimation. Examples of vegetation indices include the Normalized Difference Vegetation Index (NDVI), Vegetation Condition Index (VCI), and Vegetation Health Index (VHI) (Alahacoon & Edirisinghe, 2022). Previous research has employed various approaches, including the use of single parameters derived from satellite imagery or combinations of parameters to establish correlations with crop yield. For instance, Prasad et al. (2006) developed a crop yield estimation model by incorporating soil moisture, NDVI, temperature, and precipitation data obtained from satellite imagery (Prasad et al., 2006b). Ines et al. (2013) used AMSR-E soil moisture data and Leaf Area Index in a crop model to estimate yield (Ines et al., 2013). Torre et al. (2021) used different models to estimate rice yield and acknowledged the need for using different climatic variables and indices and machine learning techniques in crop yield estimation studies (Torre et al., 2021). Specifically for Senegal, recent studies have used precipitation-based indicators and statistical tools to predict millet yield (Fall et al., 2022). Meanwhile, other studies have used vegetation and climatic data separately or in combination to predict the yield of main crops in the country (Sarr & Sultan, 2022).

2.5 Limitations of Existing Yield Estimation Methods

Numerous studies have explored remotely sensed data, particularly vegetation indices, for estimating crop yield. However, several limitations need to be addressed to enhance the accuracy and applicability of these methods.

- Emphasizing the importance of regional-level research is imperative due to the notable variations in crop yield estimation at small scales among different districts or regions (Turvey & McLaurin, 2012). This can be a valuable tool for policy makers to make more informed decisions.
- It is important to recognize that crop yield is influenced by various factors beyond vegetation index and water availability. Other key determinants include soil characteristics, topography, nutrient availability, irrigation practices, pest and disease management, and agricultural management practices. Integrating these additional variables into crop predictive analyses can help create more comprehensive and robust models for yield estimation.

In summary, while remote sensing techniques, particularly vegetation indices, have shown promise for estimating crop yield, addressing limitations such as conducting localized research, developing crop-specific correction factors, and incorporating additional relevant variables will enhance the precision and reliability of these prediction models. By bridging these gaps, remote sensing can significantly improve agricultural decision-making and food security on a regional and global scale.

This study utilizes the SMAP-Sentinel soil moisture product, which combines SMAP Active L-band and Passive L-band Synthetic Aperture Radar (SAR), as well as the Copernicus C-band SAR to provide soil moisture data at a 1 km resolution under various weather conditions (Das et al., 2019; Singh & Das, 2022). In addition, it considers all potential factors that may influence crop yield, various demand, drought, and physiographical indices and variables were employed. Finally, the crop life cycle was categorized into planting, growing, and harvesting seasons to enhance accuracy, as the variables differ across each phase.

3.0 MILLET YIELD ESTIMATIONS IN SENEGAL: UNVEILING THE POWER OF REGIONAL WATER STRESS ANALYSIS AND ADVANCED PREDICTIVE MODELING

3.1 Introduction

Effective agricultural production has the potential to reduce food insecurity in developing countries facing rising poverty, including many countries in Africa (Christiaensen & Martin, 2018). Not only does agriculture provide food, but also it is a large source of employment. Davis et al. (2017) conducted a comprehensive study on income sources among households in Sub-Saharan Africa. They discovered that more than 80% of rural households rely on the agricultural industry for their livelihoods, irrespective of the agricultural sector's contribution to the Gross Domestic Product (GDP) in different countries. However, this major source of income and employment has been threatened by local and global crises (e.g., the COVID-19 pandemic) and environmental stressors (e.g., megadrought).

Climate change and variability have caused most developing countries to experience food insecurity in the past few decades due to over-dependence on rainfed agriculture (Bedeke, 2022). Meanwhile, adopting CSA interventions can effectively mitigate the adverse effects of climate change (Lipper et al., 2014). Among these interventions, drought and flood monitoring and forecasting systems can help understand the magnitude and timing of water shortages. The aforementioned systems hold promise in supporting decision-makers with vital information to refine national and regional plans to mitigate the consequences of extreme environmental events (UN-Environment, 2017). However, these interventions alone are insufficient for a precise assessment of food insecurity. This is because they only provide an indirect correlation between water imbalance effects and agricultural yield. On the other hand, direct yield estimation methods, which assess the effects of extreme climatic events on crop yield, can provide a more effective basis for policymakers when considering alternative strategies to tackle food insecurity (Ren et al., 2008).

Yield estimation can be accomplished through several means. Conventional forecasting systems heavily rely on field observations such as rainfall and temperature data obtained through weather stations or in-situ sensors (Bastiaanssen & Ali, 2003). The data from field observations can be used to establish crop growth and statistical sampling models (Ren et al., 2008). However, these methods generally cannot be expanded to a large area mainly due to a lack of observed data and the complexity of setting up and operating crop growth models. To address these shortcomings, recent crop yield prediction techniques use remote sensing data (Bastiaanssen & Ali, 2003) or combine ground and space observations (Singh & Das, 2022). As technology advances, remotely sensed products have proved to be reliable, easily accessible, and uniformly available even in the most remote areas. The use of recently developed remote sensing satellite imagery allows the production of high-resolution data, which is readily available and not constrained by geographical or environmental factors. This approach overcomes the limitations inherent in traditional data acquisition processes that can be hindered by such parameters (Sui et al., 2018). As these products continually improve in quality and frequency, there is a high probability that the predictability of forecasting models using these products is also improving.

Traditional remote sensing-based yield estimation techniques typically rely on a single parameter extracted from satellite images to construct a correlation between said parameter and crop yield (Panek & Gozdowski, 2020; Lopresti et al., 2015; Meroni et al., 2013). Nevertheless, recent research has started to integrate multiple parameters in an effort to establish a more comprehensive correlation between these aggregated parameters and crop yield. For instance, Prasad et al. (2006) used NDVI (Normalized Difference Vegetation Index), temperature, soil moisture, and precipitation data from satellite imagery to develop a crop yield estimation model (Prasad et al., 2006). Ines et al. (2013) used AMSR-E soil moisture data and Leaf Area Index (LAI) in a crop model to estimate yield (Ines et al., 2013). As an important aspect of ensuring accurate yield prediction using multi-spectral images, incorporating temporal, spectral, and spatial components has also been studied in crop yield estimation studies (Qiao et al., 2021).

Despite these improvements in crop yield estimations, many of these methods have been created for developing countries with limited scalability to other regions. In addition, the predictability of some previous studies is undermined by incomplete datasets attributable to adverse weather conditions. However, advancements in technology have yielded new satellite imagery products that can help bridge gaps in data while concurrently furnishing more precise and higher-resolution information. Therefore, our goal in this study is to use technological advancement in remote sensing to estimate crop yield in data-scared regions such as Senegal. The problem is further aggravated due to a lack of reliable ground observation data that is usually aggregated at a large scale, such as the department (administrative unit) level. In order to address these challenges, the following objectives should be considered: 1) examining the sensitivity of remotely sensed products to crop yield variabilities in different regions and 2) developing the regional crop yield predictive models. The novelty of this work is the contribution of sub-seasonal and unaggregated crop information in order to improve the overall model predictability in case of data scarcity in developing regions of the world.

3.2 Materials and Methods

3.2.1 Study area

This study is Senegal, a country in West Africa. Senegal has a dry tropical climate, with an average annual rainfall of about 1200 (mm) in the south and about 300 mm in the arid zones.

Between 1991 and 2020, the mean annual temperature was about 28.91 degrees Celsius (World Bank Group, 2021). The soil type is primarily sandy, which allows crops like millet, sorghum, corn, and groundnuts to grow very well in most of its agricultural departments. Rice is also grown well in the irrigated areas of Senegal. With slightly over 3 million hectares of arable land nationwide, only about 2% is irrigated (Global Yield Gap Atlas, 2016). Senegal's agriculture sector is very significant, as it takes up about 43% of the country's total area, contributes about 17% to the GDP, and offers employment to about 70% of its population (Feed The Future, 2015). Millet production, like the production of other cereals, experiences yearly variability influenced by diverse weather patterns and economic pressures (Debieu et al., 2018; Kane et al., 2016). However, as of 2016, millet had emerged as the second-largest contributor to the national food production system, following groundnuts, and accounting for roughly 20% of the total agricultural activities (Feed The Future, 2015). Furthermore, in 2018, it accounted for 30% of all cereal production in the country (Kane et al., 2023).

3.2.2. Modeling Overview

Our research proposes three hypotheses to enhance the accuracy of millet yield prediction. The first hypothesis suggests that incorporating soil moisture-based indices derived from high-resolution satellite-based soil moisture retrievals can improve the prediction. The second hypothesis suggests that a prediction model based on individual stages of millet growth will be more effective than a model based on the average of all stages. Lastly, we hypothesize that accounting for the area of millet production at the department level during different growth stages can improve the overall model predictability. To test these hypotheses, we introduce a modeling procedure, which is presented schematically in Figure 1 and briefly explained in the following section.

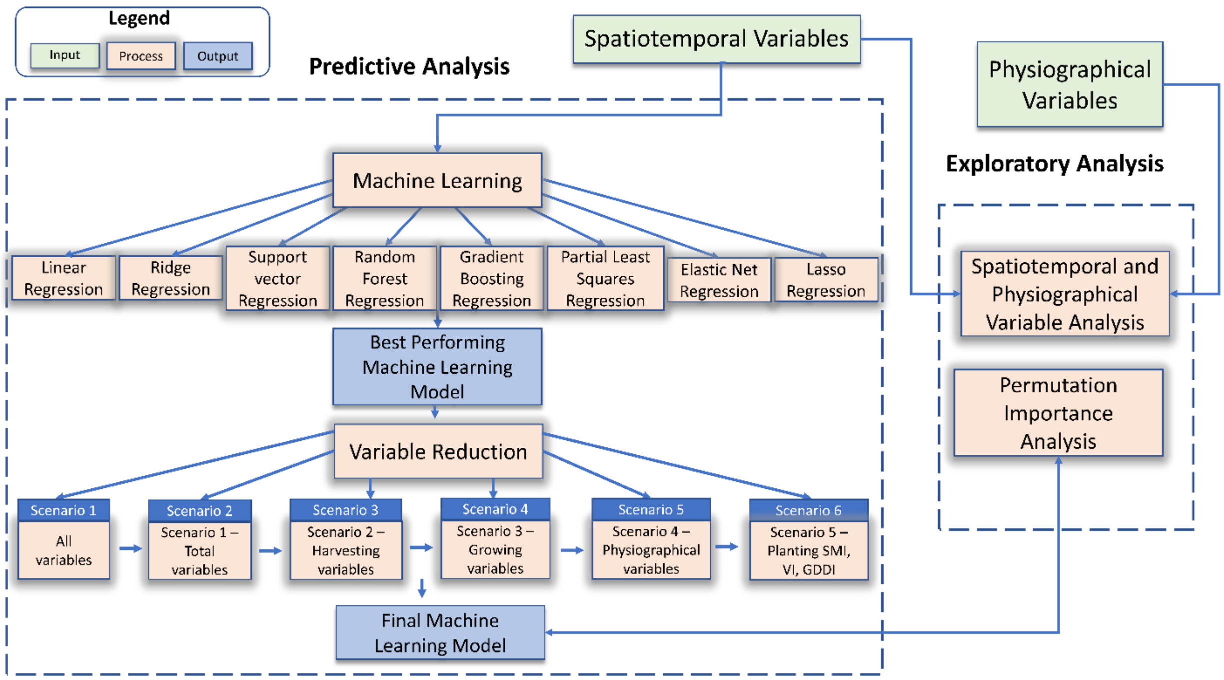

Figure 1. Overview of the modeling process

Crop yield is affected by both spatiotemporal and physiographical variables. Therefore, in the first step, we obtained remotely sensed data on temperature, soil moisture, precipitation, as well as vegetation cover to develop different spatiotemporal variables. We also utilized physiographical variables derived from a combination of remote sensing and ground observations data pertaining to soil characteristics. These variables were used to perform Predictive and Exploratory Analysis. Thirsty six spatiotemporal variables, collected from all Senegal departments, went through an initial series of checks to examine their interdependency, validity, and correlation with crop yield through the preprocessing stage. This information is then fed to a series of machine learning (ML) and nested cross-validation techniques to identify the most robust model. In the next step, we are performing variable reduction strategies to reduce the burden of future data collection and improve computational time. Variables are systematically reduced to identify the best combination of variables for the optimal ML model. Finally, the optimal model is subjected to permutation importance analysis to determine the significance of each feature. This involves shuffling the values of each feature in the validation dataset and comparing the model's performance on the shuffled and original datasets. The importance scores of each feature are averaged over multiple iterations for a more reliable estimate.

3.2.3 Data Input

3.2.3.1 Crop Yield Data

Five-year (2016 – 2020) crop yield data for millet were obtained for 43 departments in Senegal (Vieira Junior et al., 2023). Millet production is predominantly in the western region, with the highest production in the southwest and the lowest in the north (Figure 2.a). Figure 2.b shows the millet production zones clustered using k-means method and the country's average yield value of 43 departments (Figure 2). The planting period of millet across the country ranges from June to July; this is the period between the beginning and end of the planting exercise. August is the main growing period of millet in Senegal; this is the period that the crop attains most of its vegetative growth and achieves maximum canopy cover. The harvesting of millet ranges from September to November; this is the period when the crop is removed from the field (FAO, 2020). The harvesting period seems spatially longer than the planting and growing periods because millet can be harvested from the time the grains mature while the crop is still using water and nutrients for its growth-producing fresh grains until the crop stops growing and loses water-producing dry grains.

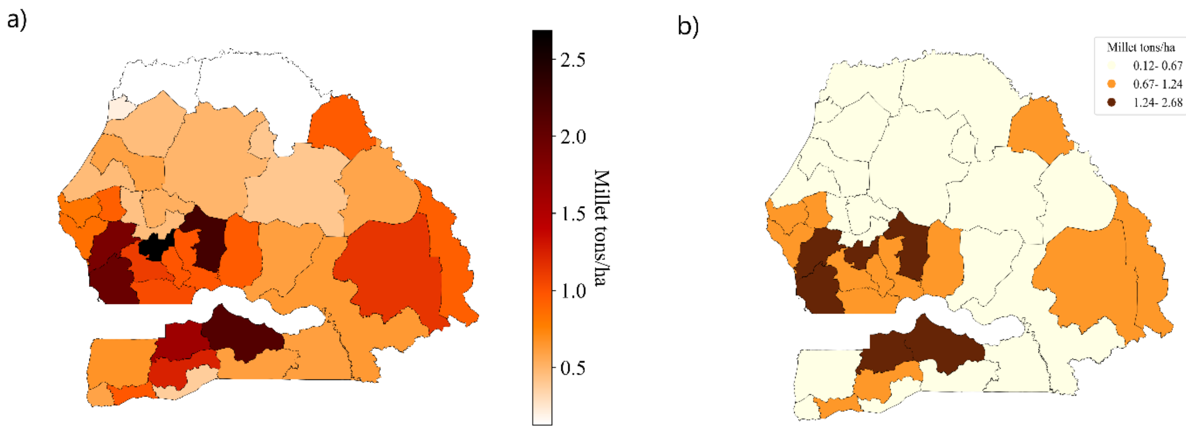

Figure 2. a) average Millet production across Senegal; b) production zones based on annual average yield

3.2.3.2 Spatiotemporal and Physiographical Variables

The following remote sensing products were employed to assess the impact of various environmental stressors on crop productivity at different stages of development. The source and further details of these remote sensing products are presented in Table 1. Also, since the reported yield data was at the department level while remote sensing data have varying spatial scales, it was necessary to upscale the remote sensing products to the department level. The upscaling was achieved by averaging all cells of a department using the ArcGIS version 10.8.1 zonal statistic tool. Additionally, the remote sensing data was only extracted for areas with agricultural land use. Then, all indices were calculated/extracted for three millet development stages: planting (June-July), growing (August), and harvesting (September to November) periods (FAO, 2020). In summary, eight remote sensing products were utilized to estimate yield during the planting, growing, harvesting, and development seasons, resulting in 38 predictor variables.

Table 1: Remote sensing products used for the yield prediction

Index	Product	Type	Publisher	Spatial Scale	Temporal Scale	Source
Soil Moisture Index (SMI)	SMAP/Sentinel-1 L2 Radiometer/Radar 30-Second Scene EASE-Grid Soil Moisture, Version 3 (SPL2SMAP_S)	Secondary	National Aeronautics and Space Administration (NASA)	1 km	12 days	https://nsidc.org/data/spl2smap_s/versions/3
Vegetation Index (VI)	SPOT/VEGETATION, PROBA-V NDVI version 2.2	Secondary	Copernicus Global Land Service	1 km	10 days	https://land.copernicus.eu/global/products/ndvi
Growing Degree Days Index	Copernicus ECMWF	Secondary	Copernicus European Centre for Medium-Range Weather Forecasts (ECMWF)	1 km	Hourly	https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land-monthly-means?tab=overview
Water Demand Index (WDI)	• SMAP/Sentinel-1 L2 Radiometer/Radar 30-Second SceneEASE-Grid Soil Moisture, Version 3 (SPL2SMAP_S) • SPOT/VEGETATION, PROBA-V NDVI version 2.2 • Copernicus ECMWF	Secondary	Singh & Das, (2022)	1 km	12 days	https://www.sciencedirect.com/science/article/pii/S0048969722029904

Table 1 (cont'd)

Soil Moisture Deficit Index (SMDI)	• Regional Hydrological Extreme Assessment System (RHEAS)	Secondary	Feed the Future: Senegal Drought and Crop Watch	5 km	Daily	Andreadis et al. (2017) https://rshydroag.egr.msu.edu/
Standardized Precipitation Index (SPI3)	• Regional Hydrological Extreme Assessment System (RHEAS)	Secondary	Feed the Future: Senegal Drought and Crop Watch	5 km	Daily	Andreadis et al. (2017) https://rshydroag.egr.msu.edu/
Standardized Runoff Index (SRI3)	• Regional Hydrological Extreme Assessment System (RHEAS)	Secondary	Feed the Future: Senegal Drought and Crop Watch	5 km	Daily	Andreadis et al. (2017) https://rshydroag.egr.msu.edu/
Drought Severity Index (DSI)	• Regional Hydrological Extreme Assessment System (RHEAS)	Secondary	Feed the Future: Senegal Drought and Crop Watch	5 km	Daily	Andreadis et al. (2017) https://rshydroag.egr.msu.edu/
Precipitation	• CHIRPS dataset	Primary	CHIRPS	5 km	Daily	https://www.chc.ucsb.edu/data/chirps
Temperature	• NCEP Reanalysis at 2.5 degrees	Primary	NOAA NCEP	5 km	Hourly	https://psl.noaa.gov/data/gridded/data.ncep.reanalysis.html
Nitrogen	• Global SoilGrids250m dataset	Secondary	ISRIC - World Soil Information	250 m	5 years	https://www.isric.org/explore/soilgrids
Sand	• Global SoilGrids250m dataset	Secondary	ISRIC - World Soil Information	250 m	Decades to Centuries	https://www.isric.org/explore/soilgrids

Table 1 (cont'd)

Silt	• Global SoilGrids250m dataset	Secondary	ISRIC - World Soil Information	250 m	Decades to Centuries	https://www.isric.org/explore/soilgrids
Clay	• Global SoilGrids250m dataset	Secondary	ISRIC - World Soil Information	250 m	Decades to Centuries	https://www.isric.org/explore/soilgrids
Soil Organic Carbon	• Global SoilGrids250m dataset	Secondary	ISRIC - World Soil Information	250 m	30 years	https://www.isric.org/explore/soilgrids
Rainfall Erosivity	• Global Rainfall Erosivity Map	Secondary	European Soil Data Centre (ESDAC)	1 km	1 to 60 minutes	https://esdac.jrc.ec.europa.eu/content/global-rainfall-erosivity

The remotely sensed products and observation data used in this study are generally categorized into spatiotemporal and physiographical groups. The spatiotemporal variables were carefully selected to capture the environmental stressors that change over time and space. Meanwhile, the physiographical variables are limited to those that represent local conditions.

3.2.3.2.1 Spatiotemporal Variables

Spatiotemporal variables show the frequency and location of the change in landscape (Meng et al., 2019). The spatiotemporal variables that are considered in this study were calculated or observed, and these include Soil Moisture Index (SMI), Vegetation Index (VI), Growing Degree Days Index (GDDI), Water Demand Index (WDI), Soil Moisture Deficit Index (SMDI), Standardized Precipitation Index (SPI), Standardized Runoff Index (SRI), and Drought Severity Index (DSI).

Soil Moisture Index (SMI): SMI is derived using soil moisture observations. Soil moisture is an important aspect that helps to determine water, energy, and carbon fluxes globally (Das et al., 2011). This study used the SMAP/Sentinel-1 active-passive soil moisture product at 1 kmresolution, SMAP L2_SM_SP version 3 (Das et al., 2020), to derive the SMI. Since SMAP L2_SM_SP dataset has a revisit interval of about 6 to 12 days (Das et al., 2019), we found approximately full coverage over the entire Senegal within 12 days. Thus, all available SMAP L2_SM_SP coverage over Senegal within 12 days were composite to estimate SMI using Eq.1 (Singh & Das, 2022) as:

$$\mathrm{SMI} = 10 - \left(\frac{SM_t}{SM_{max}}\right) \times 10 \tag{1}$$

where, SM_t is the soil moisture for day *t*, and SM_{max} is the saturated water content or soil porosity. The soil's particle density (ρ_{particle}) and soil bulk density (ρ_{bulk}) are used to compute the SM_{max} by Eq. 2:

$$SM_{max} = (1 - \rho_{bulk}/\rho_{particle}) \tag{2}$$

where, SMI is a dimensionless indicator in the range of 0–10, here, lower soil moisture values produce higher SMI (i.e., high water stress condition) and vice-versa.

Vegetation Index (VI): Vegetation indices like NDVI, Vegetation Condition Index (VCI), and Vegetation Health Index (VHI) have been widely used to study the amount and health of vegetation over time (Kogan et al., 2012; Shammi & Meng, 2021). This study used NDVI-derived vegetation index (VI) to characterize the amount of vegetation at a specific time period and location. The SPOT/VEGETATION, PROBA-V NDVI version 2.2 product at 1 km resolution is used to compute the VI (Toté et al., 2020). The VI for each day is a dimensionless indicator in the range of 0–10, which is calculated by dividing NDVI_t (NDVI of a particular day) by NDVI_{max} (the maximum NDVI value in a 20-year period (2000 – 2020), and then multiplying by 10, as per the Eq. 3 given by Singh & Das (2022):

$$\text{VI} = \left(\frac{NDVI_t}{NDVI_{max}}\right) \times 10 \quad (3)$$

The higher the NDVI, the higher the VI (i.e., high amount of vegetation), and vice-versa.

Growing Degree Days Index (GDDI): GDD is used to determine biological activities throughout a crop's development and phenology stages, which are expressed as heat units (Mcmaster & Wilhelm, 1997), derived by subtracting the base temperature, which is the temperature below which plant development activity becomes insignificant, from the average temperature (Singh & Das, 2022). The Copernicus ECMWF ERA5 Hourly Temperature dataset is used to compute the GDD (Hersbach et al., 2023). The GDDI is calculated using Eq. 4 as follows (Singh & Das, 2022):

$$\text{GDDI} = \frac{GDD}{Temperature\ Growth\ Range} \times 10 \quad (4)$$

where, GDDI is the daily GDD, and the *Temperature Growth Range* is the optimum temperature for crop development, calculated as the difference between the lower and upper base temperatures (T_{base}). Major cereal crops have an average temperature growth range of 10 degrees Celsius to 30 degrees Celsius (Anandhi, 2016; Singh & Das, 2022). However, based on the crop growth characteristics in Senegal, we used a temperature growth range of 12 degrees Celsius to 35 degrees Celsius for millet (ICRISAT, 1984). GDDI is a dimensionless index, which varies between 0 and 10, where a GDDI of 10 represents the highest potential heat unit (high crop water stress condition) to support plant activities.

Water Demand Index (WDI): The Water Demand Index is a dimensionless product that measures the amount of water required for crops, which can help identify areas where water is limited and potentially cause stress on agricultural production. WDI can be calculated based on factors influencing water demand, such as temperature, solar radiation, and atmospheric humidity (Mishra & Singh, 2013). As with prior study (Singh & Das, 2022), we have incorporated GDDI, VI, and SMI into the calculation of WDI (Eq. 5), to provide information on crop growth and development, vegetation cover, and soil moisture conditions and improve the accuracy of yield prediction:

$$\mathrm{WDI} = \mathrm{SMI}\ \times \mathrm{VI}\ \times \mathrm{GDDI} \tag{5}$$

A high value of WDI represents a high water demand, which indicates water stress conditions; a low WDI represents a low water demand, indicating the absence of water stress conditions. The theoretical maximum value of WDI is 1000, and the minimum is zero.

Soil Moisture Deficit Index (SMDI): A widely employed measure for evaluating agricultural drought is the Standardized Moisture Deficit Index (SMDI), as introduced by Narasimhan and Srinivasan in 2005. The SMDI for the study period was collected from the Senegal Drought and Crop Watch System, which was derived using root-zone soil moisture simulations of the well-known Variable Infiltration Capacity (VIC) model (Feed the Future, 2023). The SMDI used in this study ranges from -4 to +4, representing dry and wet conditions, respectively.

Standardized Precipitation Index (SPI): The SPI is a well-known meteorological drought index developed by McKee et al. (1993) that measures the magnitude of short and medium-term drought based on the amount of precipitation within a specific timescale: 3, 6, 9, 12, 24, and 48 months. This study used SPI3, which was derived using three months of precipitation to characterize short- and medium-term water deficit/excess conditions, primarily important in the context of agriculture. We used Climate Hazards Group InfraRed Precipitation with Station (CHIRPS) data at 5 km spatial resolution to estimate SPI3, which was further re-scaled to 1 km resolution. The SPI3 indicates how much above or below the median the 3-month precipitation is, ranging from -3 to +3, where positive values of the SPI3 imply precipitation greater than the median, which means there is no drought, and negative values imply that precipitation is lower than the median, and therefore there is a drought (Kubiak-Wójcicka & Juśkiewicz, 2020; Svoboda et al., 2012).

Standardized Runoff Index (SRI): The SRI is a drought index that is widely used to study hydrological drought, which is expressed as the concept employed by (McKee et al., 1993) for the SPI in defining SRI as the unit standard normal deviate linked to the percentile of hydrologic runoff accumulated over a specific period (Kubiak-Wójcicka & Juśkiewicz, 2020; Shukla & Wood, 2008). This study used the SRI3, developed on a 3-month timescale in the range of - 3 to +3. The positive values of the SRI3 indicate the absence of drought, whereas negative values of the SRI3 indicate the presence of drought (Kubiak-Wójcicka & Juśkiewicz, 2020). The VIC model simulated SRI values for the study period were obtained from then Senegal Drought and Crop Watch System at 5 km resolution, which was further re-scaled at 1 km resolution (Feed The Future, 2023).

Drought Severity Index (DSI): The DSI indicates the agricultural drought severity derived using root zone soil moisture expressed as a percentile of the soil moisture climatology as described in Andreadis et al. (2005).

DSI is expressed as a percentage between 0 and 100%, where values close to 100% represent extreme drought (water stressed) conditions. The VIC model simulated DSI was obtained from the Senegal Drought and Crop Watch System at 5 km resolution, which was further re-scaled at 1 km resolution using a linear interpolation technique (Feed The Future, 2023).

The SPI3, SRI3, SMDI, and DSI were obtained from the Regional Hydrological Extreme Assessment System (RHEAS), a software framework that streamlines water resources simulations by assimilating diverse datasets, including satellite observations (Andreadis et al., 2017). The core of RHEAS is a spatially enabled PostgreSQL/PostGIS database that automatically ingests various datasets. The software allows for nowcast and forecast simulations, and in this study, it demonstrated its ability to produce data products in different contexts.

2.3.2.2 Physiographical Variables

Physiographical variables are the environmental characteristics that affect plant growth, processes, and physiology. It is very important to note that crop yield is not only affected by natural occurrences like drought and floods but also by existing characteristics of the soil on which the crops are grown. This study has incorporated the critical physiographical variables for crop growth such as Nitrogen Content; Rainfall Erosivity; Sand, Clay, and Silt contents; and Soil Organic Carbon.

Nitrogen Content: Among the major nutrients essential for plant growth (e.g., nitrogen, phosphorus, potassium), nitrogen is vital in crop growth and development (Hofman & Cleemput, 2004). In addition, the proper application of nitrogen-rich fertilizer as a part of agricultural management practices can improve crop yield (Smil, 2011). Nitrogen goes through a series of processes and exists in different forms in the soil, but plants can only uptake the mineral form of nitrogen (nitrate and ammonium) (Robertson & Vitousek, 2009). In this study the soil nitrogen content (cg/kg) was obtained from the Global SoilGrid250 m dataset (Hengl et al., 2017).

Sand, Clay, and Silt Contents: It is very important to know the soil texture information of a particular region's sand, silt, and clay content, as this determines the type of crop grown in that region. Different crops have different growth in different soil textures; for example, maize does well in sandy loam soils, whereas rice does well in soils with more clay and silt (USA Rice, 2020). On the other hand, millet grows well in well-drained loamy soils (Oelke et al., 1990). Incorporating this element in our study will ensure that the soil type for a particular region is considered in the estimated crop yield. The soil texture information is reported in (g/kg) and was obtained from the Global SoilGrid250 m dataset (Hengl et al., 2017).

Soil Organic Carbon (SOC): Soil Organic Carbon is the carbon component in the soil's organic matter (Department of Agriculture and Food-Western Australia, 2014). The changes in SOC happen after many years, and the change is usually very small, primarily observed in the topsoil. However, it is still important to quantify the SOC as it is vital to know how much carbon stocks are available and to identify any effects a management practice may have. Soils with higher amounts of SOC tend to have better aeration, drainage, microbe support, and resistance to erosion, and are often associated with improved yields (Corning et al., 2016). This study obtained SOC (dg/kg) from the Global SoilGrid250m dataset (Hengl et al., 2017).

Rainfall Erosivity: Rainfall erosivity is the ability of rainfall to cause soil loss downslope, and its estimation is based on the storm energy and intensity (Nearing et al., 2017). Soil erosion caused by water is regarded as the primary cause of soil degradation (Panagos et al., 2017), so it is vital to identify the erosivity level in a particular region, as this affects soil fertility and crop yield. For this study, the rainfall erosivity (mm ha^{-1} h^{-1} yr^{-1}) dataset of Senegal was obtained from the Global Rainfall Erosivity map at 1 km spatial resolution developed by (Panagos et al., 2017).

3.2.3. Data Analysis

3.2.3.1 Bias Correction

After conducting an extensive analysis of crop cultivation in every department of Senegal between 2016 and 2020, we found that multiple crops were grown in the same department and could not be differentiated during data acquisition. As a result, the remotely sensed data, particularly the NDVI, which measures the reflectance of vegetation in a given area, did not distinguish millet from other crops. This led to elevated NDVI readings, even during millet harvesting periods, which were attributed to other crops grown in the same department as millet. Moreover, the growing vegetation could have influenced soil moisture readings, which are essential in calculating the WDI. To reduce the margin of error, we identified the exact size of the cultivated area where millet was grown in each department and established a Bias Correction factor (Tables S1 to S3). By considering only the area in which millet is grown, we can more accurately compute the WDI with respect to the planting, growing, and harvesting periods. This correction factor will improve the accuracy of similar calculations and provide a more precise understanding of the agricultural conditions in Senegal.

3.2.3.2 Exploratory Analysis: Spatiotemporal and Physiographical Variable Analysis

We aim to investigate the individual effects of spatiotemporal and physiographic variables on millet crop yields as well as the relationships between these variables using Spearman's rank correlation and linear mixed effects model, as they have different strengths and weaknesses. With Spearman's rank correlation, we examined the correlation between all remote sensing indices and crop yield to understand their annual variabilities. Meanwhile, we statistically identified the most relevant indices to crop yield regardless of temporal variability with the linear mixed effects model.

By investigating the independent effects of these factors, we can better understand how they influence millet crop yields.

Data were analyzed using the R statistical software (R Core Team, 2022). We employed a linear mixed effects model to test the effect of independent variables (WDI, SMI, GDDI, VI, SMDI, SRI3, SPI3, DSI, sand content, silt content, clay content, nitrogen content, SOC, and rainfall erosivity (during the planting, growing, and harvesting stages of plant growth) on millet yield using the package *'lmerTest"* (Kuznetsova et al., 2017). Independent variables were considered fixed effects, and departments and years were considered random effects. All independent variables were continuous, and we standardized the variables (mean centered at 0 and standard deviation of 1) because the means of some variables were different by many orders of magnitude, leading to scaling issues.

Model residuals were assessed visually by plotting model residuals against predicted values to assess constant variance, and the responses were log-transformed to meet the model assumption. Additionally, a QQ-plot was used to visually evaluate the normality of the model residuals, and it was found that the model satisfied the assumptions of normality. Outliers were identified using the function *outlierTest* (Fox & Weisberg, 2019). Significant mean shifting studentized residuals were identified based on Bonferroni adjusted *p* values and were removed from the model. Next, a backward elimination stepwise regression was used to identify candidate predictors of millet yield using package *"stats"* (R Core Team, 2022). Starting with the current model, including all predictors, one predictor was dropped at a time, and AIC values were computed for each model. The candidate model with the lowest AIC was used as a final model. Partial residuals plotted against each significant predictor suggested that higher-order polynomial regression was unnecessary.

Each significant predictor of millet yield was then categorized into four different categories based on quantiles to assess mean differences in millet yield. The model with categorical predictors was subjected to ANOVA. Post hoc mean separation tests were performed on significant predictors based on tukey adjusted *p* values at the alpha level of 0.05 using the *"emmeans"* package (Lenth et al., 2023). Finally, back-transformed estimated marginal means are reported.

3.2.3.3 Predictive Analysis: Crop Yield Estimations

In this task, we aimed to predict millet yield using the spatiotemporal and physiographical features dataset. The task started with preprocessing, in which we applied *MinMax* and *Robust* scaling methods, respectively, to the input and output variables to ensure that the differences would not influence our model in the scale of the features (Pedregosa et al., 2011). Then to find the best hyperparameters for the ML models and to evaluate their performance, we used nested cross-validation with the following configurations: 1) Outer cross-validation loop: This loop is responsible for estimating the model's performance. We divided the dataset into multiple, equally sized folds (in our case, 5). In each iteration of the loop, one of these folds was set aside as "test set," while the remaining folds were used for training the model. This process was repeated until each fold had been used as the test set once. The average performance across all iterations provided an unbiased estimate of the model's performance and 2) Inner cross-validation loop: This loop is responsible for selecting the best hyperparameters for the model. For each iteration in the outer loop, we performed another cross-validation process on the training data (again using 5 folds). This time, however, we used a grid search to try different combinations of hyperparameters for our ML models. The hyperparameter combination that resulted in the best average performance was chosen as the best set of hyperparameters for that iteration.

After finding the best hyperparameters using nested cross-validation, we trained the model on the entire dataset and saved it for further model evaluation. We compute the error indexes, including the Nash–Sutcliffe model efficiency coefficient (NSE), the coefficient of determination (R^2), Root mean Square Error (RSME), the normalized root mean squared error by variable range (nRMSE), and Willmot's index of agreement (d) for each iteration of the outer loop and then average them for evaluating the model performance. We repeat the above-mentioned procedure to select the best unbiased model among eight regression models for millet yield prediction. The models that we selected with their hyperparameter configurations are presented in Table 2.

Table 2. Comparison of regression models and their hyperparameter configurations

Model Name	Parameters Configuration for Grid Search	Hyperparameters Description	Model Reference
Linear Regression	N/A	No hyperparameters	Galton (1886)
Ridge Regression	alpha: logspace (-4, 4, 9)	alpha: Regularization strength (L2 penalty)	(Hoerl & Kennard, 1970)
Lasso Regression	alpha: logspace (-4, 4, 9)	alpha: Regularization strength (L1 penalty)	(Tibshirani, 1994)
Elastic Net Regression	alpha: logspace (-4, 4, 9), l1_ratio: linspace (0.1, 1, 10)	alpha: Regularization strength (L1 and L2 penalty), l1_ratio: The mix between L1 and L2 penalty	(Zou & Hastie, 2005)
Partial Least Squares Regression	n_components: range (1, 11)	n_components: Number of components to keep in the reduced dimensionality space	(Wold et al., 2001)
Support Vector Regression	C: logspace (-3, 3, 7),	kernel: Kernel type, C: Penalty parameter,	(Drucker et al., 1996)

Table 2 (cont'd)

	epsilon: logspace (-3, 3, 7), kernel: ['linear', 'rbf']	epsilon: Tube width within which no penalty is applied on the training data	
Random Forest Regression	n_estimators: [50, 100, 200], max_depth: [None, 10, 20], min_samples_split: [2, 5, 10], min_samples_leaf: [1, 2, 4], max_features: ['sqrt']	n_estimators: Number of trees, max_depth: The maximum depth allowed for trees. min_samples_split: The minimum number of samples required to split a node. min_samples_leaf: The minimum number of samples required at each leaf node. max_features: The number of features to consider when searching for the best split	(Breiman, 2001)
Gradient Boosting Regression	n_estimators': [50, 100, 200], max_depth': [3, 6, 9], min_samples_split': [2, 5, 10], min_samples_leaf': [1, 2, 4], max_features': ['sqrt'], learning_rate': [0.01, 0.1]	n_estimators: The number of boosting stages. max_depth: The maximum depth allowed for trees. min_samples_split: The minimum number of samples required to split a node. min_samples_leaf: The minimum number of samples required at each leaf node. max_features: The number of features to consider when searching for the best split. learning_rate: The	(Friedman, 2001)

Table 2 (cont'd)

	learning rate, also known as shrinkage.

3.2.3.4 Exploratory Analysis: Permutation Importance Analysis

Permutation importance is a method for assessing the relative significance of features in an ML model by quantifying the change in model performance when a particular feature's values are randomly shuffled (Breiman, 2001). The underlying principle is that if a feature is crucial, scrambling its values should significantly impact the model's performance, while shuffling an insignificant feature should have minimal or no effect. In some cases, an increase in model performance after shuffling a feature's values may occur, indicating that the feature introduces noise to the model, thereby degrading its performance. Overall, this technique offers an intuitive way to comprehend the importance of each feature in a model by directly examining the consequences of shuffling the feature values on the model's performance.

Herein, we employ the *Scikit-learn* library (Pedregosa et al., 2011) to compute permutation importance. Using the fitted ML model, we evaluated each feature by randomly shuffling its values and then assessing the model's performance on the shuffled dataset with the same scoring function utilized during model training and cross-validation. Subsequently, the change in performance is compared to that of the original performance. This process was repeated ten times for each feature, and the average performance change was calculated and reported as feature importance.

3.3 Results and Discussion

3.3.1 Explanatory Analysis: Spatiotemporal and Physiographical Variable Analysis

3.3.1.1 Assessing the Effect of Temporal Variability on the Relationship between Crop Yield and Remote Sensing Indices using Spearman's Rank Correlation

The Spearman's Rank correlation coefficients between millet yield and different spatiotemporal and physiographical variables for all departments in Senegal are presented in Figure 3.

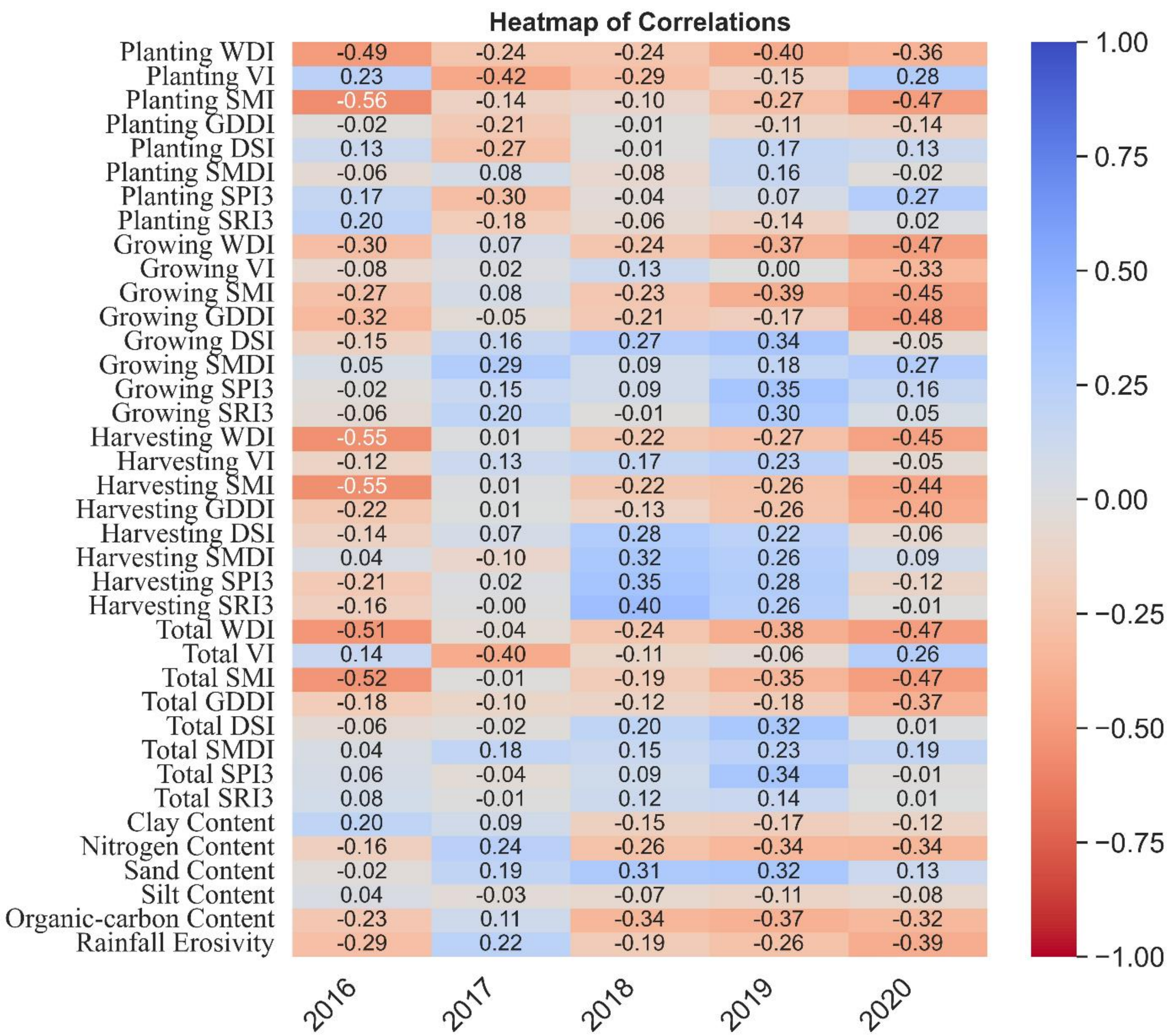

Figure 3. Spearman's Rank Correlation Coefficient analysis of millet yield and spatiotemporal/physiographical variables across departments in Senegal

Planting indices: Specific planting indices, namely WDI, SMI, and GDDI, were discovered to have a negative impact on millet crop yield due to increased water demands, unsuitable vegetation conditions, elevated temperature, and significant soil moisture deficits during planting. However, the remaining planting indices exhibited inconsistent trends, indicating complex relationships that require further investigation. This inconsistency was especially noticeable in 2016 and 2020, which corresponded to the rainy planting period as indicated by drought indices.

Growing indices: The association between millet crop yield and different growing indices is more intricate and diverse than planting indices. Certain indices primarily exhibit a negative correlation with crop yield (namely WDI, SMI, and GDDI), whereas others manifest irregular patterns, and a few are predominantly positively linked. These findings imply that the impact of these indices on crop yield throughout the growing period is contingent on specific contextual factors and other pertinent variables.

Harvesting indices: During the harvesting period, there is a lack of a discernible *p* trend in the indices' correlation. Nonetheless, WDI and SMI exhibit a robust negative correlation in the earlier and later years, indicating that increased values of these variables could have resulted in reduced yields. This timeframe corresponds to the dry harvesting season, as evidenced by the presence of negative drought indices. Additionally, all drought indices (DSI, SMDI, SPI3, and SRI3) displayed a considerable positive correlation in 2018 and 2019, implying that elevated values of these variables could have contributed to high millet yields.

Physiographical indices: The correlation coefficients between various physiographical indices and millet yield showed fluctuating trends from 2016 to 2020 that are harder to explain as the values of these indices are constant during the study period. This suggests that it is crucial to account for spatiotemporal variations alongside physiographical indices when characterizing the anomalies and fluctuations in millet yield. Over the period of three years (2018-2020), various indices, such as clay and silt content, soil organic carbon, and rainfall erosivity, exhibited predominantly negative correlation coefficients. However, there was an exception with sand, which demonstrated a general increasing trend between 2016 and 2019, followed by a decrease in 2020.

Overall summary: The study shows that the relationship between indices and millet crop yield is complex and varies depending on the stage of crop development. Overall, the correlations vary from -0.56 (Planting SMI) to +0.40 (Harvesting SRI3). However, regardless of the growth stage, all variables describing crop water demands (WDI, VI, SMI, and GDDI) are negatively correlated with yield. In contrast, all drought indices (DSI, SMDI, SPI3, and SRI3) are positively correlated with millet yields indicating no drought conditions. Nonetheless, when examining the temporal trends, the results show the occurrence of both wet and dry years during the study period. The physiographical parameters are negatively correlated with yield except for the sand content. Meanwhile, the temporal element adds more complexity and inconsistent trends. Therefore, additional analysis is required to enhance the comprehension of these relationships, accounting for other influential factors and adopting modeling techniques that accommodate complex and nonlinear relationships.

3.3.1.2 Assessing the Relationship between Crop Yield and Remote Sensing Indices Using Spearman's Rank Correlation

Stepwise regression through backward elimination following a linear mixed effect model showed that VI, GDDI, SPI3, and SRI3 in the planting period, VI in the growing period, WDI, VI, SMI, GDDI, SMDI, SPI3, and SRI3 in the harvesting period, and rainfall erosivity and soil organic carbon unvarying of time had a significant effect on millet yield (Table S4). Below is the yield prediction model with the lowest AIC in which 78.6% of the variation in the millet yield was explained by the following equation (Eq. 6):

$$yield_i \sim N\left(\alpha_{j[i],k[i]}, \delta^2\right) \tag{6}$$

$$\begin{aligned} \alpha_j \sim N\,(&\beta_0^\alpha + \beta_1^\alpha\,(Planting\ VI) + \beta_2^\alpha\,(Planting\ GDDI) + \beta_3^\alpha\,(Planting\ SRI3) \\ &+ \beta_4^\alpha\,(Planting\ SPI3) + \beta_5^\alpha\,(Growing\ VI) + \beta_6^\alpha\,(Harvesting\ WDI) \\ &+ \beta_7^\alpha\,(Harvesting\ VI_{harvesting}) + \beta_8^\alpha\,(Harvesting\ SMI) \\ &+ \beta_9^\alpha\,(Harvesting\ GDDI) + \beta_{100}^\alpha\,(Harvesting\ SMDI) \\ &+ \beta_{111}^\alpha\,(Harvesting\ SPI3) + \beta_{122}^\alpha\,(Harvesting\ SRI3) \\ &+ \beta_{133}^\alpha\,(Rainfall\ Erosivity) + \beta_{144}^\alpha\,(SOC), \delta_{\alpha j}^2\ for\ Departments\ j = 1, \dots, J) \end{aligned}$$

$\alpha_k \sim N(\mu_{\alpha k}, \delta^2{}_{\alpha k})$, for year $k = 1, \dots, K$

where, all variables were modeled as predictors, and the year and department were modeled as random effects. j and k are department and year, respectively; N is a normal or Gaussian distribution, α is a random effect representing the effect of the department (j) or year (k) on yield, β is the model coefficient, μ is the mean value of the random effects, and δ^2 is unaccounted variation, which is the overall variance of the random effects not explained by the model.

In general, some indices have a positive relationship with planting SPI3, growing VI, harvesting WDI, harvesting SRI3, and Rainfall Erosivity. In contrast, others have a negative relationship, like planting VI, planting GDDI, planting SRI3, harvesting VI, harvesting SMI, harvesting GDDI, harvesting SMDI, harvesting SPI3, and SOC.

The results of our analysis of variance indicate that certain variables have a strong correlation with millet yield. Specifically, during the planting period, SPI3 and SRI3, and during the growing period, VI, as well as SOC, showed a significant correlation with millet yield at the 0.1% significance level. SPI3 values above the median indicate above-average precipitation, while high SRI3 values indicate high hydrologic runoff, suggesting the absence of drought according to both indices. In addition, we observed that high values of planting SPI3 and harvesting SRI3 were associated with higher values of millet yields; however, the confidence interval for planting SPI3 (0.43 – 1.33) is wider than the confidence interval for growing VI (0.15 – 0.50), demonstrating greater uncertainty in the effect of planting SPI3 on millet yield.

The planting period is crucial for millet as it includes plant emergence and early stages of development when adequate moisture is necessary for germination and early growth. The young plants are sensitive to dry spells, and high runoff levels can cause erosion, making them fragile. Our findings align with Fall et al. (2021), who reported that very high precipitation could cause crop damage when associated with runoff, which explains why high values of SRI could be associated with low yield, as very high precipitation may have resulted in high surface runoff which could easily erode germinating plants. We also found that VI, derived from NDVI, had a strong correlation with millet yield during the growing period (August), consistent with Fall et al. (2021) and Panek & Gozdowski (2020), who reported the highest correlation between NDVI and cereal yield during the early stages of crop development.

Our findings indicate a relationship between SOC levels and millet yields, with moderate SOC levels being associated with higher millet yields. Conversely, elevated SOC levels are correlated with reduced millet yields. This observation agrees with Oldfield et al. (2019), who reported a decline in maize yield when SOC concentrations exceeded 2%. The consistency between these studies underscores the potential influence of SOC on crop productivity and the importance of optimizing SOC levels for various crop types (Oldfield et al., 2019).

3.3.2 Predictive Analysis: Crop Yield Estimations

3.3.2.1 Predictive Crop Yield Model Selection

Based on the results from Table 3, it is evident that the Random Forest Regressor model has the best predictive performance, as it has positive scores for all metrics (i.e., NSE, R^2, RMSE, nRMSE, and *d*). This indicates that the model can predict the millet yield more reliably compared to the other ML models after extensive nested 5-fold cross-validation. Prasad et al. (2021), Jeong et al. (2016), and Sakamoto (2020) found similar results, where Random Forest Regressor model performed reliably and faster in predicting cotton yield (Prasad et al., 2021), wheat yield (Jeong et al., 2016), and soybean yield (Sakamoto, 2020). On the other hand, the Linear Regression model has the worst predictive performance, as it has negative scores for all metrics, which is also similar to what Jeong et al. (2016), where the Linear Regression model performed poorly compared to Random Forest Regressor model. This indicates that the model is not able to predict the millet yield accurately. The other models have mixed results, with some performing better than others depending on the metric. Therefore, the Random Forest Regressor model is adopted for further investigations.

Table 3. Average performance and the best-unbiased hyperparameters

Regression Model	Optimized Hyperparameters				
	NSE	RMSE	nRMSE	*d*	
Linear Regression	-10.495	1.417	0.539	0.514	No parameters to tune for Linear Regression.
Ridge Regression	-0.022	0.597	0.245	0.626	alpha=0.0001
Lasso Regression	-0.016	0.594	0.244	0.625	alpha=0.0001, max_iter=10000

Table 3 (cont'd)

ElasticNet Regression	-0.018	0.595	0.244	0.625	alpha=0.0001, l1_ratio=0.9, max_iter=10000
Partial Least Squares	0.055	0.583	0.238	0.591	n_components=1
Support Vector Regression	-0.046	0.618	0.251	0.606	C=10, kernel='linear'
Random Forest Regressor	0.161	0.554	0.224	0.574	max_features='sqrt', n_estimators=50, random_state=42
Gradient Boosting Regressor	0.014	0.588	0.24	0.617	max_depth=6, max_features='sqrt', min_samples_leaf=4, random_state=42

3.3.2.2 Predictive Crop Yield Model Implementation and Evaluation of its Performance

Table 4 and S5 present the results of the Random Forest Regressor models, which were trained using different sets of independent variables. The models' performance was evaluated using several statistical indices, including NSE, R^2, d, RMSE, and nRMSE. The best model was selected based on the highest NSE, R^2, and *d* values and the lowest RMSE and nRMSE values. According to this criterion, the best model was identified as the one using all independent variables except Department, Harvesting, Total, Growing, Physiographical, and Planting SMI, VI, and GDDI (Figure 4a). This model exhibited the highest NSE (0.831), the lowest RMSE (0.308) and nRMSE (0.077), and the highest R^2 (0.831) and *d* (0.942) values. Conversely, the worst-performing model used all independent variables except Bias Correction factor and Department (Figure 4b), which had the lowest NSE (0.730), the highest RMSE (0.391) and nRMSE (0.098), and the lowest R^2 (0.730) and *d* (0.896) values. The exclusion of certain variables, including Department, Harvesting, Total, Growing, Physiographical, and some climate-related variables (SMI, VI, and GDDI) during the planting period, improved the Random Forest performance model. These variables may have had less influence on the response variable, or their exclusion may have reduced multicollinearity or noise in the data (Schroeder, 1990; Senaviratna & Cooray, 2019). However, these findings are dataset-specific and may not necessarily apply to other crops and regions.

Table 4. The best-performing Random Forest Regressor models for different sets of variables

Independent Variables	NSE	RMSE	nRMSE	R^2	*d*
All variables except Bias Correction Harvesting, Total, Growing, and Physiographical	0.781	0.351	0.088	0.781	0.921
All variables except Bias Correction, Department, Harvesting, Total, Growing, Physiographical, and Planting SMI, VI, and GDDI	0.796	0.340	0.085	0.796	0.928
All variables except Bias Correction	0.756	0.371	0.093	0.756	0.909
All variables except Bias Correction and Department	0.730	0.391	0.098	0.730	0.896
All variables except Bias Correction, Department, and Physiographical	0.729	0.391	0.098	0.729	0.900
All variables except Bias Correction, Harvesting and Total	0.790	0.344	0.086	0.790	0.925
All variables except Bias Correction, Harvesting, Total, and Growing	0.798	0.337	0.085	0.798	0.928
All variables except Bias Correction, Department, Harvesting, Total, and Growing	0.767	0.362	0.091	0.767	0.915
All variables except Bias Correction, Department, Harvesting, and Total	0.735	0.387	0.097	0.735	0.902
All variables except Bias Correction and Total	0.756	0.371	0.093	0.756	0.907
All variables except Bias Correction, Department and Total	0.735	0.387	0.097	0.735	0.901
All variables except Harvesting, Total, Growing, and Physiographical	0.794	0.341	0.086	0.794	0.926
All variables	0.763	0.366	0.092	0.763	0.916
All variables except Department	0.740	0.383	0.096	0.740	0.900
All variables except Department and Physiographical	0.746	0.378	0.095	0.746	0.909
All variables except Harvesting and Total	0.793	0.342	0.086	0.793	0.925
All variables except Harvesting, Total, and Growing	0.820	0.319	0.080	0.820	0.938
All variables except, Department, Harvesting, Total, and Growing	0.795	0.340	0.086	0.795	0.927
All variables Department, Harvesting, and Total	0.768	0.362	0.091	0.768	0.916
All variables except Total	0.766	0.363	0.091	0.766	0.912

Table 4 (cont'd)

All variables except Department and Total	0.746	0.379	0.095	0.746	0.904
All variables except Department, Harvesting, Total, Growing, Physiographical and Planting SMI, VI, and GDDI	0.831	0.308	0.077	0.831	0.942

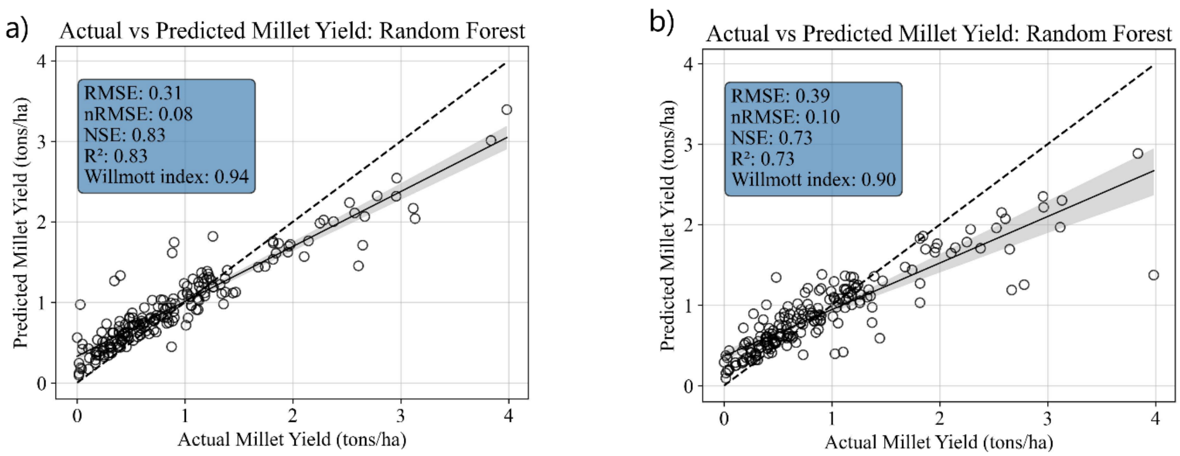

Figure 4. Comparing the predicted vs. actual millet yield for (a) the best- and (b) worst-performing Random Forest Regressor models

Meanwhile, the relatively lower statistical indices associated with the average-performing Random Forest Regressor models (refer to Table S5) could suggest the presence of other influential factors, currently unaccounted for or unidentified, that significantly impact millet yield predictions. These factors may not be entirely observable or quantifiable through the remote sensing products employed in this study.

3.3.2.3 Understanding the Key Predictors for Millet Crop Yield Using Permutation Importance Analysis:

Figure 5 illustrates the relative significance of the independent variables contributing to the optimal performance of the Random Forest Regressor model. The prominence of the WDI within the optimal model underscores the pivotal role of water availability in crop production, particularly during the planting phase. This suggests that any alterations in water accessibility, whether instigated by climatic extremes, farming management practices, or other elements, could profoundly affect millet yields. Singh and Das (2022) also found that WDI affected crop yield significantly (Singh & Das, 2022). The Bias Correction factor emerges as the second most influential variable. This high ranking infers substantial variability in millet cultivation areas across different departments, a factor of paramount importance when production data is only available in aggregated forms, such as at departmental or district levels.

The combined contribution of the indices related to drought and flood conditions (Planting SMI, Planting SPI3, Planting DSI, Planting SRI3, and Planting SMDI) accounts for approximately 49% of the variance, thereby signifying the substantial impact of extreme events on millet yield. Moreover, incorporating SPI3 and SRI3, representing precipitation and runoff data over three months, insinuates that recent climatic conditions during the planting phase can significantly influence overall crop yields. This is in line with other studies that have used these indices, SPI specifically, to monitor or predict yield (Mohammed et al., 2022; Tigkas et al., 2018). This underlines the necessity of incorporating temporal dynamics into agricultural modeling and planning. Consequently, identifying these crucial periods could aid in planning interventions or mitigation strategies to optimize yield.

In summary, these findings emphasize the intricate nature of agricultural yield predictions, which are influenced by a multitude of interconnected environmental factors. They further underscore the necessity of maintaining comprehensive, spatially and temporally explicit data pertaining to these factors to enhance the precision and applicability of these predictive models.

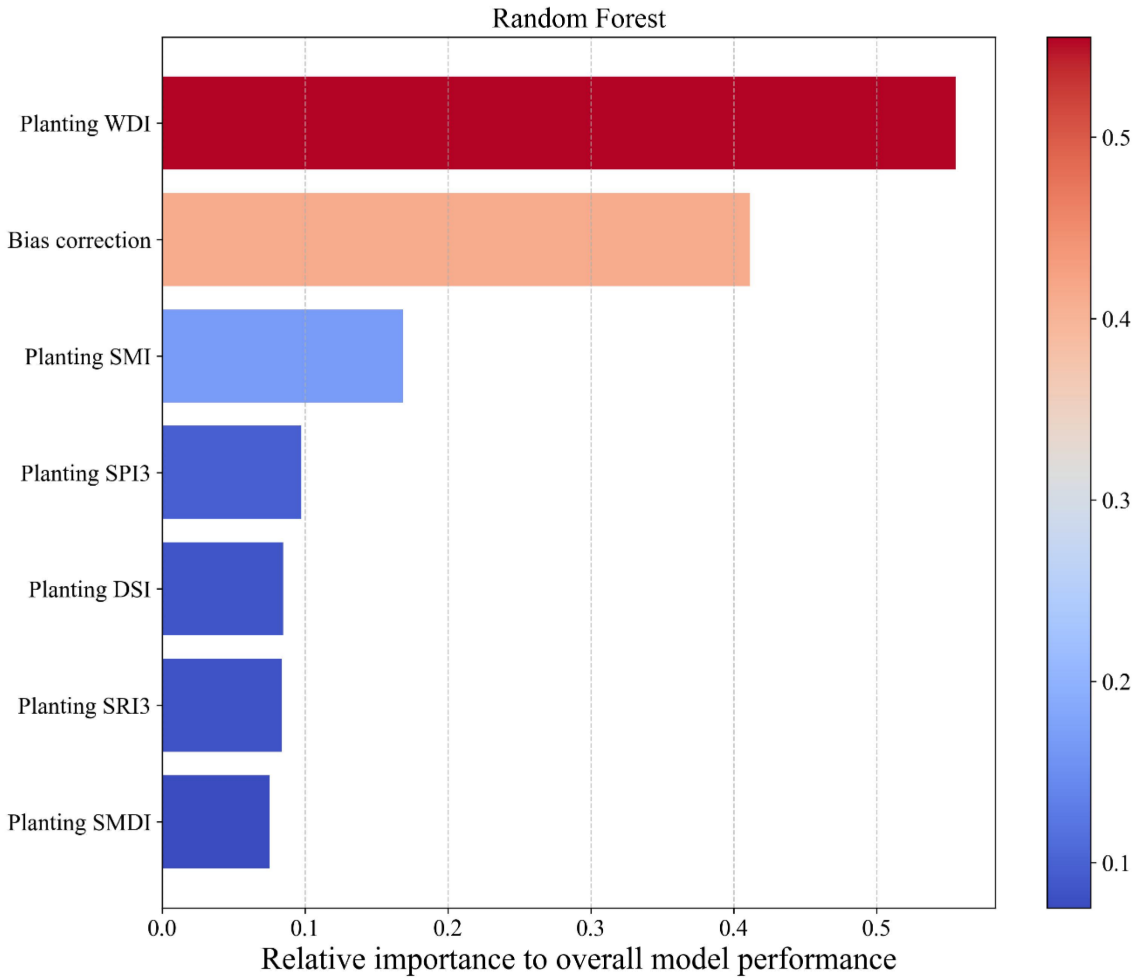

Figure 5. The relative importance of independent variables in the overall performance of the optimal Random Forest Regressor model

3.4 Conclusion

In conclusion, this study aimed to address the challenges of estimating crop yield in data-scarce regions such as Senegal by using technological advancements in remote sensing and ML. The research proposed three hypotheses to improve the accuracy of millet yield prediction; Initially, we assumed that integrating indices derived from satellite-based high-resolution soil moisture retrievals could boost the yield model predictive capability. Subsequently, we put forth the idea that a prediction model centered on distinct millet growth phases might yield greater efficacy compared to a model that relies on the average of all stages. Finally, we hypothesize that factoring in the millet production area at the departmental level during various growth stages can enhance the model's overall predictive accuracy. The findings of this research offer compelling support for the notion that integrating soil moisture-related indices, considering discrete millet growth stages, and incorporating a millet area Bias Correction factor can enhance the overall predictive performance of the model.

The Random Forest algorithm emerged as the most robust approach, outperforming other ML models and maintaining its superiority throughout the variable reduction process, which systematically minimized the set of independent variables to pinpoint the ideal combination for the optimal ML model. The analysis also highlighted the importance of water availability in seed germination and early plant development, as shown by SMI and WDI affecting model performance significantly during the planting season, as well as the significance of the Bias Correction factor in predicting millet yield. Indeed, the Bias Correction factor ranks as the second most significant variable in forecasting millet yields, highlighting its importance when production data is solely accessible in aggregated formats, such as departmental or district levels.

When it comes to maximizing millet yield, it is essential to consider the impact of drought and flood conditions. In fact, these factors account for almost half of the variance in yield. It is also important to pay attention to recent climatic conditions during the planting phase, as they can have a significant influence on overall crop yields. Specifically, SPI3 and SRI3 are key indicators to consider. By recognizing these crucial periods, we can better plan interventions and mitigation strategies to optimize yield. This underscores the need for agricultural modeling and planning that takes into account temporal dynamics.

The findings have significant implications for enhancing crop yield prediction in regions lacking adequate data and facilitating the advancement of more precise and efficient crop management strategies. The intricate and interconnected nature of environmental factors that impact crop yields highlights the importance of maintaining comprehensive data that is both spatially and temporally specific. This data is crucial in enhancing the precision and significance of predictive models, which can aid in creating interventions and mitigation tactics to optimize yield. Such measures will benefit both producers and consumers by promoting sustainable and productive agricultural practices.

4.0 OVERALL CONCLUSION

The United Nations Sustainable Development Goals (SDGs) aim to combat food insecurity by 2030. Nevertheless, numerous developing nations, Senegal included, are encountering challenging obstacles in pursuing this objective, primarily due to the emergence of various pressing issues, such as climate change and the COVID-19 global pandemic. The primary objective of this study was to assess the efficacy of cutting-edge technological advancements and products in addressing the adverse effects of these challenges. Specifically, our aim was to aid policy makers in formulating planning strategies at regional and national levels by leveraging advanced machine learning techniques and utilizing high-resolution spatial and temporal remote sensing products. The sections below summarize some of our findings:

- Soil moisture is critical for crop growth, and integrating these indices may significantly enhance yield prediction. Therefore, we hypothesized that using high-resolution satellite data to derive soil moisture-based indices can improve crop yield prediction, which was proven to be valid. Nevertheless, expanding the implementation of this approach to a national level necessitates government assistance, given that smallholder farmers may not possess the financial means and technical resources to implement it on an individual farm basis.
- In this study, we examined the hypothesis that a prediction model focused on individual stages of millet growth would outperform a model based on the average of all stages. The findings revealed that predictions based on individual stages produced superior results compared to predictions for the entire growing season. Notably, precipitation-based indices had a significant impact on crop yield prediction during the planting season, emphasizing the importance of water for germination and early plant development. As a result, it is

crucial to prioritize specific indices at different stages of millet growth, focusing on the most critical factors.

- Our hypothesis that including the area of millet production at the department level during different growth stages would improve the model's accuracy was proven to be correct. This is especially important as the current land use map cannot differentiate between different crops. Consequently, our dependence on agricultural census data becomes essential to augment the overall predictability of the model.

In light of the study's findings, there are several policy interventions and agricultural strategies that can be contemplated to enhance crop yield prediction and alleviate food insecurity in regions vulnerable to the effects of climate change and extreme events. The subsequent policies and interventions are presented as follows:

- National Policy: In light of the substantial computational and technical endeavors required to assess crop production on a national level, it is imperative for governmental bodies and humanitarian organizations to take action and provide assistance to small-scale farmers who might lack the necessary resources. In order to enhance the accessibility of modeling outcomes to the broader public, various dissemination techniques and platforms can be employed, supported by subsidies in the form of human and financial resources. It is recommended to ensure the availability of this information at local levels, such as Extension offices, as it facilitates the distribution of information and allows for model refinement. Extension officers, due to their direct engagement with smallholder farmers and familiarity with local data quality, possess valuable insights that can contribute to the accuracy and relevance of the model.

- Local Policy: Departments should be equipped with the necessary resources and expertise to analyze and interpret the collected data and model results. Moreover, Extension officers should possess comprehensive knowledge to effectively communicate advanced technologies to producers, provide guidance on suitable farming practices, and emphasize the crucial factors influencing crop yield. Achieving this can be facilitated through the implementation of workshops, training programs, or field days.
- Targeted Irrigation and Water Management: Considering the importance of water during germination and early plant development, policies focusing on irrigation and water management can be implemented, including using the WDI to assist in irrigation scheduling.
- Capacity Building and Training: In order to empower farmers to harness advanced agricultural technologies effectively, it is imperative to implement comprehensive training programs. These programs will focus on enhancing their capacity to interpret and utilize remote sensing indices and crop prediction results. By training them with this knowledge, farmers will be better equipped to make informed decisions and adapt their farming practices in alignment with the latest advancements. This policy aims to foster a proactive approach towards agricultural innovation and ensure sustainable growth in the farming sector.
- Research and Development Funding: To promote progress in agriculture, it is crucial to allocate funds towards research and development, with a specific focus on advancing remote sensing technologies for crop yield prediction. This investment will play a vital role in the advancement of more precise and efficient crop prediction models. Additionally, it will facilitate the identification of sustainable agricultural practices that can enhance

productivity while minimizing environmental impact. By prioritizing these endeavors, this policy seeks to drive innovation in agriculture and support the long-term sustainability of the sector.

In conclusion, this study demonstrates the value of incorporating soil moisture-based indices, considering individual stages of millet growth, accounting for production areas at the department level, and implementing bias correction factors in improving crop yield prediction. These findings provide valuable insights for policy makers and agricultural stakeholders, emphasizing the need for targeted interventions and technological support to enhance food security in regions vulnerable to climate change and extreme events. By implementing these policies and interventions, governments can support farmers, enhance agricultural productivity, and ultimately contribute to achieving the United Nations Sustainable Development Goal of combating food insecurity by 2030.

5.0 FUTURE WORK

Based on the findings of this study, there are several suggestions for future work that can help us better understand and use millet yield prediction models in the face of changing land use and climate conditions. These recommendations are as follows:

- Conducting long-term monitoring of soil moisture, climate variables, and crop yield data can provide valuable insights into the dynamic relationships between these factors. Long-term data collection and analysis will help identify trends, patterns, and potential changes in the relationships over time, enabling the development of more accurate and robust crop yield prediction models.
- While this study focused on millet yield prediction in Senegal, future research can expand the scope to include other crops and regions. Different crops may exhibit varying sensitivities to extreme climate factors, and understanding these dynamics for a range of crops will contribute to the development of comprehensive crop yield prediction models. Additionally, studying different regions with varying climatic conditions will provide insights into the transferability and generalizability of the findings.
- Consideration of socioeconomic factors, such as market prices, input costs, and agricultural policies, can provide a more holistic understanding of the factors influencing crop yield. In addition, integrating socioeconomic data into the crop yield prediction models can help identify additional drivers of yield variability and inform policy decisions to improve food security.
- Involving farmers in the research process through participatory approaches can enhance the relevance and applicability of crop yield prediction models. Engaging farmers in data

collection, model development, and decision-making processes will ensure that the models address their specific needs and can be effectively implemented at the farm level.

By pursuing these future research recommendations, we can advance our knowledge and practical applications in crop yield prediction, ultimately contributing to more effective agricultural planning, improved food security, and sustainable agricultural systems.

BIBLIOGRAPHY

Action Against Hunger. (2022). *SENEGAL*. https://www.actionagainsthunger.org/countries/africa/senegal

Aghelpour, P., Bahrami-Pichaghchi, H., & Kisi, O. (2020). Comparison of three different bio-inspired algorithms to improve ability of neuro fuzzy approach in prediction of agricultural drought, based on three different indexes. *Computers and Electronics in Agriculture*, *170*(February), 105279. https://doi.org/10.1016/j.compag.2020.105279

Agidew, A. meta A., & Singh, K. N. (2017). The implications of land use and land cover changes for rural household food insecurity in the Northeastern highlands of Ethiopia: The case of the Teleyayen sub-watershed. *Agriculture and Food Security*, *6*(1), 1–14. https://doi.org/10.1186/s40066-017-0134-4

Aiguo, D., & NCAR. (2019). *Palmer Drought Severity Index (PDSI) | NCAR - Climate Data Guide*. https://climatedataguide.ucar.edu/climate-data/palmer-drought-severity-index-pdsi

Alahacoon, N., & Edirisinghe, M. (2022). A comprehensive assessment of remote sensing and traditional based drought monitoring indices at global and regional scale. *Geomatics, Natural Hazards and Risk*, *13*(1), 762–799. https://doi.org/10.1080/19475705.2022.2044394

Alcamo, J., Flörke, M., & Märker, M. (2007). Future long-term changes in global water resources driven by socio-economic and climatic changes. *Hydrological Sciences Journal*, *52*(2), 247–275. https://doi.org/10.1623/hysj.52.2.247

Andreadis, K. M., Das, N., Stampoulis, D., Ines, A., Fisher, J. B., Granger, S., Kawata, J., Han, E., & Behrangi, A. (2017). The regional hydrologic extremes assessment system: A software framework for hydrologic modeling and data assimilation. *PLoS ONE*, *12*(5), 1–22. https://doi.org/10.1371/journal.pone.0176506

Arneth, A., Brown, C., & Rounsevell, M. D. A. (2014). Global models of human decision-making for land-based mitigation and adaptation assessment. *Nature Climate Change*, *4*(7), 550–557. https://doi.org/10.1038/nclimate2250

AWF. (2021). *How Does Water Use in the United States Compare to That in Africa?* https://www.awf.org/blog/how-does-water-use-united-states-compare-africa

Baba, A., Tayfur, G., Gunduz, O., Howard, K. W. ., Friedel, M. J., & Chambel, A. (2011). *Climate Change and its Effects on Water Resources*. Springer. https://link.springer.com/book/10.1007/978-94-007-1143-3

Breiman, L. (2001). Random Forests. *2021 IEEE International Conference on Consumer Electronics and Computer Engineering, ICCECE 2021*, 542–545. https://doi.org/10.1109/ICCECE51280.2021.9342376

Busby, J. W., Smith, T. G., & Krishnan, N. (2014). Climate security vulnerability in Africa mapping 3.01. *Political Geography*, *43*, 51–67. https://doi.org/10.1016/j.polgeo.2014.10.005

Chadburn, B. (2020). 5 Common Challenges with Remote Sensing and How to Tackle Them. *Northwestern Analystics*. https://www.northwesternanalytics.com/2020/11/16/5-common-challenges-with-remote-sensing-and-how-to-tackle-them/

Choi, M., Jacobs, J. M., Anderson, M. C., & Bosch, D. D. (2012). Evaluation of drought indices via remotely sensed data with hydrological variables. *Journal of Hydrology*, *476*, 265–273. https://doi.org/10.1016/j.jhydrol.2012.10.042

Chong, H., & Sunding, D. (2006). Water markets and trading. *Annual Review of Environment and Resources*, *31*, 239–264. https://doi.org/10.1146/annurev.energy.31.020105.100323

CIAT;BFS/USAID. (2016). Climate-smart agriculture in Senegal. CSA country profiles for Africa series. *International Center for Tropical Agriculture*, 20. https://climateknowledgeportal.worldbank.org/sites/default/files/2019-06/SENEGAL_CSA_Profile.pdf

CIESIN. (2000). *What is Sustainable Land Management?* http://www.ciesin.org/lw-kmn/slm/slm.html

Cooper, M. W., Brown, M. E., Hochrainer-Stigler, S., Pflug, G., McCallum, I., Fritz, S., Silva, J., & Zvoleff, A. (2019). Mapping the effects of drought on child stunting. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(35), 17219–17224. https://doi.org/10.1073/pnas.1905228116

CSSR. (2016). *Highlights of the U . S . Global Change Research Program Climate Science Special Report*. 10–34. https://doi.org/10.7930/J0DJ5CTG.U.S.

Dai, M., Huang, S., Huang, Q., Leng, G., Guo, Y., Wang, L., Fang, W., Li, P., & Zheng, X. (2020). *Assessing agricultural drought risk and its dynamic evolution characteristics _ Enhanced Reader.pdf*.

Danandeh Mehr, A., Rikhtehgar Ghiasi, A., Yaseen, Z. M., Sorman, A. U., & Abualigah, L. (2022). A novel intelligent deep learning predictive model for meteorological drought forecasting. *Journal of Ambient Intelligence and Humanized Computing*. https://doi.org/10.1007/s12652-022-03701-7

Das, N. N., Entekhabi, D., Dunbar, R. S., Chaubell, M. J., Colliander, A., Yueh, S., Jagdhuber, T., Chen, F., Crow, W., Neill, P. E. O., Walker, P., Berg, A., Bosch, D. D., Caldwell, T., Cosh, M. H., Collins, C. H., Lopez-baeza, E., & Thibeault, M. (2019). *The SMAP and Copernicus Sentinel 1A / B microwave active-passive high resolution surface soil moisture product*. *233*(August). https://doi.org/10.1016/j.rse.2019.111380

Day, A. (2019). *World Water Demand Will Increase 55% by 2050*. Save The Water.

https://savethewater.org/water-demand-to-increase-55-globally-by-2050/

de Castro, A. I., Six, J., Plant, R. E., & Peña, J. M. (2018). Mapping crop calendar events and phenology-related metrics at the parcel level by object-based image analysis (OBIA) of MODIS-NDVI time-series: A case study in central California. *Remote Sensing*, *10*(11). https://doi.org/10.3390/rs10111745

dela Torre, D. M. G., Gao, J., & Macinnis-Ng, C. (2021). Remote sensing-based estimation of rice yields using various models: A critical review. *Geo-Spatial Information Science*, *24*(4), 580–603. https://doi.org/10.1080/10095020.2021.1936656

Dhawale, R., & Paul, S. K. (2018). A comparative analysis of drought indices on vegetation through remote sensing for Latur region of India. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *42*(5), 403–407. https://doi.org/10.5194/isprs-archives-XLII-5-403-2018

Diatta, A. A., Min, D., & Jagadish, S. V. K. (2021). Drought stress responses in non-transgenic and transgenic alfalfa—Current status and future research directions. In *Advances in Agronomy* (1st ed., Vol. 170). Elsevier Inc. https://doi.org/10.1016/bs.agron.2021.06.002

Drucker, H., Burges, C. J., Kaufman, L., Smola, A., & Vapnik, V. (1996). *Support Vector Regession Machines. Advances in neural information processing systems*.

Dutta, D., Kundu, A., Patel, N. R., Saha, S. K., & Siddiqui, A. R. (2015). Assessment of agricultural drought in Rajasthan (India) using remote sensing derived Vegetation Condition Index (VCI) and Standardized Precipitation Index (SPI). *Egyptian Journal of Remote Sensing and Space Science*, *18*(1), 53–63. https://doi.org/10.1016/j.ejrs.2015.03.006

Dzavo, T., Zindove, T. J., Dhliwayo, M., & Chimonyo, M. (2019). Effects of drought on cattle production in sub-tropical environments. In *Tropical Animal Health and Production* (Vol. 51, Issue 3, pp. 669–675). https://doi.org/10.1007/s11250-018-1741-1

El Fartassi, I., Milne, A. E., El Alami, R., Rafiqi, M., Hassall, K. L., Waine, T. W., Zawadzka, J., Diarra, A., & Corstanje, R. (2023). Evidence of collaborative opportunities to ensure long-term sustainability in African farming. *Journal of Cleaner Production*, *392*(August 2022), 136170. https://doi.org/10.1016/j.jclepro.2023.136170

Esfahanian, E., Nejadhashemi, A. P., Abouali, M., Adhikari, U., Zhang, Z., Daneshvar, F., & Herman, M. R. (2017). Development and evaluation of a comprehensive drought index. *Journal of Environmental Management*, *185*, 31–43. https://doi.org/10.1016/j.jenvman.2016.10.050

Esfahanian, E., Nejadhashemi, A. P., Abouali, M., Daneshvar, F., Alireza, A. R., Herman, M. R., & Tang, Y. (2016). Defining drought in the context of stream health. *Ecological Engineering*, *94*, 668–681. https://doi.org/10.1016/j.ecoleng.2016.06.110

Fall, C. M. N., Lavaysse, C., Kerdiles, H., Dramé, M. S., Roudier, P., & Gaye, A. T. (2021).

Performance of dry and wet spells combined with remote sensing indicators for crop yield prediction in Senegal. *Climate Risk Management*, *33*(October 2020), 1–27. https://doi.org/10.1016/j.crm.2021.100331

FAO. (2008). *An Introduction to the Basic Concepts of Food Security*. 1–3. www.foodsec.org/docs/concepts_guide.pdf

FAO. (2017). *Lake Chad Basin: a crisis rooted in hunger, poverty and lack of rural development*. https://reliefweb.int/report/nigeria/lake-chad-basin-crisis-rooted-hunger-poverty-and-lack-rural-development

FAO. (2018). *Transforming Food and Agriculture To Achieve the SDGs*.

FAO. (2019). The State of Food and Agriculture. Moving Forward on Food Loss and Waste. In *Routledge Handbook of Religion and Ecology*. https://www.fao.org/3/ca6030en/ca6030en.pdf

FAO. (2023). *Sustainable Land Management*. https://www.fao.org/land-water/land/sustainable-land-management/en/

Feed the Future. (2023). *Senegal Drought and Crop Watch*. https://rshydroag.egr.msu.edu/

Findell, K. L., Berg, A., Gentine, P., Krasting, J. P., Lintner, B. R., Malyshev, S., Santanello, J. A., & Shevliakova, E. (2017). The impact of anthropogenic land use and land cover change on regional climate extremes. *Nature Communications*, *8*(1), 1–9. https://doi.org/10.1038/s41467-017-01038-w

Food and Agriculture Organization of the United Nations., International Fund for Agricultural Development, UNICEF, World Food Programme, & World Health Organization. (2019). *FAO, IFAD, UNICEF, WFP and WHO. (2019). The state and food security and nutrition in the world 2019. Safeguardin against economic slowdowns and downturns. Rome: FAO. Retrieved from http://www. fao.org/3/ca5162en/ca5162en.pdf*.

Friedman, J. (2001). Greedy Function Approximation : A Gradient Boosting Machine Author (s): Jerome H . Friedman Source : The Annals of Statistics , Vol . 29 , No . 5 (Oct ., 2001), pp . 1189-1232 Published by : Institute of Mathematical Statistics Stable URL : http://www. *The Annals of Statistics*, *29*(5), 1189–1232. https://www.jstor.org/stable/2699986

Gao, H., Bohn, T. J., Podest, E., McDonald, K. C., & Lettenmaier, D. P. (2011). On the causes of the shrinking of Lake Chad. *Environmental Research Letters*, *6*(3). https://doi.org/10.1088/1748-9326/6/3/034021

Garen, D. C. (1993). REVISED SURFACE-WATER SUPPLY INDEX FOR WESTERN UNITED STATES By David C . Garen , x Member , ASCE with the Colorado Climate Center at Colorado State University to examine. *Water Supply*, *119*(4), 437–454.

Ghaleb, F., Mario, M., & Sandra, A. N. (2015). Regional landsat-based drought monitoring from

1982 to 2014. *Climate*, *3*(3), 563–577. https://doi.org/10.3390/cli3030563

GISGeography. (2022). *Passive vs Active Sensors in Remote Sensing*. https://gisgeography.com/passive-active-sensors-remote-sensing/

Hayes, M. J., Svoboda, M. D., Wardlow, B. D., Anderson, M. C., & Kogan, F. (2012). Drought monitoring: Historical and current perspectives. *Remote Sensing of Drought: Innovative Monitoring Approaches*, 1–19. https://doi.org/10.1201/b11863

Hoerl, A. E., & Kennard, R. W. (1970). Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*, *12*(1), 55–67. https://doi.org/10.1080/00401706.1970.10488634

Holtz, L., & Golubski, C. (2021). *Addressing Africa's extreme water insecurity*. https://www.brookings.edu/blog/africa-in-focus/2021/07/23/addressing-africas-extreme-water-insecurity/

Howard, Guy, Bartram, Jamie, & WHO. (2003). *Domestic Water Quantity , Service Level and Health*.

Huang, S., Huang, Q., Leng, G., & Liu, S. (2016). A nonparametric multivariate standardized drought index for characterizing socioeconomic drought: A case study in the Heihe River Basin. *Journal of Hydrology*, *542*, 875–883. https://doi.org/10.1016/j.jhydrol.2016.09.059

Husain, N., & Trak, T. H. (2018). Climate Change Due to Human activities, and its Health impact. *International Journal of Scientific and Research Publications (IJSRP)*, *8*(9), 234–237. https://doi.org/10.29322/ijsrp.8.9.2018.p8132

Ines, A. V. M., Das, N. N., Hansen, J. W., & Njoku, E. G. (2013). Remote Sensing of Environment Assimilation of remotely sensed soil moisture and vegetation with a crop simulation model for maize yield prediction. *Remote Sensing of Environment*, *138*, 149–164. https://doi.org/10.1016/j.rse.2013.07.018

IPCC. (2019). Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems [P.R. Shukla, J. Skea, E. Calvo Buendia, V. Masson-Delmot. *Research Handbook on Climate Change and Agricultural Law*, 423–448.

Jackson, R. B., Carpenter, S. R., Dahm, C. N., Diane, M., Naiman, R. J., Postel, S. L., Running, S. W., Naiman, R. J., Postel, S. L., & Running, S. W. (2018). *Water in a Changing World Published by : Wiley on behalf of the Ecological Society of America Stable URL : http://www.jstor.org/stable/3061010 Linked references are available on JSTOR for this article : Issues in Ecology WATER IN A CHANGING WORLD*. *11*(4), 1027–1045.

Jacobs, F. (2010). *95 - The Incredible Shrinking Lake (Chad, That Is)*. https://bigthink.com/strange-maps/95-the-incredible-shrinking-lake-chad-that-is

Jones, B. A., & Fleck, J. (2020). Shrinking lakes, air pollution, and human health: Evidence from California's Salton Sea. *Science of the Total Environment*, *712*, 136490. https://doi.org/10.1016/j.scitotenv.2019.136490

Jyotsna, B. (2017). *Top 50 Applications of Remote Sensing, GIS, Geography*. https://www.geographynotes.com/topography/remote-sensing/top-50-applications-of-remote-sensing-gis-geography/4895

Karalliyadda, S. M. C. B., Kazunari, T., & Fujimura, M. (2023). Managing rain-fed uplands of cascaded tank village systems: What stakeholders really suggest? *Agricultural Systems*, *208*(March), 103643. https://doi.org/10.1016/j.agsy.2023.103643

Kelly, E., Lee, K., Shields, K. F., Cronk, R., Behnke, N., Klug, T., & Bartram, J. (2017). The role of social capital and sense of ownership in rural community-managed water systems: Qualitative evidence from Ghana, Kenya, and Zambia. *Journal of Rural Studies*, *56*, 156–166. https://doi.org/10.1016/j.jrurstud.2017.08.021

Keyantash, J. (2002). An Evaluation of a Drought. *American Meteorological Society*, *August*, 1167–1180.

Ki-moon and Patrick Verkooijen. (2021). *Start planning for a world with a lot less water*. https://edition.cnn.com/2021/03/22/opinions/world-water-day-un-ki-moon-verkooijen/index.html

Kiprop, V. (2020). *The Largest Lakes in Africa*. https://www.worldatlas.com/articles/the-largest-lakes-in-africa.html

Kogan, F., Goldberg, M., Schott, T., & Guo, W. (2015). Suomi NPP/VIIRS: improving drought watch, crop loss prediction, and food security. *International Journal of Remote Sensing*, *36*(21), 5373–5383. https://doi.org/10.1080/01431161.2015.1095370

Krasilnikov, P., Taboada, M. A., & Amanullah. (2022). Fertilizer Use, Soil Health and Agricultural Sustainability. *Agriculture (Switzerland)*, *12*(4), 16–20. https://doi.org/10.3390/agriculture12040462

Krishnamurthy R, P. K., Fisher, J. B., Schimel, D. S., & Kareiva, P. M. (2020). Applying Tipping Point Theory to Remote Sensing Science to Improve Early Warning Drought Signals for Food Security. *Earth's Future*, *8*(3), 1–14. https://doi.org/10.1029/2019EF001456

Kropp, I., Nejadhashemi, A. P., Deb, K., Abouali, M., Roy, P. C., Adhikari, U., & Hoogenboom, G. (2019). A multi-objective approach to water and nutrient efficiency for sustainable agricultural intensification. *Agricultural Systems*, *173*(March), 289–302. https://doi.org/10.1016/j.agsy.2019.03.014

Kurian, M. (2015). *Governing the Nexus*. Governing the Nexus. https://doi.org/10.1007/978-3-319-05747-7

Lillesand, T. M., Kiefer, R. W., & Chipman, J. W. (2015). *Remote Sensing and Image Interpretation* (7th ed.).

Lumbierres, M., Méndez, P. F., Bustamante, J., Soriguer, R., & Santamaría, L. (2017). Modeling biomass production in seasonal wetlands using MODIS NDVI land surface phenology. *Remote Sensing*, *9*(4), 1–18. https://doi.org/10.3390/rs9040392

Luyssaert, S., Jammet, M., Stoy, P. C., Estel, S., Pongratz, J., Ceschia, E., Churkina, G., Don, A., Erb, K., Ferlicoq, M., Gielen, B., Grünwald, T., Houghton, R. A., Klumpp, K., Knohl, A., Kolb, T., Kuemmerle, T., Laurila, T., Lohila, A., … Dolman, A. J. (2014). Land management and land-cover change have impacts of similar magnitude on surface temperature. *Nature Climate Change*, *4*(5), 389–393. https://doi.org/10.1038/nclimate2196

Mare, F., Bahta, Y. T., & Van Niekerk, W. (2018). The impact of drought on commercial livestock farmers in South Africa. In *Development in Practice* (Vol. 28, Issue 7, pp. 884–898). https://doi.org/10.1080/09614524.2018.1493091

Matondo, J. I., Alemaw, B. F., & Sandwidi, W. J. P. (2020). Climate variability and change in Africa; Perspectives, Experiences and Sustainability. In *Acta Zoologica Lituanica* (Vol. 15, Issue 2). https://doi.org/10.1080/13921657.2005.10512382

Meng, Y., Liu, X., Wu, L., Liu, M., Zhang, B., & Zhao, S. (2019). Spatio-temporal variation indicators for landscape structure dynamics monitoring using dense normalized di ff erence vegetation index time series. *Ecological Indicators*, *107*(May), 105607. https://doi.org/10.1016/j.ecolind.2019.105607

Meroni, M., Atzberger, C., Vancutsem, C., Gobron, N., Baret, F., Lacaze, R., Eerens, H., & Leo, O. (2013). Evaluation of agreement between space remote sensing SPOT-VEGETATION fAPAR Time Series. *IEEE Transactions on Geoscience and Remote Sensing*, *51*(4), 1951–1962. https://doi.org/10.1109/TGRS.2012.2212447

Miller, K. A., Rhodes, S. L., & Macdonnell, L. J. (1997). Water allocation in a changing climate: Institutions and adaptation. *Climatic Change*, *35*(2), 157–177. https://doi.org/10.1023/A:1005300529862

Minucci, G. (2021). *Enabling Adaptive Water Management to Face Drought Risk in a Changing Climate*.

MOSTAFAZADEH-FARD, B., HEIDARPOUR, M., AGHAKHANI, A., & FEIZI, M. (2007). Effects of Irrigation Water Salinity and Leaching on Soil Chemical Properties in an Arid Region BEHROUZ. *INTERNATIONAL JOURNAL OF AGRICULTURE & BIOLOGY*. https://doi.org/10.2495/SI140041

MSU extension. (2013). *The difference between "land use" and "land cover."* https://www.canr.msu.edu/news/the_difference_between_land_use_and_land_cover

Muchuru, S., & Nhamo, G. (2019). *Sustaining African water resources under climate change : Emerging adaptation measures from UNFCCC national communications*. *1338*. https://doi.org/10.1080/20421338.2018.1550934

NASA. (2014). *What Is Climate Change?* https://www.nasa.gov/audience/forstudents/k-4/stories/nasa-knows/what-is-climate-change-k4.html

NASA. (2021). *Carbon Dioxide*. https://climate.nasa.gov/vital-signs/carbon-dioxide/

National Geographic Society. (2019). *Climate Change*. https://www.nationalgeographic.org/encyclopedia/climate-change/

Nations, U. (2020). *#Envision2030 Goal 2: Zero Hunger*. https://www.un.org/development/desa/disabilities/envision2030-goal2.html#:~:text=By 2030%2C end hunger and,sufficient food all year round.

Ndoye, A. F. . (2015). *Using climate smart agriculture to build farmers' resilience in Senegal*. https://blogs.worldbank.org/nasikiliza/using-climate-smart-agriculture-to-build-farmers-resilience-in-senegal

Nelson, S., Drabarek, D., Jenkins, A., Negin, J., & Abimbola, S. (2021). How community participation in water and sanitation interventions impacts human health, WASH infrastructure and service longevity in low-income and middle-income countries: A realist review. *BMJ Open*, *11*(12), 1–18. https://doi.org/10.1136/bmjopen-2021-053320

NRDC. (2018). *Drought: Anything You Need To Know*. https://www.nrdc.org/stories/drought-everything-you-need-know#:~:text= Causes of Drought1 Natural Causes.,also impacted by how air circulates... More

OECD. (2012). *OECD Environmental Outlook to 2050: The Consequences of Inaction - Key Facts and Figures*. https://www.oecd.org/env/indicators-modelling-outlooks/oecdenvironmentaloutlookto2050theconsequencesofinaction-keyfactsandfigures.htm

OECD. (2015). Water Resources Allocation: Sharing Risks and Opportunities. *Water Intelligence Online*, *14*. https://doi.org/10.2166/9781780407623

Our World in Data. (2023). *World Population Growth*. https://ourworldindata.org/world-population-growth

Pahalvi, H. N., Rafiya, L., Rashid, S., Nisar, B., & Kamili, A. N. (2021). Chapter 2 Microbial Bioremediation of Pesticides/ Herbicides in Soil. In *Microbiota and Biofertilizers, Vol 2: Ecofriendly Tools for Reclamation of Degraded Soil Environs* (Vol. 2). https://doi.org/10.1007/978-3-030-61010-4

Palmer, W. C. (1965a). Meteorological Drought. In *U.S. Weather Bureau, Res. Pap. No. 45* (p. 58). https://www.ncdc.noaa.gov/temp-and-precip/drought/docs/palmer.pdf

Palmer, W. C. (1965b). Standardized precipitation index. In *U.S. Weather Bureau, Res. Pap. No. 45* (p. 58). https://www.ncdc.noaa.gov/temp-and-precip/drought/docs/palmer.pdf

Panek, E., & Gozdowski, D. (2020). Analysis of relationship between cereal yield and NDVI for selected regions of Central Europe based on MODIS satellite data. *Remote Sensing Applications: Society and Environment*, *17*(December 2019), 100286. https://doi.org/10.1016/j.rsase.2019.100286

Parfitt, J., Barthel, M., & MacNaughton, S. (2010). Food waste within food supply chains: Quantification and potential for change to 2050. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1554), 3065–3081. https://doi.org/10.1098/rstb.2010.0126

Piani, L., & Paris, G. (2021). Why is there water on Earth. In *SPACE.com*. https://www.space.com/why-is-there-water-on-earth

Plessis, A. (2019). *Water as an Inescapable Risk: Current Global Water Availability, Quality and Risks with a Specific Focus on South Africa*. https://link.springer.com/chapter/10.1007/978-3-030-03186-2_4

Prasad, A. K., Chai, L., Singh, R. P., & Kafatos, M. (2006a). Crop yield estimation model for Iowa using remote sensing and surface parameters. *International Journal of Applied Earth Observation and Geoinformation*, *8*(1), 26–33. https://doi.org/10.1016/j.jag.2005.06.002

Prasad, A. K., Chai, L., Singh, R. P., & Kafatos, M. (2006b). Crop yield estimation model for Iowa using remote sensing and surface parameters. *International Journal of Applied Earth Observation and Geoinformation*, *8*(1), 26–33. https://doi.org/10.1016/j.jag.2005.06.002

Pretty, J., Benton, T. G., Bharucha, Z. P., Dicks, L. V., Flora, C. B., Godfray, H. C. J., Goulson, D., Hartley, S., Lampkin, N., Morris, C., Pierzynski, G., Prasad, P. V. V., Reganold, J., Rockström, J., Smith, P., Thorne, P., & Wratten, S. (2018). Global assessment of agricultural system redesign for sustainable intensification. *Nature Sustainability*, *1*(8), 441–446. https://doi.org/10.1038/s41893-018-0114-0

Project, T. W. (2019). *Poverty and Water*. https://thewaterproject.org/why-water/poverty#:~:text=But one of the greatest causes of poverty,these people%2C poverty is a fact of life.

Qiao, M., He, X., Cheng, X., Li, P., Luo, H., Zhang, L., & Tian, Z. (2021). Crop yield prediction from multi-spectral, multi-temporal remotely sensed imagery using recurrent 3D convolutional neural networks. *International Journal of Applied Earth Observation and Geoinformation*, *102*(July), 102436. https://doi.org/10.1016/j.jag.2021.102436

Ray, R. L., Fares, A., & Risch, E. (2018). Effects of Drought on Crop Production and Cropping Areas in Texas. In *Agricultural & Environmental Letters* (Vol. 3, Issue 1, p. 170037). https://doi.org/10.2134/ael2017.11.0037

Region, D. (2015). *Agricultural Effluent treatment and aquifer storage for agricultural use*.

Ren, J., Chen, Z., Zhou, Q., & Tang, H. (2008). Regional yield estimation for winter wheat with MODIS-NDVI data in Shandong, China. *International Journal of Applied Earth Observation and Geoinformation*, *10*(4), 403–413. https://doi.org/10.1016/j.jag.2007.11.003

Rutten, M. (2012). *Water in Africa*. https://www.ascleiden.nl/content/webdossiers/water-africa#Water in Africa: introduction

Sarkodie, S. A., & Strezov, V. (2019). Economic, social and governance adaptation readiness for mitigation of climate change vulnerability: Evidence from 192 countries. *Science of the Total Environment*, *656*, 150–164. https://doi.org/10.1016/j.scitotenv.2018.11.349

Sarr, A. B., & Sultan, B. (2022). Predicting crop yields in Senegal using machine learning methods. *International Journal of Climatology*, *July 2021*, 1817–1838. https://doi.org/10.1002/joc.7947

Shafer, A. B., & Dezman, E. . (1982). *CSU_Development_SWSI_Assess_Severity_Drought_Conditions_Snowpack_Runoff_Areas_1982.pdf*.

Shahzaman, M., Zhu, W., Ullah, I., Mustafa, F., Bilal, M., Ishfaq, S., Nisar, S., Arshad, M., Iqbal, R., & Aslam, R. W. (2021). Comparison of multi-year reanalysis, models, and satellite remote sensing products for agricultural drought monitoring over south asian countries. *Remote Sensing*, *13*(16). https://doi.org/10.3390/rs13163294

Shukla, S., & Wood, A. W. (2008). Use of a standardized runoff index for characterizing hydrologic drought. *Geophysical Research Letters*, *35*(2), 1–7. https://doi.org/10.1029/2007GL032487

Singh, G., & Das, N. N. (2022). A data-driven approach using the remotely sensed soil moisture product to identify water-demand in agricultural regions. *Science of the Total Environment*, *837*(February), 155893. https://doi.org/10.1016/j.scitotenv.2022.155893

Smedley, T. (2017). *Is the world running out of fresh water?* https://www.bbc.com/future/article/20170412-is-the-world-running-out-of-fresh-water

Smith, M. J., & Southard, J. B. (2002). WATER IS ALL AROUND YOU. *Journal of Chemical Information and Modeling*, *53*(9), 1689–1699.

Spinoni, J., Barbosa, P., De Jager, A., McCormick, N., Naumann, G., Vogt, J. V., Magni, D., Masante, D., & Mazzeschi, M. (2019). A new global database of meteorological drought events from 1951 to 2016. In *Journal of Hydrology: Regional Studies* (Vol. 22). https://doi.org/10.1016/j.ejrh.2019.100593

Stagge, J. H., Tallaksen, L. M., Xu, C. Y., & Van Lanen, H. A. J. (2014). Standardized precipitation-evapotranspiration index (SPEI): Sensitivity to potential evapotranspiration model and parameters. *IAHS-AISH Proceedings and Reports*, *363*(October), 367–373.

Sur, C., Hur, J., Kim, K., Choi, W., & Choi, M. (2015). An evaluation of satellite-based drought indices on a regional scale. *International Journal of Remote Sensing*, *36*(22), 5593–5612. https://doi.org/10.1080/01431161.2015.1101653

Suruliandi, A., Mariammal, G., & Raja, S. P. (2021). Crop prediction based on soil and environmental characteristics using feature selection techniques. *Mathematical and Computer Modelling of Dynamical Systems*, *27*(1), 117–140. https://doi.org/10.1080/13873954.2021.1882505

The World Bank. (2022). *Water Resources Management*.

Tibshirani, R. (1994). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B*, *58*, 267. https://cir.nii.ac.jp/crid/1370004236282599299%22%0A%7D

Tim Sharp. (2017). *How Big is Earth?* https://www.space.com/17638-how-big-is-earth.html

Toolkits. (2012). *Why Population Matters to Food Security*. https://toolkits.knowledgesuccess.org/toolkits/family-planning-advocacy/why-population-matters-food-security#:~:text=Increasing numbers of people often,for food and reproductive health.

Toure, S., & Fane, M. (2015). *ALLIANCE FOR CLIMATE-SMART AGRICULTURE LAUNCHED IN WEST AFRICA*. https://ccafs.cgiar.org/news/alliance-climate-smart-agriculture-launched-west-africa#:~:text=The West African Alliance for,inclusive multi-stakeholder partnership framework.

Turvey, C. G., & McLaurin, M. K. (2012). Applicability of the normalized difference vegetation index (NDVI) In index-based crop insurance design. *Weather, Climate, and Society*, *4*(4), 271–284. https://doi.org/10.1175/WCAS-D-11-00059.1

UN-Environment. (2017). *Climate Change Adaptation Technologies for Water*.

UN News. (2021). *Climate and weather related disasters surge five-fold over 50 years, but early warnings save lives - WMO report*. https://news.un.org/en/story/2021/09/1098662

UN Water. (2019). *WHO and UNICEF launch updated estimates for water, sanitation and hygiene*. https://www.unwater.org/who-and-unicef-launch-updated-estimates-for-water-sanitation-and-hygiene/

UN WATER. (2020). *Water and climate change: executive summary*. 1–12.

UN WATER News. (2020). *UN World Water Development Report 2020 'Water and Climate Change''.'* https://www.unwater.org/world-water-development-report-2020-water-and-climate-change/

UNCCD. (2019). *FACTSHEET: Drought and its socio-economic impacts*. *49*(0), 5–7.

UNESCO. (2019). Water Security and the Sustainable Development Goals (Series I). In *Global Water Security Issues (GWSI) Series*. http://unesco-iwssm.org/%0Ahttp://unesco-iwssm.

UNWA. (2003). The Africa Water Vision for 2025. *United Nations Water Africa*, 1–34.

US EPA. (2021). *LAND USE - What are the trends in land use and their effects on human health and the environment?* https://www.epa.gov/report-environment/land-use#:~:text="Land use" is the term,practiced at a given place.

US EPA. (2022). *Adaptation Actions for Water Management and Ecosystem Protection*. https://www.epa.gov/arc-x/adaptation-actions-water-management-and-ecosystem-protection

US EPA. (2023). *Climate Impacts on Agriculture and Food Supply*. https://climatechange.chicago.gov/climate-impacts/climate-impacts-agriculture-and-food-supply

USAID. (2022). *Agriculture and Food Security*. https://www.usaid.gov/what-we-do/agriculture-and-food-security

USDA. (2011). Millet has many faces. *USDA/FAS Food for Progress Assistance Report*. https://gain.fas.usda.gov/Recent GAIN Publications/Millet has many faces_Dakar_Senegal_11-23-2011.pdf

USDA. (2022). *Food Security*. https://www.usda.gov/topics/food-and-nutrition/food-security#:~:text=Given population growth and rising,will need to almost double.

USDA. (2023). *Biulding Community Resilience to Climate Change in Senegal*. https://www.climatehubs.usda.gov/hubs/international/topic/building-community-resilience-climate-change-senegal

USDA ERS. (2022). *Measurement*. https://www.ers.usda.gov/topics/food-nutrition-assistance/food-security-in-the-u-s/measurement/#:~:text=Food insecurity is measured at,food intake and skipped meals.

USGS. (2021). *What is Remote sensing and What is it used for?* https://www.usgs.gov/faqs/what-remote-sensing-and-what-it-used?qt-news_science_products=0#qt-news_science_products

Wageningen University and Research. (2023). *Food security: sufficient safe and healthy food for everyone*. https://www.wur.nl/en/show/food-security-11.htm

Waigwa, S. (2007). *Tanzania: Inland Lakes Shrinking Faster Than Ever*. https://allafrica.com/stories/200702260079.html#:~:text=Arusha — Lake Jipe has shrunk by

60,are among the shrinking water bodies in Tanzania.

Weghorst, K. (1996). The Reclamation Drought Index: Guidelines and Practical Applications. In *North American Water and Environment Congress & Destructive Water* (pp. 637–642). http://cedb.asce.org/cgi/WWWdisplay.cgi?99523

West, H., Quinn, N., & Horswell, M. (2019). Remote sensing for drought monitoring & impact assessment: Progress, past challenges and future opportunities. *Remote Sensing of Environment*, *232*(November 2018), 111291. https://doi.org/10.1016/j.rse.2019.111291

WFP. (2019). *From Senegal to Marrakech: What Climate-Smart Villages Can Teach the World*. https://www.wfpusa.org/articles/from-senegal-to-marrakech-what-climate-smart-villages-can-teach-the-world/

WHO. (2011). Safe Drinking-water from Desalination. *World Health Organization*, *WHO/HSE/WS*, 28. https://www.who.int/water_sanitation_health/publications/2011/desalination_guidance_en.pdf%0Ahttp://apps.who.int/iris/bitstream/10665/70621/1/WHO_HSE_WSH_11.03_eng.pdf%0Ahttp://apps.who.int/iris/bitstream/10665/70621/1/WHO_HSE_WSH_11.03_eng.pdf

WHO. (2019). *World hunger is still not going down after three years and obesity is still growing – UN report*. https://www.who.int/news/item/15-07-2019-world-hunger-is-still-not-going-down-after-three-years-and-obesity-is-still-growing-un-report#:~:text=More than 820 million people are hungry globally&text=An estimated 820 million people,of increase in a row.

Winkler, K., Fuchs, R., Rounsevell, M., & Herold, M. (2021). Global land use changes are four times greater than previously estimated. *Nature Communications*, *12*(1), 1–10. https://doi.org/10.1038/s41467-021-22702-2

Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, *58*(2), 109–130. https://doi.org/10.1016/S0169-7439(01)00155-1

World Bank Group. (2017). *IMPROVED BUSINESS AND INVESTMENT CLIMATE IN WEST AFRICA*.

World Bank Group. (2021a). *Senegal*. https://climateknowledgeportal.worldbank.org/country/senegal/vulnerability#:~:text=Droughts in Senegal are concentrated,drought events affected the country.

World Bank Group. (2021b). *Senegal*.

World Economic Forum. (2020). *Global hunger fell for decades, but it's rising again*. https://www.weforum.org/agenda/2020/07/global-hunger-rising-food-agriculture-organization-report/#:~:text=Writer%2C Formative Content-,Almost 690 million people in the world were undernourished in,slowdowns and extreme weather events.

World Food Summit. (1996). *Report of the World Food Summit*. https://www.fao.org/3/w3548e/w3548e00.htm#AnnexV

Xianfeng, L., Xiufang, Z., Yaozhong, P., Shuangshuang, L., Yanxu, L., & Yuqi, M. (2016). *Agricultural drought monitoring : Progress , challenges , and prospects*. *26*(6), 750–767. https://doi.org/10.1007/s11442-016-1297-9

Zhao, J., Huang, S., Huang, Q., Wang, H., Leng, G., & Fang, W. (2020). Time-lagged response of vegetation dynamics to climatic and teleconnection factors. *Catena*, *189*(April 2019), 104474. https://doi.org/10.1016/j.catena.2020.104474

Zhao, Q., Chen, Q., Jiao, M., Wu, P., Gao, X., Ma, M., & Hong, Y. (2018). The temporal-spatial characteristics of drought in the loess plateau using the remote-sensed TRMM precipitation data from 1998 to 2014. *Remote Sensing*, *10*(6). https://doi.org/10.3390/rs10060838

Zhu, N., Xu, J., Li, W., Li, K., & Zhou, C. (2018). A comprehensive approach to assess the hydrological drought of inland river basin in Northwest China. *Atmosphere*, *9*(10). https://doi.org/10.3390/atmos9100370

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, *67*(2), 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x

APPENDIX

Table S1. Bias Correction factor based on each department's expected millet cultivated area during the planting period (June and July)

Department	Expected Millet Planting (%)	Expected Planting of Other Crops (%)
Bakel	25.68	74.32
Bambey	56.40	43.60
Bignona	14.28	85.72
Birkelane	27.57	72.43
Bounkiling	20.59	79.41
Dagana	9.76	90.24
Diourbel	59.94	40.06
Fatick	49.89	50.11
Foundiougne	33.15	66.85
Gossas	34.70	65.30
Goudiri	30.69	69.31
Goudomp	18.96	81.04
Guinguineo	59.36	40.64
Kaffrine	26.69	73.31
Kanel	60.04	39.96
Kaolack	38.68	61.32
Kebemer	28.21	71.79
Kedougou	0.00	100.00
Kolda	18.11	81.89
Koumpentoum	29.50	70.50
Koungheul	28.74	71.26
Linguere	45.70	54.30
Louga	23.64	76.36
Malem Hodar	25.11	74.89
Matam	57.69	42.31
Mbacke	48.14	51.86
Mbour	54.29	45.71
Medina Yoro Foulah	22.50	77.50
Nioro	46.56	53.44
Oussouye	3.06	96.94
Podor	9.46	90.54
Ranenou	74.81	25.19
Rusfique	0.00	100.00
Saint-Louis	6.84	93.16
Salemata	0.00	100.00
Saraya	0.00	100.00
Sedhiou	12.62	87.38
Tamba	22.06	77.94
	38.80	61.20

Table S1 (cont'd)

Thies		
Tivaouane	37.96	62.04
Velingara	13.19	86.81
Ziguinchor	9.23	90.77

Table S2. Bias Correction factor based on each department's expected millet cultivated area during the growing period (August)

Department	Expected Planting of Other Crops (%)	Expected Millet Growing (%)	Expected Growing of Other Crops (%)
Bakel	2.78	25.68	71.54
Bambey	0.00	56.40	43.60
Bignona	32.45	14.28	53.28
Birkelane	0.52	27.57	71.91
Bounkiling	23.13	20.59	56.27
Dagana	77.75	9.76	12.49
Diourbel	0.00	59.94	40.06
Fatick	4.95	49.89	45.15
Foundiougne	3.29	33.15	63.56
Gossas	0.00	34.70	65.30
Goudiri	1.64	30.69	67.67
Goudomp	33.18	18.96	47.86
Guinguineo	0.00	59.36	40.64
Kaffrine	0.03	26.69	73.28
Kanel	11.58	60.04	28.38
Kaolack	0.29	38.68	61.03
Kebemer	0.00	28.21	71.79
Kedougou	43.70	0.00	56.30
Kolda	36.15	18.11	45.74
Koumpentoum	0.20	29.50	70.31
Koungheul	0.08	28.74	71.18
Linguere	0.00	45.70	54.30
Louga	0.59	23.64	75.76
Malem Hodar	0.10	25.11	74.79
Matam	31.84	57.69	10.47
Mbacke	0.00	48.14	51.86
Mbour	0.00	54.29	45.71
Medina Yoro Foulah	19.11	22.50	58.39
Nioro	1.00	46.56	52.43
Oussouye	89.77	3.06	7.17
Podor	69.52	9.46	21.02
	0.00	74.81	25.19

Table S2 (cont’d)

Ranenou			
Rusfique	0.00	0.00	100.00
Saint-Louis	3.81	6.84	89.35
Salemata	21.00	0.00	79.00
Saraya	41.41	0.00	58.59
Sedhiou	35.32	12.62	52.05
Tamba	1.00	22.06	76.94
Thies	0.00	38.80	61.20
Tivaouane	0.00	37.96	62.04
Velingara	19.40	13.19	67.41
Ziguinchor	67.28	9.23	23.49

Table S3. Bias Correction factor based on each department's expected millet cultivated area during the harvesting period (September, October, and November)

Department	Expected Growing of Other Crops (%)	Expected Millet Harvesting (%)	Expected Harvesting of Other Crops (%)
Bakel	13.31	25.68	61.01
Bambey	22.13	56.40	21.47
Bignona	54.23	14.28	31.49
Birkelane	36.21	27.57	36.22
Bounkiling	33.27	20.59	46.14
Dagana	58.76	9.76	31.48
Diourbel	24.25	59.94	15.81
Fatick	26.66	49.89	23.44
Foundiougne	33.17	33.15	33.68
Gossas	40.53	34.70	24.77
Goudiri	16.52	30.69	52.80
Goudomp	44.44	18.96	36.60
Guinguineo	25.72	59.36	14.92
Kaffrine	30.02	26.69	43.29
Kanel	9.62	60.04	30.34
Kaolack	25.82	38.68	35.50
Kebemer	47.42	28.21	24.37
Kedougou	44.02	0.00	55.98
Kolda	41.06	18.11	40.83
Koumpentoum	33.75	29.50	36.75
Koungheul	32.12	28.74	39.14
Linguere	31.54	45.70	22.76
Louga	49.03	23.64	27.33
Malem Hodar	35.99	25.11	38.91
Matam	21.81	57.69	20.50
	31.49	48.14	20.37
	20.02	54.29	25.69

Table S3 (cont’d)

Mbour			
Medina Yoro Foulah	38.86	22.50	38.64
Nioro	29.10	46.56	24.33
Oussouye	64.17	3.06	32.77
Podor	46.35	9.46	44.19
Ranenou	4.43	74.81	20.76
Rusfique	31.72	0.00	68.28
Saint-Louis	61.92	6.84	31.24
Salemata	20.90	0.00	79.10
Saraya	40.27	0.00	59.73
Sedhiou	37.16	12.62	50.21
Tamba	29.77	22.06	48.17
Thies	36.75	38.80	24.45
Tivaouane	39.30	37.96	22.74
Velingara	32.94	13.19	53.86
Ziguinchor	56.65	9.23	34.12

Table S4. Analysis of Variance results for significant variables. The *p-value* at 0.1% is extremely significant, 1% is highly significant, and 5% is significant

Predictors	***Estimates***	***Confidence Interval***	***p-value***
(Intercept)	-0.36	-0.69 – -0.04	0.028
Planting VI	-0.27	-0.48 – -0.06	0.012
Planting GDDI	-0.40	-0.68 – -0.12	0.005
Planting SPI3	0.88	0.43 – 1.33	<0.001
Planting SRI3	-0.81	-1.25 – -0.38	<0.001
Growing VI	0.33	0.15 – 0.50	<0.001
Harvesting WDI	2.69	0.47 – 4.92	0.018
Harvesting VI	-0.30	-0.51 – -0.09	0.005
Harvesting SMI	-2.76	-4.88 – -0.64	0.011
Harvesting GDDI	-0.71	-1.28 – -0.14	0.015
Harvesting SMDI	-0.30	-0.58 – -0.03	0.032

Table S4 (cont'd)

Harvesting SPI3	-0.73	-1.45 – -0.01	0.048
Harvesting SRI3	0.92	0.27 – 1.56	0.005
Rainfall Erosivity	0.53	0.04 – 1.03	0.034
Soil Organic Carbon	-0.79	-1.22 – -0.36	<0.001

Table S5. The average-performing Random Forest Regressor models for different sets of variables

Independent Variables	NSE	RMSE	nRMSE	R^2	*d*
All variables except Bias Correction Harvesting, Total, Growing, and Physiographical	0.283	0.619	0.189	0.283	0.633
All variables except Bias Correction, Department, Harvesting, Total, Growing, Physiographical, and Planting SMI,VI, and GDDI	0.096	0.700	0.213	0.096	0.545
All variables except Bias Correction	0.234	0.638	0.196	0.234	0.617
All variables except Bias Correction and Department	0.161	0.664	0.204	0.161	0.592
All variables except Bias Correction, Department, and Physiographical	0.138	0.678	0.208	0.138	0.568
All variables except Bias Correction, Harvesting and Total	0.246	0.629	0.193	0.246	0.636
All variables except Bias Correction, Harvesting, Total, and Growing	0.281	0.614	0.189	0.281	0.676
All variables except Bias Correction, Department, Harvesting, Total, and Growing	0.227	0.635	0.196	0.227	0.651
All variables except Bias Correction, Department, Harvesting, and Total	0.167	0.666	0.204	0.167	0.583
All variables except Bias Correction and Total	0.239	0.634	0.195	0.239	0.619
All variables except Bias Correction, Department and Total	0.167	0.666	0.204	0.167	0.570
All variables except Harvesting, Total, Growing, and Physiographical	0.320	0.602	0.185	0.320	0.669
All variables	0.250	0.633	0.194	0.250	0.622
All variables except Department	0.192	0.653	0.201	0.192	0.606
All variables except Department and Physiographical	0.139	0.678	0.208	0.139	0.568
All variables except Harvesting and Total	0.234	0.635	0.195	0.234	0.626

Table S5 (cont’d)

All variables except Harvesting, Total, and	0.318	0.594	0.183	0.318	0.698

Table S5 (cont’d)

All variables except Department, Harvesting, Total, and Growing	0.303	0.604	0.186	0.303	0.682
All variables Department, Harvesting, and Total	0.212	0.647	0.198	0.212	0.612
All variables except the Total	0.260	0.625	0.192	0.260	0.643
All variables except Department and Total	0.206	0.648	0.199	0.206	0.613
All variables except Department, Harvesting, Total, Growing, Physiographical and Planting SMI, VI, and GDDI	0.198	0.661	0.201	0.198	0.590