IDENTIFYING DRIVERS FOR MAIZE RESPONSE TO FERTILIZER IN GHANA

Anselme K. K. Kouame[a], Isaac N. Kissiedu[a], Williams K. Atakora[a], Khalil El Mejahed[b], Prem S. Bindraban[a,*]

[a] International Fertilizer Development Center (IFDC), Muscle Shoals, AL 35662, USA.
[b] Mohammed VI Polytechnic University (UM6P), AgroBioSciences, Lot 660, Hay Moulay Rachid, 43150 Ben Guerir, Morocco.

[*] Corresponding authors: E-mail addresses: pbindraban@ifdc.org (P.S. Bindraban).

ABSTRACT

CONTEXT: Maize is the main cereal crop produced in Ghana, but its yield is severely affected by several biotic and abiotic factors. Increasing the overall productivity of maize is essential to ensure food security and lift farmers out of poverty.

OBJECTIVE: Therefore, this study was done to quantify the effect of these factors on maize yield response to fertilizer using 978 data points from on-farm and on-station trials so that intervention practices can be identified.

METHODS: The Quantitative Evaluation of the Fertility of Tropical Soils (QUEFTS) model and five regression algorithms to build the maize yield model – Ordinary Least Squares Regression (OLSR), Multivariate Linear Regression (MLR), Stepwise Multiple Linear Regression (SMLR), and Random Forest Regression (RFR) – were used to analyze the data.

RESULTS AND CONCLUSIONS: The results show that the QUEFTS model cannot significantly explain yield variability at on-station and on-farm levels (adj. R^2=3% and adj. R^2=22%, respectively). However, the MLR, SMLR, and RFR models explained more than 60% of the variability in maize yield. SMLR and RFR show that the type and rate of fertilizer applied, temperature, variety, and root zone depth are significant factors in explaining maize yield variability. They reveal that soil physical properties explain more of the yield variability (adj. R^2=32%) on-station than environmental parameters (adj. R^2=1%), with soil chemical properties explaining the highest percentage (adj. R^2=36%). On-farm, environmental covariates (adj. R^2=33%) explain more of the variability in yield response than physical (adj. R^2=25%) and chemical (adj. R^2=19%) soil variables. Detailed analytics pinpointed that high temperature and high precipitation, combined with shallow rooting depth (<50 cm), were key factors reducing the efficiency and effectiveness of on-farm fertilizer application.

Calculation of nutrient use efficiency shows an average partial factor productivity of 29-80 kg grain and an average agronomic efficiency of 13-46 kg grain per kg of applied N.

SIGNIFICANCE: While fertilizer recommendations are generally based on soil chemical properties, the results of this study indicate that methods to arrive at such recommendations should take soil physical properties and climatic variables into consideration as well.

Keywords: Variability, QUEFTS, Regression, Random Forest, Nutrient use efficiency.

1. INTRODUCTION

Global demand for food will continue to increase for at least 50 years (Cicin-Sain, 2018; EU, 2019; Tilman et al., 2011), and climate change is not helping matters. Agricultural production in sub-Saharan Africa must at least triple to meet this growing food demand (Godfray et al., 2010; Rahman et al., 2021). Moreover, the agri-food system plays a central role in the sub-Saharan African countries' economies (FAO/OECD, 2018) and is at the core of at least 12 of the 17 United Nations 2030 Sustainable Development Goals (SDGs). In Ghana, it has been the fastest-growing sector of the economy for several decades (Diao et al., 2019). Thus, agricultural growth is the main driver of poverty reduction and the largest source of employment for rural communities, mainly smallholder farmers with 2 hectares of land or less (USAID, 2022). However, farmers face changing and increasingly unpredictable weather conditions, drastically reducing soil fertility, and typically use local or inbred crop varieties. Bationo et al., (2018) reported that soil nutrient depletion rates of about 35 kg N, 4 kg P, and 20 kg K per hectare are worrisome and prevalent in all agroecological zones (AEZs) in Ghana, with nitrogen and phosphorus being the most deficient nutrients (Zingore et al., 2015). As a result, yields obtained by smallholder farmers are far below the potentially attainable yields, hampering agricultural production and jeopardizing economic development and food security (Adzawla et al., 2021). One solution is to increase fertilizer application by farmers. However, it is increasingly understood that crop response to fertilizer applied in many areas of Africa, including Ghana, is depressed by a variety of soil degradation problems and many other factors, such as crop variety, soil organic matter, and soil depth (Guilpart et al., 2017; Kpotor et al., 2014; Leenaars et al., 2018; Sadras and Calvino, 2001; Tetteh et al., 2016). Furthermore, high variability in climatic conditions (rainfall and temperature) causes uncertainties in agricultural productivity, with profound impacts on the ecology, economy, and the social welfare of rural farmers (Kyei-Mensah et al., 2019; Onduru and Du Preez, 2007).

Despite current low crop productivity, Ghana could intensify production and significantly close current yield gaps of major cereals (Bationo et al., 2018; van Loon et al., 2019), since it has been estimated that, on average, only about 20% of the potential maize yield is being achieved across Ghana (GYGA, 2021). For example, addressing nutrient deficiencies by applying fertilizer alone would help to reduce the yield gap to 50% of attainable yield (Mueller et al., 2012). However, Abunyewa and Mercer-Quarshie, (2004) and Bationo et al., (2018) reported that the maize grain yield rarely exceeds 1 t ha^{-1} in farmers' fields even with application of N, P, and K compound fertilizer, which demonstrates that many other factors contribute to nutrient use efficiency (Genuer and Poggi). Subsequently, Bua et al., (2020) found that grain yield response to fertilization is highly variable, ranging from a mere 500 kg ha^{-1} up to 8,000 kg ha^{-1}. This large variability depresses farmers' incentive and ability to purchase fertilizers in subsequent seasons (Njoroge, 2019; Roobroeck et al., 2021).

Capturing observed variabilities in maize yield can be done using model-based approaches. Various studies around the world have shown that application of system models (Wallach et al., 2018) can be useful in determining and prioritizing the relative importance of factors that contribute to yield variability (Jeong et al., 2016; Lamos-Díaz et al., 2020; Nevavuori et al., 2020; Paudel et al., 2021; Timsina et al., 2021). Models such as the QUantitative Evaluation of the Fertility of Tropical Soils (QUEFTS) have been advocated by several studies for estimating field-specific N, P, and K recommendations (Ren et al., 2015; Tabi et al., 2007; Tittonell et al., 2008; Wijayanto and Prastyanto, 2012; Xu et al., 2013). QUEFTS is not commonly used as a model to assess maize yield variability but could be used for this purpose, as it considers soil chemical properties (pH, organic carbon, available and total phosphorus, exchangeable potassium, and organic nitrogen) and fertilizer application (Njoroge, 2019; Onduru and Du Preez, 2007). The model assumes that all other production factors are optimal and does not consider the variety of crop used, the physical properties of the soil, or climatic variables. To estimate the effects of these factors not taken into account in the QUEFTS model, it is important to complement it with multivariate analyses through mathematical models of linear regression (Atiah et al., 2021; Kihara et al., 2016; van Loon et al., 2019). Thus, ordinary least squares regression (OLSR), multivariate linear regression (MLR), stepwise multiple linear regression (SMLR), and random forest regression (RFR) methods were applied along with QUEFTS to quantify, explain, and predict maize yield variability in Ghana.

2. MATERIALS AND METHODS

2.1. Data sites and sources

2.1.1. Yield and fertilizer

Maize yields from 1,818 fertilizer-yield response data points of on-farm and on-station trials conducted between 1990 and 2017 were collected from scientific papers and local institutional reports. Of the 1,818 data points, 840 were removed from the raw database due to a lack of information on geographic coordinates, dates of trials, and varieties used, as well as a lack of metadata on the proportions of the chemical elements N, P, and K in the NPK fertilizers applied at these sites. In addition, all data points in which the chemical elements N, P, and K in the fertilizers were combined with micronutrients (boron, zinc, and magnesium) were excluded from the study because one of the models used in this study (QUEFTS) cannot model or explain their effects on maize yield. At the end of the data cleaning and preparation, 978 data points were available for use in this study. Treatments were heterogeneous and included organic fertilizers alone, organic fertilizer in combination with inorganic fertilizers, and inorganic fertilizers alone. The combination of N, P, and K in the treatments included a control (no fertilizer), treatments with PK, NK, and NP, and treatments with NPK. Organic fertilizers were converted into N, P_2O_5, and K_2O (Adjei-Nsiah, 2012; Badu et al., 2019; Fening et al., 2009;

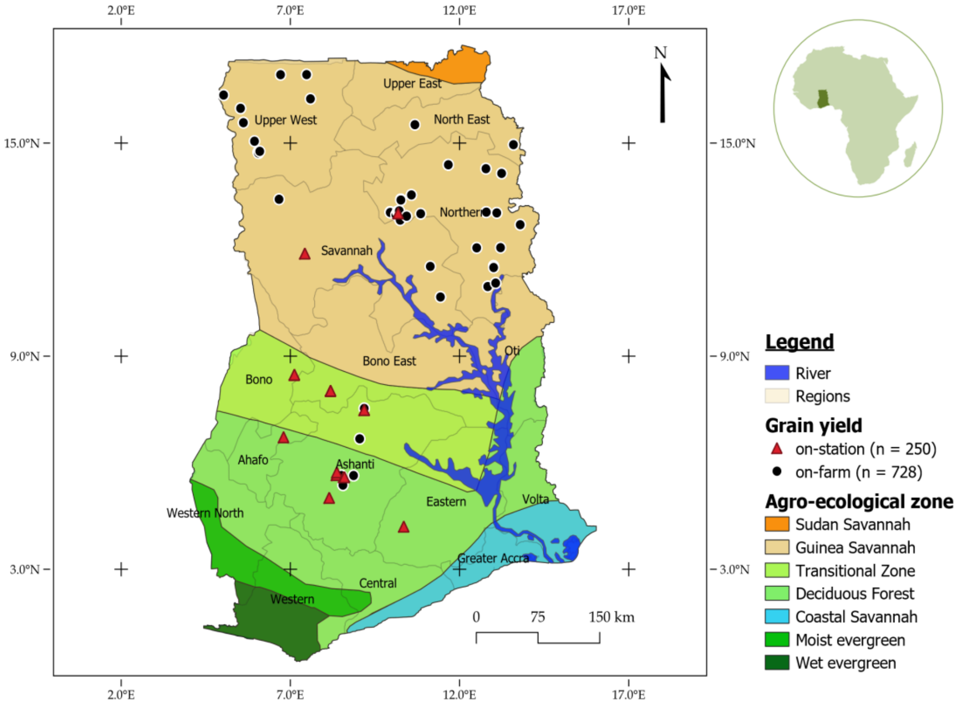

Figure 1: Map of Ghana showing AEZs and spatial distribution of on-farm and on-station trial data points

Kanton et al., 2016). The trials were conducted in three of Ghana’s AEZs: Guinea Savanna, Transitional , and Semi-Deciduous Forest (Figure 1).

2.1.2. Quantitative covariates used in the models

Table 1 presents the covariates used in the models (QUEFTS, OLSR, MLR, SMLR and RFR). Soil chemical and physical properties data were obtained from the African SoilGrids (ISRIC), at a resolution of 250 m for the 0-30 cm topsoil. Rainfall and temperature datasets are from the WorldClim (Fick and Hijmans, 2017) database, at a spatial resolution of 1 km² and elevation data (altitude) with a resolution of 30 meters from Shuttle Radar Topography Mission (SRTM) Digital Terrain Elevation. The coordinates of the location of each trial were overlaid on the maps for extracting soil chemical and physical properties, climatic factors, and elevation.

Monthly rainfall and temperature data for the years that matched the trial years and growing seasons of the yield data points were used. Temperature and rainfall were aggregated based on suggested maize planting times in Ghana, considering agroclimatic conditions (Adu et al., 2014), assuming that maize was harvested four months after its planting date. The on-station trials took place on deeper soils (132 cm) than the on-farm trials (76 cm). Most of the on-farm trials were in the Northern Region and the on-station trials in the Middle Belt.

Table 1: Quantitative variables (soil property, climatic factor, and altitude) used as input data in the models and their source

Covariate’s name	Abbreviation
Digital elevation (m) (1)	altitude
Precipitation (mm) (2)	rain
Maximum temperature (°C) (2)	tmax
Minimum temperature (°C) (2)	tmin
Africa SoilGrids - Root zone depth (cm) (3)	rzd
Africa SoilGrids - Root zone plant available water holding capacity aggregated (cm) (3)	rzwhc
Africa SoilGrids - Silt content (%) (3)	silt
Africa SoilGrids - Sand content (%) (3)	sand
Africa SoilGrids - Clay content (%) (3)	clay
Africa SoilGrids - Soil pH in H_2O (3)	pH
Africa SoilGrids nutrients - Extractable phosphorus (P) (mg kg^{-1}) (3)	Olsen
Africa SoilGrids - Cation exchange capacity (CEC) (mmol$_{(+)}$ kg^{-1}) (3)	cec
Africa SoilGrids - Exchangeable potassium (K+) (mmol kg^{-1}) (3)	Ex.K
Africa SoilGrids nutrients - Total phosphorus (P) (mg kg^{-1}) (3)	Pt
Africa SoilGrids nutrients - Total nitrogen (N) (g kg^{-1}) (3)	Ntot
Africa SoilGrids - Soil organic carbon (SOC) (g kg^{-1}) (3)	Corg

(1) https://srtm.csi.cgiar.org/srtmdata/ (2) (Fick and Hijmans, 2017).(3) (Hengl et al., 2015) (Hengl et al., 2017) .

Table 2: Descriptive characteristics statistics of all the maize yield data point location variables data presented in Table 1

	altitude (m)	rain (mm)	tmin (ºC)	tmax (ºC)	rzd (cm)	rzwhc (cm)	sand (%)	clay (%)	silt (%)	pH	Olsen (mg.kg^{-1})	cec (mmol$_{(+)}$.kg^{-1})	Ex.K (mmol.kg^{-1})	Pt (mg.kg^{-1})	Norg (g.kg^{-1})	Corg (g.kg^{-1})
on-farm (n^1=728)																
Min^2	115	498	19.0	27.0	40.0	9.00	56.6	6.11	6.50	5.40	2.70	2.40	1.88	112	0.26	3.00
Max^3	390	810	24.0	33.0	150	13.0	74.8	22.5	39.3	6.30	9.92	17.5	3.64	468	1.15	16.0
Mean	207	640	22.8	31.0	76.4	10.3	65.0	14.0	25.3	6.06	3.95	6.32	2.82	215	0.55	5.94
SD^4	63.6	51.7	0.88	1.18	28.1	1.00	5.58	4.96	6.65	0.18	1.61	3.06	0.33	82.5	0.24	2.89
CV^5(%)	31%	8%	4%	4%	37%	10%	9%	35%	26%	3%	41%	48%	12%	38%	44%	49%
on-station (n^1=250)																
Min^2	176	459	21.9	29.7	10.0	9.00	40.0	6.70	6.50	5.40	2.62	4.20	2.12	176	0.69	4.00
Max^3	292	710	23.5	32.3	150	13.0	84.1	27.0	52.0	6.10	8.50	82.9	4.85	266	1.21	15.0
Mean	258	619	22.3	30.7	132	10.6	67.3	15.7	17.0	6.01	4.40	17.4	2.84	216	0.87	8.02
SD^4	18.1	78.1	0.40	0.47	29.3	0.86	6.31	3.96	6.63	0.15	0.70	24.4	0.54	24.6	0.17	4.27
CV^5(%)	7%	13%	2%	2%	22%	8%	9%	25%	39%	2%	16%	140%	19%	11%	20%	53%

n^1: number of trials. Min^2: minimum. Max^3 : maximum. SD^4 : standard deviation. CV^5: coefficient of variation.

2.1.3. Maize variety used in linear regression

Eighteen maize varieties, consisting of one local variety, seven improved open-pollinated varieties, four inbred lines, one improved normal maize, one quality protein maize, and four commercial hybrid varieties, were used.

Table 3: Varieties used in the study and their characteristics

	Variety	On-station	On-farm	Characteristic
Genotype	Abontem	no	yes	YOPV[1]/QPM[2], STR[3], Earliness
	Aburohema	no	yes	YOPV[1]/ QPM[2]
	Akposoe	no	yes	YOPV[1]/ QPM[2], DT[4], Earliness
	Dorke SR	yes	no	OPNM[5]
	Entry 5	no	yes	Inbred line, DT[4]
	Entry 6	no	yes	Inbred line, DT[4]
	Entry 70	no	yes	Inbred line, DT[4]
	Entry 85	no	yes	Inbred line, DT[4]
	Etubi	no	yes	WQPHM[6], DT[4]
	GH 110	no	yes	SCH[7], DT[4]
	Golden Jubilee	no	yes	YOPV[1]/ QPM[2]
	Local variety	no	yes	OPNM[5]
	Mamaba	yes	yes	WQPHM[6], DT[4]
	Obatanpa	yes	yes	WOPV[8]/QPM[2]
	Omankwa	no	yes	WOPV[8]/ QPM[2], DT[4]
	Pannar53	no	yes	Hybrid
	QPM	no	yes	QPM[2]
	Wang Daata	no	yes	YOPV[1]/ QPM[2]

(1) YOPV: Yellow Open-Pollinated Variety. (2) QPM: Quality Protein Maize. (3) STR: Striga hermontica Resistance. (4) DT: Drought Tolerant. (5) OPNM: Open-Pollinated Normal Maize. (6) WQPHM: White Quality Protein Hybrid Maize. (7) SCH: Single Cross Hybrid. Yes: the variety was used. No: the variety was not used.

These names of the varieties were taken from the papers from which the 978 data points on yield response to fertilizer in on-farm and on-station trials were obtained. The characteristics of these varieties are summarized in Table 3 (Adu et al., 2014; Kpotor et al., 2014; Sallah et al., 2004; USAID/IFDC, 2015).

2.2. Data analysis pipeline

First, a descriptive analysis was conducted to characterize the observed yields from on-farm and on-station trials. Second, predictive yield modeling using the QUEFTS model was performed to assess how soil chemical properties and fertilizers explained the observed yield variability. Third, OLSR, MLR, SMLR based upon Akaike's Information Criterion algorithm, and RFR as predictive models were performed to quantify the individual and combined effects of the other independent variables presented in Table 1, including the fertilizers and maize varieties used. Finally, agronomic efficiency (AE) and partial factor productivity (PFP) were calculated to characterize nutrient use efficiencies in the three agroecological zones.

2.3. Models

2.3.1. QUEFTS

The QUEFTS model was implemented using the R package "Rquefts" (Hijmans et al., 2021) to calculate maize yield, with the maximum yield set to 10,000 kg ha^{-1}. To run QUEFTS, the model parameters must follow the prescription of (Janssen et al., 1990; Sattaria et al., 2014). For this, the soil-available P in P-Mehlich3 and the exchangeable K in K-NH4Ac extracted from the ISRIC maps were transformed into POlsen and KMehlich3 extractable according to the transfer functions from Sawyer and Mallarino, (1999).

2.3.2. Regression

To quantify the effect of maize variety (Table 3) on yield, knowing that the explanatory variable "variety" is a categorical variable at 18 levels (genotype), an MLR with categorical predictors of more than two classes was performed. As the regression required numerical inputs, the categorical variables were recoded into a set of binary variables using "Dummy encoding" (Hannay, 2019; Kassambara, 2018). Thus, the effect of variety on maize yield was measured in two ways. First, the overall effect of the variety on yield was quantified, and then, the effect of each variety on maize yield was measured. The R programming language "drop1" function was used to highlight the overall effect of the categorical variable variety on yield, after simulating the multivariate linear regression model between varieties and yield that detailed the effects of each variety (n-1) on maize yield. The default option of the R package used the first level (genotype) of the variety factor as a reference in the model to interpret the other genotypes concerning this level (genotype) (Kassambara, 2018). In this study, the "relevel()" function of the R programming language was used to define the Obatanpa variety as the reference genotype, so it does not appear in the summary tables of the linear regression with categorical variables (variety) where the effects of the genotypes are highlighted.

RFR models of maize yield in response to fertilizers, soil physical and chemical properties, maize varieties, and environmental factors were built, trained, and tested using the R packages “caret” (Kuhn et al., 2021), “ranger” (Wright and Ziegler, 2017) for variable importance ranking, and “AICcmodavg” (Mazerolle, 2020) for best model selections in the SMLR models. To train the RFR model, 70% of the data were used; the remaining 30% of the data were used for final RFR model testing. The partitioning of the data into training and test data was done randomly using the R “caret” package. Empirically, the estimated models were given by:

$$y(s) = f\,[\text{rain, tmax,tmin, genotype,altitude, Corg, Norg, ...,Clay}](s) + \varepsilon(s)$$

Where **y** represents maize yield at location **s** at on-farm or on-station trials that is modeled in two components: a trend function **f** and an error model **ε**. In the OLSR, MLR, SMLR, RFR, and QUEFTS models analyses, where fertilizer was considered as a covariate, all zero values of fertilizer application rate were replaced with the unit **1** to prevent an error in the models.

2.4. Nutrient use efficiency

AE and PFP were calculated for each data point from on-station and on-farm and the average was used in the function of AEZ. AE and PFP of each data point were computed following Dobermann, (2007) and Drechsel et al., (2015):

$$\mathrm{AE_X} = \frac{\sum_{i=1}^{n}\left(\frac{y_t - y_0}{F}\right)}{n}\ ,\ \mathrm{PFP_X} = \frac{\sum_{i=1}^{n}\left(\frac{y_t}{F}\right)}{n}$$

where **x** is the AEZ, **n** is the number of observed trials for which we have values for N, P, K, and NPK, y_t is the yield (kg ha^{-1}) for a fertilized data point, $\mathbf{y_0}$ is the yield (kg ha^{-1}) for the control data point, and F is the amount (kg ha^{-1}) of fertilizer (N, P, K, or NPK) applied.

2.5. Model validation

The performance of models was assessed using the coefficient of determination (R^2) and the adjusted coefficient of determination (adj. R^2) (Krause et al., 2005). These were used to evaluate deviations between model-simulated and observed yields and to quantify the explained portion of the yield response variability. Data and models were simulated, analyzed, and visualized using R (R Core Team, 2022) packages in RStudio© software (RStudio Team, 2022).

3. RESULTS

3.1. Observed grain yield values

On-farm yields ranged from 11 to 5,030 kg ha^{-1} in the control plots, while yields in fertilized plots ranged from 141 to 8,230 kg ha^{-1} (Figure 2A). For the on-station trials, yields ranged from 300 to 5,030 kg.ha^{-1} in the unfertilized plots and from 400 to 6,030 kg ha^{-1} in the fertilized plots (Figure 2B). The average yield of all on-station trials was higher than the average yield of all on-farm trials at 2,600 kg ha^{-1} and 2,001 kg ha^{-1}, respectively. The analysis of variance (ANOVA) statistics showed that there was a significant difference ($p<0.05$) in yield among on-station and on-farm trials. The average yield of all control treatments was 1,564 kg ha^{-1} for on-station trials and 1,194 kg ha^{-1} for on-farm trials. The overall average yield of the fertilized plots was 2,710 kg ha^{-1} in the on-station trials and 2,500 kg ha^{-1} in the on-farm trials. The ANOVA test showed a small difference ($p<0.1$) between the yields of the on-station and the on-farm trials. On-farm trials conducted in 2001 with the Obatanpa maize variety gave the highest yields in this study, between 6,700 and 8,230 kg ha^{-1} (Figure 2A, data point circled in red).

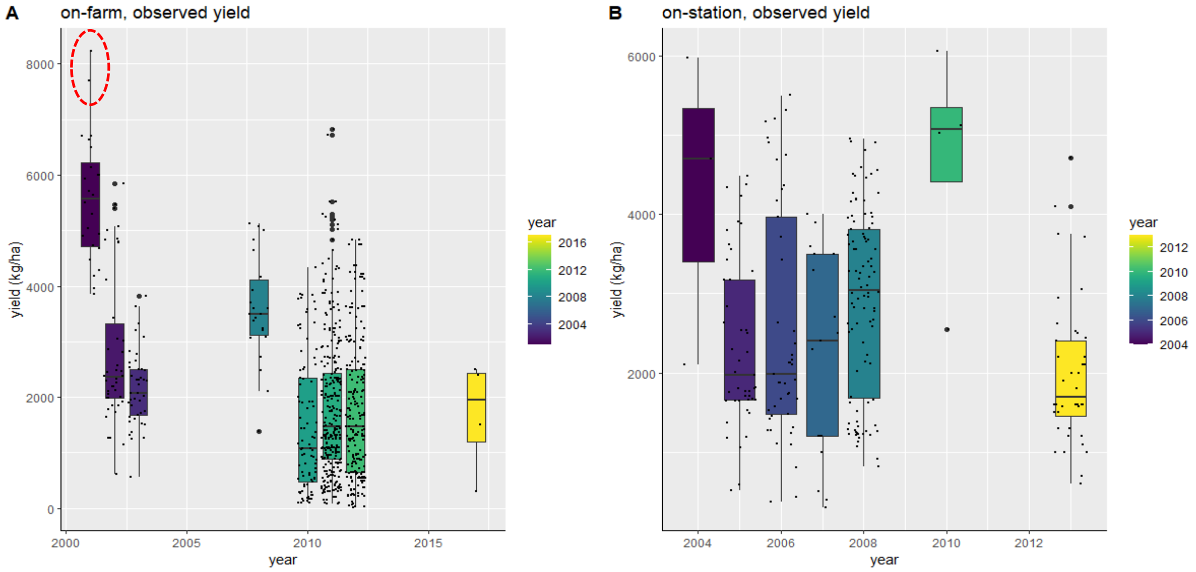

Figure 2: Interactive heatmap and box plot showing the evolution and dispersion of maize grain yields. The red dotted oval surrounds the highest on-farm yields of 2001.

3.2. QUEFTS model-estimated maize yield

The average observed yield (AOY) and average predicted yield (APY) by the QUEFTS model were 2,001 kg ha^{-1} and 2,420 kg ha^{-1}, respectively, in the on-farm trials and 2,600 kg ha^{-1} and

3,309 kg ha^{-1}, respectively, in on-station trials. For both on-farm and on-station trials, the APY was higher than the AOY.

Figure 3 shows the relationship between yield predicted by QUEFTS (Yq) and observed yield (Yo) on-station and on-farm across three scenarios: Figure 3A and 3D display the linear regression between Yq and Yo when no fertilizer was applied, Figure 3B and 3E present the linear regression between Yq and Yo when fertilizer was applied, and Figure 3C and 3D show the linear regression between Yq and Yo when no distinction was made among trials. Indeed, there were weak positive correlations between Yo and Yq in on-farm ($r=0.47$, $p<0.05$) and on-station ($r=0.17$, $p<0.05$) trials when no distinction was made between the fertilization status of the trials (Figure 3C and 3F), with an adj. R^2 of 22% and 3%, respectively. At the on-farm level, the linear regression in Figure 3D showed a weakly positive and significant correlation between Yo and Yq ($r=0.33$, $p<0.05$), with an adj. R^2 of 11%; however on-station, the linear regression showed a weakly nonsignificant negative correlation between Yo and Yq ($r=-0.23$, $p=0.27$), with a very low adj. R^2 (Figure 3A). When only the fertilized trials were considered,

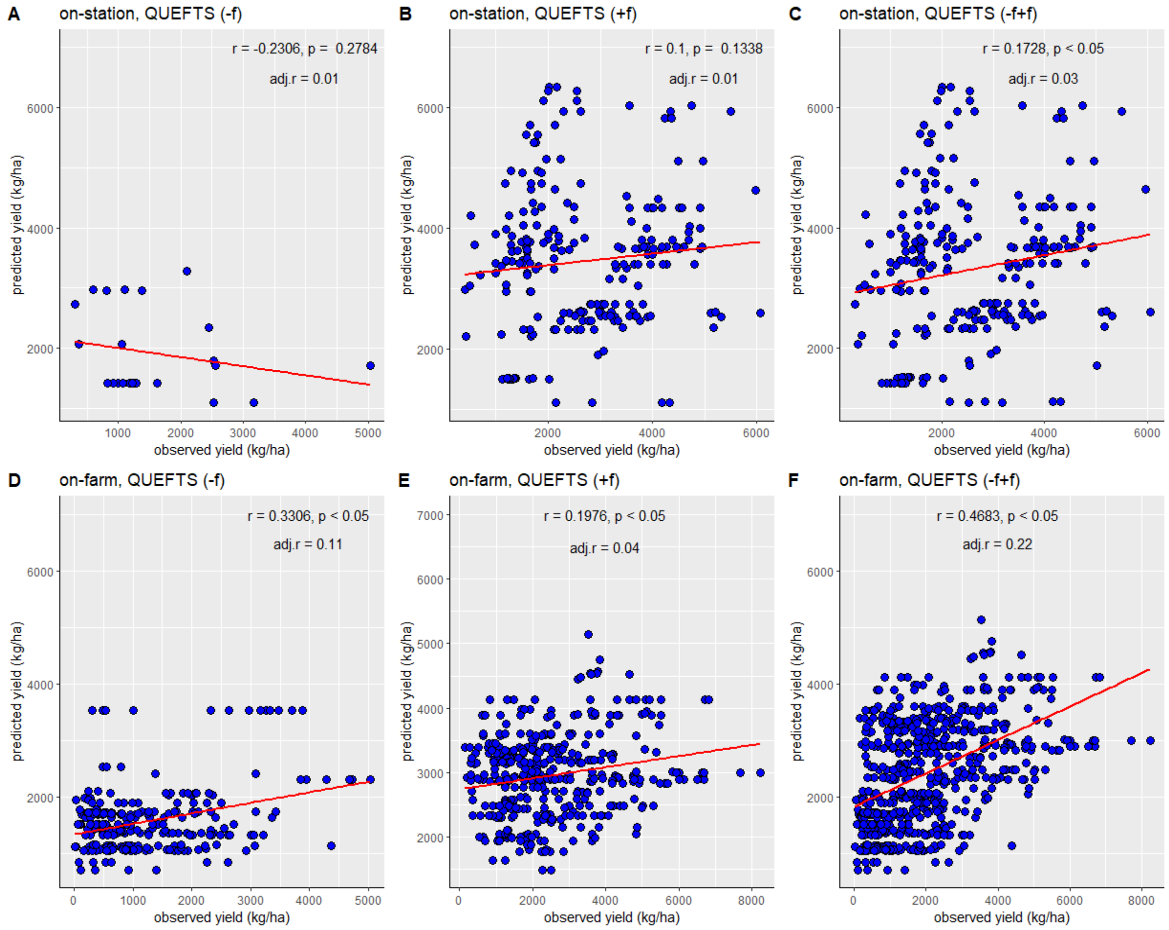

Figure 3: Relationship between observed yield and QUEFTS model-predicted maize grain yield. The red solid line represents the linear regression line and the adj.r is adjusted R^2. QUEFTS: QUEFTS model-simulated yield without considering fertilizer among the explanatory variables; QUEFTS (+f): QUEFTS model-simulated yield with fertilizer among the explanatory variables; and QUEFTS (-f+f): QUEFTS model-simulated considering all treatments.

the linear regression between Yq and Yo was weakly positive and significant on-farm, with an adj. R^2 of 4% ($r=0.19$, $p<0.05$), and weakly positive and not significant on-station, with adj. R^2 of 1% ($r=-0.1$, $p=0.13$).

Table 4 shows that linear regressions between Yq and Yo considering the years in which the trials were conducted often explained more than 30% of the variability in yield response.

Table 4: Summary of fitted adjusted R^2 values from linear regressions between observed yield and QUEFTS-predicted yield based on-station and on-farm trial years

Year	Number of observations	adj. R^2	p-value	Source
2001	24	50%	***	on-farm
2002	42	33%	***	on-farm
2003	42	4%		on-farm
2004	3	92%		on-station
2005	45	-2%		on-station
2006	45	-2%		on-station
2007	17	34%	**	on-station
2008	96	81%	***	on-station
2008	22	28%	**	on-farm
2010	89	39%	***	on-farm
2011	292	17%	***	on-farm
2012	213	32%	***	on-farm
2013	40	-1%		on-station
2017	4	98%	**	on-farm

*** and ** indicate significance at 0.1% and 1%, respectively; a blank cell indicates no significance.

In on-farm trials, the QUEFTS model explained 50%, 33%, 28%, 39%, 17%, and 32% of maize yield variability in 2001, 2002, 2008, 2010, 2011, and 2012, respectively. In on-station trials, 34% and 81% of yield variability in 2007 and 2008, respectively, were explained by the QUEFTS model. In the remaining years, the relationship between Yq and Yo were not statistically significant.

3.3. OLSR, MLR and SMLR models-estimated maize yield variability

The overall effect of the variety on maize yield was 1%, which was not significant ($p>0.05$). Table 5 shows the individual effect of each variety used in the on-station trials. The use of Mamaba and Dorke SR was significantly ($p<0.05$ and $p<0.1$, respectively) associated with an average drop in yield of 13 kg ha^{-1} and 321.30 kg ha^{-1} compared to Obatanpa, understanding that the y-intercept gives an average yield for Obatanpa of 2,713 kg ha^{-1}.

Table 5: Summary of the characteristics of the multiple linear regression model between the nominal categorical variable variety and the yield at on-station trials

	Number of observations	Estimate	Std.Error	t.value	Pr(>|t|)	
(Intercept)		2713.16	102.67	26.425	< 2e-16	***
Mamaba	8	-13.16	462.03	-0.028	0.9773	*
Dorke SR	88	-321.30	170.27	-1.887	0.0603	.

***, *, and . indicate significance at 0.1%, 5%, and 10%, respectively.

The elevation and the climatic factors each explained 5% and 1%, respectively, of the yield response, while fertilizer explained 15% of the yield variability (Table 6). When comparing groups of explanatory variables, the soil physical properties and the soil chemical properties each explained up to 32% and 36%, respectively, of the variability in yield. The combined effect of maize variety and climatic variables through model 7 is weak (adj. R^2=3%).

Adding the altitude variable to model 7 increased the adj. R^2 from 3% to 14%. The combined effect of variety, climatic factors, and altitude explained the variability in yield response at the same level as fertilizer application. The combined effect of variety, altitude, and soil chemical properties in model 9 and with the addition of soil physical properties in model 10 explained 45% of the yield response in the two models, while fertilizers combined with variety and altitude, in model 11 explained 24%.

The combined effect of soil chemical properties and climatic variables and the combined effect of soil physical properties and climatic factors explained 44% and 43%, respectively, of the variability in maize yield. However, the same combination of climatic variables and applied fertilizer in model 14 explained only 16% of the variability in maize yield. The combination of physical and chemical soil properties explained 45% of the yield response. The combination of fertilizer and soil physical properties in model 16 and fertilizer and soil chemical properties in model 17 explained 50% and 57% of the yield variability, respectively. Model 17, run with the same soil chemical properties and fertilizers used in the QUEFTS model, explained significantly more yield response variability at 57% vs. 3% with QUEFTS (Figure 3C and 3F).

The combination of all covariates in model 18 greatly explained the variability observed in yield response at 68%. Model 18 indicates that considering all covariates adds to the explanation for the variability in maize yield response to fertilizer. However, “not applicable” (NA) results were obtained for the coefficients of the sand, clay, and silt covariates. Thus, by using SMLR, including all explanatory variables, better models were obtained. Multiple linear regressions resulting from the choices of the AIC algorithm, one simulated with all explanatory variables plus fertilizer (AIC+f) and the other simulated with all covariates without fertilizer (AIC-f), showed that the model derived from AIC+f explained 68% of the observed variability, compared to 44% for the model derived from AIC-f. The model derived from AIC+f could provide a more reliable means of predicting yield, thus explaining the highest proportion of the variability in yield response in on-station trials.

Table 6: Statistics of the inferential characteristics of the explanatory variables in Table 1, variety, and fertilizer on the response of maize yield through simple and multiple linear regression models for on-station trials

Model nb[a]	Variety	Altitude	Climatic variable			Physical soil factors					Chemical soil factors							Fertilizer			Yield	Linear-Fit		p
			Rain	tmin	tmax	rzd	rzwhc	sand	clay	Silt	pH	POlsen	CEC	Ex.K	Pt	Norg	Corg	FN	FP	FK		R^2	Adj.R^2	
1	X																				+ ϵ^{e} =	1%	1%	
2		X																			+ ϵ =	5%	5%	***
3			X	X	X																+ ϵ =	2%	1%	
4						X	X	X	X	X											+ ϵ =	33%	32%	***
5											X	X	X	X	X	X	X				+ ϵ =	38%	36%	***
6																		X	X	X	+ ϵ =	16%	15%	***
7	X		X	X	X																+ ϵ =	5%	3%	
8	X	X	X	X	X																+ ϵ =	16%	14%	***
9	X	X				X	X	X	X	X											+ ϵ =	46%	45%	***
10	X	X									X	X	X	X	X	X	X				+ ϵ =	48%	45%	***
11	X	X																X	X	X	+ ϵ =	26%	24%	***
12			X	X	X	X	X	X	X	X											+ ϵ =	45%	43%	***
13			X	X	X						X	X	X	X	X	X	X				+ ϵ =	47%	44%	***
14			X	X	X													X	X	X	+ ϵ =	18%	16%	***
15						X	X	X	X	X	X	X	X	X	X	X	X				+ ϵ =	47%	45%	***
16						X	X	X	X	X								X	X	X	+ ϵ =	52%	50%	***
17[b]											X	X	X	X	X	X	X	X	X	X	+ ϵ =	58%	57%	***
18	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	+ ϵ =	70%	68%	***
SMLR - f[c]					X		X				X	X		X		X	X				+ ϵ =	46%	44%	***
SMLR + f[d]	X	X	X	X	X	X	X				X	X		X	X	X	X	X	X		+ ϵ =	70%	68%	***

FN: nitrogen fertilizer. FP: phosphorus fertilizer; FK: potassium fertilizer. *** indicates significance at 1%; a blank cell indicates no significance. nb[a]: model number. 17[b]: model number 17 uses only covariates used in the QUEFTS model. SMLR-f[c]: stepwise multilinear regression model without considering fertilizers among the covariates. SMLR+f[d]: stepwise multilinear regression model considering fertilizers among the covariates. ϵ[e]: model residual error.

The variety variable explained 16% of the variability in yield overall. This covariate explained more of the yield variability in the on-farm trials than in the on-station trials (Table 6and Table 8). The overall effect of variety is statistically significant ($p>0.05$). Table 7 shows the amount of yield variability explained by the individual effect of each variety used in the on-farm trials.

Table 7: Summary of the characteristics of the multiple linear regression model between the nominal categorical variables variety and the yield at on-farm trials

	Number of observations	Estimate	Std.Error	t.value	Pr(>|t|)	
(Intercept)		2370.2	84.77	27.961	< 2e-16	***
Mamaba	64	420.92	185.66	2.267	0.023679	*
GH 110	6	1718.13	546.08	3.146	0.001723	**
Pannar53	123	-663.41	146.23	-4.537	6.70E-06	***
Abontem	6	1494.8	546.08	2.737	0.00635	**
Aburohema	70	-751.89	179.25	-4.195	3.08E-05	***
Akposoe	6	951.46	546.08	1.742	0.081883	.
Entry 5	6	-1646.87	546.08	-3.016	2.66E-03	**
Entry 6	6	-1853.54	546.08	-3.394	0.000726	***
Entry 70	6	-1170.2	597	-1.96	5.04E-02	.
Entry 85	6	-1483.54	546.08	-2.717	0.006755	**
Etubi	34	-372.12	241.96	-1.538	0.124498	
Golden Jubilee	6	1378.13	546.08	2.524	0.011831	*
Local variety	6	259.8	546.08	0.476	0.634404	
Omankwa	64	-891.94	185.66	-4.804	1.90E-06	***
QPM	72	-1209.2	177.31	-6.82	1.95E-11	***
Wang Daata	4	-695.2	666.12	-1.044	0.296999	

***, **, *, and . indicate significance at 0.1%, 1%, 5%, and 10%, respectively; a blank cell indicates no significance.

The use of Mamaba, GH 110, Abontem, and Golden Jubilee were significantly ($p<0.05$) associated with an average yield increase of 420.92 kg ha^{-1}, 1,718.13 kg ha^{-1}, 1,494.8 kg ha^{-1}, and 1,378.13 kg ha^{-1} compared to Obatanpa, understanding that the y-intercept corresponds to an average yield of 2,370.2 kg.ha^{-1} for Obatanpa. However, the use of Pannar53, Aburohema, Entry 5, Entry 6, Entry 85, Omankwa, and QPM were significantly ($p<0.05$) associated with an average yield decrease of 663.41 kg ha^{-1}, 751.89 kg ha^{-1}, 1,646.87 kg ha^{-1}, 1,853.54 kg ha^{-1}, 1,483.54 kg ha^{-1}, 891.94 kg ha^{-1}, and 1,209.2 kg ha^{-1} compared to Obatanpa. The variety Akposoe is weakly significantly ($p<0.1$) associated with a mean yield increase of 951 46 kg ha^{-1} over Obatanpa, while the local variety and Etubi were not significantly associated with a change in the average yield. Entry 70 was weakly significantly ($p<0.1$) associated with a mean yield decrease of 1,170.2 kg ha^{-1} compared to Obatanpa. In the on-farm trials, the overall effect of the variety variable helped to further explain yield variability.

Indeed, when the variety variable is combined with altitude, climatic factors, physical and chemical factors of the soil, and even fertilizers, through models 7, 8, 9, 10, and 11, there was a significant increase in the adj. R^2 (Table 8). Physical soil factors and chemical soil factors explained less yield variability in on-farm trials than in on-station trials (Table 8). But fertilizer explained more of the variability in on-farm trials, with a 28% adj. R^2, compared to on-station. The combined effect of the climatic variable with physical soil factors in model 12, chemical soil factors in model 13, and fertilizer in model 14 explained 37%, 33%, and 47% of yield variability, respectively, in on-farm trials. The combination of physical soil factors and chemical soil factors in model 15 explained 30% of yield variability, and the combined effect of fertilizer with physical soil factors in model 16 and with chemical soil factors in model 17 explained 42% and 34 of yield variability, respectively. The difference in explanation of maize yield variability between model 17 and the QUEFTS model, considering only soil chemical properties and fertilizer, was large at 37% and 22% adj. R^2, respectively.

The implementation of the AIC algorithm through SMLR, in which all the covariates were used with fertilizers (AIC+f) and without fertilizers (AIC-f), produced the final best models that explained 62% and 46%, respectively, of the yield variability in on-farm trials.

Table 8: Statistics of the inferential characteristics of the explanatory variables in Table 1, variety, and fertilizer on the response of maize yield through simple and multiple linear regression models for on-farm trials

			Climatic variable			Physical soil factors					Chemical soil factors							Fertilizer				Linear-Fit		
Model nb[a]	Variety	Altitude	Rain	tmin	tmax	rzd	rzwhc	sand	clay	Silt	pH	POlsen	CEC	Ex.K	Pt	Norg	Corg	FN	FP	FK	Yield	R^2	Adj.R^2	p
1	X																				$+ \epsilon^{e} =$	18%	16%	***
2		X																			$+ \epsilon =$	12%	12%	***
3			X	X	X																$+ \epsilon =$	33%	33%	***
4						X	X	X	X	X											$+ \epsilon =$	25%	25%	***
5											X	X	X	X	X	X	X				$+ \epsilon =$	19%	19%	***
6																		X	X	X	$+ \epsilon =$	28%	28%	***
7	X		X	X	X																$+ \epsilon =$	44%	42%	***
8	X	X	X	X	X																$+ \epsilon =$	44%	42%	***
9	X	X				X	X	X	X	X											$+ \epsilon =$	38%	36%	***
10	X	X									X	X	X	X	X	X	X				$+ \epsilon =$	34%	32%	***
11	X	X																X	X	X	$+ \epsilon =$	45%	43%	***
12			X	X	X	X	X	X	X	X											$+ \epsilon =$	38%	37%	***
13			X	X	X						X	X	X	X	X	X	X				$+ \epsilon =$	34%	33%	***
14			X	X	X													X	X	X	$+ \epsilon =$	47%	47%	***
15						X	X	X	X	X	X	X	X	X	X	X	X				$+ \epsilon =$	31%	30%	***
16						X	X	X	X	X								X	X	X	$+ \epsilon =$	43%	42%	***
17[b]											X	X	X	X	X	X	X	X	X	X	$+ \epsilon =$	38%	37%	***
18	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	$+ \epsilon =$	64%	62%	***
SMLR - f[c]	X		X		X		X	X		X		X		X		X					$+ \epsilon =$	48%	46%	***
SMLR + f[d]	X			X	X		X	X		X		X		X		X		X	X	X	$+ \epsilon =$	63%	62%	***

FN: nitrogen fertilizer. FP: phosphorus fertilizer; FK: potassium fertilizer. ∗∗∗ indicates significance at 1%; a blank cell indicates no significance. nb[a]: model number. 17[b]: model number 17 uses only covariates used in the QUEFTS model. SMLR-f[c]: stepwise multilinear regression model without considering fertilizers among the covariates. SMLR+f[d]: stepwise multilinear regression model considering fertilizers among the covariates. ϵ[e]: model residual error.

3.4. RFR model-estimated maize yield variability

The RFR model was trained on 70% of the volume of total on-farm and on-station trials data. This allowed classification of the individual contribution of each variable in explaining the variability of maize yield. Figure 4 shows that maximum temperature, silt, organic nitrogen, and the type of fertilizer used played the most important role in explaining maize yield in on-farm trials. However, when fertilizer was added to the model variables, the most important factor in explaining yield variability was found to be nitrogen fertilizer, followed by maximum temperature, genotype, rainfall, and silt. For on-station trials, Figure 4 shows that POlsen was the most important factor explaining yield variability, far ahead of organic carbon, rain, and cation exchange capacity (CEC). However, as in on-farm trials, when fertilizer was added to the model, the most important factor in explaining yield variability was found to be nitrogen fertilizer, followed by POlsen, phosphorus fertilizer, and rain.

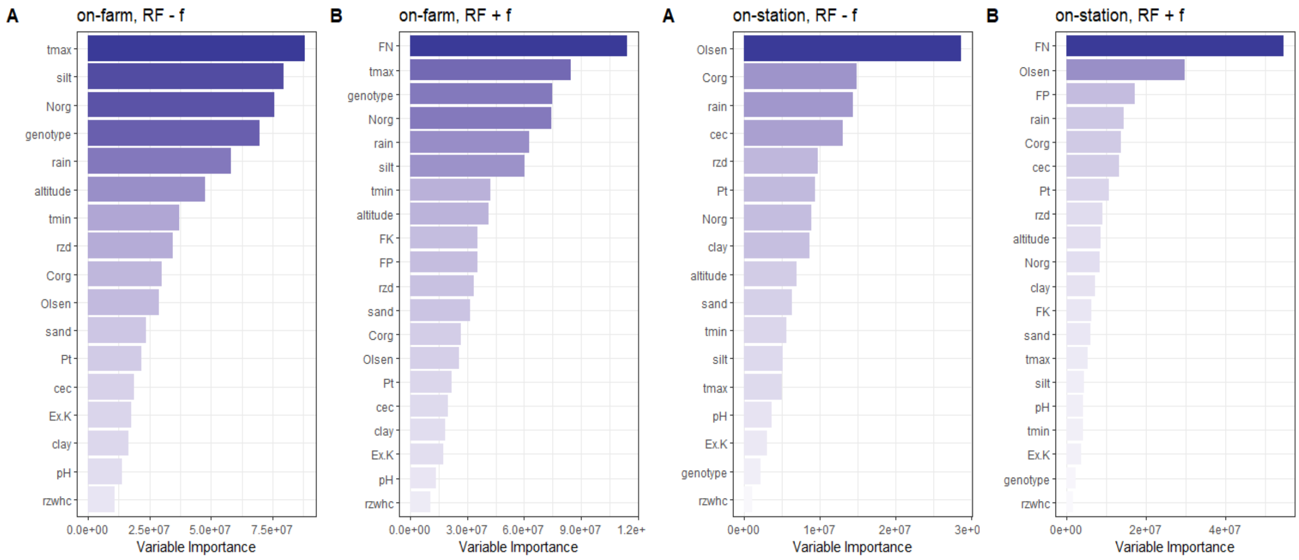

Figure 4: Ranking of the importance of factors explaining maize yield variability. RF-f: Random Forest model simulated without considering fertilizer among the explanatory variables. RF+f: Random Forest model simulated with fertilizer among the explanatory variables. Variable importance is computed using the mean decrease in Gini index and expressed relative to the maximum.

3.4.1. Predicted maize yield

The RFR model was used to explain yield variability in the on-farm and on-station trials. Of all data points, 70% were used to train the RFR model and 30% were used to test and validate the model. Table 9 shows that RFR explained more yield variability of the on-station trials than the on-farm trials.

Table 9: Summary of the random forest regression adj. R^2

	on-station data			on-farm data		
	train (n=178)	test (n=72)	p	train (n=513)	test (n=213)	p
RFR	78%	81%	***	65%	66%	***

*** indicates significance at 1%.

The adj. R^2 were 78% and 82% for the on-station trial training and testing data points and 65%

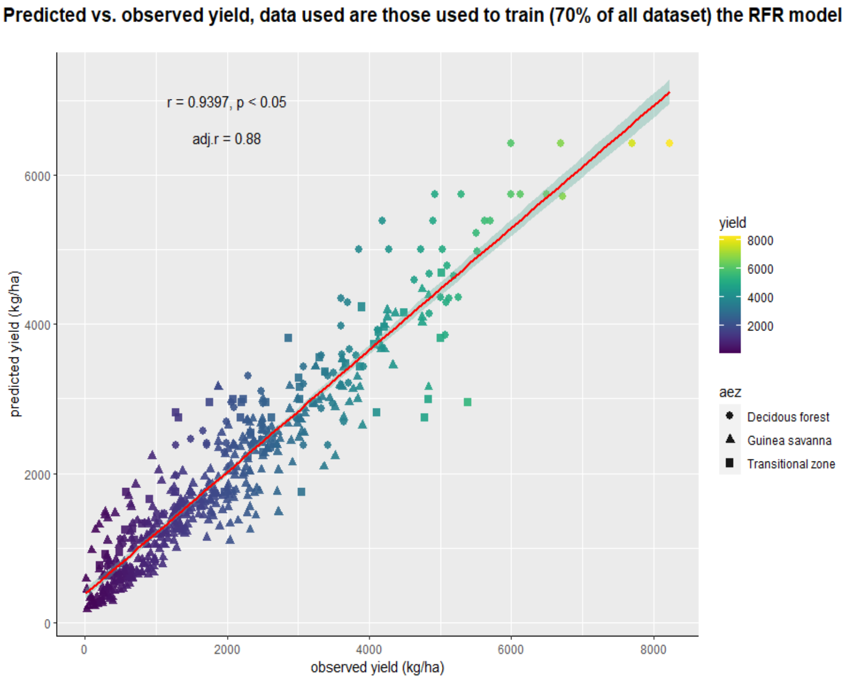

Figure 5: Relationship between observed yield (data used to train the model, representing 70% of total data) and maize grain yield predicted by the RFR model.

and 66% for on-farm trial training and testing data points, respectively. The observed trial data points and their associated predicted data were highly correlated ($p<0.001$) (Figure 5 and Figure 6). In addition, the correlations between observed and predicted yields were high,

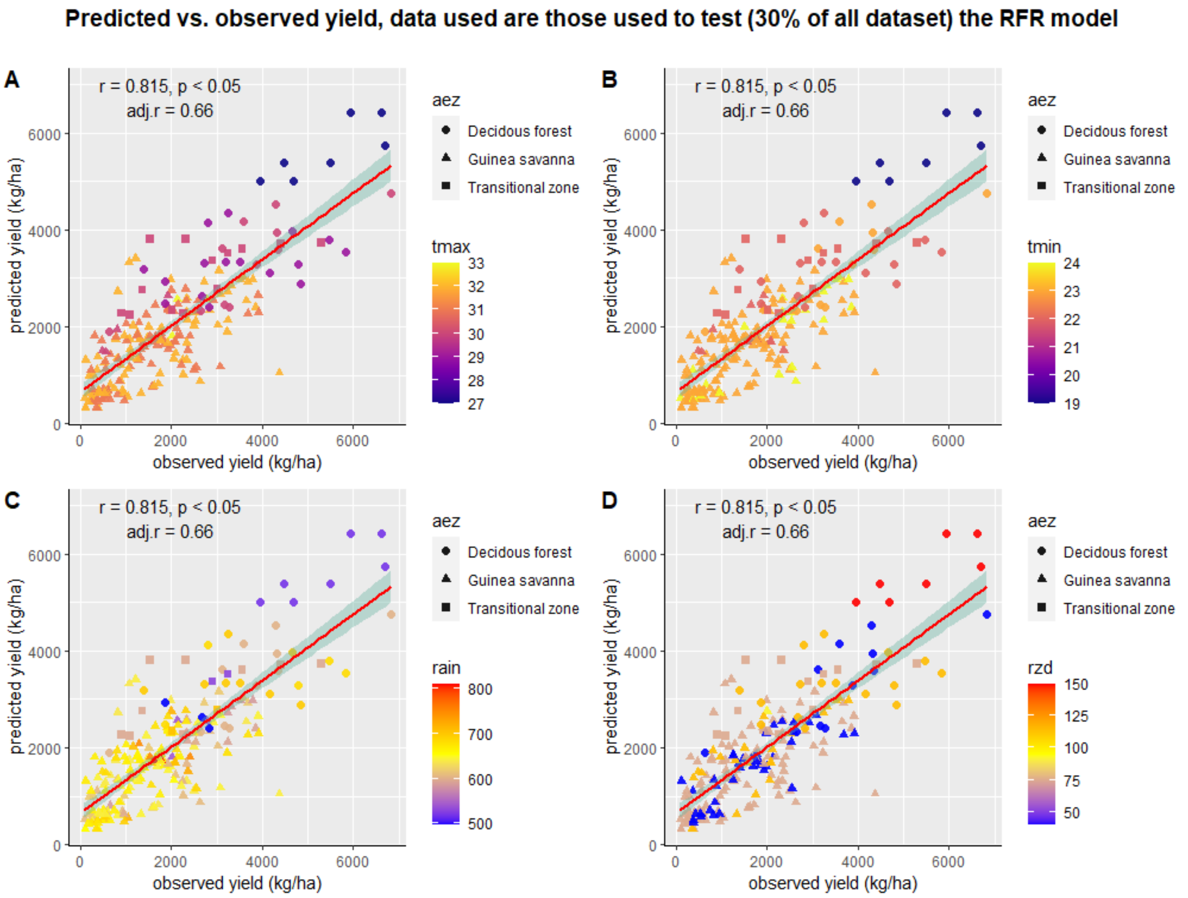

Figure 6: Relationship between observed yield (data for testing the model, representing 30% of the total data) and grain yield of maize predicted by the RFR model as a function of agroecological zone (Semi-Deciduous Forest, Guinea Savanna, and Transitional), (A) tmax, (B) tmin, (C) rainfall, and (D) rzd.

when the same training data set (70%) was used to test the RFR model; this showed a low yield gap between the observed yield (8,230 kg ha^{-1}) of the Obatanpa variety and its predicted value (6,424 kg ha^{-1}). Figure 6 reveals the strong positive correlation between observed and predicted yields in the on-farm trials and highlights the negative effects of temperature (Figure 6A and 6B) and rainfall (Figure 6C) on maize yield as these variables increase and the positive effect of root zone depth (Figure 6D) on yield as this increases. On the other hand, temperature and rainfall were found to have a more negative effect on maize yield in the Guinea Savanna than in the Transitional and Semi-Deciduous Forest AEZs. The root zone depth appeared to be shallow (<50 cm) in several data points from the Semi-Deciduous Forest AEZ but presented a yield up to 4,000 kg ha^{-1}. In addition, in the data points where root zone depth was above 125 cm, the yields were above 5,000 kg ha^{-1}.

3.4.2. Nutrient use efficiency among Agroecological zones

Partial factor productivity (PFP) and agronomic efficiency (AE) were calculated for each of the individual yield response data points from the fertilizer trials. Figure 7 shows the linear regression between fertilizer rate (nitrogen, phosphorus, and potassium) applied to on-farm and on-station trials data points with PFP (Figure 7A, 7B, and 7C) and the AE (Figure 7D, 7E, and 7F) across the three AEZs. Overall, the relationships between PFP, AE, and applied fertilizer rates were negatively correlated. Several negative values of PFP and AE were also calculated at several individual yield response data points in the fertilizer trials. Most of these negative PFP and AE values were from the on-farm trial data points in the Guinea Savanna. Most of these negative AE values were also observed when the fertilizer rate (N, P, and K) was between 20 kg ha^{-1} and 50 kg ha^{-1}. The highest values (342-379 kg kg^{-1}) of PFP when nitrogen fertilizers are applied were observed in the on-farm trials located in the Transitional zone with a fertilizer application of 7 kg ha^{-1}. The on-farm trials located in the Semi-Deciduous Forest registered the second-highest values (202-310 kg kg^{-1}). The AEs calculated at the locations in which nitrogen fertilizer was applied showed the highest values (120 kg kg^{-1}) in the on-station trials located in the Transitional zone, followed by the Semi-Deciduous Forest

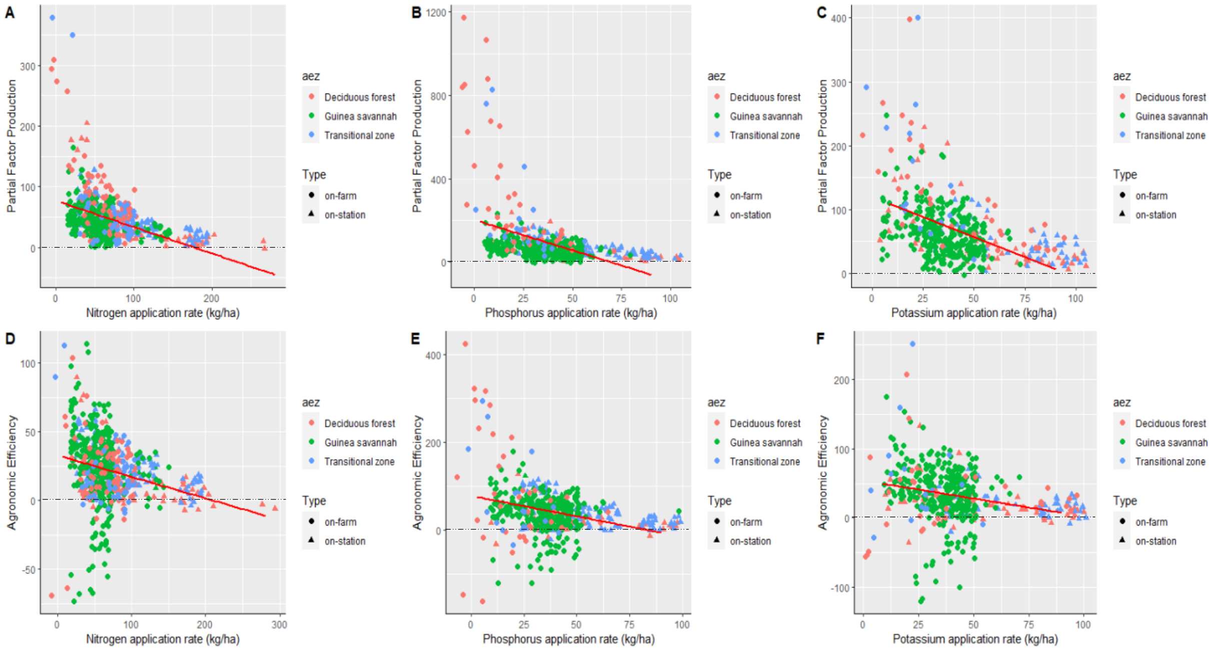

Figure 7: Relationship between nutrient use efficiencies (partial factor productivity and agronomic efficiency) of each data point of fertilizer yield response and fertilizer rate applied (kg ha^{-1}) as a function of agroecological zone (Semi-Deciduous Forest, Guinea Savanna, and Transitional) of on-farm and of on-station trials. The solid blue line represents the linear regression line.

(311 kg kg^{-1}) and the Guinea Savanna (124-161 kg kg^{-1}). When phosphorus fertilizers were applied, the highest PFP and AE values were from on-farm trials located in the Semi-Deciduous Forest, and the lowest AE values were from on-farm trials and were in the Semi-Deciduous Forest. When potassium fertilizers were applied, the highest PFP and AE values were from on-farm trials located in the Transitional zone, and the lowest PFP and AE values were from on-farm trials in the Guinea Savanna.

Table 10 shows the average rate of fertilizer use, PFP, and AE by AEZ. The average rate of fertilizer use was estimated only for trials in which fertilizer was used in each AEZ. On-farm, the ANOVA test showed that there was no significant difference in the rate of NPK used among the three AEZs. On-station, there was no significant difference in PFP among the three AEZs, and the AE of nitrogen and phosphorus was significantly different among the three AEZs. However, on-farm, the ANOVA revealed a significant difference in PFP between the three AEZs, and the AE was significantly different between the three AEZs only for phosphorus.

Table 10: Summary of average yield, average fertilizer use rate, and nutrient use efficiency in three agroecological zones

	On-station								On-farm							
	Transitional zone	Nb°	Guinea savanna	Nb°	Deciduous forest	Nb°	F-value	p	Transitional zone	Nb°	Guinea savanna	Nb°	Deciduous forest	Nb°	F-value	p
Fertilizer Use (kg ha^{-1})																
NPK (ā)	255	32	0	0	206	25	4.218	*	118	16	123	306	125	36	0.19	
N (ā)	92	126	90	6	111	61	2.727	.	70	46	56	320	65	84	11.42	***
P (ā)	53	118	33	6	56	37	3.118	*	21	16	35	307	21	36	3.12	*
K (ā)	73	48	0	0	53	61	17.54	***	25	16	35	307	43	36	17.54	***
PFP (kg kg^{-1})																
NPK (ā)	14	32	0	0	20	24	2.663		36	16	18	306	38	36	38.57	***
N (ā)	45	126	29	6	34	60	2.762	.	56	46	40	320	78	84	32.56	***
P (ā)	69	118	71	6	60	36	0.874		233	16	62	307	331	36	101.50	***
K (ā)	46	48	0	0	58	61	2.713		150	16	62	307	119	36	48.71	***
AE (kg kg^{-1})																
NPK (ā)	6	32	0	0	9	25	1.231		13	16	10	306	9	36	0.43	
N (ā)	25	126	26	6	13	61	12.37	***	18	46	23	320	20	84	1.19	

P (ā)	39	118	61	6	21	37	7.435	***	81	16	36	307	87	36	15.81	***
K (ā)	18	48	0	0	22	61	0.388		56	16	36	307	30	36	2.09	

Nb°: number of observations. (ā): average. PFP: partial factor productivity. AE: agronomic efficiency, p: P-value. ∗∗∗, and ∗ indicate significance at 0.1% and 5%; a blank cell indicates no significance.

4. DISCUSSION

4.1. Grain yield

The difference in average yield observed in on-station and on-farm trials could be attributed to various biotic and abiotic constraints (Mugwe et al., 2009; Onduru and Du Preez, 2007). On-farm trials reached the highest yields of 6,000 kg ha^{-1}, or even 8,000 kg ha^{-1} in 2001, which could be explained by the low minimum and maximum temperatures (19°-28° C) at the trial locations, sufficient precipitation during the crop growth period (540 mm), the high topographic altitude (390 m), and the deep root zone of 150 cm for storing the rainwater. However, several studies on plant breeding have shown that the agronomic potential of certain varieties, such as Obatanpa and quality protein maize (QPM), is between 4,000 and 5,000 kg ha^{-1} (Adu et al., 2014; Sallah et al., 2007; USAID/IFDC, 2015).The potential yields attributed to these varieties are yield averages from several research station trials (Adu et al., 2014; Sallah et al., 2007). Therefore, yields above these averages could be obtained in on-farm or on-station trials under optimal agronomic practices and climatic conditions. In addition, Adu-Gyamfi et al., (2019) reported yields of 6,500 kg ha^{-1} with Obatanpa, well above its potential yield, in the Guinea Savanna AEZ. Therefore, the high yields were not discarded from our analyses.

4.2. Effect of soil chemical properties on maize yield

4.2.1. QUEFTS model

The QUEFTS model weakly explained the observed variability in maize yields in on-farm and on-station trials in some locations and moderately explained high yield variability in other locations (Figure 3; Table 4). The low explanatory power of the QUEFTS model in some control data point trials (Figure 3A and 3D) suggests that factors other than the soil chemical properties considered in QUEFTS (pH, Corg, Ntot, Olsen, Pt, Ex.K) contribute significantly to the variability of maize yield (Onduru and Du Preez, 2007). In addition, when fertilizer was considered in the factors to explain maize variability, the low adj. R^2 of the linear regression (Figure 3B and 3E) shows that soil chemical properties and fertilizer are not the only factors that contribute most to the variability of maize yield. Thus, the QUEFTS model is unable to improve the understanding of native soil nutrient supply, nutrient requirements, nutrient recovery efficiency, and subsequent fertilization recommendations for maize for the wide range of soil, climate, and management conditions in Ghanaian AEZs, as also reported by Debtanu et al., (2006). Previous studies in Kenya and Benin have shown that the QUEFTS model alone was not sufficient to explain actual yields and that soil fertility was therefore not the only factor limiting crop production (Mulder, 2000; Onduru and Du Preez, 2007). However, the high variability (adj. R^2>30%) captured by QUEFTS when the linear regression between Yq and Yo is done as a function of year (Table 4) highlights that the variability of maize yield is not only spatial but also temporal. The strong spatio-temporal dynamics of soil chemical properties plays an important role in explaining the variability of maize yield by the QUEFTS model, which is based partly on empirical evidence and partly on theoretical relationships (Janssen et al., 1990). These models usually are site-specific, according to Smaling and Janssen, (1993). For example, high coefficients of variation (CV>30%, Table 2) of soil chemical properties (Corg, Ntot, Olsen, Pt, Ex.K) over time and space prevented QUEFTS from accurately capturing all the variability. In addition, climatic factors, and soil physical properties, such as temperature and soil depth, which are not considered as input variables for the model, also contribute to the weak explanation of yield variabilities.

4.2.2. Linear regression Models

Contrary to the QUEFTS, the linear regression models revealed the importance of soil chemical properties in maize production for on-station trials, as corroborated by Braimoh and Vlek, (2006). The portion of yield variability explained by soil chemical properties was about 36%. Indeed, according to Braimoh and Vlek, (2006), maize yields will be less than 520 kg ha^{-1} under repeated cropping on soils with poor chemical qualities. Yeboah et al., (2016) also showed that improving soil chemical properties, notably N, P, K, CEC, and pH, had a pronounced effect on maize grain yields. Therefore, knowledge about the condition of the land and its evolution over time is necessary to promote land management practices to maintain or improve productivity and sustainable use of natural resources (Bindraban et al., 2000).

In on-farm trials, soil chemical properties explained 19% of yield variability, while QUEFTS explained 10%. This means that soil chemical properties also played an important role in maize production in on-farm trials, but with other factors having a larger impact on yield compared to on-station trials.

For on-farm and on-station trials, MLR and SMLR, run with the same soil chemistry and fertilizer as used in QUEFTS, explained significantly more variability in yield response, i.e., 37-57% (adj. $R^2_{\text{model 17}}$) vs. 4-22% (adj. R^2_{QUEFTS}). A model such as QUEFTS has response curves and internal interactions (Janssen and Guiking, 1990; Janssen et al., 1990), whereas an empirical model (MLR, SMLR) directly explains and predicts yield without many internal processes. For these reasons, the response rates explained with the empirical models are higher than the rates with the QUEFTS model (Bonilla-Cedrez et al., 2021).

4.3. Effect of soil physical properties on maize yield

Soil physical properties explained 32% of the yield variability in on-station trials. Indeed, modifying soil structure, such as through ridging, can affect aeration, compaction, water movement, soil temperature, resistance to erosion, and plant root growth, according to Tueche, (2014). These agronomic practices directly or indirectly affect the root zone depth and root zone water-holding capacity, variables that were considered in our analysis.

In on-farm trials, soil physical properties explained 25% of yield variability. This finding reveals that soil physical properties played a major role in the observed low yield. In addition to the importance of soil physical factors revealed by MLR in on-station and on-farm trials (Table 6 and Table 8), the SMLR and RFR further explore the fact that soil physical factors, such as the percentage of silt and sand in the soil as well as the water-holding capacity of the soil as a function of soil depth, played key roles in determining current maize yield (Figure 4). The soil silt and clay content played a key role in the variation in fertilizer effects on maize grain productivity (Roobroeck et al., 2021; Sileshi et al., 2010; Sileshi et al., 2022). Mobilian and Craft, (2021) reported that soil texture affects other soil properties, such as bulk density, water-holding capacity, permeability, and porosity. For example, soils that primarily consist of sand particles have high permeability and low water-holding capacity compared to soils with higher silt and clay content. The combined effect of a soil's physical and chemical properties is very important. It explained 30% and 45% of yield variability in on-farm and on-station trials, respectively. Improving soil physical and chemical properties increased maize yield by 6.7% in sandy loam soil in the Brong-Ahafo region (Akolgo et al., 2020). Moreover, according to Hamarashid et al., (2010), soil texture and chemical soil components had a marked influence on the structure and activity of the microbial population and mineralization of carbon.

4.4. Effect of climatic factors on maize yield

In on-station trials, the low adj. R^2 (1%; $p>0.05$) suggests that there was no significant effect of environmental factors on maize yield. However, in on-farm trials, environmental covariates had a significant effect on maize yield. They even influenced maize yield variability more than soil physical and chemical properties. Indeed, the adj. R^2 of MLR (Table 8) for environmental factors is 33%, compared to 19% and 25% for those of physical and chemical soil properties, respectively. According to Acquah and Kyei, (2012) and Cudjoe et al., (2021), maize yield increases with increasing rainfall in appropriate amounts and distribution, and it decreases with increasing temperature. But the findings in this study revealed a negative correlation between rainfall and maize yields (Figure 6C). This negative trend could be explained by the low rooting depth (Figure 6D) of the soils and, therefore, the low soil storage capacity for rainwater. Figure 6C shows that the rainfall in the northern and upper west region during the growing season was high; however, Figure 6D demonstrates that the root zone depth in these regions is shallow.

4.5. Effect of altitude on maize yield

Elevation contributed 12% to the explanation of yield variability in the on-farm trials, compared to only 5% in on-station trials. Indeed, one of the on-farm trials was at a higher altitude (with the high yields), which gave a higher range in altitudes, thus resulting in the 12% explanation; however, the on-station trials did not have such a high range in elevations, so the explained portion is much less at 5%.

4.6. Effect of variety on maize yield

In on-station trials, very little overall effect (adj. R^2=1%) of variety on yield was observed (Table 5, Table 6 and Figure 4) because the number of varieties used (3) was small. However, in on-farm trials, in which many more varieties (17) were used, the combined effect of all varieties of maize explained up to 16% of the yield variability. Moreover, the RFR model ranked it as significant in predicting maize yield. The combined effect of maize variety and fertilizer explained much of the yield variability (43%;Table 8), comparable to the combined effect of maize variety and climatic factors (42%). These results show that the climatic conditions and the amount of fertilizer applied contribute significantly to the expression of the genetic potential of the maize varieties (Kpotor et al., 2014). Bawa and Tang, (2021) evaluated the response of maize varieties to different nitrogen fertilizer rates and found that the varieties responded differently to the applied nitrogen rate. The use of different maize varieties (QPM, hybrid, inbred, or local), with different genetic potential, induces a variable yield response depending on the climatic zone and whether the fertilizer is applied, at a blanket rate or a rate specific to the dynamic state of the soil. Compared to Obatanpa, some varieties contributed to an average yield increase of 54%, while other varieties contributed to an average yield decrease of 51% (Table 7). According to Kpotor et al., (2014), genotypes were significantly different for grain yield because they were developed from different parental lines, belong to different maturity groups, are season-specific, or were developed individually to meet specific breeding objectives.

4.7. Effect of fertilizer on maize yield

The increased adj. R^2 values at 37%, 42%, and 47% in MLR models (Table 8), when fertilizer is combined with soil chemical properties, soil physical properties, and climatic variables, respectively, reveal that fertilizer application has a significant effect on explaining maize yields in Ghana, as the three covariates used alone in MLRs explain 19%, 25%, and 12%, respectively. The soil physical properties combined with fertilizer revealed a significant maize yield variability explanation (adj. R^2=42%) and showed the crucial role of soil texture and structure in maize yield augmentation. This result is corroborated by Martins et al., (2018) and Zheng et al., (2003), who stressed the importance of considering soil texture in applying commercial fertilizers. Hence, the effect of fertilizers on grain yield can be improved by altering soil physical properties, since physical soil properties and fertilizer together explained more of the yield response than each did separately.

In on-station and on-farm trials, the comparison of the AIC+f vs. AIC-f and RF+f vs. RFR-f models evidenced the strong positive effect of fertilizer in explaining maize yield variability. Indeed, without fertilizer (AIC-f), the best model resulting from the SMLR explained 46% and 44% of the maize yield variability in the on-farm and on-station trials, respectively, while the inclusion of fertilizer in the AIC+f model explained 62% and 68%, respectively, with nitrogen and potassium being the most and least important components in the yield responses, respectively. Several studies have shown that the application of organic and/or inorganic fertilizer improves the yields of crops, such as maize, in Ghana (Adjei-Nsiah et al., 2007; Boateng et al., 2006; Bua et al., 2020; MacCarthy et al., 2012).

By considering all covariates in the MLR model, the adj. R^2 was 62% and 68% for the on-farm and on-station trials, respectively, which shows that the variability of maize yield is the result of the combined effect of several factors. Thus, improved crop yields are obtained when soil fertility, whether chemical and/or physical, farm management, and climatic conditions (rainfall, temperature, solar radiation) are adequate and managed synchronously (Ravensbergen et al., 2021; Sadras and Calvino, 2001).

The RFR model explains up to 60% of the maize yield variability for on-station trials and on-farm trials. An evaluation of the capacity of SMLR and RFR to quantify yield variability revealed they performed comparably well on-farm and on-station, which could be attributed to the diversity of soil physical and chemical properties, climatic factors, and maize genotype. The two models provided a more accurate method for explaining maize yield variability and predicting yield than QUEFTS.

4.8. Nutrient use efficiency

The negative correlation between PFP, AE, and applied fertilizer rates across the AEZs means that the higher the applied fertilizer rate, the less efficiently maize uses the applied fertilizer. This is due to several factors. The shallow soil depth, combined with heavy rainfall, may cause the applied fertilizer to run off and no longer be available for good plant growth (Figure 6C and 6D). In addition, the negative AE and PFP values from the on-farm trials were in the Guinea Savanna AEZ of Ghana (Figure 7) with shallow soils (<75 cm). Providing more nutrients will not make them available to the maize throughout its vegetative growth, so the application of nitrogen, phosphorus, and potassium fertilizers at 20-50 kg ha^{-1} does not give as high an NUE. The highest PFP was obtained in the on-farm trials located in the Transitional zone and the Semi-Deciduous Forest with a low fertilizer application rate of 4-8 kg ha^{-1}. In this AEZ, the deeper soil depth, low temperature, and high rainfall contribute to the efficient use of the nutrients present in the soil, but this also leads to soil nutrient depletion. However, the lowest AE values of 4-8 kg kg^{-1} from on-station trials located in the Semi-Deciduous Forest zone with the highest rate of nitrogen fertilizer applied (281 kg ha^{-1}) showed that the deeper soil and supply of more nutrients through the native soil fertility, consequently, may not give a high NUE. Similar AE and PFP can be attained at lower rates of fertilizer application, and the trials on-station use too much fertilizer.

5. CONCLUSION

Increasing fertilizer application in Ghana is very important, but site-specific application to maximize yield and minimize yield variability may be a better approach that will help to motivate farmers in applying appropriate amounts of fertilizers. Soil fertility was not the only determining factor for maize production in Ghana AEZs, as has been touted in arriving at fertilizer recommendations. Other factors identified through the MLR, SMLR, and RFR models are as important as soil nutrients and applied fertilizers. RFR and SMLR, through the AIC algorithm considering maize variety, altitude, weather variables, soil physical and chemical properties, and fertilizers, explained over 60% of the maize yield variability for both the on-station and on-farm trials, suggesting that attention should also be given to factors such as variety, temperature, and root zone depth. Therefore, fertilizer recommendations cannot be arrived at based on soil chemical properties alone but should consider these additional variables as well.

6. DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

7. REFERENCES

Abunyewa, A. A., and Mercer-Quarshie, H. (2004). Response of maize to magnesium and zinc application in the semi-arid zone of West Africa. *Asian Journal of Plant Sciences*, *3*(1), 1-5. https://doi.org/10.3923/ajps.2004.1.5

Acquah, H. De-G., and Kyei, C.K. (2012). Estimating the effect of climatic variables and crop area on maize yield in Ghana. *Journal of Social Development Sciences*, *3*(9), 313-321. https://doi.org/10.22610/jsds.v3i9.715

Adjei-Nsiah, S. (2012). Response of Maize (Zea mays L.) to Different Rates of Palm Bunch Ash Application in the Semi-deciduous Forest Agro-ecological Zone of Ghana.

Applied and Environmental Soil Science, *2012*, 870948. https://doi.org/10.1155/2012/870948

Adjei-Nsiah, S., Kuyper, T.W., Leeuwis, C., Abekoe, M.K., Cobbinah, J., Sakyi-Dawson, O., and Giller, K.E. (2007). Farmers' agronomic and social evaluation of productivity, yield and N2-fixation in different cowpea varieties and their subsequent residual N effects on a succeeding maize crop. *Nutr Cycl Agroecosyst*, *80*(3). https://doi.org/10.1007/s10705-007-9133-3

Adu-Gyamfi, R., Agyin-Birikorang, S., Tindjina, I., Ahmed, Shafik M., Twumasi, A. Duoduaa, Avornyo, V. K., and Singh, U. (2019). One-Time Fertilizer Briquettes Application for Maize Production in Savanna Agroecologies of Ghana. *Agronomy Journal*, *111*(6), 3339-3350. https://doi.org/https://doi.org/10.2134/agronj2019.04.0292

Adu, G. B., Abdulai, M. S., Alidu, H., Nustugah, S. K., Buah, S. S., Kombiok, J. M., Obeng-Antwi, K., Abudulai, M., and Etwire, P. M. (2014). *Recommended Production Practices for Maize in Ghana*. CSIR/AGRA. https://www.researchgate.net/publication/270511013Recommended_Production_Practices_for_Maize_in_Ghana

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 Journal*, *1*. https://doi.org/10.1016/j.bioeco.2021.100019

Akolgo, G.A., Kemausuor, F., Awafo, E.A., Amankwah, E., Atta-Darkwa, T., Essandoh, E.O., Bart-Plange, A., and Maia, C.M.B.de F. (2020). Biochar as a Soil Amendment Tool: Effects on Soil Properties and Yield of Maize and Cabbage in Brong-Ahafo Region, Ghana. *Open Journal of Soil Science*, *10*(03), 91-108. https://doi.org/10.4236/ojss.2020.103005

Atiah, W.A., Amekudzi, L.K., Akum, R.A., Quansah, E., Antwi-Agyei, P., and Danuor, S.K. (2021). Climate variability and impacts on maize (Zea mays) yield in Ghana, West Africa. *Quarterly Journal of the Royal Meteorological Society*. https://doi.org/10.1002/qj.4199

Badu, E., Kaba, J.S., Abunyewa, A.A., Dawoe, E.K., Agbenyega, O., and Barnes, R.V. (2019). Biochar and inorganic nitrogen fertilizer effects on maize (Zea mays L.) nitrogen use and yield in moist semi-deciduous forest zone of Ghana. *Journal of Plant Nutrition*, *42*(19), 2407-2422. https://doi.org/10.1080/01904167.2019.1659347

Bationo, A., Fening, J.O., and Kwaw, A. (2018). Assessment of Soil Fertility Status and Integrated Soil Fertility Management in Ghana. In A. Bationo, S. Youl, J.O. Fening, D. Ngaradoum, and F. Lompo (Eds.), *Improving the Profitability, Sustainability and Efficiency of Nutrients Through Site Specific Fertilizer Recommendations in West Africa Agro-Ecosystems* (Vol. 1, pp. 93-138). Springer International Publishing AG 2018. https://doi.org/10.1007/978-3-319-58789-9_7

Bawa, A., and Tang, Y. (2021). Yield and Growth Response of Maize (Zea mays L.) to Varietal and Nitrogen Application in the Guinea Savanna Agro-Ecology of Ghana. *Advances in Agriculture*, *2021*, 1-8. https://doi.org/10.1155/2021/1765251

Bindraban, P.S., Stoorvogel, J.J., Jansen, D.M. , Vlaming, J., and Groot, J.J.R. (2000). Land quality indicators for sustainable land management: proposed method for yield gap and soil nutrient balance. *Agriculture, Ecosystems and Environment*(81), 103-112. https://doi.org/10.1016/S0167-8809(00)00184-5

Boateng, S.A., Zickermann, J., and Kornahrens, M. (2006). Poultry manure effect on growth and yield of maize. *West African Journal of Applied Ecology*, *9*(1). https://doi.org/10.4314/wajae.v9i1.45682

Bonilla-Cedrez, C., Chamberlin, J., and Hijmans, R.J. (2021). Fertilizer and grain prices constrain food production in sub-Saharan Africa. *Nature Food*, *2*(10), 766-772. https://doi.org/10.1038/s43016-021-00370-1

Braimoh, A.K., and Vlek, P.L.G. (2006). Soil quality and other factors influencing maize yield in northern Ghana. *Soil use and management*, *22*(2), 165-171. https://doi.org/10.1111/j.1475-2743.2006.00032.x

Bua, S., El-Mejahed, K., MacCarthy, D., Adogoba, D.S., Kissiedu, I.N., Atakora, W.K., Fosu, M., and P.S., Bindraban. (2020). *Yield responses of maize to fertilizers in Ghana* (IFDC FERARI n°2, Issue. IFDC. https://ifdc.org/wp-content/uploads/2020/09/FERARI-Policy-Brief-2-Yield-Responses-of-Maize-to-Fertilizers-in-Ghana.pdf

Cicin-Sain, I. (2018). *Population growth and food: A systems perspective*. Retrieved 01 Jully 2021 from https://sustainablefoodtrust.org/articles/population-growth-and-food-a-systems-perspective/

Cudjoe, G.P., Antwi-Agyei, P., and Gyampoh, B.A. (2021). The Effect of Climate Variability on Maize Production in the Ejura-Sekyedumase Municipality, Ghana. *Climate*, *9*(10). https://doi.org/10.3390/cli9100145

Debtanu, M., Das, D. K., and Pathak, H. (2006). Simulation of fertilizer requirement for irrigated wheat in eastern India using the QUEFTS model. *Archives of Agronomy and Soil Science*, *52*(4), 403-418. https://doi.org/10.1100/tsw.2006.43

Diao, X., Hazell, P., Kolavalli, S., and Resnick, D. (2019). Ghana's economic and agricultural transformation: Past performance and future prospects. 277. https://doi.org/10.1093/oso/9780198845348.001.0001

Dobermann, A. (2007). *Nutrient use efficiency – measurement and management* IFA International Workshop on Fertilizer Best Management Practices, Brussels, Belgium. https://www.cabdirect.org/cabdirect/abstract/20083154896

Drechsel, P., Heffer, P., Magen, H., Mikkelsen, R., Singh, H., and Wichelns, D. (2015). Managing water and nutrients to ensure global food security, while sustaining ecosystem services. In P. Drechsel, P. Heffer, H. Magen, R. Mikkelsen, and D. Wichelns (Eds.), *Managing water and fertilizer for sustainable agricultural intensification* (First edition ed., pp. 1-7). Horgen, Switzerland: International Potash Institute (IPI); Georgia, USA: International Plant Nutrition Institute (IPNI); Colombo, Sri Lanka: International Water Management Institute (IWMI); Paris, France: International Fertilizer Industry Association (IFA). https://www.iwmi.cgiar.org/2015/03/managing-water-and-fertilizer-for-sustainable-agricultural-intensification/

EU. (2019). Global food supply and demand Consumer trends and trade challenges. *European Union (EU) Agricultural Markets Briefs*, https://ec.europa.eu/info/sites/default/files/food-farming-fisheries/farming/documents/market-brief-food-challenges-sep2019_en.pdf.

FAO/OECD. (2018). *Food Security and Nutrition: Challenges for Agriculture and the Hidden Potential of Soil*. FAO. http://www.fao.org/3/CA0917EN/ca0917en.pdf

Fening, J.O., Yeboah, E., Gyapong, T.A., and Gaizie, E. (2009). On farm evaluation of the contribution of three green manures to maize yield in the semi-deciduous forest zone of Ghana. *African Journal of Environmental Science Technology*, *3*(9). https://www.ajol.info/index.php/ajest/article/view/46071

Fick, S.E., and Hijmans, R.J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302-4315. https://doi.org/10.1002/joc.5086

Genuer, R. , and Poggi, J.-M. (2020). *Random Forests with R*. Springer. https://doi.org/https://doi.org/10.1007/978-3-030-56485-8

Godfray, H.C.J., Beddington, J.R., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., Pretty, J., Robinson, S., Thomas, S.M., and Toulmin, C. (2010). Food security: the challenge of feeding 9 billion people. *Sci. (80-.). 327*(5967), 812-818. https://doi.org/10.1126/science.1185383

Guilpart, N., Grassini, P., van Wart, J., Yang, H., van Ittersum, M.K., van Bussel, L.G.J., Wolf, J., Claessens, L., Leenaars, J.G.B., and Cassman, K.G. (2017). Rooting for food security in Sub-Saharan Africa. *Environmental Research Letters*, *12*(11), 1-8. https://doi.org/10.1088/1748-9326/aa9003

GYGA. (2021). *Agriculture in Ghana*. University of Nebraska-Lincoln and Wageningen University & Research. Retrieved 24th March 2021 from https://www.yieldgap.org/ghana

Hamarashid, N. H., Othman, M. A., and Hussain, Mohamed-Amin H. (2010). Effects of soil texture on chemical compositions, microbial populations and carbon mineralization in soil. *The Egyptian Society of Experimental Biology*, *6*(1), 59-64. https://www.bibliomed.org/?mno=187211

Hannay, K. (2019). *Introduction to Statistics and Data Science, Chapter 12: Regression with Categorical Variables*. https://faculty.nps.edu/rbassett/_book/regression-with-categorical-variables.html

Hengl, T., Heuvelink, G.B., Kempen, B., Leenaars, J.G., Walsh, M.G., Shepherd, K.D., Sila, A., MacMillan, R.A., Mendes de Jesus, J., Tamene, L., and Tondoh, J.E. (2015). Mapping Soil Properties of Africa at 250 m Resolution: Random Forests Significantly Improve

Current Predictions. *PLoS One*, *10*(6), 1-26. https://doi.org/10.1371/journal.pone.0125814

Hengl, T., Leenaars, J.G.B., Shepherd, K.D., Walsh, M.G., Heuvelink, G.B.M., Mamo, T., Tilahun, H., Berkhout, E., Cooper, M., Fegraus, E., and Wheeler, I. (2017). Soil nutrient maps of Sub-Saharan Africa: assessment of soil nutrient content at 250 m spatial resolution using machine learning. *Nutr. Cycl. Agroecosystems*, 1-26. https://doi.org/10.1007/s10705-017-9870-x

Hijmans, R. J., Pypers, P., and Wolff, J. (2021). Rquefts: Quantitative Evaluation of the Native Fertility of Tropical Soils. *R package version 1.2-1 (2021)*. https://cran.r-project.org/web/packages/Rquefts/index.html

ISRIC. *Africa SoilGrids nutrients*. ISRIC. Retrieved 12 Apr. 2021 from https://data.isric.org/geonetwork/srv/eng/catalog.search#/search

Janssen, B. H., and Guiking, F. C. T. (1990). Modelling the response of crops to fertilizers. In M. L. van Beusichem (Ed.), *Plant nutrition - physiology and application* (pp. 699-703). Kluwer Academic Publishers.

Janssen, B.H., Guiking, F.C.T., van der Eijk, D., Smaling, E.M.A., Wolf, J., and van Reuler, H. . (1990). A system for quantitative evaluation of the fertility of tropical soils (QUEFTS). *Geoderma*, *46*, 299-318. https://doi.org/10.1016/0016-7061(90)90021-Z

Jeong, J.H., Resop, J.P., Mueller, N.D., Fleisher, D.H., Yun, K., Butler, E.E., Timlin, D.J., Shim, K.M., Gerber, J.S., Reddy, V.R., and Kim, S.H. (2016). Random Forests for Global and Regional Crop Yield Predictions. *PLoS One*, *11*(6), e0156571. https://doi.org/10.1371/journal.pone.0156571

Kanton, R.A.L., Prasad, P.V.V., Mohammed, A.M., Bidzakin, J.K., Ansoba, E.Y., Asungre, P.A., Lamini, S., Mahama, G., Kusi, F., and Sugri, I. (2016). Organic and Inorganic Fertilizer Effects on the Growth and Yield of Maize in a Dry Agro-Ecology in Northern Ghana. *Journal of Crop Improvement*, *30*(1), 1-16. https://doi.org/10.1080/15427528.2015.1085939

Kassambara, A. (2018). *Machine learning essentials: Practical guide in R*. 1/Ed. ISBN: 1986406857, https://www.datanovia.com/en/product/machine-learning-essentials-practical-guide-in-r/?url=/5-bookadvisor/54-machine-learning-essentials/

Kihara, J., Nziguheba, G., Zingore, S., Coulibaly, A., Esilaba, A., Kabambe, V., Njoroge, S., Palm, C., and Huising, J. (2016). Understanding variability in crop response to fertilizer and amendments in sub-Saharan Africa. *Agric. Ecosyst. Environ.*, *229*, 1-12. https://doi.org/10.1016/j.agee.2016.05.012

Kpotor, P., Akromah, R., Ewool, M.B., Kena, A.W., Owusu-Adjei, E., and Tuffour, H.O. (2014). Assessment of the Relative Yielding Abilities and Stability of Maize (Zea mays L.) Genotypes under Different Levels of Nitrogen Fertilization across Two Agro-Ecological Zones in Ghana. *International Journal of Scientific Research in Agricultural Sciences*, *1*(7), 128-141. https://doi.org/10.12983/ijsras-2014-p0128-0141

Krause, P., Boyle, D.P., and Bäse, F. (2005). Comparison of different efficiency criteria for hydrological model assessment. *Adv. Geosci.*(5), 89-97. https://doi.org/10.5194/adgeo-5-89-2005

Kuhn, M., Jed, W., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., and Kenkel, B. (2021). Classification and Regression Training. https://github.com/topepo/caret/

Kyei-Mensah, C., Kyerematen, R., and Adu-Acheampong, S. (2019). Impact of Rainfall Variability on Crop Production within the Worobong Ecological Area of Fanteakwa District, Ghana. *Advances in Agriculture*, *2019*, 1-7. https://doi.org/10.1155/2019/7930127

Lamos-Díaz, H., Puentes-Garzón, D.E., and Zarate-Caicedo, D.A. (2020). Comparison Between Machine Learning Models for Yield Forecast in Cocoa Crops in Santander, Colombia. *Revista Facultad de Ingeniería*, *29*(54), e10853-e10853. https://doi.org/10.19053/01211129.v29.n54.2020.10853

Leenaars, J.G.B., Claessens, L., Heuvelink, G.B.M., Hengl, T., Ruiperez Gonzalez, M., van Bussel, L.G.J., Guilpart, N., Yang, H., 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

MacCarthy, D.S., Vlek, P.L.G., and Fosu-Mensah, B.Y. (2012). The Response of Maize to N Fertilization in a Sub-Humid Region of Ghana: Understanding the Process Using a Crop Simulation Model. In J. Kihara, D. Fatondji, J. W. Jones, G. Hoogenboom, R. Tabo, and A. Bationo (Eds.), *Improving Soil Fertility Recommendations in Africa using* *the Decision Support System for Agrotechnology Transfer (DSSAT)*. Springer Science + Business Media. https://doi.org/10.1007/978-94-007-2960-5_5

Martins, J.D.L., Soratto, R.P., Fernades, A., and Dias, P.H. (2018). Phosphorus fertilization and soil texture affect potato yield. *Revista Caatinga*, *31*(3), 541-550. https://doi.org/10.1590/1983-21252018v31n302rc

Mazerolle, M. J. (2020). *Model selection and multimodel inference using the AICcmodavg package*. https://cran.r-project.org/web/packages/AICcmodavg/vignettes/AICcmodavg.pdf

Mobilian, C., and Craft, C.B. (2021). Wetland Soils: Physical and Chemical Properties and Biogeochemical Processes. *Reference Module in Earth Systems and Environmental Sciences, Elsevier*. https://doi.org/10.1016/B978-0-12-819166-8.00049-9

Mueller, N.D., Gerber, J.S., Johnston, M., Ray, D.K., Ramankutty, N., and Foley, J.A. (2012). Closing yield gaps through nutrient and water management. *Nature*, *490*(7419), 254-257. https://doi.org/10.1038/nature11420

Mugwe, J., Mugendi, D., Kungu, J., and Muna, M.M. (2009). Maize Yields Response to Application of Organic and Inorganic Input under on-Station and on-Farm Experiments in Central Kenya. *Experimental Agriculture*, *45*(1), 47-59. https://doi.org/10.1017/S0014479708007084

Mulder, I. (2000). *Soil Fertility: QUEFTS and Farmer's Perceptions* (Vol. Working Paper No 30). IIED. https://www.researchgate.net/publication/228856865_Soil_fertility_QUEFTS_and_farmers%27_perceptions

Nevavuori, P., Narra, N., Linna, P., and Lipping, T. (2020). Crop Yield Prediction Using Multitemporal UAV Data and Spatio-Temporal Deep Learning Models. *Remote Sensing*, *12*(23), 4000. https://www.mdpi.com/2072-4292/12/23/4000

Njoroge, K.S. (2019). *Explaining variability in maize yield responses to nutrient applications in smallholder farms of western Kenya* (Publication Number 978-94-6395-164-7) Wageningen University & Research]. The Netherlands. https://doi.org/10.18174/503185

Onduru, D.D., and Du Preez, C.C. (2007). Spatial and temporal aspects of agricultural sustainability in the semi-arid tropics: a case study in Mbeere district, Eastern Kenya. *Tropical Science*, *47*(3), 134-148. https://doi.org/10.1002/ts.207

Paudel, D., Boogaard, H., de Wit, A., Janssen, S., Osinga, S., Pylianidis, C., and Athanasiadis, I.N. (2021). Machine learning for large-scale crop yield forecasting. *Agricultural Systems*, *187*. https://doi.org/10.1016/j.agsy.2020.103016

R Core Team. (2022). R: A language and environment for statistical computing. e. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rahman, A.N., Larbi, A., Kotu, B., Asante, M.O., Akakpo, D.B., Mellon-Bedi, S., and Hoeschle-Zeledon, I. (2021). Maize–legume strip cropping effect on productivity, income, and income risk of farmers in northern Ghana. *Agronomy Journal*, 1-12. https://doi.org/10.1002/agj2.20536

Ravensbergen, A.P.P., Chamberlin, J., Craufurd, P., Shehu, B.M., and Hijbeek, R. (2021). Adapting the QUEFTS model to predict attainable yields when training data are characterized by imperfect management. *Field Crops Research*, *266*. https://doi.org/10.1016/j.fcr.2021.108126

Ren, T., Zou, J., Wang, Y., Li, X. K., Cong, R.H., and Lu, J.W. (2015). Estimating nutrient requirements for winter oilseed rape based on QUEFTS analysis. *Journal of Agricultural Science*, *154*(3), 425-437. https://doi.org/10.1017/S0021859615000301

Roobroeck, D., Palm, C.A., Nziguheba, G., Weil, R., and Vanlauwe, B. (2021). Assessing and understanding non-responsiveness of maize and soybean to fertilizer applications in African smallholder farms. *Agric Ecosyst Environ*, *305*, 107165. https://doi.org/10.1016/j.agee.2020.107165

RStudio Team. (2022). RStudio: Integrated Development for R. . *RStudio (2021.09.2), PBC, Boston, MA URL http://www.rstudio.com/.*

Sadras, V.O., and Calvino, P.A. (2001). Quantification of grain yield response to soil depth in soybean, maize, sunflower, and wheat. *Agronomy Journal*, *93*(3), 577-583. https://doi.org/10.2134/agronj2001.933577x

Sallah, P.Y.K., Abdulai, M.S., and Obeng-Antwi, K. (2004). Genotype X environment interactions in three maturity groups of maize cultivars. *African Crop Science Journal*, *12*(2), 95-104. https://doi.org/https://doi.org/10.4314/acsj.v12i2.27667

Sallah, P.Y.K., Twumasi-Afriyie, S., Ahenkora, K., Asiedu, E.A., Obeng-Antwi, E., Osei-Yeboah, S., Frimpong-Manso, P.P., Ankomah, A., and Dzah, B.D. (2007). Agronomic potentials of quality protein maize hybrids developed in Ghana. *Journal of*

Agricultural Science and Applications, *40*(1), 81-89. https://doi.org/10.4314/gjas.v40i1.2157

Sattaria, S.Z., van Ittersuma, M.K. , Bouwmanb, A.F., Smitd, A.L., and Janssena, B.H. (2014). Crop yield response to soil fertility and N, P, K inputs in different environments: Testing and improving the QUEFTS model. *Field Crops Research*, *157*, 35-46. https://doi.org/10.1016/j.fcr.2013.12.005

Sawyer, J.E., and Mallarino, A.P. (1999). Differentiating and understanding the Mehlich 3, Bray, and Olsen soil phosphorus tests. *Agronomy Conference Proceedings and Presentations*, *12*. https://dr.lib.iastate.edu/handle/20.500.12876/4349

Sileshi, G.W., Akinnifesi, F.K., Debusho, L.K., Beedy, T., Ajayi, O.C., and Mong'omba, S. (2010). Variation in maize yield gaps with plant nutrient inputs, soil type and climate across sub-Saharan Africa. *Field Crops Research*, *116*(1-2), 1-13. https://doi.org/10.1016/j.fcr.2009.11.014

Sileshi, G.W., Kihara, J., Tamene, L., Vanlauwe, B., Phiri, E., and Jama, B. (2022). Unravelling causes of poor crop response to applied N and P fertilizers on African soils. *Experimental Agriculture*, 1-17. https://doi.org/10.1017/S0014479721000247

Smaling, E.M.A., and Janssen, B.H. (1993). Calibration of QUEFTS, a model predicting nutrient uptake and yields from chemical soil fertility indices. *Geoderma*, *59*(1-4), 21-44. https://doi.org/10.1016/0016-7061(93)90060-X

Tabi, F.O., Diels, J., Ogunkunle, A.O., Iwuafor, E.N.O., Vanlauwe, B., and Sanginga, N. (2007). Potential nutrient supply, nutrient utilization efficiencies, fertilizer recovery rates and maize yield in northern Nigeria. *Nutrient Cycling in Agroecosystems*, *80*(2), 161-172. https://doi.org/10.1007/s10705-007-9129-z

Tetteh, F., Larbi, A., Nketia, K.A., Senayah, J.K., Hoeschle-Zeledon, I., and Abdulrahman, N. (2016). *Suitability of soils for cereal cropping in Northern Ghana*. International Institute of Tropical Agriculture. https://www.researchgate.net/publication/335581532_Suitability_of_soils_for_cereal_cropping_in_northern_Ghana/link/5d6e6bbf299bf16522f2b9b6/download

Tilman, D., Balzer, C., Hill, J., and Befort, B.L. (2011). Global food demand and the sustainable intensification of agriculture. *Proc Natl Acad Sci U S A*, *108*(50), 20260-20264. https://doi.org/10.1073/pnas.1116437108

Timsina, J., Dutta, S., Devkota, K.P., Chakraborty, S., Neupane, R.K., Bishta, S., Amgain, L.P., Singh, V.K., Islam, S., and Majumdar, K. (2021). Improved nutrient management in cereals using Nutrient Expert and machine learning tools: Productivity, profitability and nutrient use efficiency. *Agricultural Systems*, *192*. https://doi.org/10.1016/j.agsy.2021.103181

Tittonell, P., Corbeels, M., van Wijk, M. T., Vanlauwe, B., and Giller, K. E. (2008). Combining Organic and Mineral Fertilizers for Integrated Soil Fertility Management in Smallholder Farming Systems of Kenya: Explorations Using the Crop-Soil Model FIELD. *Agronomy Journal*, *100*(5), 1511-1526. https://doi.org/10.2134/agronj2007.0355

Tueche, Jacques Roberto. (2014). *Relationship between soil physical properties and crop yields in different cropping systems in southern Cameroon* University of Hohenheim]. http://opus.uni-hohenheim.de/volltexte/2014/1000/pdf/Thesis_JRTueche_06_08_2014_for_print_final_good2.pdf

USAID. (2022). *Agriculture and food security*. USAID. Retrieved 18/02/2022 from https://www.usaid.gov/ghana/agriculture-and-food-security

USAID/IFDC. (2015). Seed Guide-Recommended Commercial Maize, Rice and Soybean Varieties Available for Northern Ghana. In United States Agency for International Development (Ed.), (pp. 28): Feed the Future Ghana Agriculture Technology Transfer Project.

van Loon, M.P., Adjei-Nsiah, S., Descheemaeker, K., Akotsen-Mensah, C., van Dijk, M., Morley, T., van Ittersum, M.K., 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

Wallach, D., Makowski, D., Jones, J. W., and Brun, F. (2018). Working with dynamic crop models. In: Academic Press-Elsevier.

Wijayanto, Y., and Prastyanto, E. (2012). A study of using QUEFTS model for establishing site specific fertilizer recommendation in maize on the basis of farmer fields. *AGRIVITA, Journal of Agricultural Science*, *33*(3), 273-278. https://media.neliti.com/media/publications/53171-EN-a-study-of-using-quefts-model-for-establ.pdf

Wright, M.N., and Ziegler, A. (2017). ranger: A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software*, *77*(1), 1-17. https://doi.org/10.18637/jss.v077.i01

Xu, X., He, P., Pampolino, M. F., Chuan, L., Johnston, A. M., Qiu, S., Zhao, S., and Zhou, W. (2013). Nutrient requirements for maize in China based on QUEFTS analysis. *Field Crops Research*, *150*, 115-125. https://doi.org/10.1016/j.fcr.2013.06.006

Yeboah, E., Asamoah, G., Kofi, B., and Abunyewa, A.A. (2016). Effect of Biochar Type and Rate of Application on Maize Yield Indices and Water Use Efficiency on an Ultisol in Ghana. *Energy Procedia*, *93*, 14-18. https://doi.org/10.1016/j.egypro.2016.07.143

Zheng, Z., Parent, L.E., and MacLeod, J.A. (2003). Influence of soil texture on fertilizer and soil phosphorus transformations in Gleysolic soils. *Canadian journal of soil science*, *83*(4), 395-403. https://doi.org/10.4141/S02-073

Zingore, S., Mutegi, J., Agesa, B., Tamene, L., and Kihara, J. (2015). Soil degradation in sub-Saharan Africa and crop production options for soil rehabilitation. *Better Crop.*, *99*(1), 24-26. https://cgspace.cgiar.org/handle/10568/68702