FERARI

FERTILIZER RESEARCH & RESPONSIBLE IMPLEMENTATION

Estimation of Water-Limited Maize Yield using the LINTUL-2 Model and Spatial Analysis of the Yield Gap in Ghana

IFDC FERARI Research Report No. 13

K.B.D. Simperegui, P.S. Bindraban, K.K.A. Kouame, D.H. Peluffo-Ordóñez, and W.K. Atakora

IFDC

MOHAMMED VI POLYTECHNIC UNIVERSITY

College of Computing

Estimation of Water-Limited Maize Yield using the LINTUL-2 Model and Spatial Analysis of the Yield Gap in Ghana

K.B.D. Simperegui[1,2], P.S. Bindraban[2*], K.K.A. Kouame[2], D.H. Peluffo-Ordóñez[1], W.K. Atakora[2]

[1]*Mohammed VI Polytechnic University, Morocco*
[2]*International Fertilizer Development Center, USA*
[*]*Correspondence: pbindraban@ifdc.org*

This report is based on the thesis by K.B.D. Simperegui, submitted to the College of Computing of UM6P for obtaining the diploma of master's degree in Data Science. This thesis was conducted under the Fertilizer Research and Responsible Implementation (FERARI) program.

Citation:

Simperegui, K.B.D., P.S. Bindraban, K.K.A. Kouame, D.H. Peluffo-Ordóñez, and W.K. Atakora. 2023. Estimation of Water-Limited Maize Yield using the LINTUL-2 Model and Spatial Analysis of the Yield Gap in Ghana. IFDC FERARI Research Report No. 13.

Acknowledgments

This project has been carried out thanks to the participation and collaboration of several individuals and institutions. We would like to express our sincere gratitude to:

- The OCP Foundation for the Scholarship for Education at Mohammed VI Polytechnic University (UM6P), as well as the Fertilizer Research and Responsible Implementation (FERARI) program for their financial and technical support throughout this internship period.
- Professor Ahmed Ratnani, Director of Al-Khwarizmi, and the entire school staff.
- Professor Diego Peluffo, our UM6P supervisor, for his support, kindness, valuable advice, recommendations, and guidance that greatly contributed to refining our ideas during moments of doubt. A heartfelt thank you.
- Dr. Prem Bindraban, FERARI program Director, our IFDC supervisor, for his availability and technical contribution that greatly enriched our understanding of crop modeling and research methodology. It was a true pleasure learning from you.
- Mr. Anselme Kouame, for his availability, rigor, valuable advice, readings, and observations which allowed us to enhance the scientific quality of this document while deepening our knowledge of crop modeling. Thank you very much.
- Mr. Kwesie Benjamin for his technical support and guidance during our introduction to mapping and remote sensing. Thank you for your kindness and simplicity.
- We express our deep gratitude to our family for their unwavering support throughout our academic journey.
- Our fellow students and all those whose names have not been mentioned here. Your contributions have been invaluable, and we are grateful to you.

Thank you!

CONTENTS

TABLES

FIGURES

ABBREVIATIONS

AEZ	agroecological zone
AfBD	African Development Bank
ANOVA	analysis of variance
APSIM	Agricultural Production Systems Simulator
BS	base saturation
DLAI	reduction of LAI due to leaf death
DRAIN	drainage rain
DSSAT	Decision Support System for Agrotechnology Transfer
DTEFF	daily effective temperature
DTR	daily total radiation
ET	evapotranspiration
FAO	Food and Agriculture Organization of the United Nations
FE	iron
FLV	allocation coefficient to leaves
FPAR	PAR fraction in DTR
FRT	allocation coefficient to roots
FRT	root growth rate
FRTMOD	root partitioning
FSHMOD	shoot partitioning
FSO	allocation coefficient to storage organs
FST	allocation coefficient to steam
GDP	gross domestic product
GLAI	growth of LAI
GLV	growth rate of leaf dry matter
GTOTAL	total daily produced dry matter
IDW	inverse distance weighting
IFPRI	International Food Policy Research Institute
IPAR	intercepted photosynthetically active radiation
IRLS	iterative reweighted least squares
IRRIG	irrigation
K	constant light extinction coefficient
LAI	leaf area index
LAICR	critical leaf area index
LINTUL	Light Interception and Utilization
LUE	light use efficiency
MiDA	Millennium Development Authority
MoFA	Ministry of Food and Agriculture of Ghana
mt/ha	metric ton per hectare
N	nitrogen
N20	nitrous oxide

PAR	photosynthetically active radiation
PEVAP	potential evaporation
PFJ	Planting for Food and Jobs
PTRAN	potential transpiration
RAIN	rain
RDRDV	leaf death due to ageing
RDRSH	leaf death due to shading
RGRL	relative growth rate of LAI during exponential growth
RNINTC	intercepted rain
RRME	robust regression based on M-estimators
RUNOFF	runoff
RWLV	relative weight of leaves
RWRT	relative weight of roots
RWSO	relative weight of storage organs
RWST	relative weight of steam
S	sulfur
SLA	specific leaf area
SSA	sub-Saharan African
TIN	triangulated irregular network
TRAN	actual transpiration
TRANRF	transpiration reduction factor
TRANSCO	transpiration coefficient
TSUM	temperature summation
VIF	variance inflation factor
WC	water content
WCAD	water content at air dryness
WCCR	critical soil water level
WCFC	water content at field capacity
WCWET	critical soil water content for transpiration reduction due to waterlogging
WCWP	water content at wilting point
Zn	zinc

SUMMARY

Maize holds a significant position within Ghana's cereal production, contributing to 45% of the total cereal production. Despite this, the average maize yield of 2.4 metric tons per hectare (mt ha^{-1}) between 2017 and 2019 falls well below its potential range of 5-6 mt ha^{-1}. To comprehensively grasp the dynamics of the maize yield gap in Ghana, we employed the light use efficiency (LINTUL-2) crop model alongside statistical and geospatial analyses. This allowed us to assess the variability of maize water-limited potential yield and yield gap across 10 designated study sites, extending our evaluation to a national scale. Utilizing random forest regression, followed by ridge regression, we endeavored to uncover the principal drivers behind maize yield gap in Ghana.

Our findings reveal a water-limited yield gap ranging from 18% to 74% across the 10 study sites and diverse fertilizer treatments. The combined approach of random forest and ridge regression, explaining 87% of the yield gap variability (RMSE = 472.6 kg ha^{-1}), highlights noteworthy trends. Notably, at a 5% confidence level, soil organic matter, soil carbon content, base saturation, and soil nitrogen content emerge as the most influential factors, explaining 13.81%, 13.80%, 11.56% and 10.25% of the maize yield gap variability under water limited conditions, respectively. The Ridge Regression underscores the significance of soil organic matter, base saturation, soil nitrogen content, nitrogen application, phosphorus application, potassium application, and sulfur application for reducing the maize yield gap. Our research also emphasis the potential of sulfur application as a secondary nutrient to effectively decrease the maize yield gap, particularly when integrated with macronutrients (NPK) and the kriging interpolation reveals high potential for maize production in the northern part of the country.

CHAPTER 1: INTRODUCTION

1.1 Background

Agriculture plays a crucial role in global food security and economic development, particularly in Africa where it serves as a source of employment and income for a large portion of the rural population. According to Oxford Business Group (2021), Oxford Business Group agriculture constitutes the majority of the active workforce in Africa and represents around 15% of sub-Saharan African (SSA) gross domestic product (GDP). It is also vital for global economic growth, providing raw materials for various industries and contributing to poverty reduction (De Janvry and Sadoulet, 2010). Maize (*Zea mays* L.) is a pivotal cereal crop in SSA and contributes significantly to ensuring food security and nutrition in the region, making it an indispensable crop (Shah et al., 2016). In 2019, the Food and Agriculture Organization of the United Nations (FAO) reported that West Africa's global maize production reached 26 million metric tons (mt), accounting for roughly 37% of its overall cereal production (FAOSTAT, 2021).

In Ghana, maize cultivation is very useful (GSS, 2018), and according to the African Development Bank (AfBD), maize is the most important cereal crop in the country, accounting for approximately 45% of total cereal production (MoFA, 2020). Furthermore, maize is a key crop for food security in Ghana, as it is grown by small farmers and used for food consumption and livestock and poultry feed (Andam et al., 2017; MiDA, 2010; MoFA, 2010). Given the importance of maize in the Ghanaian diet, the government has taken various measures to improve the yield per hectare (ha). These measures include the establishment of programs such as Planting for Food and Jobs (PFJ), launched in 2017, which aims to boost agricultural production, create jobs in the agricultural sector, and improve the country's food security by subsidizing agricultural inputs (certified seeds, fertilizer subsidy, extension service delivery) and output markets for farmers (Mabe et al., 2018; MoFA, 2019b). Between 2017 and 2019, the initiative led to an average improvement of approximately 40% in maize yield reaching an average of 2.4 mt ha^{-1} compared to the period of 2013 to 2016, when the national average was 1.73 mt ha^{-1}, as reported by the Ministry of Food and Agriculture (MoFA) in 2020 and FAOSAT in 2021. However, despite these coordinated actions and an improvement in total maize production, the current yield is still far from the attainable yield (5-6 mt ha^{-1}) (IFPRI, 2014; Scheiterle and Birner, 2018), thus maintaining the gap in maize production (Adzawla et al., 2021).

In their study on the yield response of maize to fertilizer in Ghana, Bua et al. (2020) revealed that yields vary greatly between and within agroecological zones (AEZs), and yield ranges from 2 to 10 mt ha^{-1}. It is evident that the increase in maize yield cannot be attributed solely to normal application rates of fertilizers, as other factors also play an important role (Bua et al., 2020; A.K. Kouame et al., 2023; K.K.A. Kouame et al., 2021). These additional factors include root zone depth, soil properties, weather, and elevation, as reported by A.K. Kouame et al. (2023). Furthermore, maize growing in Ghana is essentially rainfed, and small-scale farmers use traditional agricultural techniques to produce their crops (SRID/MoFA, 2017; World Bank, 2017; Worqlul et al., 2019). In addition, within the different AEZs, there is a high variability of precipitation (Manzanas et al., 2014). The Guinea Savannah region, compared to the Forest zone and the Transition zone, is characterized by unimodal rainfall pattern with an average annual precipitation of 1,100 millimeters (mm) (Darfour and Rosentrater, 2016b). The average annual precipitation for Forest and Transition zones is respectively 1,500 mm and 1,300 mm.

1.2 Problem Statement

Given the significant gap between actual maize yields and their potential yields, estimated to be around 17% and 98% by Boullouz et al. (2022), it is crucial to identify the primary factors responsible for this yield gap. In Ghana, various models for crop simulation, such as the Decision Support System for Agrotechnology Transfer (DSSAT) (Jones et al., 2003), the Agricultural Production Systems Simulator (APSIM) (Keating et al., 2003) and CROPGRO-Peanut, have been used in different studies, including those conducted by MacCarthy et al. (2018), Danquah et al. (2020), and Naab et al. (2004), to evaluate maize yield gaps under different conditions. DSSAT and APSIM, along with AquaCrop (Hsiao et al., 2009) and InfoCrop (Aggarwal et al., 2006), are by far the most commonly used models in crop yield simulation (Fayaz et al., 2021). However, since these models were developed in Europe or the United States, they require a large amount of data for calibration to be usable in other regions (Mourice et al., 2014), which makes their use difficult in regions with very little data. As an alternative to this limitation, summary models that capture the essence of major processes and reduce the need for detailed processes, such as Light Interception and Utilization (LINTUL) or Environmental Policy Integrated Climate (EPIC), can be much more appropriate to use (Zhao et al., 2019).

The LINTUL model approach, designed by Spitters and Schapendonk (1990), was based on the assumption that plant growth is directly proportional to the amount of intercepted light in optimal conditions. LINTUL-1 was developed to simulate potential yields by taking into account various factors, including an adequate supply of water and nutrients, a pest, disease, and weed-free environment, as well as prevailing weather conditions (van Oijen and Leffelaar, 2008b). LINTUL-2 assumes all conditions to be optimal except for limited water (Farré et al., 2000; Spitters and Schapendonk, 1990), and LINTUL-3 assumes all conditions to be optimal except for nitrogen (N) and water as the limiting factors (Shibu et al., 2010). The LINTUL models have been used in multiple studies conducted in SSA nations, including Togo, Nigeria, and South Africa, to assess the impact of specific variables on the yield of cassava and potatoes (Adiele et al., 2021; Ezui et al., 2018; Machakaire et al., 2016). Based on the literature review and the literature search, it was observed that the LINTUL crop model has not been sufficiently applied to analyze the key factors affecting the maize yield gap in Ghana. Boullouz et al. (2022) used LINTUL-1 to quantify and explain maize yield gaps across Ghana. With the implementation of the PFJ initiative since 2017, one could assume better adoption of fertilizers and better management of farmlands. However, this assumption cannot be made for water as a limiting factor. Several studies have shown that maize production can be reduced by 20-40% in case of drought (Lobell et al., 2011). According to Shen et al. (2020), the effect of drought on maize yield depends on the growth stage of the crop, as maize has varying water requirements throughout its growth stages. During the reproductive growth period, it requires about 8-9 mm of water per day per plant, which is crucial for grain production (Aslam et al., 2015). Over the entire season, maize requires an average of 450-600 mm of water, with a production of 15 kg of grains per millimeter of consumed water (Du Plessis, 2003; Rockström, 2003). Ultimately, maize consumes approximately 250 liters of water per kilogram of biomass production, i.e., about 500 liters per kilogram of grains, until maturity. This highlights the importance of water for maize growth to significantly reduce the yield gap, which is estimated at 6 mt ha^{-1} (Boullouz et al., 2022; IFPRI, 2014; Scheiterle and Birner, 2018). Therefore, a yield gap analysis of maize considering water as a limiting factor is essential. To this end, the LINTUL-2 model will be used in this study to estimate the water-limited potential yield

in Ghana, and the yield gap determined by subtracting the actual maize yield from the water-limited potential yield will be spatially modeled.

1.3 Activity Statement

To estimate farm-level maize yields gap in Ghana, a literature review was conducted to gain a better understanding of the estimation of water-limited maize yield using the LINTUL-2 model and spatial analysis of the yield gap in Ghana. Data collection has been carried out to build an accurate prediction model for maize water-limited potential yield using LINTUL-2. After the determination of the water-limited potential yield, the calculation of the yield gap has been performed, followed by the assessment of maize yield gap variability using a random forest regression. Secondly, a ridge regression has been used to evaluate the effect of these key variables on the maize yield gap. Random forest can detect linear and non-linear relationships between the yield gap and its covariates (Mohapatra et al., 2020), making it ideal for assessing the most influential variables that are related to maize yield gaps in Ghana, but not their quantitative impact on yield variability. We therefore used a regression model to better understand the dynamics of the maize yield gap in our trials. However, as the first attempt using multilinear regression (MLR) did not meet the necessary validity assumptions, we proceeded with a ridge regression analysis to explore the direction of correlations between our covariates and the yield gap. By utilizing these methods, we can accurately evaluate the water-limited potential yield of maize, calculate the yield gap, and identify key variables that contribute to yield gap, leading to the formulation of policies to address the yield gap in Ghana.

1.4 Objectives

- Assess whether the meteorological data (remote sensing data) currently available on a global scale are sufficient to justify their use in modeling data poor environments.
- Develop an accurate prediction model for maize yield under water-limited conditions in Ghana using LINTUL-2.
- Determine the key factors that are related to maize water-limited yield gap in Ghana.
- Assess the spatial variation of maize water-limited potential yield and yield gap in Ghana.
- Formulate policies for decision-makers to address the yield gap in Ghana.

1.5 Hypotheses

- LINTUL-2 is capable of accurately estimating the maize water-limited potential yield across Ghana.
- Random forest and ridge regression can identify the main factors that contribute significantly to the yield gap of maize in Ghana.
- Maize water-limited potential yield and yield gap have a large spatial variation in across different regions of Ghana.

1.6 Justification Statement

Maize is an important cereal to ensure the food security in Ghana. Unfortunately, the national production falls short of meeting the growing demand of the population, primarily due to yield gaps between the potential and actual yields. Identifying the key factors that influence maize yield and its variability across Ghana would allow for the formulation of precise recommendations to

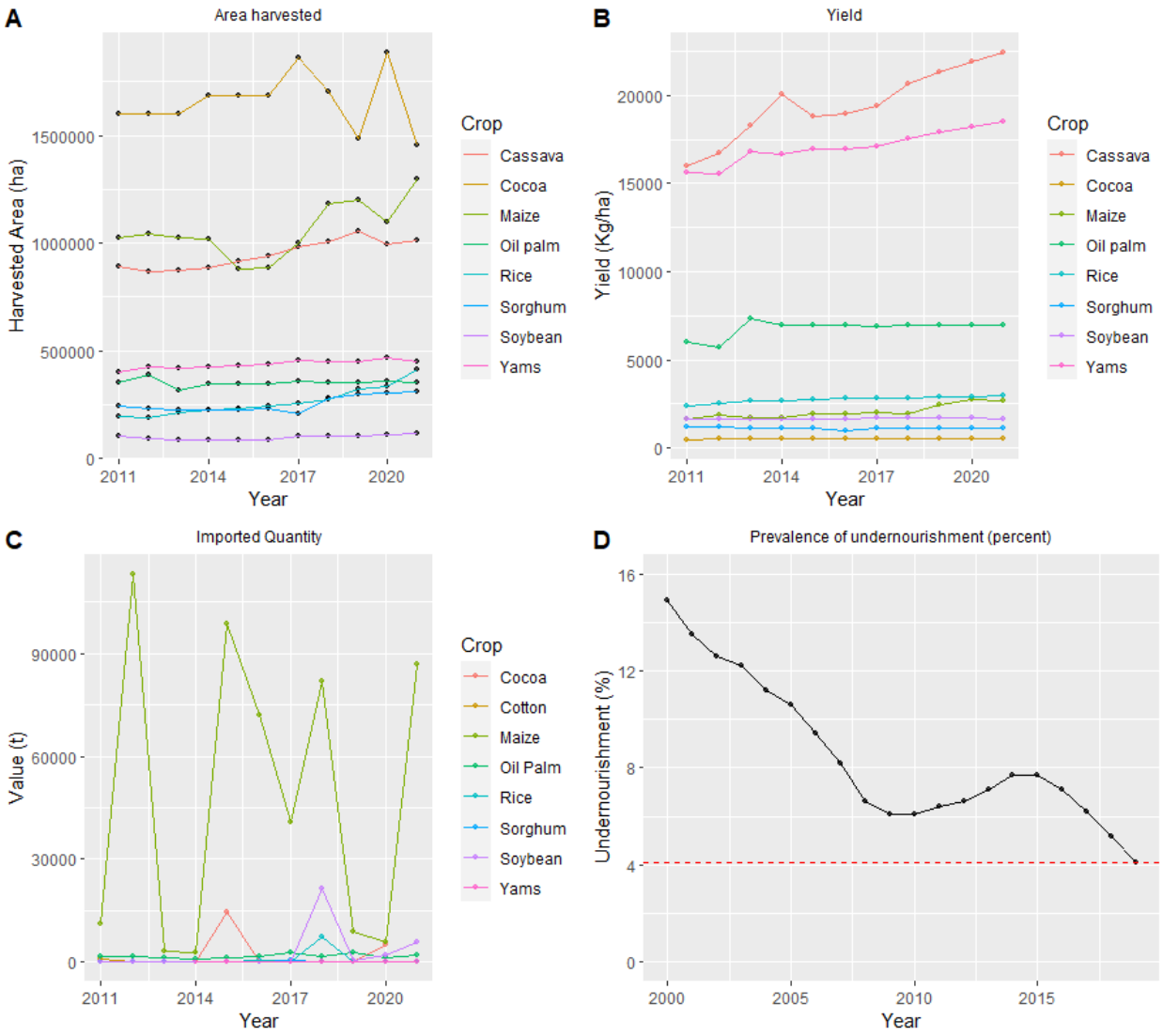
reduce these gaps and achieve food security in the country, while also contributing to the achievement of the Sustainable Development Goals (SDGs) related to food. One major obstacle in this research is the scarcity of reliable data, which is essential for accurate crop modeling. To overcome this, we leveraged two different remote sensing datasets and evaluated their effectiveness in representing actual data. By identifying the dataset that best aligns with ground truth data, we can enhance the reliability and applicability of the research, paving the way for more targeted and effective agricultural interventions in Ghana.

CHAPTER 2: LITERATURE REVIEW

2.1 Agriculture in Ghana

Agriculture is one of the main pillars of Ghana's development (Addo and Amponsah, 2018; Bua et al., 2020; Darfour and Rosentrater, 2016b). It occupies over 57% of the country's land area, employs 38.3% of the active population, and contributes approximately 20% to the country's GDP (MoFA, 2019a). The crops grown in Ghana include maize, oil palm, rubber, coconut, pineapple, groundnut, soybean, sorghum, millet, rice, and yam (Darfour and Rosentrater, 2016a; Quaye, 2008). The agriculture sector in Ghana is mostly dominated by smallholder farmers (90%) whose land size does not exceed 2 ha and who still practice traditional agriculture (Adzawla et al., 2021; Scheiterle and Birner, 2018). It also manages to cover 51%, 60%, and 50% of the country's respective needs for cereals, fish, and meat (Darfour and Rosentrater, 2016a; SARI, 1996; Scheiterle and Birner, 2018).

Figure 1 provides an overview of Ghana's agriculture sector, including information on imports and the number of undernourished people. Over the last decade, maize and cocoa have had average land areas of over 1 million ha and 1.5 million ha, respectively. However, despite occupying the most significant proportion of cultivated land, they have the lowest yields of 5.2 and 2.06 mt ha^{-1}, respectively. This low maize productivity is the main reason for high maize imports compared to other agricultural products (MoFA, 2020). Figure 1D shows a significant decrease in the proportion of undernourished people, reaching 4.1% in 2021, or about 1.3 million people, which is still significant.



Source: Author creation from FAOSTAT/FAO (2023).

Figure 1. Overview of Ghana's agriculture sector

2.2 Overview of Maize

Maize cultivation plays a significant role in Ghana's agricultural sector, providing food, income, and livelihoods for millions of people. It is grown throughout the country, with Coastal Savannah zone, Forest zone, Transition zone and Guinea Savannah zone being the primary production areas (Darfour and Rosentrater, 2016a). Farmers typically grow maize in rotation with other crops to maintain soil fertility and reduce pest and disease pressure (Darfour and Rosentrater, 2016a). Both rainfed and irrigated systems are used for maize cultivation, with the majority of production being rainfed and traditional (Scheiterle and Birner, 2018).

However, maize cultivation in Ghana faces various challenges, including a poor soil fertility, pest and disease pressure, post-harvest losses ranging from 5% to 70% (Darfour and Rosentrater, 2016b), and the impact of climate change. To overcome production limitations in specific areas,

several maize varieties have been developed (Adu et al., 2014). For example, the Obatanpa variety is an open-pollinated variety that has good protein quality and can produce an average yield of 4.6 mt ha^{-1}. The Omankwa variety is tolerant to drought and Striga and has a good protein quality, yielding an average of 4.5 mt ha^{-1} (Adu et al., 2014). Planting of certified seed is recommended between late May to early July in Sudan Savannah, late May to late June in Guinea Savannah, mid-March to late April in Transition Savannah, early March to late April in Forest Savannah, and late March to late April in Coastal Savannah for the major season. Planting dates are between mid-July to early September and mid-July to early September in Transition and Forest, respectively, for the minor season. Planting should be at 75 cm between every two rows and 40 cm between stands for early maturing varieties and 80 cm between every two rows and 40 cm between stands with 2-4 seeds per hole of 5-7 cm deep (Adu et al., 2014).

According to Chapoto and Ragasa (2013), the application of 1 kg of nitrogen per hectare increases the maize yield by 22 kg in Ghana. This demonstrates the significant role of nitrogen in maize yield. In fact, due to the low nitrogen content in soil, application of fertilizer is highly recommended to improve soil productivity. However, the current fertilizer application rate in Ghana is only 44 kg N per hectare (Chapoto and Ragasa, 2013), which is lower than the recommended rate of 61 kg (Essel et al., 2020).

2.3 Crop Model in Ghana

Several years ago, empirical models, particularly regression analysis, were commonly used in crop production studies (Oteng-Darko et al., 2013). These models were useful in establishing the relationship between relevant variables and extrapolating based on the relations detected in the training dataset (Paustian et al., 2019). However, for many years, mechanistic models have been used in some studies to simulate the growth of real crops by describing the soil-plant-atmosphere system using mathematical equations development (Jame and Cutforth, 1996). These models are able to predict the final state of total biomass or crop yield by providing valuable quantitative information about the significant processes involved in crop growth and development, such as leaf, root, stem, and grain development (Jame and Cutforth, 1996). Furthermore, unlike empirical models, mechanistic model parameters have a physical meaning, which facilitates understanding and interpretation (Cytiva, n.d.).

Various mechanistic models, such as DSSAT (Jones et al., 2003), APSIM (Keating et al., 2003; McCown et al., 1996), AquaCrop (Hsiao et al., 2009; Raes et al., 2009; Steduto et al., 2009), and LINTUL (Spitters and Schapendonk, 1990), have been developed to simulate crop growth and yield formation processes based on meteorological data and soil characteristics. In Ghana, crop models have been used to assess maize variability and yield gaps in Ghana AEZs with DSSAT (MacCarthy et al., 2018), the implications of different sets of climate variables on regional maize yield simulations with LINTUL-5 (Srivastava et al., 2020), the yield gap in smallholder farming with APSIM (Danquah et al., 2020), and the determinants of yield gap by combining LINTUL-1 with statistical and geospatial analyses (Boullouz et al., 2022).

All these studies results revealed that the significant determinants for closing the yield gap in Ghana are soil organic matter, soil water-holding capacity, root zone depth, rainfall, sulfur fertilizer, and nitrogen fertilizer. The use of improved fertilizer, supplementary irrigation, and enhanced farmer practices have the potential to reduce the yield gap (MacCarthy et al., 2018).

CHAPTER 3: DATA AND METHODOLOGY

3.1 Study Area

Ghana is a West African country bordered by Côte d'Ivoire to the west, Burkina Faso to the north, and Togo to the east. With a total area of approximately 238,537 square kilometers, Ghana is situated between 4° and 12° north latitude and 4° and 3° west longitude. It has a population of around 31 million people, with great ethnic diversity. Akan is the most widely spoken local language, while English is the official language (GSS, 2015). Before 2018, the country was comprised of 10 administrative regions, which were subsequently expanded to 16. These regions include Greater Accra, Central, Eastern, Western, Ashanti, Northern, Upper East, Upper West, Volta, Bono, Bono East, Savannah, North East, Oti, Western North, and Ahafo, each with its own regional capital (GSS, 2021; Kendie, 2019). All 272 districts are categorized into six different AEZs – Coastal Savannah, Evergreen Forest, Deciduous Forest, Transitional, Guinea Savannah, and Sudan Savannah – with the three main zones being the humid south, the northern Sudan, and the Sahelian northeast (GSS, 2015).

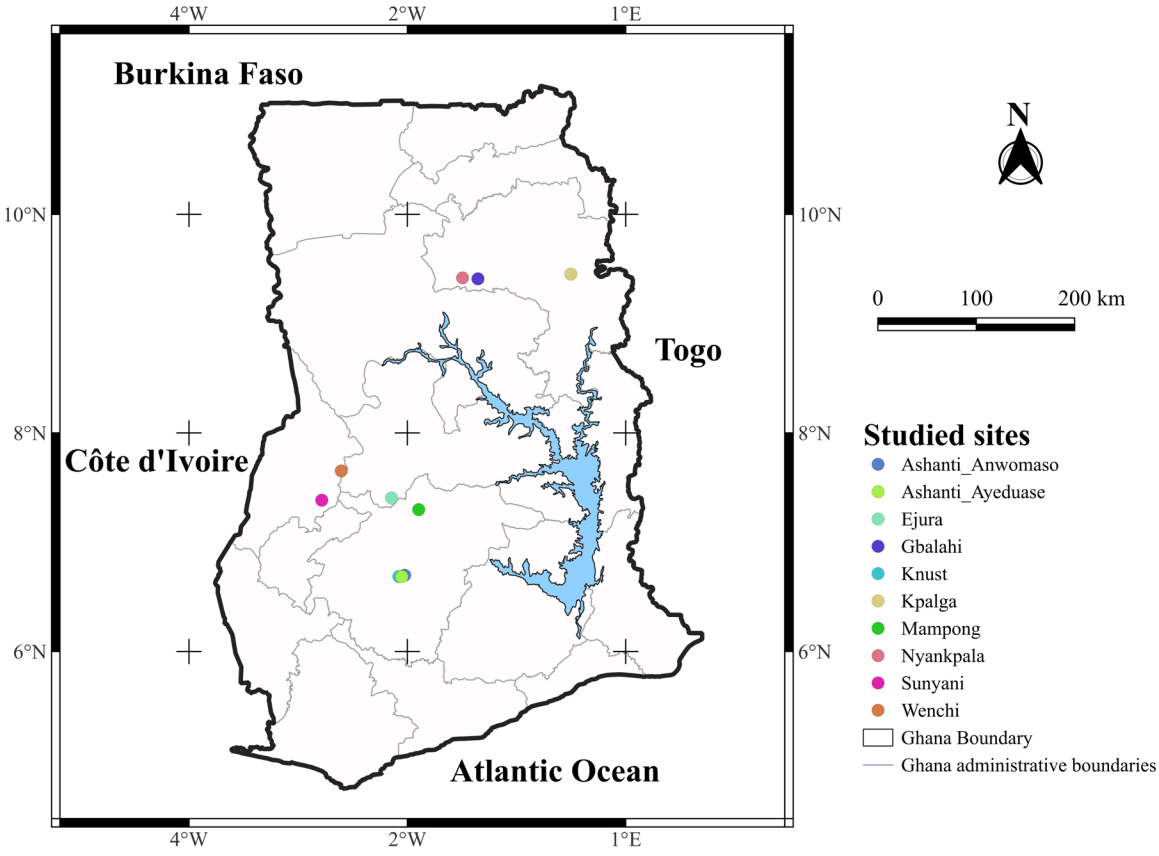


Figure 2. FERARI 2020 study sites selected for this research

In order to identify the key factors related to the water-limited maize yield gap in Ghana, we carefully selected 10 sites (as shown in Figure 2) from the Fertilizer Research and Responsible

Implementation (FERARI[1]) project trials conducted during the 2020 growing season. These selected sites were used to assess the water-limited potential yield of maize and to investigate the drivers of the yield gap. To map the potential yield and yield gap in Ghana, we used 2,236 data points selected from Boullouz et al. (2022). In their study, they selected randomly between 8 and 10 data points per district based on the vegetation band in the land use map of the country. It is important to note that in our study, we used the same study sites, random points, and datasets as Boullouz et al. (2022), by considering two additional study sites – Gbalayi and Kpalga (Figure 2) to provide a more accurate quantification of maize water-limited yield gap in Ghana. Boullouz et al. (2022) investigated this objective under optimal conditions using LINTUL-1, and our study aimed to extend their findings to water-limited conditions.

3.2 Analytical Framework of the Study

Figure 3 presents the analytic framework of this study, which involved 13 steps to achieve the research objectives. The first step in the study involved data collection from different data sources. In fact, to run the LINTUL-2 model, some weather data and soil parameters (Table 1) are required. The soil parameters were obtained from ISRIC – World Soil Information.[2] The weather data could be extracted from Google Earth Engine from ERA5 at 11 kilometer (km) resolution or from NASA[3] Prediction Of Worldwide Energy Resources (POWER) at 1/2° latitude by 5/8° longitude resolution (Table 1). As LINTUL-1 has been calibrated using the temperature and solar distribution of ERA5-Land and also on the basis of the precipitation comparison results presented below, we ran our simulations with ERA5-Land weather data. Additionally, due to the model's sensitivity to rainfall data and distribution, actual rainfall data was acquired from the Ghana Meteorological Agency (GMET) for the 10 trial locations. The GMET rainfall data was complete for Ashanti Anwomaso, Ashanti Ayeduase, Kwame Nkrumah University of Science and Technology (KNUST), Ejura, Nyankpala, Gbalayi, and Kpalga. However, at Sunyani and Wenchi, there were some missing daily rainfall values, which were supplemented with ERA5 rainfall data. Additionally, for Mampong, where the rainfall was relatively negligible, we relied on the ERA5 rainfall data. The remaining five variables (solar radiation, minimum temperature, maximum temperature, vapor pressure, and mean wind speed) were obtained from ERA5-Land to ensure consistency in data source. Since the GMET stations did not cover the entire country, weather data from ERA5-Land was collected for the 2,236 randomly selected data points to perform the spatial analysis of yield gaps in Ghana.

Table 1 gives more information about the sources of the different data used.

[1] https://ifdc.org/projects/fertilizer-research-and-responsible-implementation-ferari/
[2] Available on ISRIC SoilGrids.
[3] Available on NASA POWER.

Table 1. Data source of the variables and parameters used

Variable	Source	Author
Weather data		
Solar radiation (kJ m^{-2} d^{-1})	ERA-5	Hersbach et al. (2020)
Minimum temperature (° Celsius)	ERA-5	Hersbach et al. (2020)
Maximum temperature (° Celsius)	ERA-5	Hersbach et al. (2020)
Vapor pressure (kPa),	ERA-5	Hersbach et al. (2020)
Mean wind speed (m s^{-1})	ERA-5	Hersbach et al. (2020)
Precipitation (mm d^{-1})	GMET	Ghana Meteorological Agency
Soil data		
Water content at saturation (WCST)	ISRIC data	Leenaars et al. (2018)
Water content at field capacity (WCFC)	ISRIC data	(Poggio et al., 2021)
Permanent wilting point (WCWP)	ISRIC data	Leenaars et al. (2018)

Leenaars

Soil properties, crop growth, and yield data for the 10 selected sites were obtained from the FERARI trial datasets. After preprocessing, which involved removing data points with missing actual yield values, normalization of the yield standardizing the numerical variables using respectively "*bestNormalize()"* and the "*scale()"* functions in *R*, and unifying the treatment names applied per plot, the dataset consisted of 460 unique data points. Table 3 presents general information about the trial locations. Next, we calibrated the crop and soil parameters for the LINTUL-2 model. For parameterizing soil parameters, we used soil parameters of the upper 0-30 cm from the soil surface obtained from ISRIC for the 10 study sites and the random points. After LINTUL-2 model parameterization, potential yield, harvest index and vegetative biomass were simulated. The maize yield gap across the ten FERARI study sites was then determined.

Next, we statistically evaluated the observed maize yield and maize yield gap variability across the 10 study sites and then across the different treatments applied. This analysis was performed using analysis of variance (ANOVA) and Tukey post hoc test. Subsequently, we constructed a random forest model with fivefold cross-validation, hyperparameter tuning, and recursive feature selection on various explanatory variables (Table 2), except for the ones used in the simulation with the LINTUL-2 model (Table 1), including climate, precipitation, and soil data to explain the maize yield gap under water-limited conditions. It was implemented using the train and trainControl functions from the caret package in R. The resulting model was configured with ntrees = 500, node size = 1, and mtry = 2. Due to the violation of the validity assumptions of the linear regression model, we were not able to assess the effect of the covariates on the yield gap. So, we used the random forest model results to select the most important variables for explaining the maize yield gap, on which we performed a ridge regression using "*stats*" and "*RidgeCV*" from the libraries "*regressors*" and "*sklearn*" in *Python* to know the direction of their effect. The Pareto principle is commonly applied in model construction, which involves splitting the dataset into a training set (80%) and a testing set (20%) to prevent overfitting (Hemdan et al., 2020). However, this approach is primarily used for forecasting models. Since our study's main objective is to

comprehend the yield gap in Ghana, we will not split our data and instead use the entire dataset for model construction.

Finally, using the calibrated parameters of LINTUL-2, we simulated the maize water-limited potential yield of 2,236 randomly selected data points for 2017 and then calculated the yield gap at these data points using maize observed yield from MoFA. Based on the data points, we interpolated the potential yield and yield gap using ordinary kriging, through the QGIS plugin (Smart-Map) developed by Pereira et al. (2022), to create a geospatial analysis of the maize potential yield and yield gap across Ghana. Then, we use the semivariograms to assess the quality of the interpolations. Finally, we formulated policies for decision-makers to address the yield gap in Ghana.

Table 2. ***Independent variables used to explain the maize yield gap variability in our study site***

Variable	Meaning
N_applied*	amount of nitrogen fertilizer applied
S_applied	amount of sulfur fertilizer applied
Zn_applied	amount of zinc fertilizer applied
Fe_applied	amount of iron fertilizer applied
Potasum_applied	amount of potassium fertilizer applied
Pho_applied*	amount of phosphorus fertilizer applied
pH	pH of the soil
Soil_N*	soil nitrogen content
Mg_meq_100g	magnesium milliequivalents per 100 grams of soil
K_me_100g*	soil exchangeable potassium milliequivalents per 100 grams of soil
eCEC_meq_100g*	effective cation exchange capacity milliequivalents per 100 grams of soil
Soil_OM*	soil organic matter content
Soil_Zn*	soil zinc content
Soil_Fe*	soil iron content
Soil_C*	soil carbon content
Soil_P	soil phosphorus content
TEB_me_100g	total exchangeable bases milliequivalents per 100 grams of soil
EA_meq_100g*	activation energy milliequivalents per 100 grams of soil
BS_*	base saturation
SO4S_mg_kg	sulfate sulfur milligrams per kilogram of soil

* Variables that were retained as relevant drivers by the recursive feature selection.

Table 3. General information on the study sites

Trial Info	Ashanti Anwomaso	Ashanti Ayeduase	Ejura	KNUST	Mampong	Nyankpala	Sunyani	Kpalga	Gbalayi	Wenchi
Latitude*	6.697183	6.697183	7.40472	6.685167	7.298333	9.418917	7.401056	9.453555	9.495556	7.6518
Longitude*	-1.5504	-1.5504	-1.64447	-1.57769	-1.39581	-0.99428	-2.28891	-0.00262	-0.9719444	-2.1056
Planting date*	June 1	June 1	September 16	June 1	September 24	June 24	October 1	June 27	July 9	August 19
Nb. Reps*	4	4	3	4	3	4	1	3	4	3
Nb. Plots*	40	40	30	96	48	88	05	24	40	51
WCSAT**	0.338	0.339	0.339	0.363	0.365	0.382	0.353	0.341	0.382	0.339
WCFC**	0.114	0.117	0.117	0.133	0.136	0.149	0.127	0.121	0.149	0.117
WCWP**	0.048	0.049	0.049	0.062	0.062	0.074	0.056	0.05	0.074	0.049
WHC	0.11	0.09	0.07	0.07	0.04	0.04	0.11	0.05	0.11	0.09
Total amount during the growing season (mm)***	409.9	409.9	217.8	409.9	339.08	581.8	211.57	576.1	540.4	464.26

Source: *From FERARI 2020 trials datasets.
**From ISRIC data.
***From GMET weather data and missing values from ERA5 weather data.

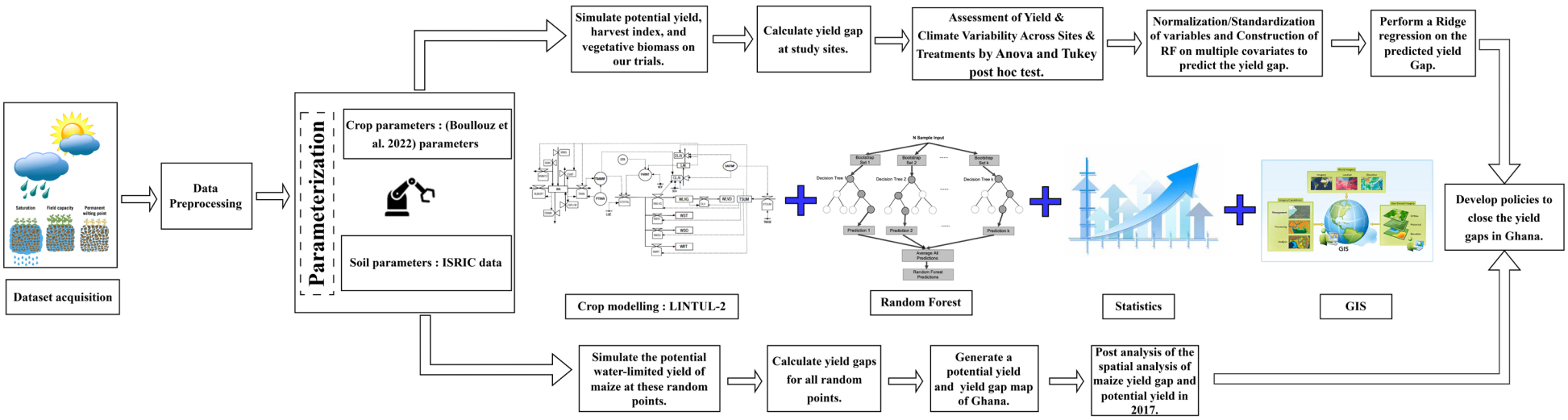


Figure 3. Analytical framework

Table 4 provides a comprehensive overview of the diverse treatments implemented across the trials during the various experiments in the studied sites of the 2020 growing season. These treatments encompassed the application of specific combinations of nitrogen (N), phosphorus (P), potassium (K), sulfur (S), iron (Fe), and zinc (Zn) to the crops. The most applied treatments across the examined sites were NPK, NPK+S, and NPK+Zn. The application rate is described in Table 4.

Table 4. Total numbers of treatment applied in the 10 different trials

Treatment	Observation	Nutrient Rate (kg ha^{-1})					
		N	P_2O_5	K_2O	S	Zn	Fe
Control	40	0	0	0	0	0	0
NKS	19	18	0	25	10	0	0
NPK	69	18-120	20-40	25-40	0	0	0
NPK+S	90	18-120	20-40	25-40	15-10	0	0
NPK+Zn	66	120	20-40	0-40	0	0-2.5	0
NPK+Zn+Fe	32	120	20-40	40	0	2.5	5
NPK+S+Zn	35	120	40	40	15	2.5	0
NPK+S+Zn+Fe	33	120	20-40	40	15	0-2.5	5
NPS	19	18	20	0	10	0	0
PK	19	0	0-20	25	10	0	0
PKS	20	0	20	25	10	0	0
PS	18	0	20	0	10	0	0

Source: FERARI 2020 trials dataset.

3.3 LINTUL-2

This section introduces the LINTUL-2 model. To understand the model, we will first present the LINTUL-1 model and then explain the additional parameters and relationships in LINTUL-2.

3.3.1 LINTUL-1 Overview

LINTUL-1 is a model that simulates the daily biomass production ($GTOTAL$, expressed by Equation 3) based on light interception ($IPAR$, $MJ\ ha^{-1}d^{-1}$) and light use efficiency (LUE, $g\ (DM)MJ^{-1}$) under optimal conditions (ample supply of water, and nutrients in a pest-, disease- and weed-free environment, under the prevailing weather conditions (van Oijen and Leffelaar, 2008a). $IPAR$ is expressed as:

$$IPAR = 0.5 \times DTR \times (1 - e^{-k \times LAI}) \quad (1)$$

The model uses the dry weights of plant organs, including leaves, stems, roots, and grains, as state variables. The daily crop biomass increment is calculated as the product of the amount of intercepted photosynthetically active radiation (IPAR) by the crop and the light use efficiency (Farré et al., 2000). LINTUL-1 GTOTAL is expressed as:

$$GTOTAL_{LINTUL\text{-}1} = \frac{dw}{dt} = LUE \times IPAR \quad (2)$$

Dry matter produced is partitioned among the various plant organs using partitioning factors defined as a measured function of thermal time. The partitioning coefficient of dry matter to the leaves (FLV), stem (FST), roots (FRT), and storage organs (FSO) follows the development stage of the crop. During the vegetative stage, most of the dry matter is allocated to the leaves, then to the roots and to the stems, and during the reproductive stage, the crop allocates the dry matter increasingly more to the storage organs (Boullouz et al., 2022). Partition coefficients for $LINTUL$-1 are expressed as follows:

$$\begin{cases} RWRT_{LINTUL\text{-}1} = GTOTAL_{LINTUL\text{-}1} \times FRT_{LINTUL\text{-}1} \\ RWST_{LINTUL\text{-}1} = GTOTAL_{LINTUL\text{-}1} \times FST_{LINTUL\text{-}1} \\ RWLV_{LINTUL\text{-}1} = GTOTAL_{LINTUL\text{-}1} \times FLV_{LINTUL\text{-}1} \\ RWSO_{LINTUL\text{-}1} = GTOTAL_{LINTUL\text{-}1} \times FSO_{LINTUL\text{-}1} \end{cases} \quad (3)$$

where $RWRT$stands for the leaf growth rate, $RWST$ is the steam growth rate, $RWRT$ is the root growth rate, and RWSO the storage organ growth rate.

3.3.2 Impact of Limited Water Availability on Crop Growth

The difference between LINTUL-2 and LINTUL-1 impacts the simulation results through three key variables. The first variable, the transpiration reduction factor ($TRANRF$), affects total crop growth and the leaf expansion rate in the juvenile stage, while the second, the relative modification of allocation to root by drought ($FRTMOD$), and third, the relative modification of allocation to shoot by drought ($FSHMOD$), variables influence respectively root and shoot partitioning. $FRTMOD$ and $FSHMOD$ are expressed as:

$$FRTMOD = MAX[1., 1./(TRANRF + 0.5)] \quad (4)$$

$$FSHMOD = 1 - FRT/(1. - FRT/FRTMOD) \quad (5)$$

The modified quantities of total daily biomass produced and the distribution of roots and shoots are expressed as follows:

$$GTOTAL_{LINTUL\text{-}2} = \frac{dW}{dt} = GTOTAL_{LINTUL\text{-}1} \times TRANRF \quad (6)$$

$$\begin{cases} FRT_{LINTUL\text{-}2} = FRT_{LINTUL\text{-}1} \times FRTMOD \\ FST_{LINTUL\text{-}2} = FST_{LINTUL\text{-}1} \times FSHMOD \\ FLV_{LINTUL\text{-}2} = FLV_{LINTUL\text{-}1} \times FSHMOD \\ FSO_{LINTUL\text{-}2} = FSO_{LINTUL\text{-}1} \times FSHMOD \end{cases} \quad (7)$$

$$\begin{cases} RWRT_{LINTUL\text{-}2} = GTOTAL_{LINTUL\text{-}2} \times FRT_{LINTUL\text{-}2} \\ RWST_{LINTUL\text{-}2} = GTOTAL_{LINTUL\text{-}2} \times FST_{LINTUL\text{-}2} \\ RWLV_{LINTUL\text{-}2} = GTOTAL_{LINTUL\text{-}2} \times FLV_{LINTUL\text{-}2} \\ RWSO_{LINTUL\text{-}2} = GTOTAL_{LINTUL\text{-}2} \times FSO_{LINTUL\text{-}2} \end{cases} \quad (8)$$

3.3.3 Soil Water Balance

The soil water balance is added to LINTUL-2 that drives water availability and the impact on crop production. It is determined by various factors such as precipitation, irrigation, runoff or run-on, percolation to or capillary rise from deeper soil layers, and evapotranspiration of water from the soil surface and the crop (van Oijen and Leffelaar, 2008b).

The actual rate of water loss from the soil surface by evaporation ($EVAP, mm\ d^{-1}$) and from the crop by transpiration ($TRAN, mm\ d^{-1}$) depends on the potential values, the water content (WC, $m^3\ H_2O\ m^{-3}$soil): the water content at saturation ($WCST$, $m^3\ H_2O\ m^{-3}$ soil), at field capacity ($WCFC$, $m^3\ H_2O\ m^{-3}$ soil), at wilting point ($WCWP$, $m^3\ H_2O\ m^{-3}$ soil), and at air dryness ($WCAD$, $m^3\ H_2O\ m^{-3}$ soil) and the soil characteristics (Nyombi, 2010). Water content affects many stages of crop development such as emergence, root growth, crop growth rate, and the allocation of biomass over roots and shoot of the crop. In fact, all those stages could be hampered if the water content is below the critical water content ($WCCR$, $m^3\ H_2O\ m^{-3}$ soil) or more than the $WCFC$ (van Oijen and Leffelaar, 2008b). $WCCR$ is expressed as:

$$WCCR = WCWP + MAX\left(0.01, \frac{PTRAN}{PTRAN + TRANCO}(WCFC - WCWP)\right) \quad (9)$$

where $PTRAN$ is the potential transpiration rate of the crop and $TRANCO$ is the transpiration coefficient. From the observations made in Figure 4, showing the relational diagram of LINTUL-2, it can be inferred that water balance of the soil involves three main sources of supply: precipitation, irrigation, and water absorption by roots as they elongate to reach new layers. However, it should be noted that a portion of this water is lost and does not penetrate the soil. This loss is due to several factors, such as interception by plant leaves, runoff, drainage, plant transpiration caused by solar radiation, and soil evaporation. The amount of water in the single rooted soil layer at time t is:

$$w_t = w_0 + \left\{\left(R + I + \frac{dw_{exp}}{dt}\right) - (R_{int} + RN + D + TRAN + EVAP)\right\}\Delta_t, \quad (10)$$

where R is the rain ($RAIN$, mm d^{-1}), I is the Irrigation (IRRIG, mm d^{-1}), $\frac{dw_{exp}}{dt}$ is the water explored by the growing roots ($EXPLOR$, mm d^{-1}), R_{int} is the intercepted rain ($RNINTC$, mm d^{-1}), RN is the runoff ($RUNOFF$, mm d^{-1}), D is the drainage rate ($DRAIN$, mm d^{-1}), and Δ_t is the time step and equals to one day.

The modeling of soil water balance is typically achieved through two primary methods: the tipping bucket approach and the Richards approach.

3.3.4 Potential and Actual Rates of Evaporation and Transpiration

Crop evapotranspiration (ET, expressed by Equation 12) is crucial in agricultural systems and the hydrological cycle, with over 90% of water used in agriculture lost by soil evaporation and crop transpiration (Ding et al., 2013). Evaporation refers to the process by which water is converted from its liquid state to its gaseous state and is influenced by factors such as temperature, humidity, wind, and solar radiation. On the other hand, transpiration is the process by which plants absorb water through their roots and release it into the atmosphere through their leaves, mainly through tiny pores called stomata (Wang and Wang, 2022). Both of these processes are essential for regulating the water balance of crops, as they enable the plants to maintain their water content and prevent excessive water loss. Furthermore, transpiration plays a vital role in the uptake of essential nutrients by the plants, as water is the medium through which nutrients are transported from the soil to the plant cells (Wang and Wang, 2022).

The Penman equation is used to determine potential evapotranspiration, including the potential transpiration ($PTRAN$, expressed by Equation 14) and the potential evaporation ($PEVAP$, expressed by Equation 13) (Farré et al., 2000; Nyombi, 2010). The actual transpiration ($TRAN$, expressed by Equation 11) and the actual evaporation ($EVAP$) rates are determined by considering the potential transpiration and the soil water status and the potential evaporation and the soil water status, respectively (Penman, 1948; van Oijen and Leffelaar, 2008b).

$$TRAN = TRANRF \times PTRAN \tag{11}$$

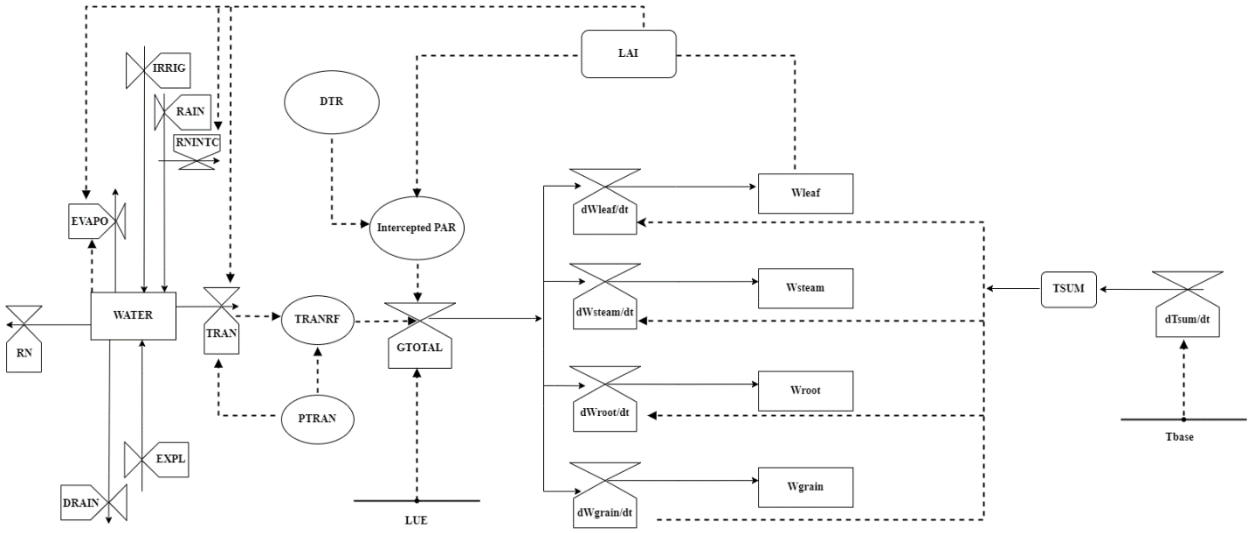
$$ET = \frac{\Delta}{\Delta + \gamma}\frac{R_{net}}{\lambda} + \frac{\gamma}{\Delta + \gamma}\frac{E_{air}}{\lambda} \tag{12}$$

$$PEVAP = e^{(1-0.5L)}\left\{\frac{\Delta}{\Delta + \gamma}\frac{R_{net\_soil}}{\lambda} + \frac{\gamma}{\Delta + \gamma}\frac{E_{air}}{\lambda}\right\} \tag{13}$$

$$PTRAN = MAX\left(0\,, e^{(1-0.5L)}\left\{\frac{\Delta}{\Delta + \gamma}\frac{R_{net\_canopy}}{\lambda} + \frac{\gamma}{\Delta + \gamma}\frac{E_{air}}{\lambda}\right\} - 0.5\,R_{int}\right) \tag{14}$$

where Δ is the slope of saturated vapor pressure curve; E_{air} is the aerodynamic term; γ is the adiabatic psychrometer; λ is the heat of vaporization, and R_{net} is the net radiation term.

Evaporation decreases as soil water content falls below field capacity ($WC < WCFC$) and stops when it becomes airdry ($WC = WCAD$). Under ample water supply, water uptake rate by the crop follows the potential transpiration rate closely. But when soil water content goes below a critical level ($WC < WCCR$), actual transpiration ($TRAN$) decreases, and the stomata closes. The critical soil water level ($WCCR$) lies between wilting point and field capacity ($WCWP < WCCR < WCFC$), which depends on crop characteristics expressed in the "transpirational constant" ($TRANCO$). Too much water in the soil can hamper crop functioning, including transpiration when soil water content is higher than the waterlogging threshold ($WC > WCWET$) (van Oijen and Leffelaar, 2008b).



Source: Author creation from Nyombi (2010) and van Oijen and Leffelaar (2008b).

Note: Wleaf, weight of leaves; Wstem, weight of stems; Wroot, weight of roots; Wgrain, weight of grain; Tbase, the low lowest temperature for the crop to grow; and Tsum, accumulated temperature.

Figure 4. Relational diagram of LINTUL-2 to simulate the water-limited potential yield

3.4 Methods of Data Analysis: Tools and Techniques

In this section, we outline the various tools and techniques employed to estimate the maize water-limited yield in Ghana.

3.4.1 Rainfall Data Comparison

A time series is a set of statistical observations arranged in chronological order. The analysis of these kind of data finds application in various domains, including weather forecasting, sales analysis, and the health domain, enabling a better understanding of underlying effects and facilitating informed decision-making (Jose, 2022). Time series comprise four different

components: trend, seasonal variations, the cyclical variations, and random fluctuations (Chatfield and Xing, 2019; Jose, 2022). The trend represents a consistent pattern in the data, while the other variations encompass changes occurring at some specific moments (Dagum and Cholette, 2006). These variations may be influenced by various unseen factors, such as increased sales during specific months of the year or fluctuations in unemployment rates between seasons. So, it is important to separate these components during analysis to avoid drawing incorrect conclusions.

Given the significance of rainfall and soil parameters in our model and the limited availability of crop modeling data, particularly in African countries, we aimed to determine whether ERA5 weather data or NASA POWER weather data is more closely aligned with actual weather data from GMET. This will help to know which data source, between NASA POWER and ERA5, to use for the random points. Our focus will be on the rainfall trend, as it offers insights into the overall evolution of the series. While several tools are used to de-seasonalize time series data, the moving average is one commonly employed method (Chatfield and Xing, 2019). However, since our data spans only one year, it may not effectively reveal differences between all components of the rainfall distribution. To address this, we employed a local polynomial regression, which combines local regression and moving average techniques. This method proves to be highly effective in removing noise from data and smoothing it out. We anticipate that the seasonal/random variation will be treated as the noisy components and will be effectively captured through the use of local polynomial regression.

In the local polynomial regression, the smoothed value of the variable is estimated by the function:

$$m(z) = \sum_{j=1}^{p} \beta_j (z - x)^j, \tag{15}$$

for z in the neighborhood of x defined by the bandwidth $h(x)$ as the window $[x - h(x), x + h(x)]$ (Fan and Gijbels, 1996; Ledolter, 2008). The bandwidth is a critical parameter, and selecting an appropriate value is crucial for obtaining accurate estimations. If the bandwidth is too small, only a few local variables will contribute to the estimation, while too high of a value can lead to estimations that are very close to the original values and consistently biased (Ruppert et al., 1995). To address this, we used the "*dpill*" function to select the appropriate bandwidth for our data.

After smoothing the rainfall data, we used two different metrics to compare ERA5 and NASA POWER rainfall data with GMET rainfall data: the person correlation coefficient and the mean absolute distance ($MADI$) derived from the mean absolute error. $MADI$ is expressed as:

$$MADI = \frac{1}{n} \sum_{i=1}^{n} \left| RS_{GMET} - RS_{platform} \right|, \tag{16}$$

where RS_{GMET} represents the smooth rainfall data from GMET, and $RS_{platform}$ represents the smooth rainfall data from either ERA5 or NASA POWER.

3.4.2 Exploratory Analysis

In order to compare the distribution of a variable among two or more groups, both parametric and non-parametric tests have been developed (Orcan, 2020). Parametric tests, unlike non-parametric ones, assume that the key variable is normally distributed. However, most statistical tests suffer from the issue of overlapping treatments that belong to different groups simultaneously (Conrado et al., 2017). To address this problem, the Scott-Knott test was developed as a statistical method to rank and compare the means of multiple groups (Jelihovschi et al., 2014). The Scott-Knott test is a hierarchical clustering algorithm proposed by Alastair J. Scott and Martin Knott in 1974, designed for situations where the groups have an equal number of observations (Jelihovschi et al., 2014).

In the present investigation, we will be using ANOVA with the Tukey post hoc test to evaluate the variability of maize yield gap across the 10 study sites and the different treatments applied. This choice is motivated by the uneven distribution of fertilizer treatments. To implement the ANOVA and Tukey post hoc test, we utilized the R packages agricolae and dplyr, which provide an efficient and user-friendly implementation of the algorithm.

3.4.3 Explanatory Analysis

Here, we will explain the different approaches used to assess the impact of the independent variables on the yield gap. They are respectively: the random forest for the most important variables selection, the multicollinearity analysis between the independent variables selected, the construction of the ridge regression, and interpretation of the results. Although MLR could potentially provide a good fit to quantify the maize yield gap in our trials, it was not a suitable choice for our study due to the violations of the assumptions of homoscedasticity, normality, and non-autocorrelation of errors, even after transforming variables. As a result, we chose ridge regression as a more suitable approach.

To achieve this, we focused exclusively on eight of the 10 stations, excluding Ejura and KNUST due to the presence of negative yield gap values. Therefore, for this yield gap analysis, these stations were excluded from consideration.

3.4.3.1 Random Forest Explanation

Random forest is an machine learning ensemble method which can solve both regression and classification problems (Breiman, 2001). This technique constructs some bag of decision trees, called "base learners" or "weak learners", $h_1(x)$, …, $h_J(x)$ (prediction of the response variable at x using the j^{th} tree), and these base learners are combined to give the "ensemble predictor" $f(x)$ (Cutler et al., 2012; Mohapatra et al., 2020). In regression, the base learners are averaged (Cutler et al., 2012). The method aims to find a prediction function $f(X)$ that predicts the real-valued response variable Y based on the real-valued input or predictor variables X.

The machine learning has multiple advantages, such as fast training and prediction, and can be used for high-dimensional problems. It can also provide measures of variable importance, missing value imputation and outlier detection (Cutler et al., 2012; Mohapatra et al., 2020).

This process is achieved by minimizing the expected value of the loss function $L(Y, f(X))$:

$$E_{XY}\left(L(Y, f(X))\right) \tag{17}$$

where L is a measure of how close $f(X)$ is to Y and the subscripts the joint distribution of X and Y. The choice of L depends on the type of work we are doing. For the regression task:

$$L(Y, f(X)) = \left(Y - f(X)\right)^2 \tag{18}$$

As said before, the final output of the model is a function of the different output of the weak learner's $h_J(x)$. This aggregation also depends on the task. For the regression:

$$f(x) = \frac{1}{J}\sum_{j=1}^{J} h_j(x) \tag{19}$$

3.4.3.2 Multicollinearity Analysis

Before the model building, it was important to avoid multicollinearity among the covariates. Multicollinearity appears when there is a high degree of correlation among the predictors (Morales-Oñate and Morales-Oñate, 2021). To check for multicollinearity, we used the Pearson correlation matrix at the significance threshold of 5%. When two variables have a correlation higher than 85%, one of them is removed from the analysis.

3.4.3.3 Ridge Regression

For decades, MLR has been used in various fields to identify linear relationships in datasets. The standard linear model assumes that:

$$y_i = \beta_1 x_{i1} + \cdots + \beta_{ip-1} x_{ip-1} + \beta_p x_{ip} + e_i \text{ for } i = 1, \ldots, n \tag{20}$$

Where n is the sample size, $x_{i1}, \ldots, x_{ip}$ are the explanatory variables, and y_i is the response variable. It is based on several assumptions, including the normality of errors, no multicollinearity among independent variables, homoscedasticity of error variance, and independence of observations (Schmidt and Finan, 2018). Meeting all the assumptions required for MLR is indeed a significant challenge during the model building process. One of the key challenges, particularly in the presence of multicollinearity, has led to the introduction of alternative approaches such as ridge regression, Stein shrinkage, least absolute shrinkage selection operator (LASSO), and elastic net (R.W. Hoerl, 2020).

A regularization term was introduced by A.E. Hoerl and Kennard (1970) in ridge regression to address the instability caused by multicollinearity. This regularization term is added to the least squares estimate, resulting in the ridge regression estimate:

$$\hat{\beta}^* = [X'X + kI]^{-1}X'Y; k \geq 0 \tag{21}$$

With I the unit matrix, k the family of the estimate, and X the matrix of the covariates (Abdelgadir and Eledum, 2016; A.E. Hoerl and Kennard, 1970). The ridge regression estimate can also be expressed as a function of the ordinary least squares estimate $(\hat{\beta})$:

$$\hat{\beta}^* = \left[I_p + k(X'X)^{-1}\right]^{-1}\hat{\beta}; k \geq 0 \tag{22}$$

By incorporating this small positive value into the diagonal, we achieve a less biased estimate that is more desirable than the ordinary least squares estimate, as it is closer to the true value.

3.4.3.4 Quality Evaluation of the Model

After building the MLR model, we evaluated the assumptions of the linear regression model, which included the normality of errors, homoscedasticity, independence of errors, linear relationship between the dependent variable and covariates, and absence of multicollinearity (Poole and O'Farrell, 1971; Schmidt and Finan, 2018). The Breusch-Pagan test was used to assess homoscedasticity with the null hypothesis "H_0: the variances are equal," the Shapiro-Wilk test for normality with the null hypothesis, "the distribution of the variable is normal," and the Durbin-Watson test for independence of errors with the null hypothesis "the independence of errors" (Palm and Iemma, 2002). After assessing all assumptions for the MLR, we used performance metrics of linear regression, including the root mean square error ($RMSE$, expressed in Equation 23), coefficient of determination (R^2, expressed in Equation 24), and adjusted coefficient of determination ($R^2{}_{adjusted}$, expressed in Equation 25), to determine how much the covariates explained the variance of the maize yield gap. While R^2 overestimates the amount of variance explained by the covariates and is sensitive to the number of covariates used in the model, for this study we used the $R^2{}_{adjusted}$, which is more robust (Karch, 2020).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{23}$$

$$R^2 = 1 - \sum_{=1}^{n}\frac{(y_k - \hat{y}_k)^2}{(y_k - \bar{y})^2} \tag{24}$$

$$R^2{}_{adjusted} = 1 - \left(\frac{n-1}{n-p-1}\right) \times (1 - R^2) \tag{25}$$

where y_k is the observed value of the dependent variable, $\hat{y}_k$ is the predicted value of the dependent variable, n is the total sample size, and p is the number of predictors.

CHAPTER 4: RESULTS AND DISCUSSION

4.1 Results of Our Analysis

This section provides an overview of the findings derived from our analysis.

4.1.1 Parametrization of LINTUL-2 Maize

4.1.1.1 Rainfall Comparison

In this analysis, we conducted a comparison of the rainfall data from ERA5 and NASA POWER for 2020, 2021, and 2022 with the actual rainfall values obtained from the GMET at three locations: Tamale, Ejura, and Kumasi. The purpose of this comparison was to determine which platform's rainfall predictions were closest to the observed values, enabling us to select the most suitable platform for downloading random points weather data.

4.1.1.1.1 Comparison of the Smooth Distributions

Figure 5 illustrates the comparison of smooth daily rainfall data for the year 2020 obtained from GMET, ERA5, and NASA POWER. At the Tamale location, the smooth rainfall distribution from NASA POWER closely resembles that of GMET, showing a positive correlation coefficient of 0.911. Moreover, the mean absolute distance between these two distributions is 0.734, which is half of the mean absolute distance observed between GMET and ERA5 smooth rainfall data. In contrast, at Kumasi, the correlation coefficients between ERA5, NASA POWER, and GMET are nearly identical. However, the mean absolute distance between ERA5 and GMET smooth rainfall is 1.231, which is approximately three-fifths of the mean absolute distance between GMET and NASA POWER smooth rainfall. Lastly, at Ejura, the NASA POWER smooth rainfall exhibits higher correlation with GMET smooth rainfall. Nevertheless, the mean absolute distance between ERA5 and GMET smooth rainfall is smaller than that observed between NASA POWER and GMET smooth rainfall.

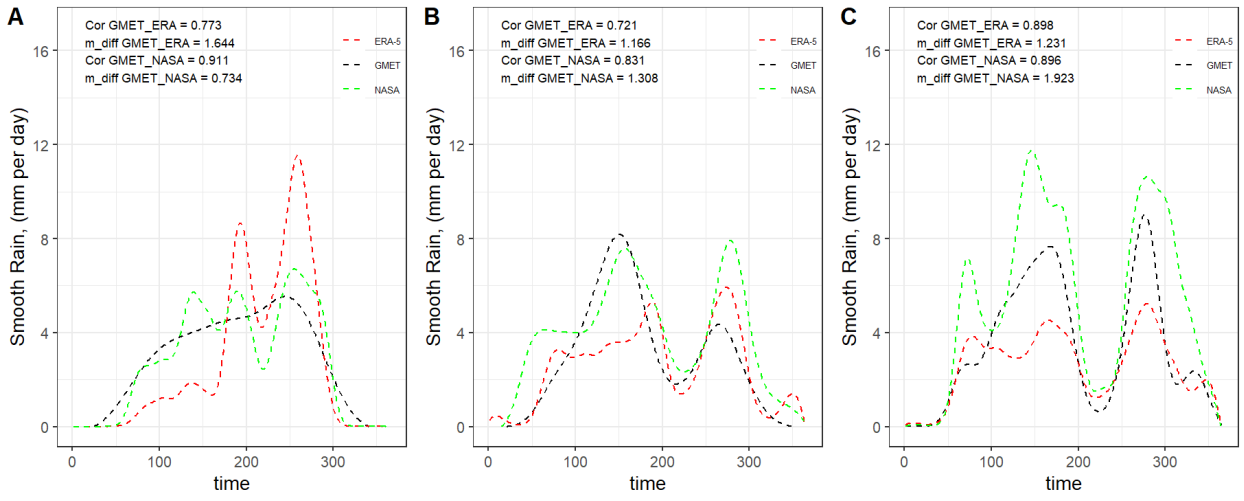


Figure 5. Comparison of 2021 daily smooth rainfall data from GMET, ERA5, and NASA POWER datasets for two locations: (A) Tamale, (B) Ejura, and (C) Kumasi

Figure 6 depicts the comparison of smooth daily rainfall data from NASA POWER, GMET, and ERA5 for the year 2021. In the Tamale location, the correlation coefficient between GMET and ERA5 smooth rainfall and GMET and NASA smooth rainfall is nearly identical, indicating similar trends in all these smooth data. Although there is a slightly smaller mean absolute distance between GMET and NASA POWER smooth rainfall, this metric remains very similar, differing only by a few dozen for ERA5 and NASA POWER.

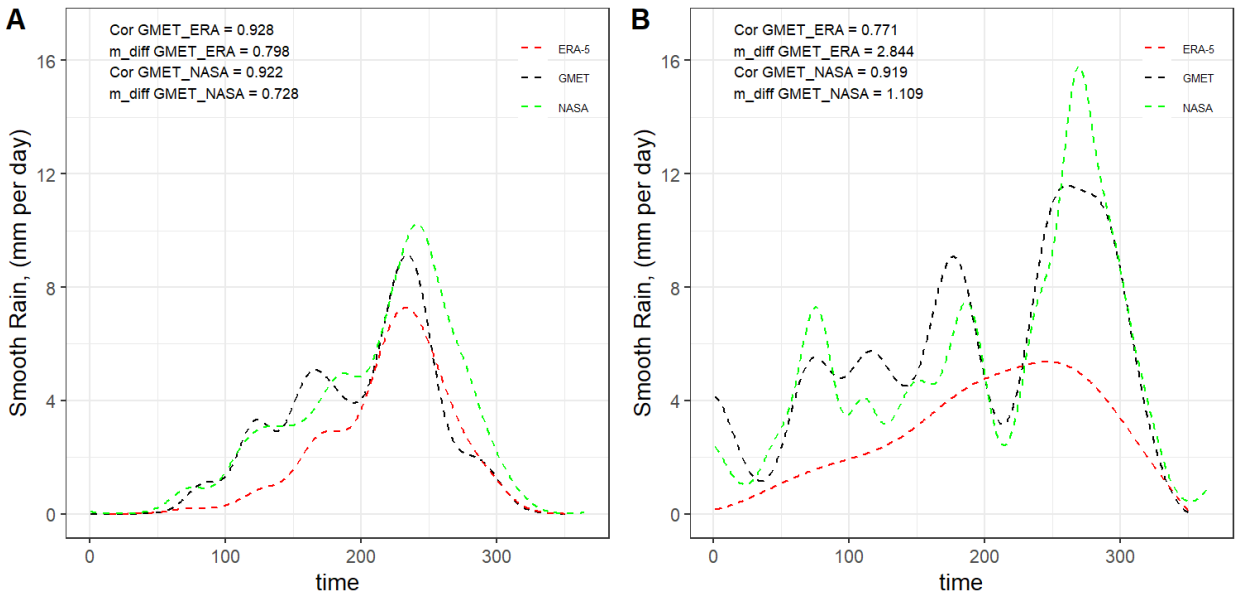


Figure 6. Comparison of 2020 daily smooth rainfall data from GMET, ERA5, and NASA POWER datasets for two locations: (A) Tamale and (B) Ejura

Figure 7 illustrates the comparison of smooth daily rainfall data between ERA5, NASA POWER, and GMET. In the regions of Ejura and Kumasi, we observed that ERA5 and GMET smooth rainfall data show a stronger correlation and have a lower mean absolute distance compared to NASA POWER and GMET smooth rainfall.

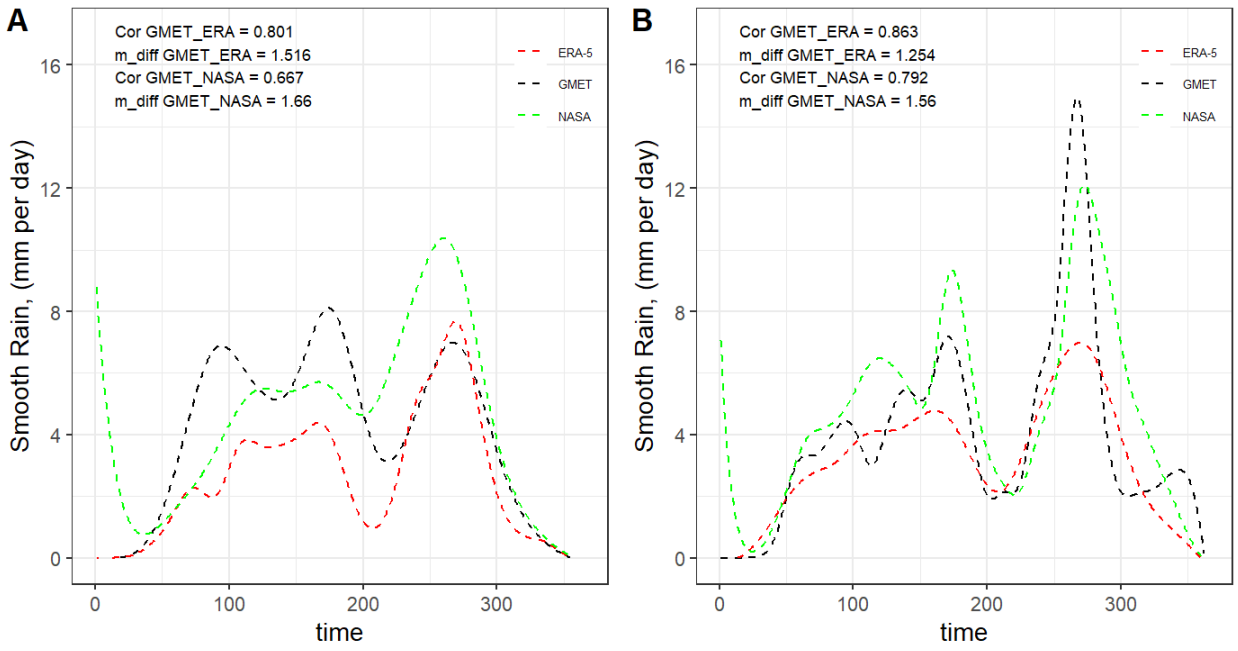


Figure 7. Comparing 2022 daily smooth rainfall data from GMET, ERA5, and NASA POWER datasets for two locations: (A) Ejura, and (B) Kumasi

Based on our observations, we have noted that in three out of the seven assessments conducted on locations where we have complete rainfall data from GMET, ERA5 smooth rainfall exhibits a stronger correlation with GMET smooth rainfall. However, it is crucial to consider the mean absolute distance as a more significant metric, as it provides insights beyond the direction of the trend. According to the mean absolute distance analysis, ERA5 smooth rainfall is found to be closer to GMET smooth rainfall than NASA POWER in four out of seven assessments. This finding is consistent with the results obtained by Gleixner et al. (2020), who reported that ERA5 significantly reduced the bias in temperature and rainfall data compared to the previous version, ERA-Interim, across the African continent. We have therefore decided to use GMET data where available and ERA5 data where no data is available.

4.1.1.1.2 Simulation Results

In addition to rainfall, soil properties, such as soil water content at saturation, field capacity, and wilting point, play a crucial role in determining the available water for plant growth. Due to data limitations, we relied on existing literature, particularly the work by Saxton et al. (1986) and soil data provided by ISRIC (Leenaars et al., 2018; Poggio et al., 2021). Figure 8 illustrates the simulation results of maize water-limited yield at ten study sites using weather data from ERA5, NASA POWER, and a combination of ERA5 and GMET rainfall, along with soil parameters from the aforementioned sources.

Upon analysis, we observe that when employing the soil parameters from Saxton et al. (1986), the predicted maize water-limited yield exceeds 6 mt ha^{-1} at three, five, and four locations for GMET+ERA5, NASA POWER, and ERA5 weather data, respectively. Conversely, when utilizing the ISRIC parameters, the predicted maize water-limited yield exceeding 6 mt ha^{-1} is only found

at three, four, and three locations for GMET+ERA5, NASA POWER, and ERA5 weather data, respectively. Application of the Saxton et al. (1986) parameters results in predicted maize water-limited yields of greater than 4-6 mt ha^{-1}, as estimated by the International Food Policy Research Institute (IFPRI ,2014) at many locations. However, due to the use of only soil texture in our trials to determine soil water characteristics from (Saxton et al., 1986) and considering that ISRIC data originates from spatial analysis of soil characteristics, we have chosen to prioritize the ISRIC data for our simulations.

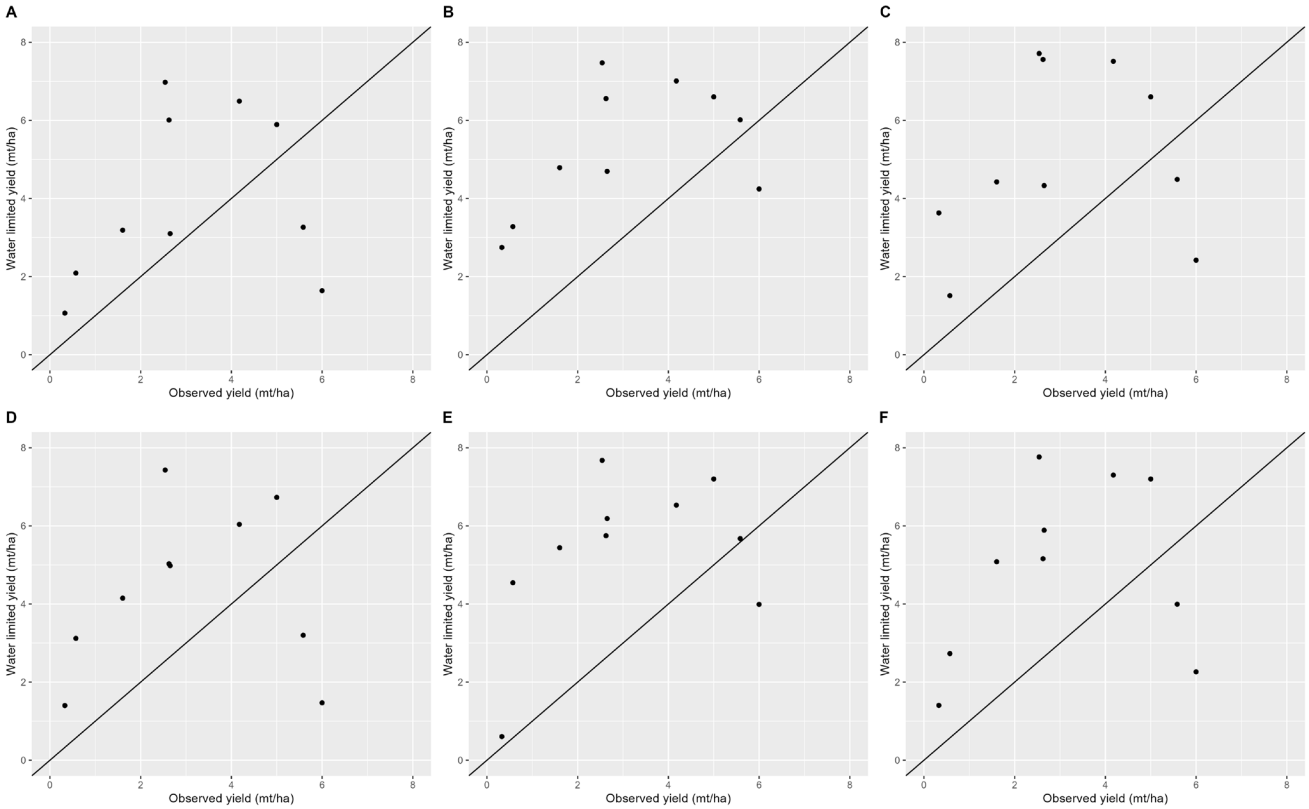


Figure 8. LINTUL-2 simulation results comparing different weather data sources and soil parameters: (A) GMET and Saxton, (B) NASA POWER and Saxton, (C) ERA5 and Saxton, (D) GMET and ISRIC, (E) NASA POWER and ISRIC, and (F) ERA5 and ISRIC

4.1.1.2 Calibrated Values Retained

Table 5 displays the calibrated values used in this study for simulating the water-limited potential yield of maize in Ghana.

Table 5. Calibrated parameters to simulate the water-limited potential yield of maize in Ghana

Parameters	Original Value	Source	Recalibrated Value	Source
Tsum Maturity	1750 GDD	Farré et al. (2000)	1796	Boullouz et al. (2022)
Tsum Anthesis	970 GDD	Farré et al. (2000)	1027	Boullouz et al. (2022)
SLA	0.016	Farré et al. (2000)	0.032	Srivastava et al. (2020)
Partitioning coefficient of dry matter		Farré et al. (2000)		Boullouz et al. (2022)
RGRL	0.009 °C d^{-1}	Farré et al. (2000)	0.02 °C d^{-1}	Srivastava et al. (2020)
ROOTDM	1.2 m	Spitters and Schapendonk (1990)	1 m	Srivastava et al. (2020)
WCSAT, WCFC, WCWP	0.55, 0.36, and 0.23, respectively	Spitters and Schapendonk (1990)	See table 3	ISRIC

WCSAT, water content at saturation; WCFC, water content at field capacity; WCWP, water content at wilting point.

4.1.2 Yield Observations in FERARI Field Experiments

Figure 9 presents the observed and model-predicted water-limited yields after simulation. The LINTUL-2 model predicts a range of water-limited potential yields from 1.4 to 7.43 mt ha^{-1}. The highest predicted yield is observed in Gbalahi, located in the northern region of Ghana. This value is slightly higher than the 4-6 mt ha^{-1} found by IFPRI for Ghana (IFPRI, 2014; Scheiterle and Birner, 2018). The other two northern locations also have yields greater than 5 mt ha^{-1}. On the other hand, the smallest predicted yield is observed in Mampong, one of the towns in the Ashanti region (Figure 9).

Across the study sites, the model predicts higher water-limited yields compared to the observed yields in eight locations (Figure 9). However, at Ejura and KNUST, the predicted yields are lower than the yields observed in the field, which suggests that our model may not generalize well or that there might be inaccuracies in the weather data for these specific locations. This discrepancy should not occur unless the rainfall data are not accurate, as the model assumes all parameters to be optimum except for water availability. In the following sections, we will investigate the reasons behind these unexpected results in these locations.

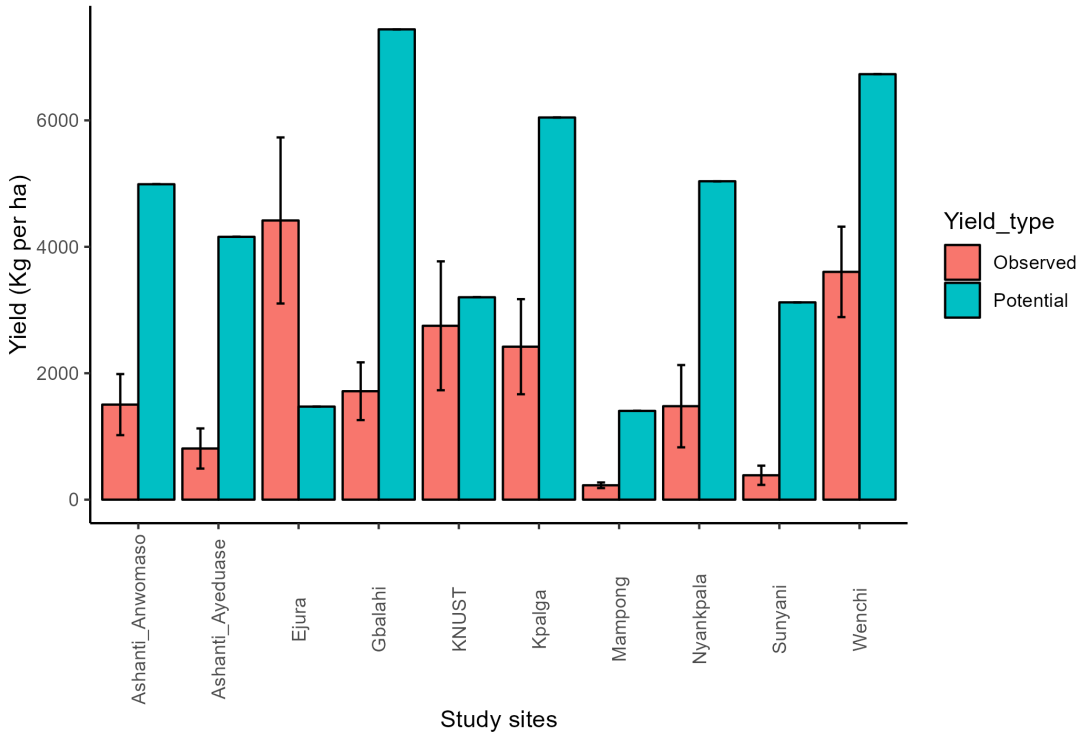
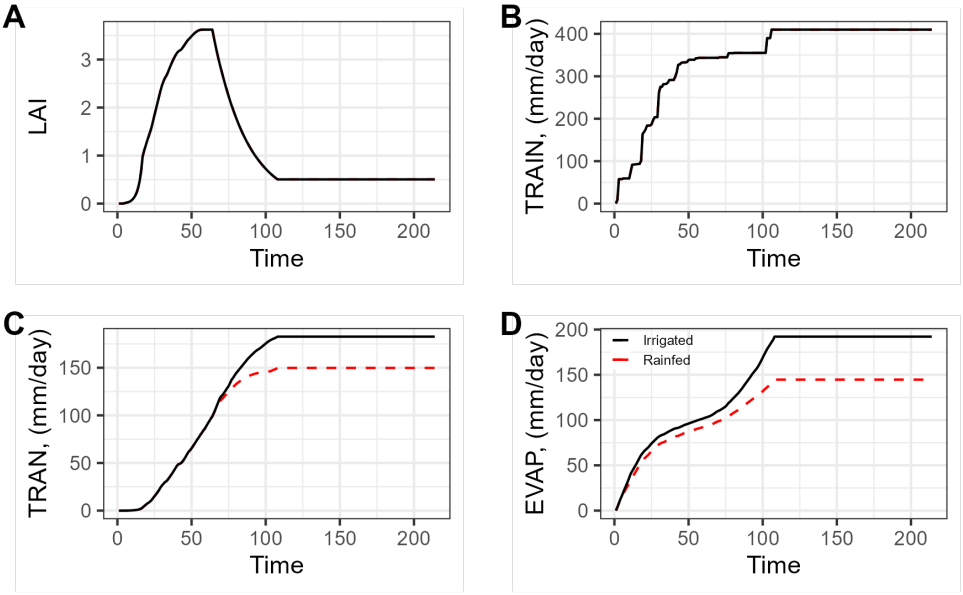


Figure 9. Water-limited potential yield and observed maize yield by study site

4.1.2.1 Ashanti Anwomaso

The predicted potential water-limited yield at Ashanti Anwomaso is 4.98 mt ha^{-1} (Figure 9). Figure 10 illustrates that the leaf area index (LAI) of maize at this location reaches 3.6, which is comparable to the LAI under irrigation conditions. This can be attributed to the favorable distribution of rainfall during the vegetative period, amounting to approximately 300 mm. This rainfall amount is not far from the required water for maize cultivation, estimated to be between 450 and 600 mm, according to Du Plessis (2003). The crop rainfed transpiration rate remains consistent with the transpiration rate under irrigation conditions until reaching an asymptotic threshold after 65 days post-sowing. The rapid attainment of this threshold compared to the plant's transpiration under irrigation can be attributed to the limited rainfall received during the period from 65 to 100 days after sowing. As a result, both the crop rainfed evapotranspiration (comprising maize transpiration and soil evaporation) and the water-limited potential yield are reduced, estimated at 8.49 mt ha^{-1} by Boullouz et al. (2022).

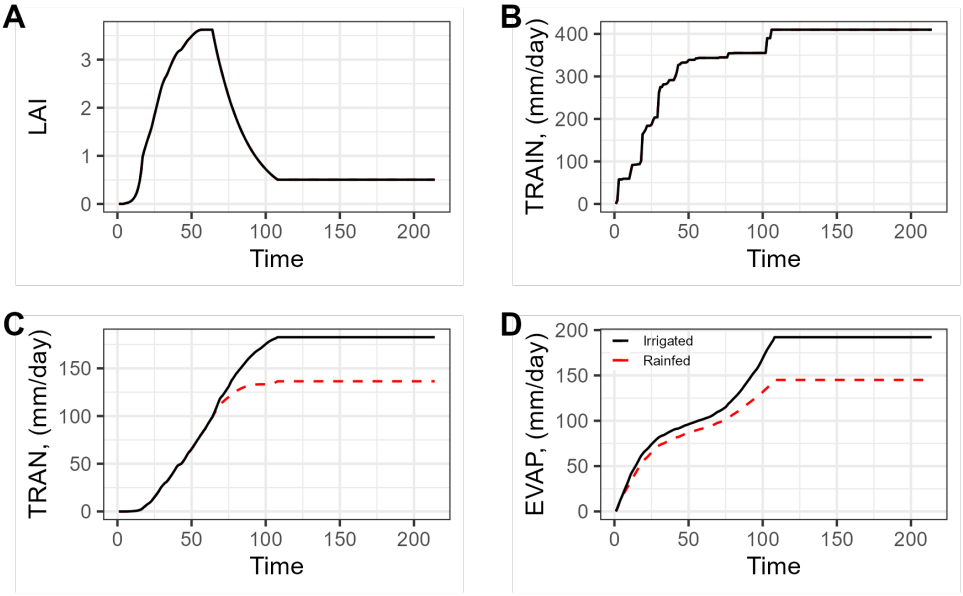


The LAI and TRAIN curves for irrigation and rainfed conditions overlap. LAI, Leaf Area Index; TRAN, Transpiration; TRAIN, Total Rain; and EVAP, Evaporation.

Figure 10. LAI (A), TRAN (B), TRAIN (C), and EVAP (D) simulation results by LINTUL-2 at Ashanti Anwomaso

4.1.2.2 Ashanti Ayeduase

At Ashanti Ayeduase, the predicted potential water-limited yield is 4.15 mt ha^{-1} (Figure 9), which is 0.84 mt ha^{-1} lower than the observed yield at Anwomaso. However, the growth patterns of maize at Ayeduase, as depicted in Figure 11, are quite similar to those observed at Anwomaso, with the only notable difference being the crop transpiration. At Ashanti Anwomaso, the crop rainfed transpiration reaches 150 mm, while at Ayeduase, it is approximately 135 mm. This difference can be attributed to the slightly lower water-holding capacity of the soil at Ayeduase compared to Anwomaso, with the latter being 2% higher (Zhang et al., 2021) (Table 3). Consequently, despite experiencing similar rainfall patterns after 65 days post-sowing, the soil at Anwomaso can retain more water, which the maize crop can utilize for transpiration and other physiological needs. This disparity in water availability accounts for the difference in yield between the two locations (He and Wang, 2019) and aligns with the findings of Boullouz et al. (2022), who reported a yield of 8.28 mt ha^{-1} for the same site.

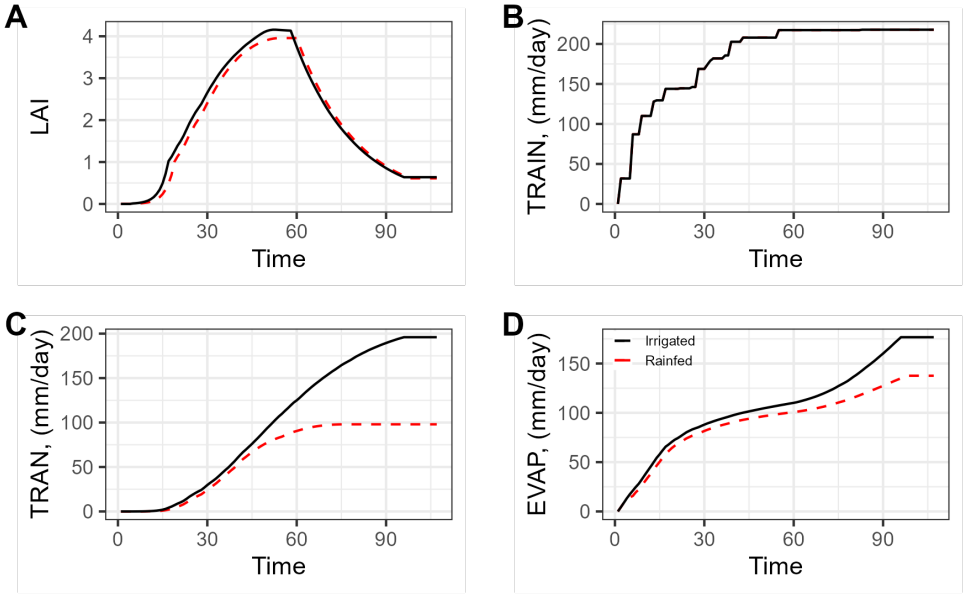


The LAI and TRAIN curves for irrigation and rainfed conditions overlap. LAI, Leaf Area Index; TRAN, Transpiration; TRAIN, Total Rain; and EVAP, Evaporation.

Figure 11. LAI (A), TRAN (B), TRAIN (C), and EVAP (D) simulation results by LINTUL-2 at Ashanti Ayeduase

4.1.2.3 Ejura

The LAI at Ejura reaches a value of 4, with a slight difference compared to the LAI of maize under irrigation conditions (Figure 12). This discrepancy can be attributed to the relatively low rainfall between 15 and 25 days after sowing. However, the crop faced a severe drought from 42 days after sowing until maturity. As a result, the yield at Ejura was significantly lower at 1.47 mt ha^{-1} (Figure 9), whereas LINTUL-1 predicts a potential yield of 8.27 mt ha^{-1} for this location (Boullouz et al., 2022). This highlights the critical role of water availability during the grain filling period compared to the vegetative period. Despite the lower LAI observed in the Ashanti locations compared to the one at Ejura, which can be attributed to higher rainfall during the vegetative phase in Ejura, the yield in the Ashanti regions surpasses that of Ejura.

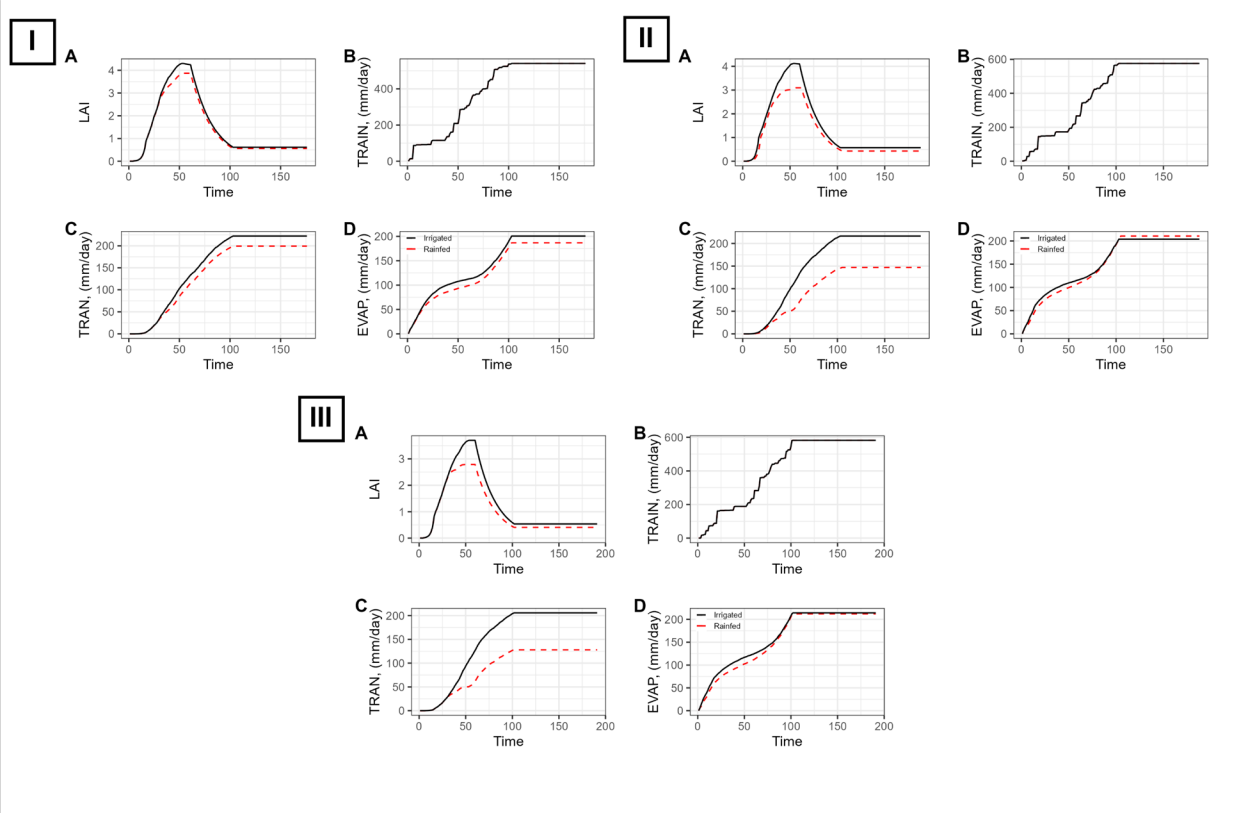


The TRAIN curve for irrigation and rainfed conditions overlap. LAI, Leaf Area Index; TRAN, Transpiration; TRAIN, Total Rain; and EVAP, Evaporation.

Figure 12. LAI (A), TRAN (B), TRAIN (C) and EVAP (D) simulation at Ejura by LINTUL-2

4.1.2.4 Gbalahi, Kpalga, and Nyankpala

Gbalahi, Kpalga, and Nyankpala are three locations situated in the northern part of the country (Figure 2). The LAI values for these trials are 4, 3, and 2.7, respectively (Figure 13). Despite experiencing a short drought period during the vegetative stage, ranging from the 10th to the 40th day after sowing at Gbalahi, the 16th to the 50th day after sowing at Kpalga, and the 21st to the 50th day after sowing at Nyankpala, these locations receive ample rainfall throughout the agricultural season, with precipitation exceeding 500 mm (Table 3). This favorable rainfall pattern contributes to the development of maize crops at these locations. The temporary drought during the vegetative stage leads to a reduction in the crop rain fed LAI, preventing it from reaching the levels observed in maize under irrigation conditions. Additionally, the crop experiences lower transpiration rates at Nyankpala and Kpalga due to the lower water availability during this period. However, rainfall resumes after this brief dry spell and continues until the maturity of the maize crops at all three locations. As a result, at Gbalahi, Kpalga, and Nyankpala, the rainfed yields reached 7.43, 6.04, and 5.03 mt ha^{-1} respectively (Figure 9). It is worth noting that although the crop rain fed LAI at Nyankpala is lower compared to Ejura, Ashanti Ayeduase, and Ashanti Anwomaso, LINTUL-2 predicts a higher yield for this location. This highlights the maize's resilience to drought during the vegetative period compared to the grain filling period (McMillen et al., 2022). McMillen et al. (2022) also highlight the effect of the variety on maize's ability to withstand drought. The presence of sufficient rainfall in these northern locations ensures an adequate soil moisture level, facilitating evaporation, particularly at Nyankpala and Kpalga, where crop transpiration rates are comparatively lower.

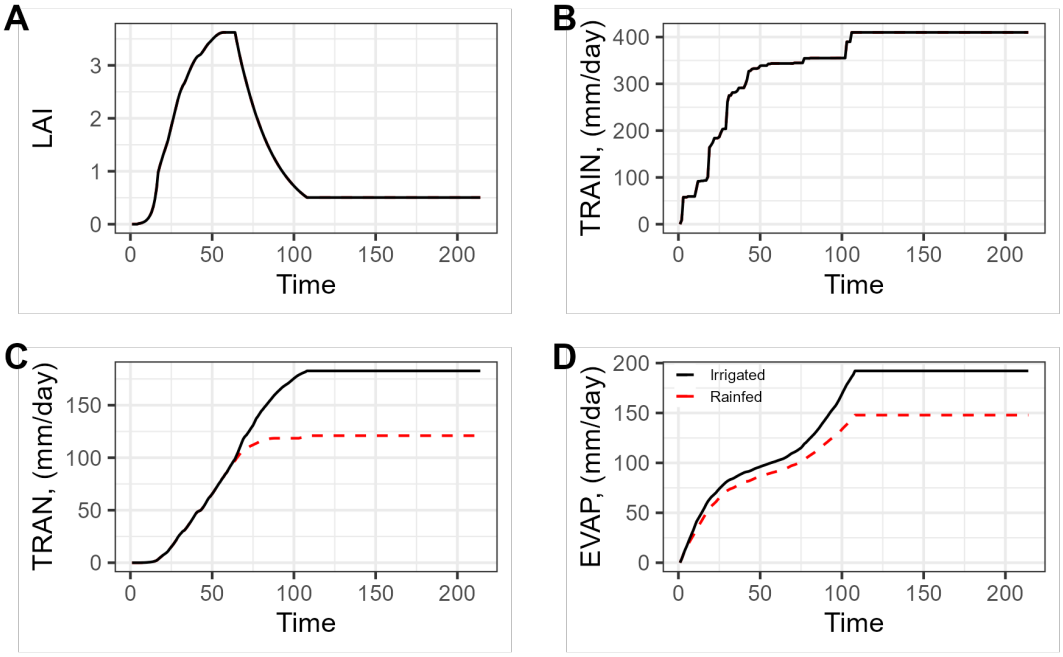


The TRAIN curve for irrigation and rainfed conditions overlap. LAI, Leaf Area Index; TRAN, Transpiration; TRAIN, Total Rain; and EVAP, Evaporation.

Figure 13. LAI (A), TRAN (B), TRAIN (C), and EVAP (D) simulation results by LINTUL-2 at (I) Gbalahi, (II) Kpalga, and (III) Nyankpala

4.1.2.5 KNUST

At the KNUST location, the water-limited yield is estimated to be 3.2 mt ha^{-1} (Figure 9). The crop rainfed LAI reaches 3.6, which is similar to the LAI observed for maize under irrigation conditions, as shown in Figure 14. However, 62 days after sowing, the transpiration of the crop under water-limited conditions significantly decreases due to the limited rainfall received from the 56th day after sowing to the 100th day after sowing. This prolonged period of reduced rainfall adversely affects both soil evaporation and crop yield. As a result, the reduced soil moisture availability during this critical period leads to a decrease in crop transpiration, causing a decline in the water-limited yield. The observed water limited yield of 3.2 mt ha^{-1} is lower compared to the potential value of 6.76 mt ha^{-1} reported by Boullouz et al. (2022) for the same site.



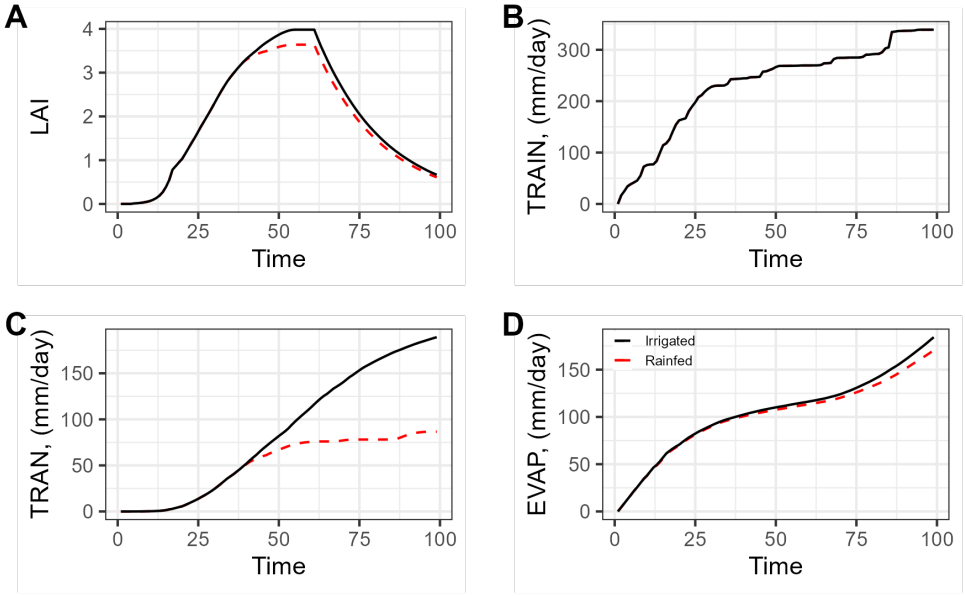
The LAI and TRAIN curves for irrigation and rainfed conditions overlap. LAI, Leaf Area Index; TRAN, Transpiration; TRAIN, Total Rain; and EVAP, Evaporation.

Figure 14. LAI (A), TRAN (B), TRAIN (C), and EVAP (D) simulation results by LINTUL-2 at KNUST

4.1.2.6 Mampong

During the vegetative stage, Mampong experiences favorable rainfall distribution, with a LAI value of 3.6, which is very close to the rainfall received by maize under irrigation conditions (Figure 15). However, 35 days after sowing, the daily rainfall distribution significantly decreases, leading to a collapse in crop transpiration and resulting in a lower yield of 1.40 mt ha^{-1} (Figure 9).

This observation in Mampong suggests that maize is particularly sensitive to water stress during the grain filling period (Barutçular et al., 2016). Unlike Nyankpala, where the crop showed a higher tolerance to drought during the vegetative stage, the maize in Mampong demonstrates a greater vulnerability to water stress during the critical grain filling period. This emphasizes the importance of sufficient water availability during this stage to ensure optimal crop performance and maximize yield potential (McMillen et al., 2022).

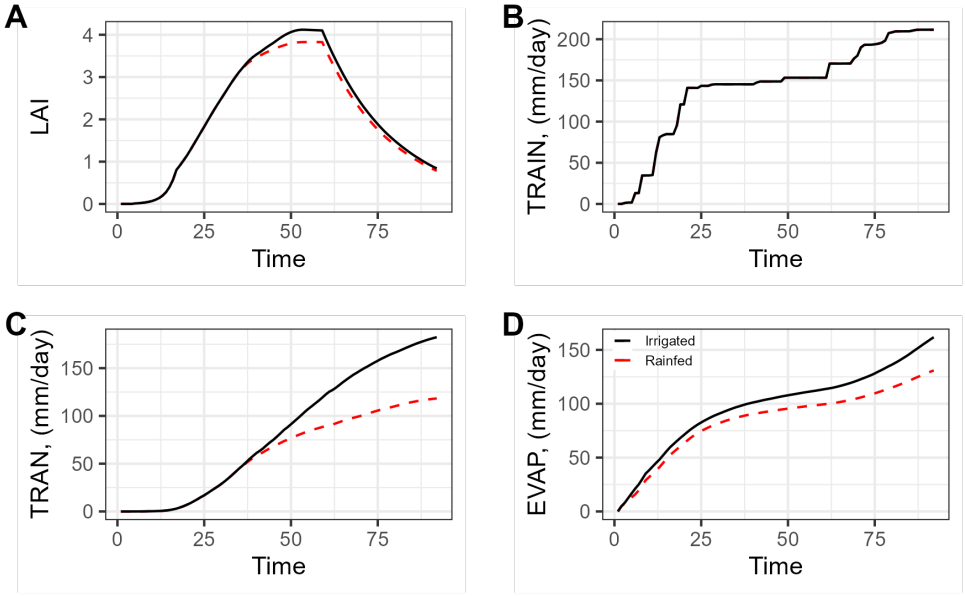


The TRAIN curve for irrigation and rainfed conditions overlap. LAI, Leaf Area Index; TRAN, Transpiration; TRAIN, Total Rain; and EVAP, Evaporation.

Figure 15. LAI (A), TRAN (B), TRAIN (C), and EVAP (D) simulation results by LINTUL-2 at Mampong

4.1.2.7 Sunyani

Figure 16 illustrates the simulation results at Sunyani. During the initial 25 days after sowing, Sunyani receives a relatively good rainfall distribution. However, the subsequent period experiences a significant scarcity of rain until it resumes 63 days after sowing. It is important to note that this trial has the lowest cumulative rainfall value among all the trials conducted. As a consequence, there is a reduction in evapotranspiration at this location, leading to a decrease in the water-limited yield, which is predicted to be 3.12 mt ha^{-1} (Figure 9).

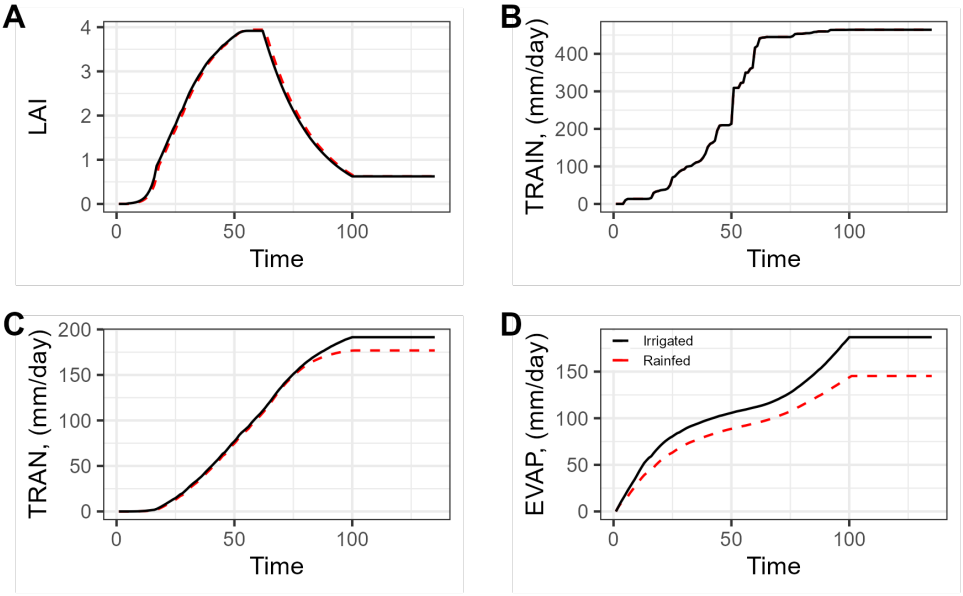


The TRAIN curve for irrigation and rainfed conditions overlap. LAI, Leaf Area Index; TRAN, Transpiration; TRAIN, Total Rain; and EVAP, Evaporation.

Figure 16. LAI (A), TRAN (B), TRAIN (C), and EVAP (D) simulation results by LINTUL-2 at Sunyani

4.1.2.8 Wenchi

At Wenchi, the water-limited yield reaches 6.73 mt ha^{-1}, the second highest yield among our trials (Figure 17). Except a little drought from the 5th to 14th days after sowing at this location, the rainfall distribution is quite good either during the vegetative stage or the grain filling stage. This results in an LAI and transpiration rate of maize quite similar to the one of maize in irrigation conditions. For the evaporation of soil, it collapses compared to the one in irrigation conditions. This could be explained by the fact that a part of the rainfall is lost either by leaf interception or runoff/drainage.



The TRAIN curve for irrigation and rainfed conditions overlap. LAI, Leaf Area Index; TRAN, Transpiration; TRAIN, Total Rain; and EVAP, Evaporation.

Figure 17. LAI (A), TRAN (B), TRAIN (C), and EVAP (D) simulation results by LINTUL-2

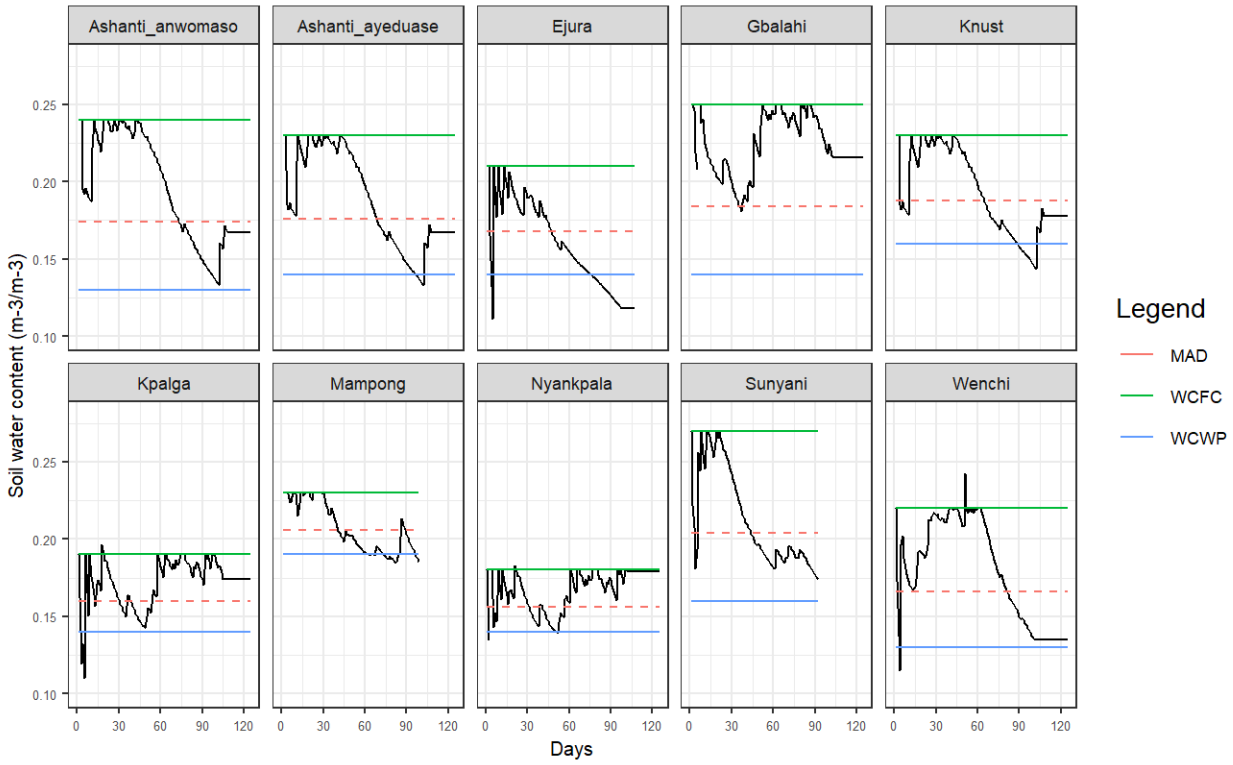
4.1.3 Water Content

Figure 18 shows the soil water content during the simulation period across trial locations. It is noteworthy that the soil water-holding capacity is higher at Gbalahi, Ashanti Anwomasso, Sunyani, Wenchi, and Ashanti Ayeduase compared to other locations. At Gbalahi, where the highest yield of 7.43 mt ha^{-1} was recorded, the soil water content remains mostly above the management allowable depletion (MAD), representing the critical threshold for water stress and growth reduction in crops (Datta et al., 2017). However, the MADD temporarily declines below the MAD approximately 36 days after sowing. This indicates that, even during the short period of drought observed from the 10^{th} to 40^{th} days after sowing, the soil retained sufficient water, to prevent water stress in maize at this location. On the other hand, at Ashanti Anwomaso, Ashanti Ayeduase, Sunyani, and Wenchi, the soil water content falls below the MAD after 72, 65, 40, and 80 days after sowing, respectively, resulting in water stress for the maize crop. However, except for Ashanti Ayeduase, the soil water content does not reach the water content at wilting point (WCWP). This discrepancy could explain the difference between the water-limited yield found and the potential yields reported by Boullouz et al. (2022) at those locations. Datta et al. (2017) reveal that soil water content below this threshold could induce a growth reduction of the crop and the death of the crop.

At Ejura and KNUST, where the water-limited yields are lower than those observed in the field, the soil water content falls below the MAD at 45 and 65 days after sowing, respectively. From that moment, maize experiences water stress at these locations and then becomes more pronounced at Ejura, where the soil water content falls significantly below the wilting point. Similar observations are made at KNUST, where the soil water content falls below the WCWP before rising again above the WCWP and the MAD at 100 days after sowing. These findings elucidate the lower yields at

these locations, which are even below the observed field values. Mampong, where the lowest yield of 1.40 mt ha^{-1} was recorded, exhibits one of the lowest soil water-holding capacities (5%). Furthermore, the soil water content falls below the MAD, leading to water stress in maize from 40 days after sowing. It even approaches or falls below the WCWP between 58 and 82 days after sowing.

Overall, the variations in soil water content across the trial locations have a significant impact on maize growth and yield, highlighting the importance of adequate soil water availability for optimal crop performance.



MAD, Management Allowable Depletion; WCFC, Water Content at Field Capacity; WCWP, Water content at Wilting Point

Figure 18. Daily soil water content variation across the study sites

4.1.4 Photosynthetically Active Radiation by Study Site

Figure 19 illustrates the distribution of PAR across the study sites. The highest levels of IPAR were observed in Kpalga, Gbalahi, and Wenchi, with values of 981.42, 974.37, and 965.38 Mj m^{-2}, respectively. Surprisingly, even in locations like Mampong and Ejura, where the simulated water-limited yield is relatively low, the IPAR values exceeded 930 Mj m^{-2}. These values are higher than those observed in the Ashanti locations of KNUST and Sunyani, where water-limited yields are higher than those of Ejura and Mampong. These findings emphasize the significant influence of soil parameters and rainfall on the yield predictions made by the LINTUL-2 model.

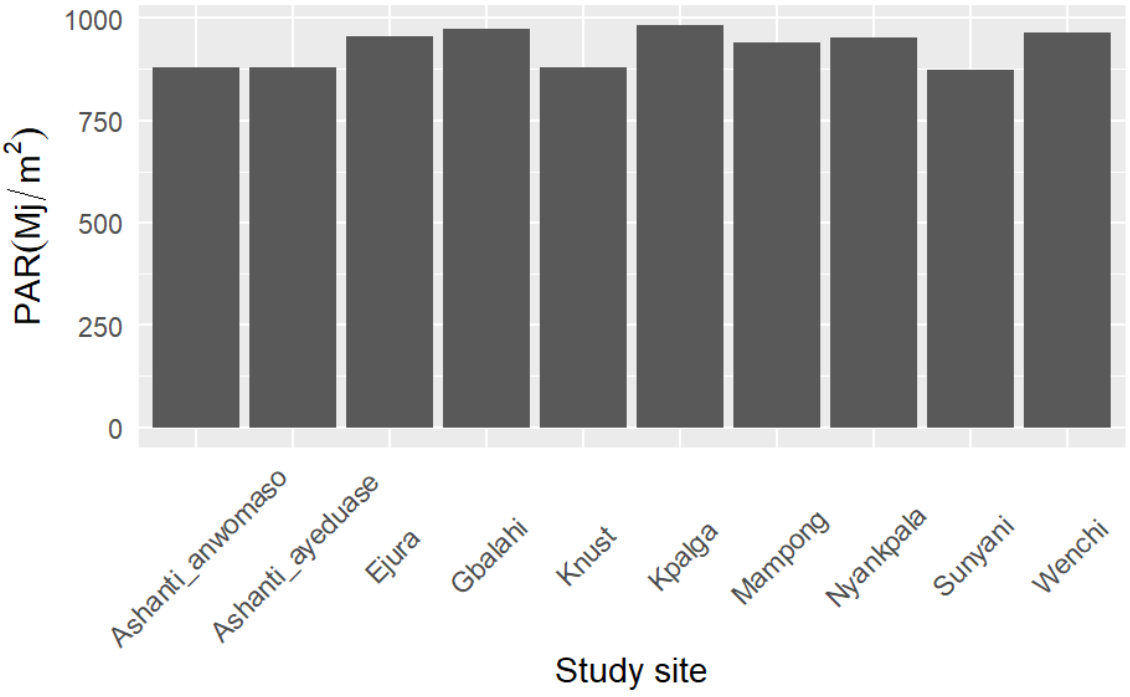


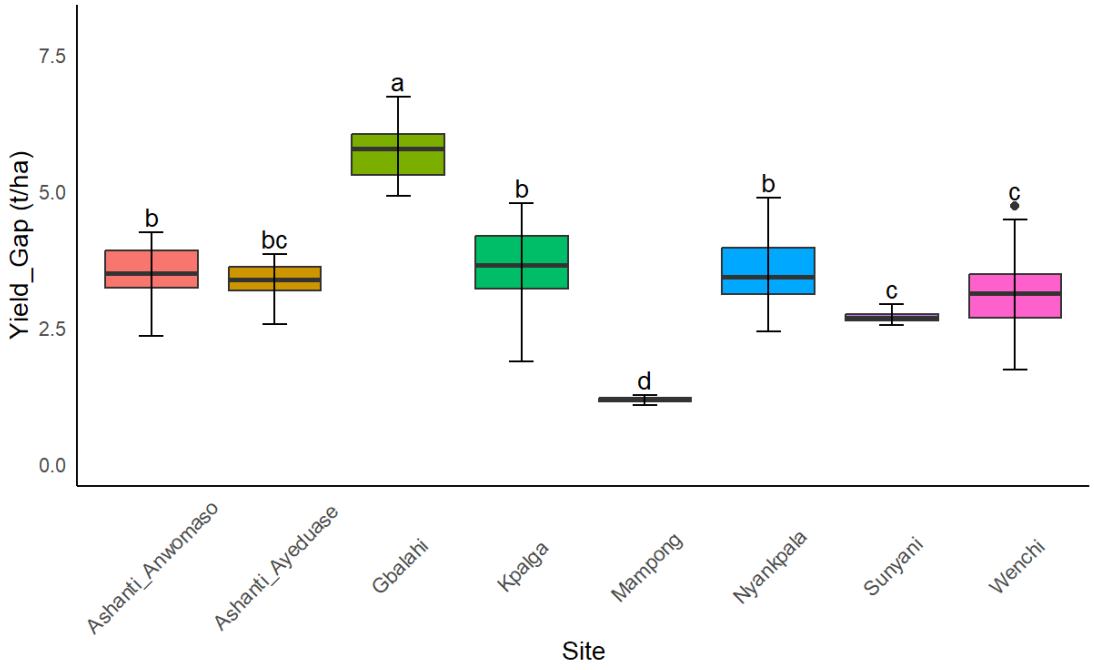
Figure 19. Photosynthetically Active Radiation by study site

4.1.5 Quantifying the Yield Gap of Maize in FERARI Locations

The locations of Ejura and KNUST were excluded from the yield gap analysis due to the lower water-limited yields of maize compared to the observed field values, which leads to a negative yield gap. We believe there may be a problem with the rainfall distribution or soil parameters of these trials. Figure 20 illustrates the variation in yield gap across the study sites. The water-limited yield gap is quantified as a percentage of the potential water-limited yield, ranging from 18.3% to 74.3%.

Based on the results of the ANOVA and Tukey post hoc test at a 5% significance level, significant differences were observed in the mean values of the yield gap across the study sites. Consequently, the locations were categorized into four distinct groups based on their yield gap values. Gbalahi is identified as the location with the highest yield gap values, which is understandable considering its highest water-limited yield among all the locations. Conversely, Mampong is categorized as

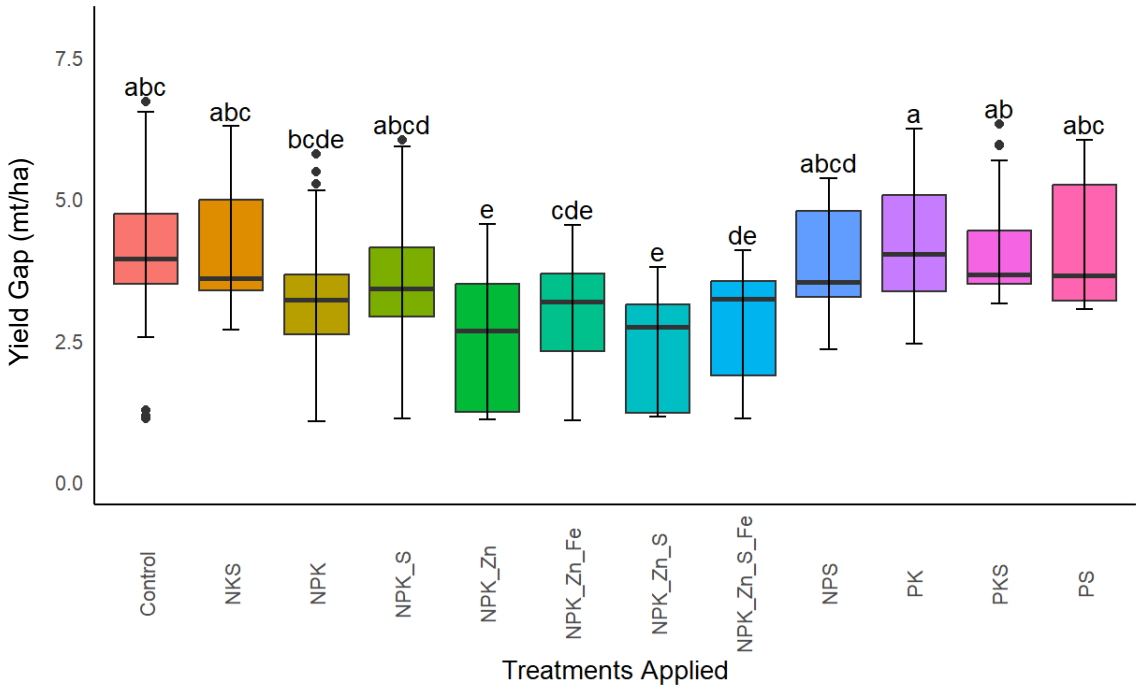
the location with the lowest yield gap values, indicating that the water-limited yield at this location is relatively close to the observed field yield.



The sites grouped together are labeled with the same letter(s).

Figure 20. Yield gap variation per study site

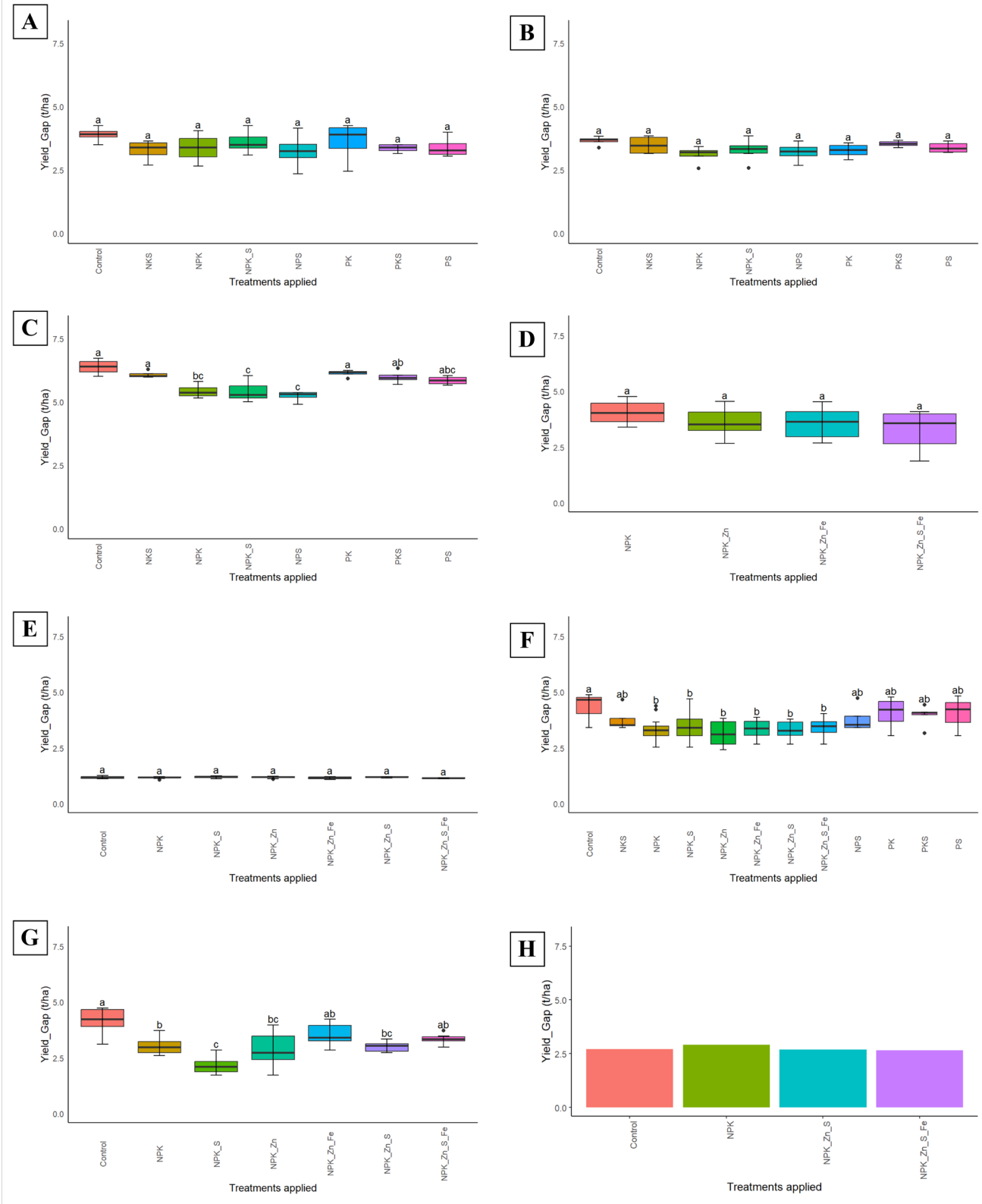
Figure 21 displays the variation in the yield gap based on different fertilizer treatments. From this figure, we observe that the results of the ANOVA and Tukey post hoc test indicate a significant difference in the mean yield gap depending on the fertilizer treatment applied, with a threshold of 5%. Specifically, fertilizer treatments involving NPK+Zn and NPK+Zn+S exhibit the lowest maize yield gap, showing the importance of micronutrients and secondary nutrients (Kugbe et al., 2019). However, treatments involving the secondary nutrient sulfur and certain macronutrients (NPKS, NPS, PKS, NKS, and PS) result in the highest yield gap. Adding a mixture of secondary nutrients or micronutrients does not always result in a lower yield gap than that obtained with a single secondary nutrient or micronutrient. For example, the use of NPK+Zn+Fe, NPK+Zn+S, or NPK+Zn+S+Fe treatments does not result in a statistically different yield gap to that of the treatment involving only zinc as micronutrient and macronutrient. Nevertheless, given that both location and fertilizer treatment have some impact on the maize yield gap, it would be interesting to see the effect when both factors are considered. This could lead to better conclusions.



The treatments grouped together are labeled with the same letter(s).

Figure 21. Yield gap variation per treatment applied

Figure 22 depicts the variation in yield gap by study site and fertilizer treatment. The results of the ANOVA and Tukey post hoc test, conducted at a significance level of 5%, indicate that there is no significant difference in yield gap among the different fertilizer application treatments at Ashanti Anwomaso, Ashanti Ayeduase, Kpalga, and Mampong. In most trials, the control treatment group exhibits the highest yield gap value, while the treatments with NPK+S or NPK+Zn exhibit the lowest yield gap across the various locations. However, at Gbalahi, where the highest water-limited yield is observed, the Anova and Tukey post hoc test reveal a significant difference in the mean yield gap values among the applied treatments. The NPK, NPS, and NPK+S treatments perform well at this location, resulting in the group with the lowest yield gap. Overall, it is evident that the treatment involving the sulfur nutrient consistently leads to the group with the lowest yield gap across the different locations.



The treatments grouped together are labeled with the same letter(s).

Figure 22. Yield gap variation by study site and treatment applied: (A) Ashanti Anwomaso, (B) Ashanti Ayeduase, (C) Gbalahi, (D) Kpalga, (E) Mampong, (F) Nyankpala, (G) Wenchi, and (H) Sunyani

4.1.6 Explaining the Maize Yield Gap

To better understand the impact of fertilizer application on closing the yield gap, additional covariates were added to the 11 retained after the recursive feature selection (Table 2), such as the amount of zinc applied (Zn_applied), amount of iron applied (Fe_applied), amount of potassium applied (Potasum_applied), and amount of sulfur applied (S_applied). This micronutrients have been identified as important factors for maize yield gap closure in Ghana (Kugbe et al., 2019).

Figure 23 presents the correlation matrix of the relevant drivers that were retained. Correlation between covariates in a regression analysis can introduce instability in the model (Abdelgadir and Eledum, 2016). The correlation matrix indicates that numerous variables are significantly correlated to each other at a significance level of 5%. For instance, the soil carbon content and soil organic matter are highly positively correlated, with a Pearson correlation coefficient of 1. Similarly, the base saturation and EA_meq_100g are strongly negatively correlated, with a Pearson correlation coefficient of -0.88. To avoid multicollinearity in the model, one variable from each correlated pair with a high Pearson correlation coefficient should be removed. To determine which variables to remove, the variable importance of the best model obtained after hyperparameter tuning was examined.

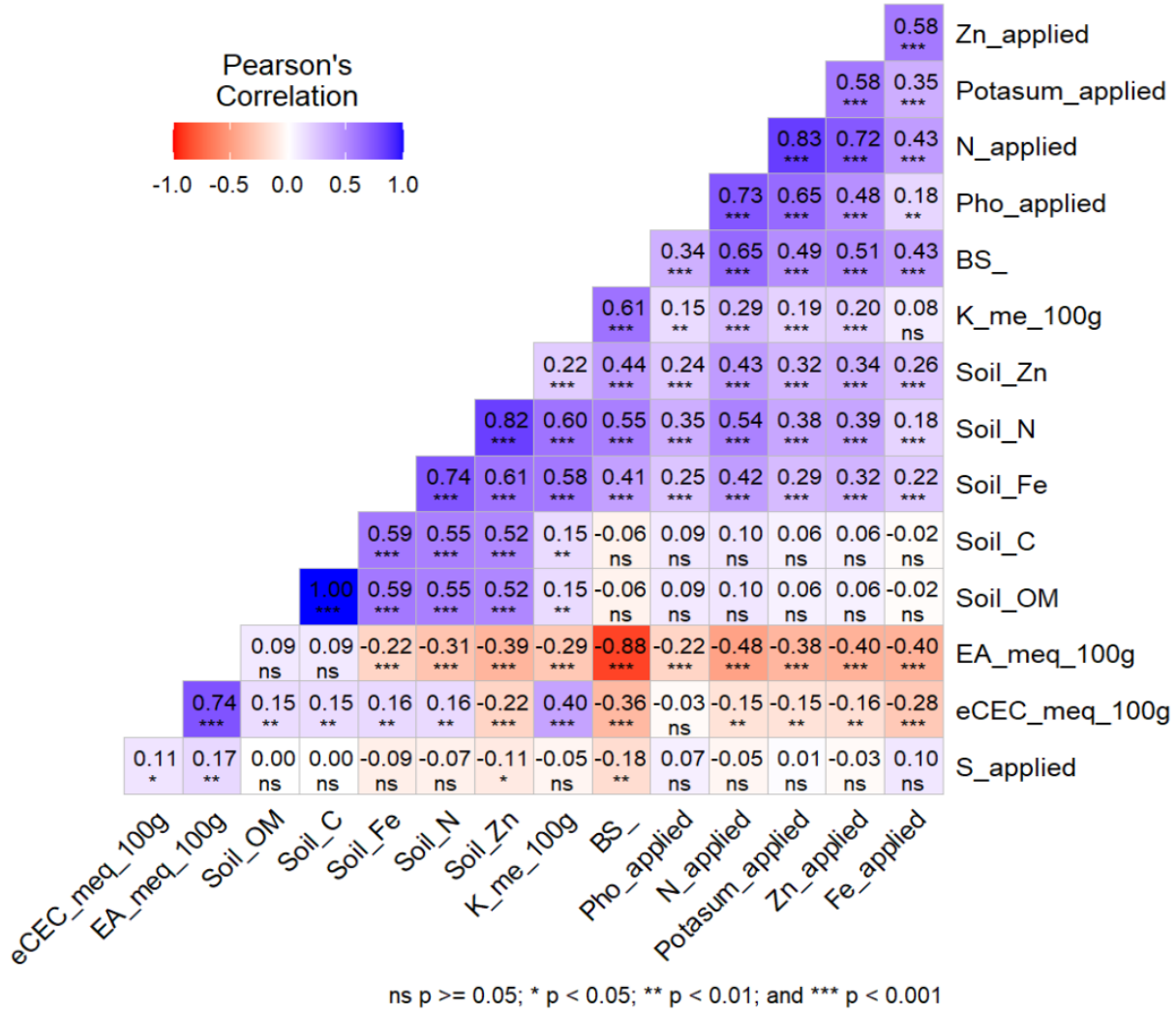


Figure 23. Correlation matrix of the relevant drivers retained

Figure 24 displays the feature importance of the model based on the optimum parameters and selected variables. From the feature importance plot, it can be observed that the most important drivers of the yield gap in the trials are the soil organic matter, soil carbon content, base saturation, and soil nitrogen content, explaining 13.81%, 13.80%, 11.56% and 10.25%, respectively, of the variability in maize yield gap under water-limited conditions.

On the other hand, the amount of iron applied appears to be less relevant in explaining the yield gap in the trials. Therefore, it can be excluded from the analysis. Additionally, since the EA_meq_100g and Soil_C variables are relatively less important compared to the BS and Soil_OM variables, they will also be removed from the analysis. The correlation matrix indicates that the information conveyed by EA_meq_100g and Soil_C is already captured by BS and Soil_OM, as they exhibit a high and significant positive correlation.

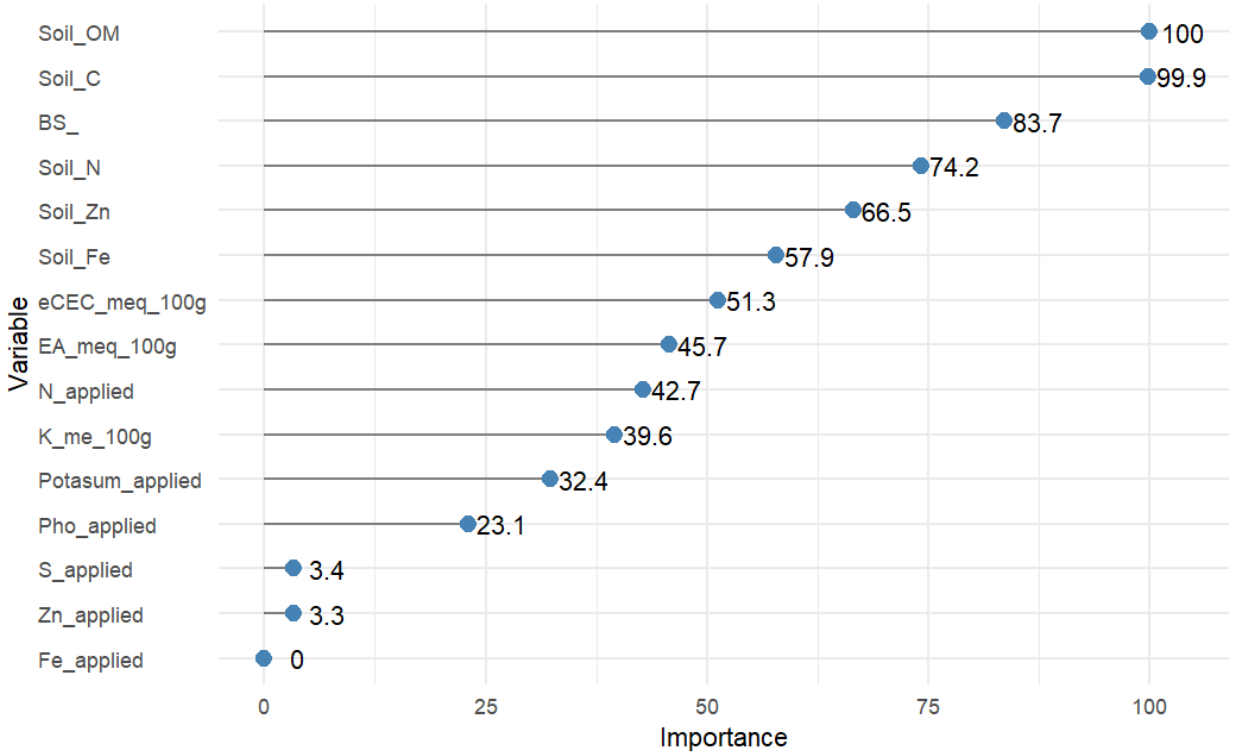
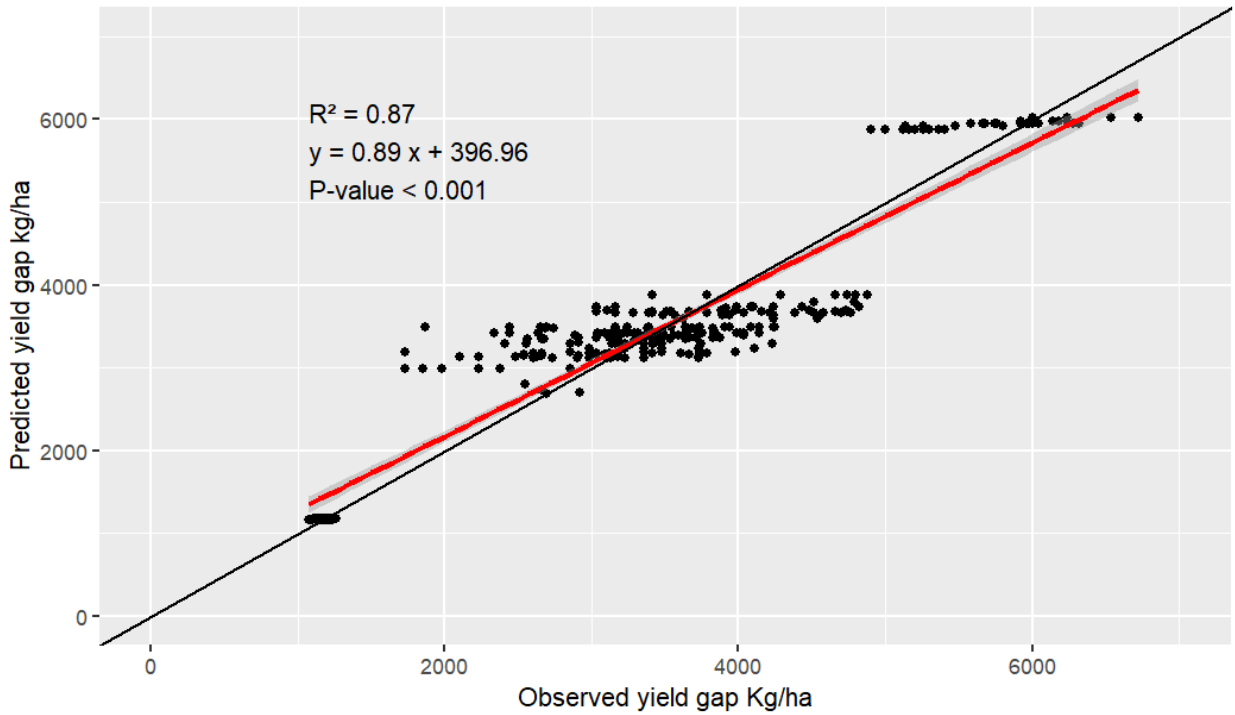


Figure 24. Importance of the covariates in explaining yield gap

The random forest regression model, using the retained covariates, demonstrates a strong performance in explaining 87% of the yield gap variability in the trials, with a RMSE of 472.63 kg ha^{-1}. This indicates that the model is highly effective in predicting the yield gap in the trials.

Figure 25 illustrates the relationship between the yield gap predicted by the model and the observed yield gap. The regression line closely aligns with the one-to-one line, indicating the model's accuracy in predicting the yield gap. However, as the random forest is considered a black box machine learning model, its interpretability is limited. In an attempt to gain a better understanding of the influence of each driver on the predicted yield gap, an MLR was initially performed. However, the validity assumptions of the MLR model, including normality, homoscedasticity, and absence of autocorrelation in errors, were not met.

Due to these limitations, a ridge regression was conducted to determine the direction of the impact of each independent variable on the predicted yield gap. While this approach does not provide detailed insights into the magnitude or significance of the effects, it helps identify the general trend and direction of the influence of each variable on the predicted yield gap.



The red line corresponds to the regression line and the black line to the one-to-one line.

Figure 25. Random forest accuracy in explaining the yield gap using soil and fertilizer variables

Table 6 shows the results of the ridge regression of the predicted yield gap, obtained from the random forest model, on the covariates. The resultant model was able to explain 97.57% of the variability of the predicted yield gap by the RF, with a RMSE of 295.49 kg ha^{-1}. Based on the results, we observe that, at a significance level of 5%, an increase in nitrogen application, soil nitrogen content, soil organic matter, phosphorus application, base saturation, potassium application, and sulfur application lead to a reduction in the yield gap (Table 6). This suggests that higher levels of these factors are associated with improved yield performance and closing of the yield gap. On the other hand, an increase in soil potassium content, effective cation exchange capacity, soil zinc content, soil iron content, and zinc application are associated with an increase in the yield gap. This implies that higher levels of these factors may contribute to a wider gap between potential and actual yields.

Table 6. Results of the ridge regression of the predicted yield by the random forest on the covariates

Covariate	Estimate	Std. Error	T value	p-value
Intercept	0.00	0.01	0.01	>0.05
N_applied	-0.16	0.02	-8.29	<0.05
Soil_N	-1.05	0.03	-41.83	<0.05
Soil_K_me_100g	0.60	0.02	32.15	<0.05
eCEC_meq_100g	0.26	0.02	17.40	<0.05
Soil_OM	-0.40	0.01	-37.19	<0.05
Soil_Zn	0.53	0.02	30.67	<0.05
Soil_Fe	0.03	0.01	2.55	<0.05
Pho_applied	-0.06	0.01	-5.92	<0.05
BS	-0.14	0.02	-5.56	<0.05
Potasum_applied	0.00	0.01	0.10	>0.05
Zn_applied	0.01	0.01	1.19	>0.05
S_applied	-0.02	0.01	-2.60	<0.05

4.1.7 Spatial Analysis of Maize Water-Limited Potential Yield and Yield Gap in Ghana

Figure 26 provides an insightful overview of the spatial distribution of maize water-limited potential yield and yield gap across Ghana. Note that larger areas without data points are due to nature reserves or waterbodies. We conducted our analysis on 2,236 randomly selected points, revealing a considerable variation in water-limited potential yield, ranging from 0.6 to 8.7 mt ha^{-1}. Consequently, the yield gap exhibited a range from 0.2 to 7.8 mt ha^{-1}.

The spatial representation of water-limited potential yield and yield gap adheres to Tobler's First Law of Geography, affirming that everything is interrelated, with nearby locations displaying higher correlation than distant ones (Tobler, 1970). Notably, the plot showcases a few instances where points with significant disparities in potential yield or yield gap overlap. This observation reinforces the suitability of our points for interpolation purposes.

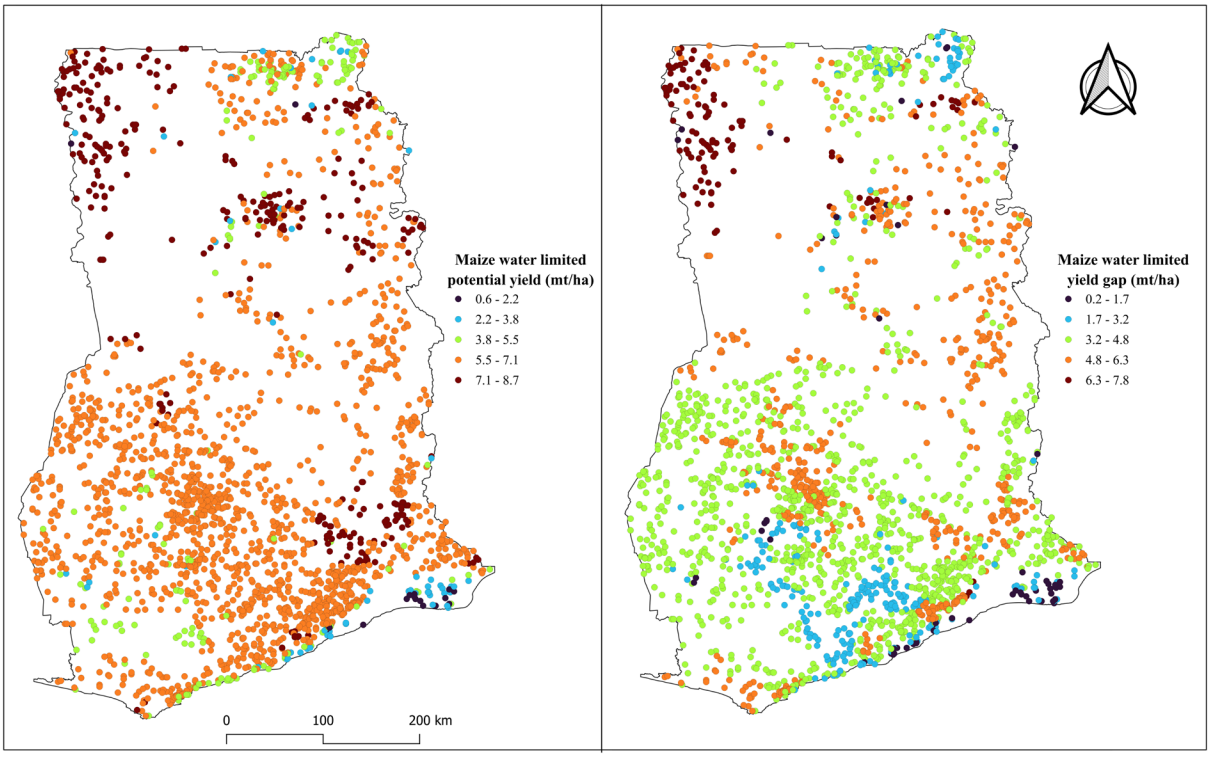


Figure 26. Overview of the spatial distribution and variability of the maize water-limited potential yield and yield gap in Ghana

Figure 27 illustrates the semivariograms resulting from the ordinary kriging interpolation for maize water-limited potential yield and yield gap. Analyzing this plot, we observed that the interpolation for maize water limited potential yield exhibits a range of 75,000 meters (m), while for the yield gap, the range is 55,000 meters. These ranges imply a spatial autocorrelation of points within distances of less than 75 km for potential yield and 55 km for yield gap interpolation. Beyond these distances, two points are no longer correlated.

The nugget and sill values of our semivariograms are 0.2 and 0.72 for potential yield, and 0.25 and 0.94 for yield gap, respectively. These values suggest that there are not substantial jumps between neighboring points due to their relatively small magnitude. However, it is worth noting that our semi variogram displays another positive slope after a distance of 300,000 m, indicating that two points with a distance greater than 300 km also exhibit spatial correlation.

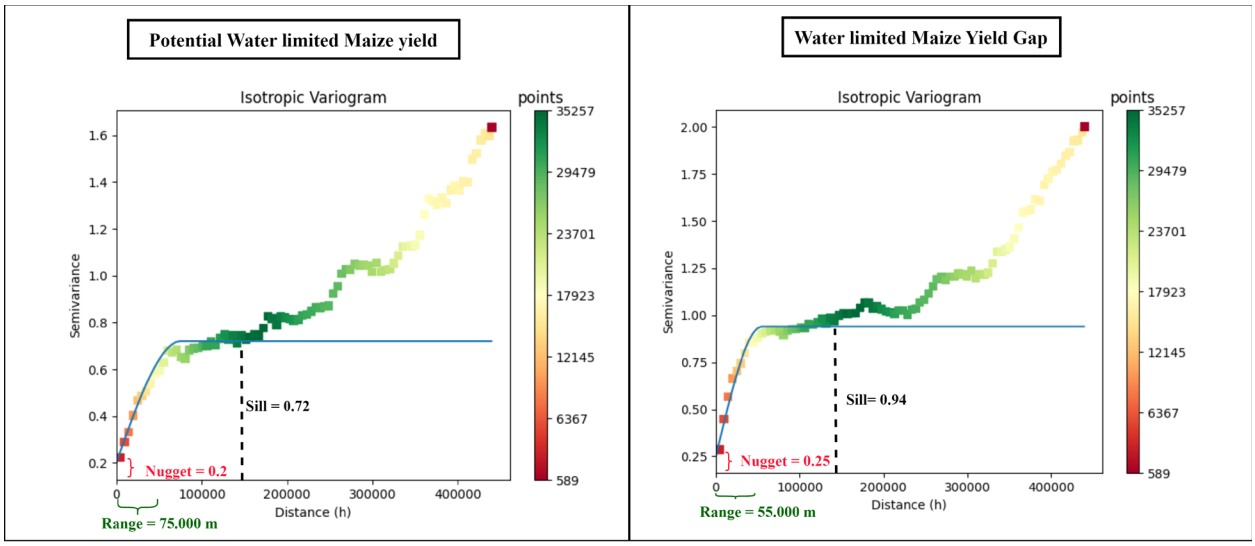


Figure 27. Semivariograms of the potential yield and yield gap

Figure 28 displays the map generated through ordinary kriging, showcasing the distribution of water-limited yield and yield gap for maize. The regions of Upper West, West Savannah, and the eastern part of the Eastern region exhibit the highest water-limited potential yield, surpassing 6 mt/ha. Conversely, the southern parts of Volta, Greater Accra, and Upper East regions demonstrate lower water-limited potential yield, ranging between 1.7 and 3 mt ha^{-1}. These variations can be attributed to the significant rainfall amounts, exceeding 357 mm, recorded at the selected random points in these areas.

The yield gap map, varying between 0.64 and 7.58 mt ha^{-1}, reveals a correlation between higher water-limited yield areas and larger yield gaps. This suggests that despite the higher potential in these locations, on-farm performance falls short, resulting in significant yield gaps. Additionally, a notable observation is that the yield gap in the southern part of Ghana is generally lower than that in the northern regions.

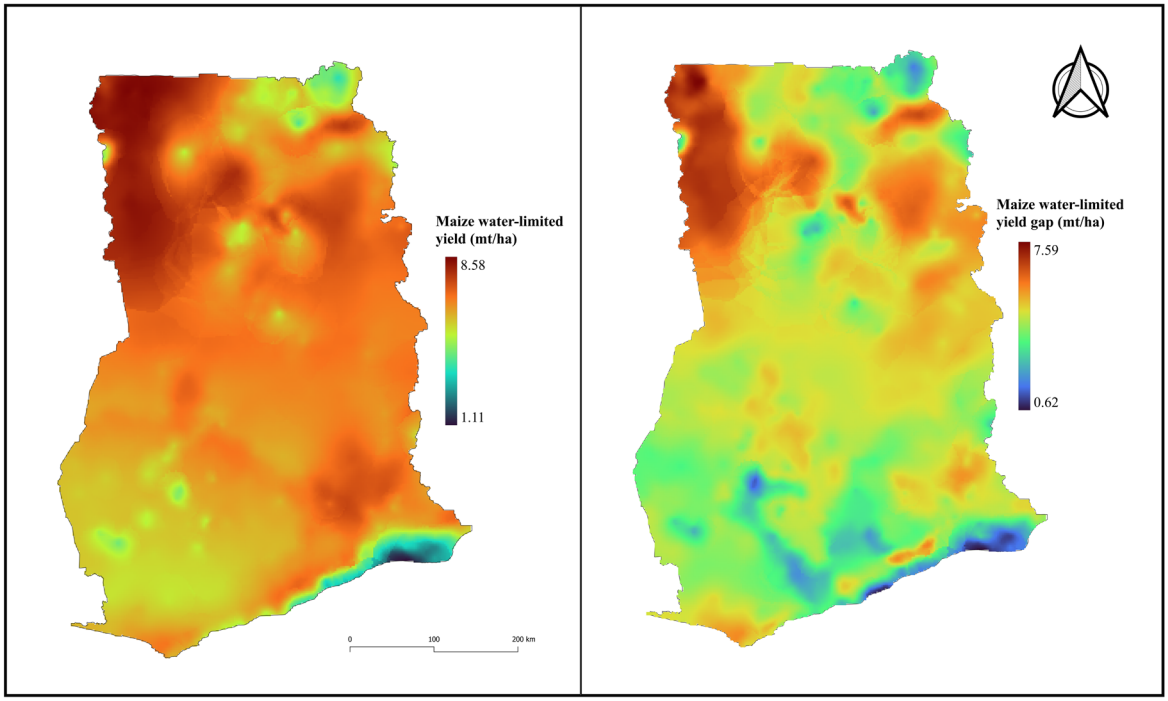


Figure 28. Map of water-limited yield and yield gap of maize based on ordinary kriging at a resolution of 2 km

4.2 Discussion of Results

4.2.1 Potential Yield of Maize in Ghana

The water-limited yield of maize in our study trials ranged from 1.4 to 7.43 mt ha^{-1}, with the highest yield observed in Gbalahi, located in Tamale metropolitan district, in the northern region of Ghana. This finding aligns with the study by MacCarthy et al. (2018),who reported that the mean water-limited yield in Tamale, a district within the northern region, varied between 3,262 kg ha^{-1} and 4362 kg ha^{-1} during different planting windows. Our results also indicated a positive correlation between water-limited yield and IPAR, supporting the findings of Gurjar et al. (2017), who emphasized the role of radiation in crop growth and photosynthesis.

However, we observed that the simulated water-limited yield by the LINTUL-2 model is more influenced by water availability rather than intercepted PAR. This is expected, considering the crucial role of water availability during the grain filling stage of maize. Khoshvaghti et al. (2014) further confirmed this relationship in their study, demonstrating that water stress during the grain filling stage can negatively impact grain length, cob diameter, grain volume, and ultimately grain yield in maize. Barutçular et al. (2016) and (McMillen et al., 2022) also emphasized on the same point in their studies.

The maize water-limited yield map displays significant differences compared to LINTUL-1 results in certain regions, while showing similarities in others. In the northern part of the country, which demonstrated a higher potential yield of around 8 mt ha^{-1} with LINTUL-1, there is now more variability in water-limited yield across the regions. Specifically, the Upper West and Western

Savannah regions maintain a high potential yield of 8 mt ha^{-1} with LINTUL-2. However, in the Upper East, North East, and Eastern Northern regions, there is a decrease in potential yield, averaging around 6 mt ha^{-1} and dropping to 3 mt ha^{-1} in the eastern part of Upper East. Regions such as Western, Central, Western North, Ahafo, Ashanti, and South Brong Ahafo, where LINTUL-1 predicted a potential yield between 5.4 and 6 mt ha^{-1}, exhibit a water-limited yield of 4-5 mt ha^{-1} with LINTUL-2. In the southern part of Volta and the eastern part of Greater Accra, where the water-limited yield is around 1.5 mt ha^{-1}, the difference was significant, with a 5 mt ha^{-1} lower yield than that predicted by LINTUL-1. For the remaining parts of the country, both LINTUL-1 and LINTUL-2 predict yields within the same range of 6.5 mt ha^{-1}.

The slight variation between LINTUL-1 and LINTUL-2 results across the country could be attributed to the high rainfall amount in the selected random points, exceeding 357 mm. As the only difference between the two crop models is the integration of soil water balance in LINTUL-2, no significant difference in the results would be observed if the rainfall distribution is good. However, other factors may contribute to the variation, such as the suitability of the calibrated values to the entire country. For instance, crop partitioning coefficients, Tsum to maturity, and anthesis values calibrated at Ejura in the Transitional zone may vary in other AEZs. Additionally, Yamba et al. (2023) highlighted the variability of rainfall and temperature across different AEZs, with the northern part of the country experiencing higher temperatures. Moreover, the quality and accuracy of the weather data used for water-limited yield prediction could also play a significant role. Despite the improvement in ERA-interim weather data quality, transitioning to ERA5 after reanalysis, the scarcity of weather stations in Africa poses challenges in obtaining high-resolution weather data (Gleixner et al., 2020; Harris et al., 2014). The situation is further exacerbated in rural areas, which are the primary agricultural areas, where very few weather stations are available (Gleixner et al., 2020).

4.3 Yield Gap of Maize in Ghana

The water-limited yield gap, expressed as a percentage of the potential water-limited yield, ranged from 18.3% to 74.3% in our study. This aligns with previous findings in Navrango, Upper East region of northern Ghana, by MacCarthy et al. (2018), who reported water-limited exploitable yield gaps between 22% and 35% and potential yield gaps ranging from 59% to 75%. Another study by van Loon et al. (2019) conducted in Nkoranza and Savelugu municipalities of Ghana found maize water-limited yield gaps ranging from 67% to 84%. Our results are consistent with these previous studies. However, it is worth noting that the range of yield gap variability in our study is three times greater than that reported in the other studies. This discrepancy can be attributed to the different assumptions and underlying crop models used in each study (Salo et al., 2016). For instance, van Loon et al. (2019) employed a hybrid-maize simulation model, while MacCarthy et al. (2018) used the DSSAT model.

In the study by Boullouz et al. (2022), the potential yield gap variation was estimated to range from 17% to 98.2% using the LINTUL-1 model. Notably, the integration of water balance in the LINTUL-2 model, as observed in our study, resulted in a reduced yield gap variation, narrowing it to a range of 81.2% to 55.96%. This emphasizes the importance of incorporating water balance considerations in maize yield prediction, particularly in rainfed production systems.

4.3.1.1 Soil Characteristics and Yield Gap

The soil organic matter, base saturation, and soil nitrogen, identified as important drivers of the yield gap, were found to have a negative correlation with the maize water-limited yield gap. An increase in these variables was associated with a decrease in the yield gap. Y. Liu et al. (2013) conducted a study that revealed an increase in soil organic matter enhances the soil's water-holding capacity, allowing it to retain more water for crop transpiration even during periods of drought. Massignam et al. (2009) found in their study that nitrogen is a crucial nutrient for improving maize yield, which explains the negative relationship observed between soil nitrogen content and the maize water-limited yield gap.

On the other hand, among the soil parameters, positive associations were observed between the yield gap and soil zinc, soil iron, soil exchangeable potassium, and effective cation exchange capacity. Numerous studies, including those by Manzeke et al. (2012), D.-Y. Liu et al. (2017), and Elsayed et al. (2022), have demonstrated the positive effect of zinc micronutrients on improving maize yield. However, the contradiction with our findings can be explained by the fact that 87.8% of the data points had soil zinc content below the critical value of 5 mg kg^{-1}, as identified by Adeoye and Agboola (1985) in Nigeria. According to that study, critical values were determined for different soil parameters that affect maize production. In our trials, we found that the maximum exchangeable potassium was 0.49 milliequivalents (mEq) of K per 100 grams (g) of soil, which is significantly lower than the critical K value of 0.6 mEq g^{-1}. This indicates that the potassium content in the soil of our study sites is below the critical threshold recommended for optimal maize production and could explain the positive relationship found between soil exchangeable potassium and the maize yield gap.

4.3.1.2 Fertilizer Application and Yield Gap

The results of the ridge regression analysis of the important variables showed that fertilizer application has a positive effect on maize yield and reduces the water-limited yield gap in our study trials. Specifically, nitrogen application, phosphate application, and sulfur application were found to decrease the maize yield gap and were identified as the most influential fertilizer drivers, as they contributed more to the explanation of the maize yield gap according to the random forest model. This finding is consistent with the study conducted by Boullouz et al. (2022), which also reported a positive effect of phosphate fertilizer application on reducing the yield gap by 8 kg ha^{-1}. Nafiu et al. (2011) found a positive relationship between NPK application and dry matter production and yield of eggplant in southwestern Nigeria. Specifically, an increase in NPK application up to 200 kg ha^{-1} had a positive effect on yield and dry matter production. Nitrogen, phosphorus, and potassium are the three major macronutrients that are essential for crop production (Zewide and Reta, 2021). Given the high soil degradation and decreasing soil fertility, the supply of these major nutrients through fertilization or land management is crucial for improving crop yield.

However, it should be noted that excessive application of macronutrients like nitrogen can have negative consequences. High nitrogen application can lead to increased nitrous oxide (N_2O) emissions, contributing to climate change (Leitner et al., 2020). Leitner et al. (2020) found that closing 50% of the yield gap would triple the baseline N_2O emissions, estimated at 0.24 kg N_2O-N ha^{-1}, while closing 75% of the yield gap would increase emissions by a factor of 5. This is particularly evident when nitrogen application exceeds 100 kg ha^{-1}. Therefore, nitrogen

should be applied carefully to avoid any complications that could harm crop yield and hinder the achievement of Sustainable Development Goal 2, which aims to create a world free of hunger by 2030.

The secondary nutrient sulfur was also found to have a negative correlation with the yield gap. The ANOVA and Tukey post hoc test, conducted at a significance level of 5%, indicated that there is a significant difference in yield gap among some locations, and treatments involving sulfur nutrient consistently showed the smallest yield gap. In three of the four locations with significant differences, the treatment involving sulfur stood out with the smallest yield gap; these were Wenchi, KNUST, and Gbalahi. Similar results were found by Agyin-Birikorang et al. (2022), who reported that the omission of sulfur decreased maize yield by 35% in the Savanna agroecological zones of Northern Ghana. Similarly, Kugbe et al. (2019) found that the application of sulfur, boron, and zinc in Tolon District in Northern Ghana increased maize yield by 30%.

Despite the positive effect of fertilization on maize yield, we observed that the micronutrient zinc tended to increase the maize water-limited yield gap in our study sites. However, its effect was not too significant in explaining the variability of the maize yield gap according to the random forest model, as indicated by the feature importance graph. Overall, we agree that fertilization has a positive effect on maize yield in Ghana, as treatments without fertilization consistently showed higher yield gaps. This finding aligns with the study by Thomas (2020), which revealed that fertilizers are a key component for improving crop yields in SSA.

CHAPTER 5: CONCLUSION AND RECOMMENDATIONS

This study focused on estimating water-limited maize yield using the LINTUL-2 model and conducting spatial analysis of the yield gap in Ghana. The results revealed that LINTUL-2 performed well in simulating water-limited yield in the majority of the locations, except for Ejura and KNUST, where poor rainfall distribution resulted in lower water-limited yield compared to field observations. Analysis of the maize yield gap in the other high-performing locations highlighted the negative correlation of nitrogen application, soil nitrogen content, soil organic matter, phosphorus application, base saturation, potassium application, and sulfur application with the yield gap. Increasing these parameters may help to close the yield gap in our study trials. On the other hand, higher soil potassium content, effective cation exchange capacity, soil zinc content, soil iron content, and zinc application were associated with an increase in the yield gap.

The ERA5 weather data has been identified as a more reliable and closer representation of observed rainfall data, making it suitable for modelling environments in Africa, where data scarcity poses significant challenges. Nonetheless, to enhance the robustness and reliability of the simulation model, collection of accurate rainfall data and comprehensive soil characteristics during the trials is recommended. Availability of such reliable data will facilitate a better understanding of the dynamics of the maize water-limited yield gap in Ghana.

Furthermore, our findings emphasize the potential of sulfur application as a secondary nutrient to reduce the maize yield gap when combined with the major macronutrients (NPK). Ghanaian authorities can therefore promote a combination of macronutrients and sulfur nutrient application to improve the maize yield and reduce the yield gap. By encouraging farmers to adopt such practices, they can significantly improve maize productivity. Moreover, focusing on strengthening agricultural activities, especially in the northern regions of the country where the observed maize yield is significantly below the water-limited potential yield, can lead to substantial improvements in maize production and food security. By implementing these strategies, Ghana can effectively address the yield gap and achieve more sustainable and productive agriculture. However, considering the potential release of nitrous oxide due to nitrogen use, it is essential to exercise caution when using mineral fertilizers and consider transitioning to organic alternatives.

CHAPTER 6: REFERENCES

Abdelgadir, G.A., and Eledum, H. (2016). A comparison study of ridge regression and principle component regression with application. *International Journal of Science and Research, 3*(8), 1-11. doi:https://edupediapublications.org/journals

Addo, A., and Amponsah, S.K. (2018). Present status and future prospects of agricultural machinery industry in Ghana. *Agricultural Mechanization in Asia, Africa and Latin America*. doi:http://hdl.handle.net/123456789/1415

Adeoye, G.O., and Agboola, A.A. (1985). Critical levels for soil pH, available P, K, Zn and Mn and maize ear-leaf content of P, Cu and Mn in sedimentary soils of South-Western Nigeria. *Fertilizer research, 6*, 65-71. https://doi.org/10.1007/BF01058165

Adiele, J., Schut, A., van den Beuken, R., Ezui, K., Pypers, P., Ano, A., Egesi, C.N., and Giller, K. (2021). A recalibrated and tested LINTUL-Cassava simulation model provides insight into the high yield potential of cassava under rainfed conditions. *European Journal of Agronomy, 124*, 126242. https://doi.org/10.1016/j.eja.2021.126242

Adu, G., Abdulai, M., Alidu, H., Nustugah, S., Buah, S., Kombiok, J., . . . Etwire, P. (2014). Recommended production practices for maize in Ghana. Accra: AGRA/CSIR.

Adzawla, W., .Atakora, W.K., Kissiedu, I.N., Martey, E., Etwire, P.M., Gouzaye, A., and Bindraban, P.S. (2021). Characterization of farmers and the effect of fertilization on maize yields in the Guinea savannah, Sudan savannah, and transitional agroecological zones of Ghana? *EFB Bioeconomy*. https://doi.org/10.1016/j.bioeco.2021.100019

Aggarwal, P.K., Kalra, N., Chander, S., and Pathak, H. (2006). InfoCrop: a dynamic simulation model for the assessment of crop yields, losses due to pests, and environmental impact of agro-ecosystems in tropical environments. I. Model description. *Agricultural Systems, 89*(1), 1-25. doi:https://doi.org/10.1016/j.agsy.2005.08.001

Agyin-Birikorang, S., Tindjina, I., Adu-Gyamfi, R., Dauda, H.W., Fugice Jr, J., and Sanabria, J. (2022). Managing essential plant nutrients to improve maize productivity in the savanna agroecological zones of northern Ghana: The role of secondary and micronutrients. *Journal of Plant Nutrition, 46*(1), 38-57. https://doi.org/10.1080/01904167.2022.2027984

Andam, K., Johnson, M., Ragasa, C., Kufoalor, D., and Das Gupta, S. (2017). A chicken and maize situation: the poultry feed sector in Ghana. Discussion Paper 01601. International Food Policy Research Institute.

Aslam, M., Maqbool, M.A., and Cengiz, R. (2015). Drought stress in maize (*Zea mays* L.) Effects, resistance mechanisms, global achievements and. Cham: Springer.

Barutçular, C., Dizlek, H., El-Sabagh, A., Sahin, T., EL-Sabagh, M., and Islam, M.S. (2016). Nutritional quality of maize in response to drought stress during grain-filling stages in Mediterranean climate condition. *Exp. Biol. Agric. Sci, 4*, 644-652. http://dx.doi.org/10.18006/2016.4(Issue6).644.652

Boullouz, M., Bindraban, P. S., Kissiedu, I.N., Kouame, A.K., Devkota, K.P., and Atakora, W.K. (2022). An integrative approach based on crop modeling and geospatial and statistical analysis to quantify and explain the maize (*Zea mays*) yield gap in Ghana. *Frontiers in Soil Science, 2*, 68. https://doi.org/10.3389/fsoil.2022.1037222

Breiman, L. (2001). Random forests. *Machine learning, 45*, 5-32. doi:https://doi.org/10.1023/A:1010933404324

Bua, S., El Mejahed, K., MacCarthy, D., Adogoba, D.S., Kissiedu, I.N., Atakora, W.K., . . . Bindraban, P. S. (2020). *Yield response of maize to fertilizers*.

Chapoto, A., and Ragasa, C. (2013). Moving in the right direction? Maize productivity and fertilizer use and use intensity in Ghana.

Chatfield, C., and Xing, H. (2019). *The analysis of time series: An introduction with R.* CRC Press.

Conrado, T.V., Ferreira, D.F., Scapim, C.A., and Maluf, W.R. (2017). Adjusting the Scott-Knott cluster analyses for unbalanced designs. *Crop Breed Appl. Biotechnol.*, *17(1)*, 1-9. http://dx.doi.org/10.1590/198470332017v17n1a1

Cutler, A., Cutler, D.R., and Stevens, J.R. (2012). Random forests. *Ensemble machine learning: Methods and applications*, 157-175. https://doi.org/10.1007/978-1-4419-9326-7_5

Cytiva. (n.d.). Mechanistic-vs-statistical-models. Retrieved from https://www.cytivalifesciences.com/en/us/solutions/bioprocessing/knowledge-center/mechanistic-vs-statistical-models

Dagum, E.B., and Cholette, P.A. (2006). The Components of Time Series. *In: Benchmarking, temporal distribution, and reconciliation methods for time series* (Vol. 186). New York: Springer.

Danquah, E.O., Beletse, Y., Stirzaker, R., Smith, C., Yeboah, S., Oteng-Darko, P., Frimpong, F., and Ennin, S.A. (2020). Monitoring and modelling analysis of maize (*Zea mays* L.) yield gap in smallholder farming in Ghana. *Agriculture, 10*(9), 420. https://doi.org/10.3390/AGRICULTURE10090420

Darfour, B., and Rosentrater, K.A. (2016a). *Agriculture and food security in Ghana.* Paper presented at the 2016 ASABE annual international meeting.

Darfour, B., and Rosentrater, K.A. (2016b). *Maize in Ghana: an overview of cultivation to processing.* Paper presented at the 2016 ASABE Annual International Meeting.

Datta, S., Taghvaeian, S., and Stivers, J. (2017). *Understanding soil water content and thresholds for irrigation management*.

De Janvry, A., and Sadoulet, E. (2010). Agricultural growth and poverty reduction: Additional evidence. *The World Bank research observer, 25*(1), 1-20. https://doi.org/10.1093/wbro/lkp015

Ding, R., Kang, S., Zhang, Y., Hao, X., Tong, L., and Du, T. (2013). Partitioning evapotranspiration into soil evaporation and transpiration using a modified dual crop coefficient model in irrigated maize field with ground-mulching. *Agricultural water management, 127*, 85-96. https://doi.org/10.1016/j.agwat.2013.05.018

Du Plessis, J. (2003). *Maize production*. Department of Agriculture Pretoria, South Africa.

Elsayed, N.S., Obaid, H., Shi, D., Lei, P., Xie, D., Ni, J., . . . Ni, C. (2022). Effect of zinc application on maize productivity and eukaryotic microorganism's diversity in a newly cultivated field. *Journal of Soil Science Plant Nutrition, 22*(3), 3697-3707. https://doi.org/10.1007/s42729-022-00920-x

Essel, B., Abaidoo, R.C., Opoku, A., and Ewusi-Mensah, N. (2020). Economically optimal rate for nutrient application to maize in the semi-deciduous forest zone of Ghana. *Journal of Soil Science Plant Nutrition, 20*, 1703-1713. https://doi.org/10.1007/s42729-020-00240-y

Ezui, K., Leffelaar, P., Franke, A., Mando, A., and Giller, K. (2018). Simulating drought impact and mitigation in cassava using the LINTUL model. *European Journal of Agronomy, 219*, 256-272. https://doi.org/10.1016/j.fcr.2018.01.033

Fan, J., and Gijbels, I. (1996). *Local polynomial modelling and its applications: monographs on statistics and applied probability 66* (Vol. 66). CRC Press.

FAOSTAT. (2021). Crops data. Food and Agriculture Organization of the United Nations. https://www.fao.org/faostat/en/#data/QC

Farré, I., Van Oijen, M., Leffelaar, P., and Faci, J. (2000). Analysis of maize growth for different irrigation strategies in northeastern Spain. *European Journal of Agronomy, 12*(3-4), 225-238. doi:https://doi.org/10.1016/S1161-0301(00)00051-4

Fayaz, A., Kumar, A., Nisar, F., Abidi, I., Rasool, F., Dar, Z., . . . Kumar, Y.R. (2021). Crop Simulation Models: A Tool for Future Agricultural Research and Climate Change. *Asian Journal of Agricultural Extension, Economics Sociology, 39*(6), 146-154.

Gleixner, S., Demissie, T., and Diro, G.T. (2020). Did ERA5 improve temperature and precipitation reanalysis over East Africa? *Atmosphere, 11*(9), 996. https://doi.org/10.3390/atmos11090996

GSS. (2021). The Ghana 2021 Population and Housing Census (PHC) *Volume 3A*.

GSS. (2015). Ghana demographic and health survey 2014. Rockville, Maryland: GSS, GHS, and ICF International.

GSS. (2018). Ghana living standards survey (GLSS7): Poverty trends in Ghana; 2005-2017. Accra: GSS.

Gurjar, G., Swami, S., Meena, N., and Lyngdoh, E. (2017). Effect of solar radiation in crop production. *Natural Resource Management for Climate Smart Sustainable Agriculture, Soil Conservation Society of India. New Delhi*.

Harris, I., Jones, P.D., Osborn, T.J., and Lister, D.H. (2014). Updated high-resolution grids of monthly climatic observations–the CRU TS3. 10 Dataset. *International Journal of Climatology, 34*(3), 623-642. https://doi.org/10.1002/joc.3711

He, D., and Wang, E. (2019). On the relation between soil water holding capacity and dryland crop productivity. *Geoderma, 353*, 11-24. https://doi.org/10.1016/j.geoderma.2019.06.022

Hemdan, E.E.-D., Shouman, M.A., and Karar, M.E. (2020). Covidx-net: A framework of deep learning classifiers to diagnose COVID-19 in x-ray images. *arXiv*. https://doi.org/10.48550/arXiv.2003.11055

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., . . . Schepers, D. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999-2049. https://doi.org/10.1002/qj.3803

Hoerl, A.E., and Kennard, R.W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics, 12*(1), 55-67. https://doi.org/10.1080/00401706.1970.10488634

Hoerl, R.W. (2020). Ridge regression: a historical context. *Technometrics, 62*(4), 420-425. https://doi.org/10.1080/00401706.2020.1742207

Hsiao, T.C., Heng, L., Steduto, P., Rojas-Lara, B., Raes, D., and Fereres, E. (2009). AquaCrop—The FAO crop model to simulate yield response to water: III. Parameterization and testing for maize. *Agronomy Journal, 101*. https://doi.org/10.2134/agronj2008.0218s

IFPRI. (2014). *Maize productivity in Ghana* (Vol. 5). International Food Policy Research Institute.

Jame, Y., and Cutforth, H. (1996). Crop growth models for decision support systems. *Canadian Journal of Plant Science, 76*(1), 9-19. https://doi.org/10.4141/cjps96-003

Jelihovschi, E.G., Faria, J.C., and Allaman, I.B. (2014). ScottKnott: A package for performing the Scott-Knott clustering algorithm in R. *TEMA, 15*, 3-17. doi:https://doi.org/10.5540/tema.2014.015.01.0003

Jones, J.W., Hoogenboom, G., Porter, C.H., Boote, K.J., Batchelor, W.D., Hunt, L., . . . Ritchie, J.T. (2003). The DSSAT cropping system model. *European Journal of Agronomy, 18*(3-4), 235-265. https://doi.org/10.1016/S1161-0301(02)00107-7

Jose, J. (2022). Introduction to time series analysis and its applications.

Karch, J. (2020). Improving on Adjusted R-squared. *Collabra: Psychology, 6*(1). https://doi.org/10.31234/osf.io/v8dz5

Keating, B.A., Carberry, P.S., Hammer, G.L., Probert, M.E., Robertson, M.J., Holzworth, D., . . . Hochman, Z. (2003). An overview of APSIM, a model designed for farming systems simulation. *European Journal of Agronomy, 18*(3-4), 267-288. https://doi.org/10.1016/S1161-0301(02)00108-9

Kendie, S. (2019). Regions and regionalism in Ghana: Reflections on a contemporary issue in Ghana.

Khoshvaghti, H., Eskandari-Kordlar, M., and Lotfi, R. (2014). Response of maize cultivars to water stress at grain filling phase. *Azarian Journal of Agriculture*.

Kouame, A.K., Bindraban, P.S., Kissiedu, I.N., Atakora, W.K., and El Mejahed, K. (2023). Identifying drivers for variability in maize (*Zea mays* L.) yield in Ghana: A meta-regression approach. *Agricultural Systems, 209*, 103667. https://doi.org/10.1016/j.agsy.2023.103667

Kouame, K.K.A., Bindraban, P.S., Kissiedu, I.N., Atakora, W.K., and El Mejahed, K. (2021). *Evaluation and Geospatial Analysis of Variability in Maize Yield Response to Fertilizer (NPK) Using Modeling in Ghana*.

Kugbe, J., Kombat, R., and Atakora, W. (2019). Secondary and micronutrient inclusion in fertilizer formulation impact on maize growth and yield across northern Ghana. *Cogent Food Agriculture, 5*(1), 1700030. https://doi.org/10.1080/01904167.2022.2027984

Ledolter, J. (2008). Smoothing time series with local polynomial regression on time. *Communications in Statistics—Theory Methods, 37*(6), 959-971. https://doi.org/10.1080/03610920701693843

Leenaars, J.G., Claessens, L., Heuvelink, G.B., Hengl, T., González, M.R., van Bussel, L.G., . . . and Cassman, K.G. (2018). Mapping rootable depth and root zone plant-available water holding capacity of the soil of sub-Saharan Africa. *Geoderma, 324*, 18-36. https://doi.org/10.1016/j.geoderma.2018.02.046

Leitner, S., Pelster, D.E., Werner, C., Merbold, L., Baggs, E.M., Mapanda, F., and Butterbach-Bahl, K. (2020). Closing maize yield gaps in sub-Saharan Africa will boost soil N2O emissions. *Current Opinion in Environmental Sustainability, 47*, 95-105. https://doi.org/10.1016/j.cosust.2020.08.018

Liu, D.-Y., Zhang, W., Yan, P., Chen, X.-P., Zhang, F.-S., and Zou, C.-Q. (2017). Soil application of zinc fertilizer could achieve high yield and high grain zinc concentration in maize. *Plant and Soil, 411*, 47-55. https://doi.org/10.1007/s11104-016-3105-9

Liu, Y., Gao, M., Wu, W., Tanveer, S.K., Wen, X., and Liao, Y. (2013). The effects of conservation tillage practices on the soil water-holding capacity of a non-irrigated apple orchard in the Loess Plateau, China. *Soil and Tillage Research, 130*(2013), 7-12. https://doi.org/10.1016/j.still.2013.01.012

Lobell, D.B., Schlenker, W., and Costa-Roberts, J. (2011). Climate trends and global crop production since 1980. *Science, 333*(6042), 616-620. https://doi.org/10.1126/science.1204531

Mabe, F.N., Danso-Abbeam, G., and Ehiakpor, D.S. (2018). Assessment of implementation of Planting For Food and Jobs (PFJ) programmes: Lessons and ways forward.

MacCarthy, D.S., Adiku, S.G., Freduah, B.S., Kamara, A.Y., Narh, S., and Abdulai, A.L. (2018). Evaluating maize yield variability and gaps in two agroecologies in northern Ghana using a crop simulation model. *South African Journal of Plant, 35*(2), 137-147. https://doi.org/10.1080/02571862.2017.1354407

Machakaire, A.T., Steyn, J.M., Caldiz, D., and Haverkort, A. (2016). Forecasting yield and tuber size of processing potatoes in South Africa using the LINTUL-potato-DSS model. *Potato Research, 59*, 195-206. https://doi.org/10.1007/s11540-016-9321-0

Manzanas, R., Amekudzi, L.K., Preko, K., Herrera, S., and Gutierrez, J.M. (2014). Precipitation variability and trends in Ghana: An intercomparison of observational and reanalysis products. *Climatic Change*. https://doi.org/10.1007/s10584-014-1100-9

Manzeke, G.M., Mapfumo, P., Mtambanengwe, F., Chikowo, R., Tendayi, T., and Cakmak, I. (2012). Soil fertility management effects on maize productivity and grain zinc content in smallholder farming systems of Zimbabwe. *Plant and Soil, 361*, 57-69. https://doi.org/10.1007/s11104-012-1332-2

Massignam, A., Chapman, S., Hammer, G., and Fukai, S. (2009). Physiological determinants of maize and sunflower grain yield as affected by nitrogen supply. *Field Crops Research, 113*(3), 256-267. https://doi.org/10.1016/j.fcr.2009.06.001

McCown, R.L., Hammer, G.L., Hargreaves, J.N.G., Holzworth, D.P., and Freebairn, D.M. (1996). APSIM: A novel software system for model development, model testing and simulation in agricultural systems research. *Agricultural Systems, 50*(3), 255-271. https://doi.org/10.1016/0308-521X(94)00055-V

McMillen, M.S., Mahama, A.A., Sibiya, J., Lübberstedt, T., and Suza, W.P. (2022). Improving drought tolerance in maize: Tools and techniques. *Frontiers in Genetics, 13*, 1001001. https://doi.org/10.3389/fgene.2022.1001001

MiDA. (2010). *Investment opportunity in Ghana: Maize, rice, and soybean.* Accra, Ghana.

MoFA. (2010). *2010 MoFA annual program review.* Accra, Ghana: MoFA.

MoFA. (2019a). Agriculture in Ghana: Facts and Figures 2018. Accra, Ghana: SRID.

MoFA. (2019b). Planting for Food and Jobs (PFJ) Performance Review (2017 and 2018) and Outlook (2019). PFJ Secretariat. Accra, Ghana: MoFA.

MoFA. (2020). *Meet the press; Presentation by the Minister For Food and Agriculture. Minist. Food Agric. Stat. Res. Inf. Dir. MoFA, Accra, Ghana*. https://www.mofa.gov.gh/site/media-centre/press-briefing/331-meet-the-press

Mohapatra, N., Shreya, K., and Chinmay, A. (2020). *Optimization of the random forest algorithm.* Paper presented at the Advances in Data Science and Management: Proceedings of ICDSM 2019.

Morales-Oñate, V., and Morales-Oñate, B. (2021). MTest: A bootstrap test for multicollinearity. *Munich Personal RePEc Archive*. https://doi.org/10.33333/rp.vol51n2.05

Mourice, S.K., Rweyemamu, C.L., Tumbo, S.D., and Amuri, N. (2014). Maize cultivar specific parameters for decision support system for agrotechnology transfer (DSSAT) application in Tanzania. *American Journal of Plant Sciences*. http://hdl.handle.net/10625/53414

Naab, J.B., Singh, P., Boote, K.J., Jones, J.W., and Marfo, K.O. (2004). Using the CROPGRO-peanut model to quantify yield gaps of peanut in the Guinean Savanna zone of Ghana. *Agronomy Journal, 96*(5), 1231-1242. https://doi.org/10.2134/agronj2004.1231

Nafiu, A., Togun, O., Abiodun, M., and Chude, V. (2011). Effects of NPK fertilizer on growth, dry matter production and yield of eggplant in southwestern Nigeria. *Agriculture Biology Journal of North America, 2*(7), 1117-1125. https://doi.org/10.5251/abjna.2011.2.7.1117.1125

Nyombi, K. (2010). *Understanding growth of East Africa highland banana: experiments and simulation.* Ph.D., Wageningen University and Research.

Orcan, F. (2020). Parametric or non-parametric: Skewness to test normality for mean comparison. *International Journal of Assessment Tools in Education, 7*(2), 255-265. https://doi.org/10.21449/ijate.656077

Oteng-Darko, P., Yeboah, S., Addy, S., Amponsah, S., and Danquah, E. O. (2013). Crop modeling: A tool for agricultural research – A review.

Oxford Business Group. (2021). Agriculture in Africa 2021.

Palm, R., and Iemma, A. (2002). Conditions d'application et transformations de variables en régression linéaire. *Notes de Statistique et d'Informatique*. https://hdl.handle.net/2268/81767

Paustian, K., Collier, S., Baldock, J., Burgess, R., Creque, J., DeLonge, M., . . . and Goddard, T. (2019). Quantifying carbon for agricultural soil management: from the current status toward a global soil information system. *Carbon Management, 10*(6), 567-587. https://doi.org/10.1080/17583004.2019.1633231

Penman, H.L. (1948). Natural evaporation from open water, bare soil and grass. *Proceedings of the Royal Society of London. Series A. Mathematical Physical Sciences, 193*(1032), 120-145. https://doi.org/10.1098/rspa.1948.0037

Pereira, G.W., Valente, D.S.M., de Queiroz, D.M., de Freitas Coelho, A.L., Costa, M.M., and Grift, T. (2022). Smart-map: An open-source QGIS plugin for digital mapping using machine learning techniques and ordinary kriging. *Agronomy Journal, 12*(6), 1350. https://doi.org/10.3390/agronomy12061350

Poggio, L., De Sousa, L. M., Batjes, N.H., Heuvelink, G., Kempen, B., Ribeiro, E., and Rossiter, D. (2021). SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. *Soil, 7*(1), 217-240. https://doi.org/10.5194/soil-7-217-2021

Poole, M.A., and O'Farrell, P.N. (1971). The assumptions of the linear regression model. *Transactions of the Institute of British Geographers*, 145-158. https://doi.org/10.2307/621706

Quaye, W. (2008). Food security situation in northern Ghana, coping strategies and related constraints. *African Journal of Agricultural Research, 3*(5), 334-342. https://edepot.wur.nl/29768

Raes, D., Steduto, P., Hsiao, T.C., and Fereres, E. (2009). AquaCrop – The FAO crop model to simulate yield response to water: II. Main algorithms and software description. *Agronomy Journal, 101*(3), 438-447. https://doi.org/10.2134/agronj2008.0140s

Rockström, J. (2003). Water for food and nature in drought–prone tropics: vapour shift in rain-fed agriculture. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 358*(1440), 1997-2009. https://doi.org/10.1098/rstb.2003.1400

Ruppert, D., Sheather, S.J., and Wand, M.P. (1995). An effective bandwidth selector for local least squares regression. *Journal of the American Statistical Association, 90*(432), 1257-1270. https://doi.org/10.1080/01621459.1995.10476630

Salo, T.J., Palosuo, T., Kersebaum, K.C., Nendel, C., Angulo, C., Ewert, F., . . . and Moriondo, M. (2016). Comparing the performance of 11 crop simulation models in predicting yield response to nitrogen fertilization. *Journal of Agricultural Science, 154*(7), 1218-1240. https://doi.org/10.1017/S0021859615001124

SARI. (1996). *Annual Report*.

Saxton, K., Rawls, W.J., Romberger, J.S., and Papendick, R. (1986). Estimating generalized soil-water characteristics from texture. *Soil Science Society of America Journal, 50*(4), 1031-1036. https://doi.org/10.2136/sssaj1986.03615995005000040039x

Scheiterle, L., and Birner, R. (2018). Assessment of Ghana's comparative advantage in maize production and the role of fertilizers. *Sustainability, 10*(11), 4181. https://doi.org/10.3390/su10114181

Schmidt, A.F., and Finan, C. (2018). Linear regression and the normality assumption. *Journal of clinical epidemiology, 98*, 146-151. https://doi.org/10.1016/j.jclinepi.2017.12.006

Shah, T.R., Prasad, K., and Kumar, P. (2016). Maize – A potential source of human nutrition and health: A review. https://doi.org/10.1080/23311932.2016.1166995

Shen, H., Chen, Y., Wang, Y., Xing, X., and Ma, X. (2020). Evaluation of the potential effects of drought on summer maize yield in the Western Guanzhong Plain, China. *Agronomy, 10*(8), 1095. https://doi.org/10.3390/agronomy10081095

Shibu, M.E., Leffelaar, P.A., Van Keulen, H., and Aggarwal, P.K. (2010). LINTUL3, a simulation model for nitrogen-limited situations: Application to rice. *European Journal of Agronomy, 32*(4), 255-271. https://doi.org/10.1016/j.eja.2010.01.003

Spitters, C., and Schapendonk, A. (1990). Evaluation of breeding strategies for drought tolerance in potato by means of crop growth simulation. In: El Bassam, N., Dambroth, M., Loughman, B.C. (eds) Genetic Aspects of Plant Mineral Nutrition. Developments in Plant and Soil Sciences, vol 42. Dordrecht: Springer. https://doi.org/10.1007/978-94-009-2053-8_24

SRID/MoFA. (2017). Technical Report. Minist. Food Agric. (MoFA)-Statistics, Res. Inf. Dir. (SRID).

Srivastava, A.K., Ceglar, A., Zeng, W., Gaiser, T., Mboh, C.M., and Ewert, F. (2020). The implication of different sets of climate variables on regional maize yield simulations. *Atmosphere, 11*(2), 180. https://doi.org/doi:10.3390/atmos11020180

Steduto, P., Hsiao, T.C., Raes, D., and Fereres, E. (2009). AquaCrop – The FAO crop model to simulate yield response to water: I. Concepts and underlying principles. *Agronomy Journal, 101*(3), 426-437. https://doi.org/10.2134/agronj2008.0139s

Thomas, M.A.H. (2020). *Improving Crop Yields in Sub-Saharan Africa-What Does the East African Data Say*. International Monetary Fund.

Tobler, W.R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography, 46*(sup1), 234-240. https://doi.org/10.2307/143141

van Loon, M.P., Adjei-Nsiah, S., Descheemaeker, K., Akotsen-Mensah, C., van Dijk, M., Morley, T., . . . and Reidsma, P. (2019). Can yield variability be explained? Integrated assessment of maize yield gaps across smallholders in Ghana. *Field Crops Research, 236*, 132-144. https://doi.org/10.1016/j.fcr.2019.03.022

van Oijen, M., and Leffelaar, P. (2008a). *Lintul-1: potential crop growth. A simple general crop growth model for optimal growing conditions (example: spring wheat)*.

van Oijen, M., and Leffelaar, P. (2008b). *Lintul-2: water limited crop growth. A simple general crop growth model for water-limited growing conditions (example: spring wheat)*.

Wang, L.K., and Wang, M.-H.S. (2022). Understanding evaporation, transpiration, evapotranspiration, precipitation and runoff volume for agricultural waste management. *Evolutionary Progress in Science, Technology, Engineering, Arts, Mathematics*, 1-81. https://doi.org/10.17613/m8tf-zd10

World Bank. (2017). Ghana: Agriculture Sector Policy Note Transforming Agriculture for Economic Growth, Job Creation and Food Security.

Worqlul, A.W., Dile, Y.T., Jeong, J., Adimassu, Z., Lefore, N., Gerik, T., Srinivasan, R., and Clarke, N. (2019). Effect of climate change on land suitability for surface irrigation and irrigation

potential of the shallow groundwater in Ghana. *Computers and Electronics in Agriculture, 157*, 110-125.

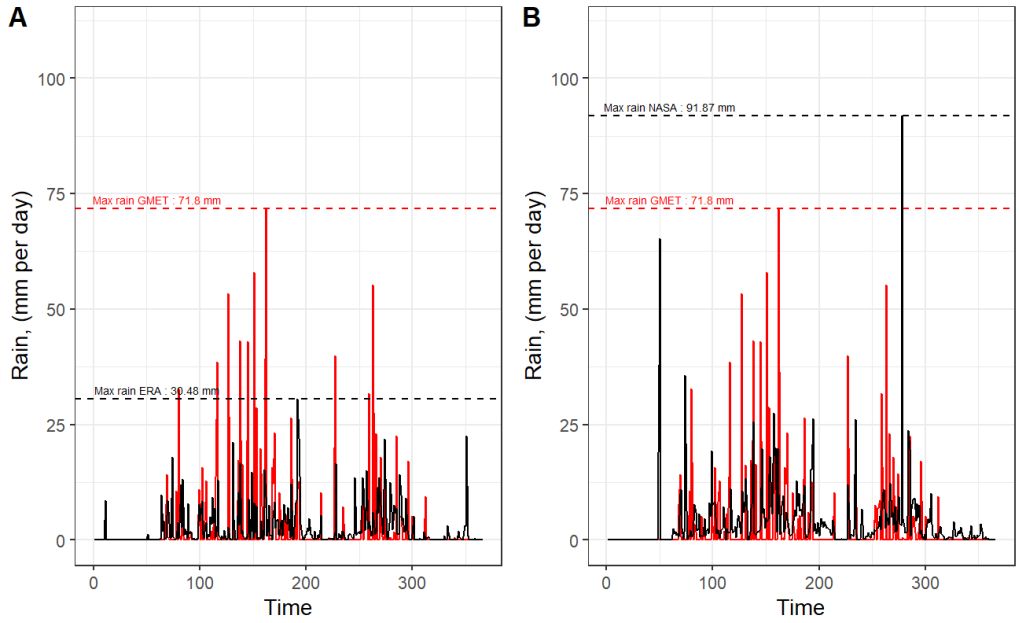
Yamba, E.I., Aryee, J.N., Quansah, E., Davies, P., Wemegah, C.S., Osei, M.A., . . . and Amekudzi, L.K. (2023). Revisiting the agro-climatic zones of Ghana: A re-classification in conformity with climate change and variability. *PLOS Climate, 2*(1), e0000023. https://doi.org/10.1371/journal.pclm.0000023

Zewide, I., and Reta, Y. (2021). Review on the role of soil macronutrient (NPK) on the improvement and yield and quality of agronomic crops. *Journal of Agriculture and Food Science, 9*(1), 7-11. https://doi.org/10.26765/DRJAFS23284767

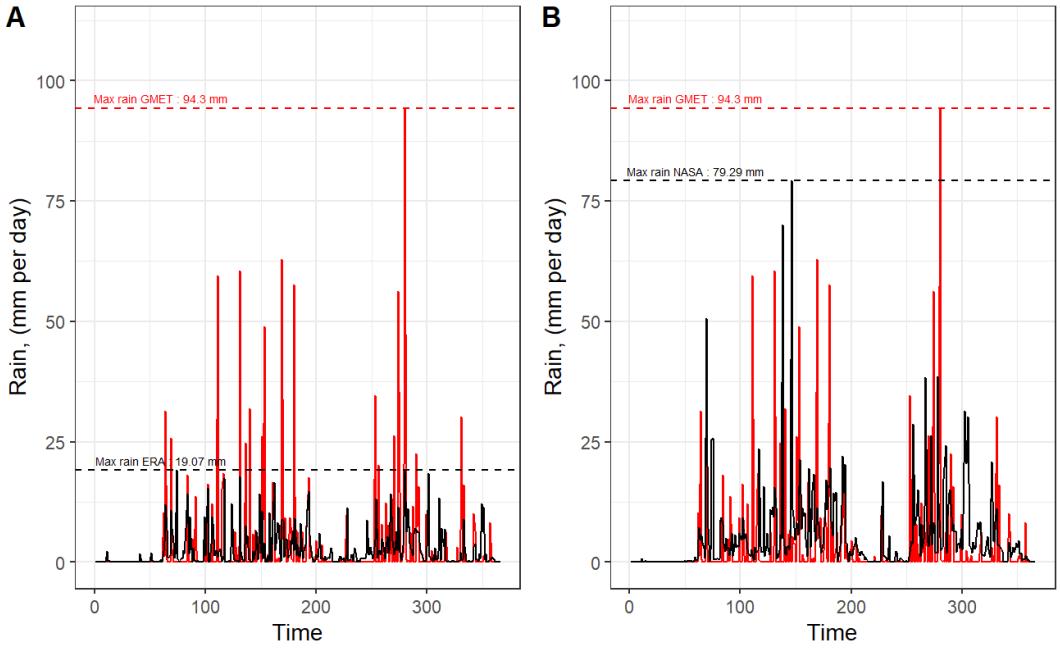
Zhang, Y.-W., Wang, K.-B., Wang, J., Liu, C., and Shangguan, Z.-P. (2021). Changes in soil water holding capacity and water availability following vegetation restoration on the Chinese Loess Plateau. *Scientific Reports, 11*(1), 9692. https://doi.org/10.1038/s41598-021-88914-0

Zhao, C., Liu, B., Xiao, L., Hoogenboom, G., Boote, K.J., Kassie, B.T., . . . and Hernandez-Ochoa, I.M. (2019). A SIMPLE crop model. *European Journal of Agronomy, 104*, 97-106. https://doi.org/10.1016/j.eja.2019.01.009

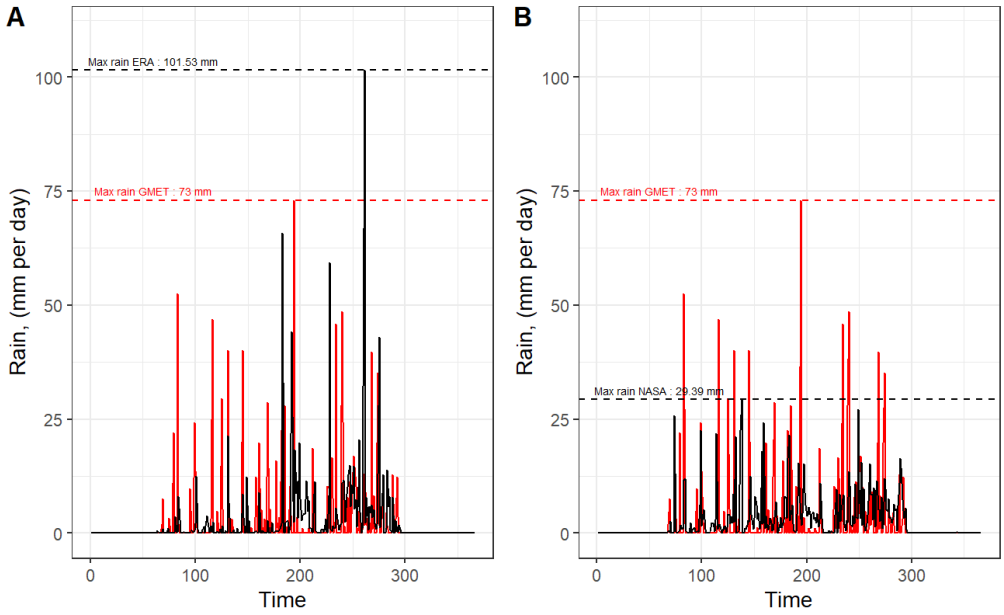
CHAPTER 7: APPENDIX



Appendix A. Ejura daily rainfall comparision (A) GMET vs ERA-5 and (B) GMET vs NASA Power



Appendix B. Kumasi daily rainfall comparision (A) GMET vs ERA-5 and (B) GMET vs NASA Power



Appendix C. Tamale daily rainfall comparison (A) GMET vs ERA-5 and (B) GMET vs NASA Power

FERARI

FERTILIZER RESEARCH & RESPONSIBLE IMPLEMENTATION

FERARI is an international public-private partnership that builds science-based approaches to site-specific fertilization for widespread adoption by farmers in Ghana for improved food and nutrition security. This calls for a transformation of the fertilizer and food systems that must be driven by evidence-based agro-technical perspectives embedded in multi-stakeholder processes.

To support this transformation, the following institutions have partnered to implement the Fertilizer Research and Responsible Implementation (FERARI) program:

- International Fertilizer Development Centre (IFDC)
- Mohammed VI Polytechnic University (UM6P)
- OCP Group
- Wageningen University and Research (WUR)
- University of Liège (ULiège)
- University of Ghana (UG)
- University for Development Studies (UDS)
- Kwame Nkrumah University of Science and Technology in Kumasi (KNUST)
- University of Cape Coast (UCC)
- University of Energy and Natural Resources (UENR)
- Akenten Appiah-Menka University of Skills Training and Entrepreneurial Development (AAMUSTED) College of Agriculture Education
- Council for Scientific and Industrial Research in Kumasi (CSIR-SRI) and in Tamale (CSIR-SARI) and its subsidiary (CSIR-SARI-Wa)

FERARI operates in conjunction with the Planting for Food and Jobs program of the Government of Ghana (GoG) to embed development efforts into national policy priorities to reach impact at scale. It trains five Ph.D. and two post-doctoral candidates and dozens of master's-level students in building the evidence base for its interventions.

FERARI conducts hundreds of fertilizer response trials on maize, rice, and soybean, on-station and also with farmers, and demonstrates them to farmer groups in the northern and middle belt of Ghana. It conducts surveys among farmers and actors in the value chain to understand the drivers for use of fertilizers and other inputs and the marketing of the produce to enhance farm productivity and income. It helps the GoG to establish a Ghana National Fertilizer Platform, developing its soil mapping expertise toward an information platform.


IFDC
Developing Agriculture from the Ground Up