Received: 27 November 2018 | Revised: 19 September 2019 | Accepted: 13 October 2019

DOI: 10.1002/fes3.189

ORIGINAL RESEARCH

Food and Energy Security | WILEY

Soil and management-related factors contributing to maize yield gaps in western Kenya

Sussy Munialo[1,2,3] | A. Sigrun Dahlin[3] | Cecilia Onyango M.[1] | W. Oluoch-Kosura[4] | Håkan Marstorp[3] | Ingrid Öborn[2,5]

[1]Department of Plant Science and Crop Protection, University of Nairobi, Nairobi, Kenya

[2]World Agroforestry Centre, Nairobi, Kenya

[3]Department of Soil and Environment, Swedish University of Agricultural Sciences, Uppsala, Sweden

[4]Department of Agricultural Economics, University of Nairobi, Nairobi, Kenya

[5]Department of Crop Production Ecology, Swedish University of Agricultural Sciences, Uppsala, Sweden

Correspondence
Sussy Munialo, Department of Plant Science and Crop Protection, University of Nairobi, P.O. Box 29053, Nairobi, Kenya.
Email: munialos@yahoo.co.uk

Funding information
Swedish Research Council Formas, Grant/Award Number: 220-2014-646

Abstract

The solution to reducing existing yield gaps on smallholder farms lies in understanding factors limiting yield in areas with agricultural intensification potential. This study applied an integrated analysis approach comprising Classification and Regression Tree (CART), generalized linear mixed model (GLMM), and factor analysis (FA), to explain soil and management-related factors influencing maize yield gaps, in order to enhance yields. The study was conducted in Mukuyu and Shikomoli in western Kenya, sites with, respectively, high and low agroecological potential regarding soil fertility. Maize yield gaps were quantified by comparing yields on the 90th percentile of farms to yields determined in 189 fields on 70 randomly sampled smallholdings. Soil and management-related factors were determined at early and late maize development stages.

Maize yield on the 90th percentile of farms in Mukuyu and Shikomoli was 5.1 and 4.8 t/ha, respectively, and the average yield gap was 1.8 and 2.6 t/ha, representing 35% and 54% unachieved yield for Mukuyu and Shikomoli, respectively. In FA, soil was revealed to be the main factor influencing maize yield gaps at both sites, rather than management-related variables. The CART method identified maize density, chlorophyll values, maize height, and depth to compact layer as consistent factors affecting yield at both sites, while GLMM identified soil texture (silt content) as important. According to CART, weed cover at early stages and maize density at late stages were the most limiting factor in maize production in Mukuyu and Shikomoli, respectively. Generalized linear mixed model analysis identified agroecology-specific factors influencing maize yield gaps as soil-available phosphorus and zinc, plus weed pressure at early maize stages in Mukuyu, and plus soil cation exchange capacity and exchangeable magnesium in Shikomoli. Through an integrated approach, it was possible to identify both consistent and agroecology-specific factors limiting crop yields. This can increase the applicability of the findings to smallholder farms.

KEYWORDS

critical yield periods, integrated approach, intensification potential, management, soil, yield gap

1 | INTRODUCTION

Despite decades of investment in new agricultural technologies, yield gaps (the difference between potential and achieved yield) for major staple food crops such as maize persist globally. The problem is particularly acute in areas of the world dependent upon rainfed agriculture and especially in sub-Saharan Africa, where agricultural productivity has stagnated (Ray, Mueller, West, & Foley, 2013). As a result, maize yield gaps exceeding 90% have been reported (Ray et al., 2013). In Kenya, maize yield is as low as 1.8 t/ha on smallholder farmers, compared with potential yield of 6.0 t/ha, resulting in yield gaps greater than 50%. This makes many households food-insecure (Oloo, Ranabhat, & Gemenet, 2013). Socioeconomic, biophysical, soil, and management-related factors causing yield gaps have been identified (van Ittersum et al., 2013). However, there is a growing need to advance understanding of soil and management-related factors directly influencing maize yield (Beza, Silva, Kooistra, & Reidsma, 2016; Cassman, Dobermann, Walters, & Yang, 2003).

Soil factors are vital for growth and development of maize during critical yield determination stages (Fischer, Byerlee, & Edmeades, 2014). Soil nutrient uptake and accumulation in maize plants starts after seedling emergence and increases as the plants advance through the critical yield determination stages, that is, ear initiation, ear determination, silking, and tasseling (O'Keeffe, 2009). Soil physical properties such as clay, silt, and sand content affect soil organic matter content and are important for water retention and availability, soil workability, soil trafficability, and nutrient supply to plants (Sherpherd, 2010). Field measurements of chlorophyll content, crop height, and crop vigor are good indicators of crop status, and can be related to soil properties to identify within-field management practices limiting yield (Duncan, Dash, & Atkinson, 2015). Weed cover and weed height are indicators of field nutrient status, water availability, and radiation use efficiency, and can be related to weed management measures (Reid et al., 2014; Sherpherd, 2010). Soil factors influencing maize yield gaps have been widely studied (Affholder, Poeydebat, Corbeels, Scopel, & Tittonell, 2013; Fermont, Asten, Tittonell, Wijk, & Giller, 2009; Okumu et al., 2011). However, only a few studies have examined the effect on crop yield of, for example, rooting depth (depth to compact layer) (van Bussel et al., 2015), which is important as it determines availability of water and nutrients (Sherpherd, 2010). Furthermore, the links between soil factors and crop status indicators such as chlorophyll content, weed height, crop vigor, and weed pressure as within-season constraints to crop yield have not been fully explored.

Many soil and management-related factors associated with maize yield gaps on smallholder farms are affected by high spatial variability in agroecological and economic conditions (Sultan, Baron, Dingkuhn, Sarr, & Janicot, 2005). It is therefore important to design site-specific soil and crop management measures, which recognize that farmers operate under diverse soil and climate conditions and are resource-constrained (Banerjee et al., 2014; van Ittersum et al., 2013). However, most existing studies on factors influencing yield gaps have been performed at global, regional, or national level and the findings are general, making it challenging to devise site-specific crop and management measures applicable on smallholder farms (van Bussel et al., 2015). A better approach is to estimate yield gaps and understand the causes at local level, before scaling the results to larger spatial areas (van Ittersum et al., 2013). This will result in crop and soil measures that are applicable on both small and large spatial scales.

Yield gaps can be quantified using simulated or measured potential yield as reference. Simulated potential yield provides a better estimate of maize yield gaps because it accounts for genotype, environment, and management interactions, unlike estimates derived from experimental stations and farmers' fields (van Ittersum et al., 2013). However, unavailability of the data needed to effectively calibrate simulation models for smallholder farming conditions limits their use. Yield gaps can also be estimated by comparing yield on the 90th percentile of farms, representing the genotype–environment interactions prevailing in smallholder production, with that on other farms within the same area, to capture local variability (Lobell, Cassman, Field, & Field, 2009). This is a more accurate approach taking consideration of smallholder farming conditions than basing estimates on optimal yield data from experimental stations (van Ittersum et al., 2013).

Linear regression and correlation methods have been widely used in yield gap studies to show specific factors influencing crop yields (Krupnik et al., 2015; Mackay et al., 2011; Neumann, Verburg, Stehfest, & Muller, 2010; Sawasawa, 2003). However, the heterogeneity in smallholder farms is likely to result in high spatial variability in yield gaps and their causes. To obtain a spatial view of the causes of yield gaps on smallholder farms, multivariate statistics such as Classification and Regression Tree (CART) are needed (Roel, Firpo, & Plant, 2007). Classification and Regression Tree models have been used to explain variables and interactions influencing crop yields in eastern India (Banerjee et al., 2014). Other methods such as factor analysis (FA) can cluster variables into common and easily interpretable factors (Yong & Pearce, 2013). This can help show consistent factors causing yield gaps over a larger spatial scale and guide policy interventions to enhance yields. The method has been used in clinical studies to obtain general conclusions on clinical conditions (Oh et al., 2016; Ohshiro & Ueda, 2018), but its use in agricultural studies is still low. Combining different methods to examine factors influencing maize yield gaps can provide complementary findings that are relevant at different spatial scales in smallholder farming systems, improving yield gap studies.

The aim of the study was to improve the understanding of consistent and site-specific factors limiting yields on smallholder farms and the causes of yield gaps in two agroecologically contrasting regions, by applying different multivariate methods. Specific objectives were to (a) assess consistent soil and crop management-related factors across agroecology affecting maize yield gaps; (b) assess specific agroecology crop and management-related factors effecting maize yield gaps; and (c) recommend approaches based on the findings for reducing maize yield gaps. It was hypothesized that studying yield gaps at two contrasting agroecologies by applying different multivariate methods will result in both consistent and specific agroecology factors influencing maize yield gaps and improve the applicability of findings on smallholder farms in enhancing yields. The study will also fill the knowledge gaps on the variability of soil properties including the possible effect of rooting depth (depth to compact layer) on maize yield gaps as suggested by van Bussel et al. (2015).

2 | MATERIALS AND METHOD

2.1 | Description of the study sites

The study was conducted in Mukuyu village (0°38′N, 35°41′E), Kakamega County (Site 1), and Shikomoli village (0°4′19N, 34°43′E), Vihiga County (Site 2), in Kenya (Figure 1). Both villages are included in the Intensification of Africa (Afrint) project (Djurfeldt, Andesron, Holmen, & Jirstrom, 2011) and were selected based on intensification potential in production of staple crops such as maize regarding agroecology, population density, and market access (Karugia, 2003).

Mukuyu is located at an altitude of 1,600 m above sea level (asl) and in the agroecological zone Upper Midland 3 (UM3). Annual rainfall ranges between 1,000 and 1,600 mm (mean 1,450 mm) and falls in a bimodal pattern, with long rains occurring in February–August and short rains in September–November, while other months are predominantly dry. Daily temperature varies between 14 and 26°C (mean 20°C). Well-drained Ferrasols are the dominant soil type, with Acrisols occurring in some places (Jaetzold, Schmidt, & Hornetz Berthhold, 2010). In the long-rain season, maize and beans are generally intercropped, but some farmers prefer to grow maize as a pure stand to facilitate weeding. Maize varieties with a growing period of 6–8 months, such as H613 and H614, are preferred due to their high yield potential (One Acre Fund, 2016). Maize is harvested in September–October, and crop residues are removed and used as animal feed. The soil is then generally cultivated for production of beans, potatoes, and vegetables, but sometimes left fallow or grazed by animals. Mukuyu village occupies an area of approximately 3.56 km^2 with an estimated population of 1,664 (KNBS, 2010). The average farm size is 1.5 ha (Djurfeldt & Wambugu, 2011).

Shikomoli village is located at an altitude of 1,400 m asl and is predominantly in the agroecological zone Upper Midland 1 (UM1). It experiences an equatorial-type climate, with annual rainfall of 1,600–2,000 mm (mean 1,700 mm) falling as long rains in February–July and short rains in August–December. Daily temperature ranges between 14 and 32°C (mean 23°C). The soils are mainly Cambisols and are sandy, stony, and moderately deep. Other soil types found in the area include Acrisols and Nitisols (Jaetzold et al., 2010). Short-season certified maize varieties with a growing season of 4–5 months, such as Duma 43 and DK 8,031, are preferred during the long-rain season (One Acre Fund, 2016). Some farmers grow indigenous maize varieties during the short-rain season, as these are believed to be drought-tolerant and more suited to this drier season. Shikomoli village occupies an area of 1.37 km^2 with an estimated population of 2,923 (KNBS, 2010). The average farm size is 0.5 ha (Djurfeldt & Wambugu, 2011).

Inorganic fertilizer, mostly diammonium phosphate (18% N, 20% P), is applied at an average rate of 135 kg/ha in Mukuyu and 72 kg/ha in Shikomoli. This is well below the 250 kg/ha recommended for maize, despite the Kenyan government initiative to subsidize fertilizer costs (Oseko & Dienya, 2015). High costs and inadequate knowledge of the amount, frequency, and timing of fertilizer application are factors impeding inorganic fertilizer use (Mavuthu, 2017; Sheahan, Black, & Jayne, 2012). Organic fertilizer utilization is low, hampered mainly by low availability in Mukuyu, where land holdings are large, and low quality in Shikomoli, resulting from poor preparation and storage methods. Family and hired labor are used in agriculture, with women providing a large share. Youth participation in agriculture is low (KNBS, 2010), and the area under maize cultivation in Mukuyu and Shikomoli is decreasing because of a shift by younger workers to enterprises considered more productive for income generation, such as sugarcane farming and tree planting for timber production (MEMR, 2013).

2.2 | Identifying and geo-referencing maize plots

The sampling units were plots growing maize in the current season. These plots belonged to 70 randomly selected households and were delineated based on present management practices, distance from the homestead, and size. Sixty of the households (30 in Mukuyu and 30 in Shikomoli) had participated in the Afrint project since 2002 (Djurfeldt et al., 2011). The remaining 10 households (five per village) were selected randomly and added to the initial sample size using the Afrint sampling design (Djurfeldt et al., 2011). The Afrint project used purposive sampling based on intensification potential concerning agroecology and market access to select

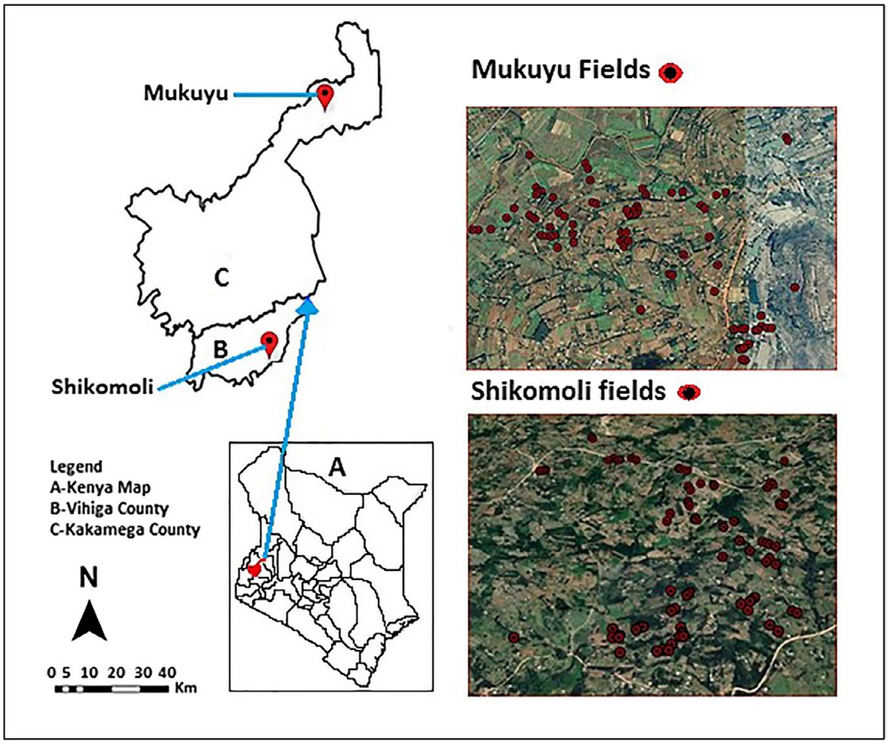

FIGURE 1 Location and distribution of households and farms in Mukuyu village, Kakamega County, and Shikomoli village, Vihiga County, western Kenya

the two villages, followed by random selection of 30 out of 150 households in every village (Karugia, 2003).

To avoid oversegmentation, the maize plots sampled had to be larger than 0.04 ha in Mukuyu and larger than 0.004 ha in Shikomoli (reflecting smaller plot sizes in Shikomoli than Mukuyu). For each household and maize plot identified, coordinates and circumference were recorded using a hand-held Garmin Global Positioning System (GPS) (GPSMAP® 62), and the area was estimated from the coordinates. The total number of maize plots identified was 170 (89 in Mukuyu and 81 in Shikomoli). A 4 m × 4 m area, hereafter referred to as the study plot, was marked out at the center of each maize plot and georeferenced.

2.3 | Measuring crop performance and weed pressure

Data on crop performance indicators and weed pressure were collected to determine management-related factors influencing maize yield gaps. The crop performance indicators measured were maize density, maize development stage, height, chlorophyll level (Soil Plant Analysis Development (SPAD) readings), vigor, and yield. The weed pressure indicators were weed cover and weed height. The measurements were conducted in the study plots at ear initiation and at silking and tasseling, corresponding to maize development stages 1 and 3, respectively, according to O'Keeffe (2009). Maize density was determined by taking a count of all maize plants, maize height was measured on 10 randomly chosen plants, and maize development stage was determined by counting the number of leaves from the base of the maize plant to the youngest fully developed leaf having a visible leaf collar (O'Keeffe, 2009). This was done early in the day, and care was taken not to include leaves within whorl, not fully expanded, and with no visible leaf collar. The majority leaf value within the study plot was then used to describe the maize development stage. The chlorophyll levels were determined using a SPAD 502 chlorophyll meter (Minolta Camera Co.) by taking readings of the youngest fully developed leaf from 15 randomly selected plants per study plot, at approximately 25% from the leaf tip and leaf base, respectively. Crop vigor was determined as the presence of disease and pests through observations on 10 randomly selected plants and using a 1–5 Likert scale (Sherpherd, 2010) where 1 = almost completely infested with pest or diseases (75%–100%), 2 = heavily infested (50%–75%), 3 = moderately infested (10%–50%), 4 = low infestation (less than 10%), and 5 = no diseases or pests. Weed coverage was assessed using an improvised mottle chart with 12 percentage levels (1%, 3%, 5%, 10%, 15%, 20%, 25%, 30%, 40%, 50%, 75%, and 90%), where 1% represents low weed infestation, and 90% represents severe weed infestation (Sherpherd, 2010).

Maize yield was determined at the end of the growing period on a dry matter basis using the method described by Tobergte and Curtis (2013). In brief, all plants in the study plots were harvested and the grain was shelled, cleaned, weighed, and recorded in kg. A subsample of approximately 200 g was oven-dried at 75°C for 24 hr and weighed to determine moisture content and to calculate yield as kg dry matter for the 4 m × 4 m study plot. The values obtained were converted to tons per ha. The grain yield was determined at 13% moisture content.

2.4 | Soil measurements, sampling, and analysis

Slope, soil erosion status, and depth to compact soil layer were recorded in the study plot, and soil samples were taken and analyzed for texture, pH, and soil nutrient status. Plot slope was determined at the start of the maize growing season, using a Likert scale (FAO, 2006), and percentage slope was determined using a modified L square. Soil erosion status was determined during maize stages 1 and 3 using a Likert scale developed by FAO (2006). Depth to soil compaction layers with resistance of 200, 300, and 500 pounds per square inch (psi) was determined when the soil was at field capacity, using a Humboldt H-4210A Portable Static Cone Penetrometer with 10 section points. In all study plots, the static cone penetrometer was pressed into the soil until the gauge read 200 psi (and then 300 and 500 psi). For each psi gauge value, a recording was made for the penetrometer sections that remained aboveground. The exact penetrometer value was computed by subtracting the values for the penetrometer sections aboveground from the values for the 10 section points. This was repeated 10 times at randomly selected points. The penetrometer values were then converted to centimeters.

In January 2016, at the start of the maize growing season, soil samples were taken to a depth of 0–20 cm at 10 randomly selected points in each plot, using a soil corer (Ø 25 mm), and bulked to one composite sample per study plot. The soil samples were air-dried and passed through a 2-mm sieve at Crops and Nutrition Laboratory Services in Nairobi. Soil texture, pH, total soil carbon (C) and nitrogen (N), and extractable soil nutrients were determined using methods described by Pansu and Gautheyrou (2006) (Table 1). The extractable soil nutrients measured were boron (B), calcium (Ca), copper (Cu), iron (Fe), potassium (K), magnesium (Mg), manganese (Mn), phosphorus (P), sulfur (S), and zinc (Zn). Cation exchange capacity (CEC) was calculated from the amounts of Mehlich-3-extracted nutrient elements (Table 1) and exchangeable acidity. In evaluating soil fertility, the soil nutrient concentrations were compared with critical values established for maize production (FAO, 2007).

2.5 | Quantifying maize yield gaps

Maize yield gap was determined by comparing yields on the 90th percentile of farms to that on other farms within the same site (Lobell et al., 2009). The 90th percentile yield was computed for each site based on yield as:

$$K\text{th} = L\left[\frac{(P - cfb)}{f}\right] U - L \tag{1}$$

where Kth is the 90th percentile; L is the lower limit of the critical value within which the 90th percentile occurs; P is the critical interval where the 90th percentile occurs, calculated as (K/100) multiplied by the number of values in the distribution; cfb is the cumulative frequency of all intervals below the critical value, but not including the critical value; f is the frequency in the critical interval; and U is the upper limit of the critical value that is not included in the critical interval.

Maps showing the frequency of plots with large and small maize yield gaps were then drawn using the Geo-statistical Analyst tool in Arc Map 10.1 (Hengl, 2007).

2.6 | Data analysis

Descriptive statistics and t test analysis in R statistics were used to assess crop performance, weed pressure, and soil properties at the two sites. The data were subjected to generalized linear mixed models (GLMMs) and CART analysis. In GLMM, the penalized quasi-likelihood (PQL) technique implemented in R statistics was used to identify significant factors influencing maize yield gaps (Hui, Mueller, & Welsh, 2016). The analysis involved setting random and fixed effects, where the random effects were the study plots and the fixed effects were the soil and crop variables determined. Classification and Regression Tree analysis employed the binary recursive partitioning technique, where variables were divided into exclusive homogeneous variables in three steps (Berk, 2008). First, the tree split the parent node (average maize yield gaps in tons/ha) into two homogeneous child nodes, which were placed to the right and left, depicting low and high yield gaps, respectively. The two child nodes were further split and the process continued, resulting in an overgrown tree that was pruned by setting the cost complexity (cp) value at 0.01. The CART analysis indicated the level of variable occurrence that resulted in large or small maize yield gaps. In each split, the left side, with "yes" as a Boolean choice, showed factors that contributed to reducing yield gaps, while the right side, with "no" as a Boolean choice, indicated factors that led to large maize yield gaps. The analysis was implemented using the recursive partitioning (rpart) package in R statistics (Therneau & Atkinson, 2015). Classification and Regression Tree and GLMM were chosen owing to their ability to handle highly skewed data (Gordon, 2013). Factors influencing maize yield gaps identified by CART and GLMM were subjected to FA in R statistics (Beaujean, 2014), using varimax rotation to regroup variables into small easily interpretable sets based on variance (Yong & Pearce, 2013).

TABLE 1 Soil analysis methods applied in the study

Soil measurement	Method
Soil pH and EC	Potentiometric method using reference and measurement electrodes with soil:water of 1:2
B, Ca, Cu, Fe, Mg, Mn, K, S, Zn	Extraction with Mehlich 3 solution containing diluted ammonium fluoride and ammonium nitrate followed by atomic emission spectrometry (ICP)
Available P	Olsen P extraction using sodium bicarbonate solution followed by colorimetric determination
Total N	Kjeldahl digestion followed by colorimetric determination
Total C	Walkley and Black method through wet oxidation by acidified dichromate in the presence of sulphuric acid
Clay, Sand, Silt	Hydrometer method using 10% sodium hexametaphosphate as the dispersing agent
Exchangeable Al	Colorimetric method using KCl (1N) extraction
Exchangeable acidity	Titration after extraction with KCl and titration with NaOH

TABLE 2 Mean, minimum, and maximum values of crop performance and weed pressure indicators at maize development stages 1 and 3 in Mukuyu and Shikomoli

	Stage 1		Stage 3		
Mukuyu	Mean	Min and Max	Mean	Min and Max	(Stages)
Maize density, plants/hectare	$44e^{3}$ a	$37e^{3}$–$56e^{3}$	$35e^{3}$a	$28e^{3}$–$48e^{3}$	**
Maize height (cm)	41 a	18–47	245 a	213–278	
Maize vigor (cm)	3.9 a	2–5	4.1 a	3–5	**
SPAD values	38 a	33–43	36 a	16–65	**
Weed coverage (%)	29 a	15–60	46 a	20–75	**
Weed height (cm)	10 a	6–12	31 a	25–36	**
Shikomoli					
Maize density, plants/hectare	$55e^{+3b}$	$37e^{+3}$–$71e^{+3}$	$32e^{+3}$a	$26e^{+3}$–$43e^{+3}$	**
Maize height (cm)	58 b	28–66	172 b	148 –209	
Maize vigor	3.8 b	1–5	4.2 b	2–5	**
SPAD values	37 a	27–40	34 a	20–54	**
Weed coverage (%)	31 a	15–50	34 b	10–50	**
Weed height (cm)	13 b	8–17	23 b	16–30	**

**Significant difference ($p \leq .001$) between key maize development stages 1 and 3. Different letters (a, b) indicate significant differences between sites. SPAD = Soil Plant Analysis Development. Max and min are the highest value and lowest value, respectively, recorded for a given variable, and mean is the average of all observations.

3 | RESULTS

3.1 | Crop performance and weed pressure at key maize development stages

There were significant differences in crop performance and weed pressure between key maize development stages, and also between sites (Table 2). Chlorophyll content decreased as maize progressed through stages 1 (ear initiation) to 3 (silking and tasseling) at both sites. Higher plant densities were recorded during early maize development (stage 1) than later (stage 3) (Table 2). Weed coverage was high at both sites during early and later stages of maize development.

3.2 | Characterization of soil physical and chemical properties

All soil properties except extractable B, Cu, Fe, S, acid saturation, clay content, and electrical conductivity (EC) differed between the sites (Table 3). At both sites, total C, total N, extractable B, K, Mg, P, and S, and CEC were below the critical values required to support important physiological processes in maize growth. Soil organic matter (calculated from total C) was on average 2.4% and 1.7%, for Mukuyu and Shikomoli, respectively, and C:N ratio was 14 and 10, respectively (Table 3). The soils in Mukuyu and Shikomoli contained shallow compacted layers. Significant differences were noted for erosion status and slope, with Shikomoli exhibiting higher erosion status and steeper slopes. However, soil properties varied widely within the villages, as indicated by the minimum and maximum values (Table 3).

3.3 | Maize yields and yield gaps in Mukuyu and Shikomoli

Maize yield ranged from 0.1 to 7.1 t/ha in Mukuyu and from 0.1 to 5.1 t/ha in Shikomoli. The 90th percentile yield was 5.1 and 4.8 t/ha in Mukuyu and Shikomoli, respectively (Figure 2). Mean measured maize yield was 3.3 t/ha and 2.2 t/ha in Mukuyu and Shikomoli, respectively.

The yield gap was significantly different ($p = .0001$) between Mukuyu (mean 1.8 t/ha) and Shikomoli (mean 2.6 t/ha), representing 35% and 54% unachieved yield at farm level compared to the 90th percentile yield. The distribution of fields with low yield gaps was greater in Mukuyu than in Shikomoli (Figure 3a,b).

3.4 | Soil and management-related factors contributing to yield gaps in Mukuyu

In Mukuyu, GLMM analysis showed that maize yield gaps were significantly affected by weed coverage and weed height in early stages of maize development, SPAD readings at stage 3, extractable Zn and P, and percentage silt (Table 4). High weed height and higher weed coverage at stage 1 contributed to large yield gaps, as shown by the positive coefficient and R-values, while high SPAD readings at stage 3, high extractable Zn and P concentrations, and high percentage of silt resulted in low yield gaps, as shown by the negative coefficient and R-values.

The CART analysis for Mukuyu showed that weed coverage at development stage 1 (WC1) was the main factor causing yield gaps (Figure 4). In the first split, the 53% of plots with weed coverage <28% (Node 1) had an average yield gap of 0.86 t/ha, whereas the 47% of plots with higher weed coverage had an average yield gap of 2.9 t/ha. Plots where depth to compact layer (CmpD) was great (≥10 cm) (Node 2) showed smaller yield gaps, as did plots with maize density at development stage 3 (MDD3) ≥32,000 ha^{-1} (Node 4). Other important variables that resulted in reduced yield gaps were relatively high soil Mn, tall maize plants at stage 3, and low weed height at stage 1. The 10% of plots that had weed coverage below 28%, an easily penetrable soil to at least 10 cm depth, maize density above 32,000 plants/ha, and extractable Mn above 83 mg/kg showed an average yield gap of −0.96 t/ha (Node 15).

Factors that increased yield gaps were weed coverage at stage 1, tall weeds (>9.6 cm) at stage 1 (WH1) (Node 3), low P availability (<17 ppm) (Node 5), and low maize density (Node 7). The 10% of plots that had weed coverage above 28%, weed height more than 9.6 cm in stage 1 (WH1), and maize density below 37,000 plants/ha at stage 3 (MDD3) showed an average yield gap of 4.6 t/ha.

In FA, the clustering of variables influencing maize yield gaps in Mukuyu showed that factors 1, 2, and 4, with a total proportion of variance of 0.42, were predominantly soil variables, while factors 3 and 5, with a total proportion of variance of 0.19, were management-related variables (Table 5).

3.5 | Factors contributing to maize yield gaps in Shikomoli

The GLMM analysis indicated that maize yield gaps in Shikomoli were significantly affected by: maize density and maize height at late maize stages, CEC, depth to compact layer, Mg concentration, and silt content (Table 6). High plant density and maize height at stage 3 and high Mg, CEC, silt percentage, and depth to compact layer reduced maize yield gaps, as shown by the negative coefficient and R-values (Table 6).

The CART analysis identified maize density as the most important factor contributing to yield gaps in Shikomoli (Figure 5). The approximately 64% of plots with maize density at harvest (MDD3) ≥32,000 ha^{-1} (Node 1) had a yield gap of 2.1 t/ha, whereas the 36% of plots with lower maize density had an average yield gap of 3.5 t/ha (Node 1). Lower yield gaps were also recorded for plots with greater depth to compact layer (CmpD3 ≥24 cm) (Node 2), sand content <53% (Node 4), EC ≥43 meq 100 g^{-1} (Node 7), Na content ≥16 ppm (Node 11), and Zn content ≥7.3 ppm (Node 13). The 4% of plots with maize density >32,000 ha^{-1}, CmpD3 >24 cm, and sand content <53% had a yield gap of −0.1 t/ha.

Factors that increased the yield gap (right side of Figure 5) were low maize density in stage 3 (MDD3), low maize height in stage 3 (MH3) (Node 3), and low chlorophyll content in stage 1 (SPAD1 <38) (Node 6). The 10% of plots that had MDD3 <32,000 ha^{-1}, maize height <219 cm, SPAD1 <38,

TABLE 3 Mean, minimum (Min), and maximum (Max) values of soil chemical and physical properties in Mukuyu and Shikomoli

	Mukuyu		Shikomoli			
Soil variables	Mean	Min and Max	Mean	Min and Max	Critical values	
Carbon (C) %	1.4*	0.8–3.4	1.01*	0.5–2	>2.7[a]	**
Nitrogen (N) %	0.1*	0.08–0.3	0.1*	0.06–0.2	>0.2[a]	**
Potassium (K) ppm	174	52–865	100	32–785	>94[a]	**
Calcium (Ca) ppm	1,032	287–3,440	687	212–1,840	>400[a]	**
Magnesium (Mg) ppm	160	57–582	106*	36–251	>120[a]	**
Phosphorus (P) ppm	20*	1.3–112	44	4.7–334	>30[a]	**
Sulphur (S) ppm	9*	1.7–24	8*	2.7–21	>20[a]	
Boron (B) ppm	0.3*	0.04–2.5	0.4*	0.02–1.6	>0.8[a]	
Copper (Cu) ppm	2.4	1.02–7	2	1–5	>1[a]	
Iron (Fe) ppm	160	60–462	148	70.4–242	>10[a]	
Manganese (Mn) ppm	98	11–291	197	61–355	>20[a]	**
Zinc (Zn) ppm	4	0.6–31	11	0.8–46	>5[a]	**
pH	5.6	4.7–7.6	5.8	4.9–6.9	>5.5[a]	
Exchange aluminum (Al) meq/100 g	0.2	0.01–1.04	0.1	0.01–0.78	<0.5[a]	**
Acid saturation %	6	0.22–32	5.8	0.5–34	<10[a]	
Cation exchange capacity (CEC) meq/100 g	10*	5–34	6*	3–13	>15[a]	**
Electrical conductivity (EC) mS/m	44*	9–140	38*	15–116	>80[a]	
Sand (Sa) %	54	30–79	66	49–84		**
Silt (Si) %	9.8	4–32	8	4–14		**
Clay (Cl) %	36	14–54	24	12–44		
Erosion status	1	1–2	2	1–3		**
Slope %	3	1–7	6	1–12	0–2[b]	**
Depth to compact layer at 500 psi (cm)	20	9–25	21	8–25	>50[b]	**

*Value below the critical level for maize growth.

**Significant difference ($p = <.0001$) between sites. Critical values ([a]NAAIAP, 2014; [b]FAO, 2007) represent extractable nutrient concentration in soil above which an economic yield response to added nutrient is unlikely. Max and min are the highest and lowest values recorded for a given soil variable, and mean is the average of all observations.

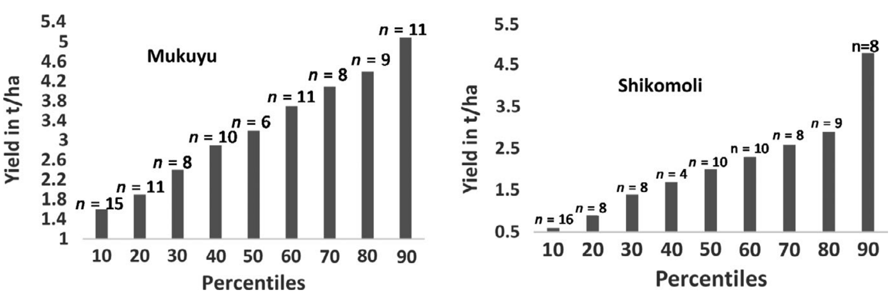

FIGURE 2 Distribution of maize yield in (left) Mukuyu and (right) Shikomoli. The *y*-axis shows average yield for each percentile bar, n = number of plots

and maize density in stage 1 (MDD1) <48,000 ha^{-1} had an average yield gap of 4.3 t/ha (Node 29).

In Shikomoli, soil variables had a total proportion of variance of 0.32 and loaded on factors 1 and 3. Management factors had a total proportion of variance of 0.24 and loaded on factors 2 and 5 (Table 7). Factor 4 had either soil or management factors, with a proportion of variance of 0.06.

3.6 | Consistent factors influencing maize yield gaps at both sites

Factor analysis showed that soil was the overarching factor influencing maize yield gap, rather than management-related factors (Tables 5 and 7). Classification and Regression Tree analysis showed that consistent factors influencing maize yield gap at both sites were maize density, SPAD value,

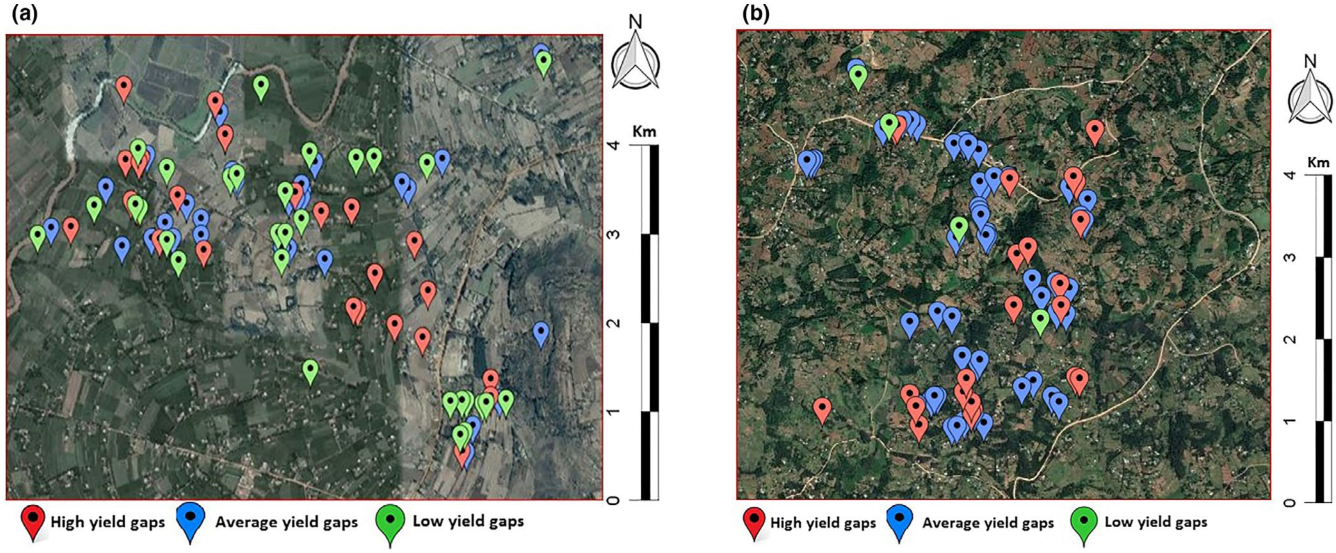

FIGURE 3 (a) Distribution of maize yield yaps in Mukuyu. Red points show high yield gaps (2.00–4.55 t/ha), blue points show average yield gaps (1.0–1.99 t/ha), and green points show low yield gaps (−2.3–0.97 t/ha). (b) Distribution of maize yield gaps in Shikomoli. Red points show high yield gaps (2.00–4.63 t/ha), blue points indicate average yield gaps (1.0–1.99 t/ha), and green points show low yield gaps (−0.01–0.0.55 t/ha)

maize height, and depth to compact layer (Figures 4 and 5). Generalized linear mixed model analysis identified silt content as a constant factor affecting maize yield gaps at both sites (Tables 4 and 6).

4 | DISCUSSION

4.1 | Consistent factors influencing maize yield gaps regardless of agroecology

The high maize yield gaps observed show potential to increase yield by 35%–54% at smallholding level in Mukuyu and Shikomoli, with high and low agroecological potential, respectively. Similarly, Gathala et al. (2011) and Liu, Yang, Hubbard, and Lin (2012) report exploitable maize yield gaps of >40% and recommend adjustments in soil and crop management measures to increase yield. Use of integrated analysis (CART, GLMM, FA) showed that maize yield gaps were influenced by factors that were consistent across agroecological zones and by site-specific factors. The study showed how an integrated approach in the analysis of maize yield gaps can provide complementary findings that are of wider relevance on smallholder farms. The study has particularly extended, as a new result, the recent findings by van Loon et al. (2019), who used an integrated approach and identified only site-specific factors influencing maize yield gaps on smallholder farms.

The high proportion of variance attributable to soil properties at both Mukuyu and Shikomoli, as demonstrated by FA (Tables 5 and 7), suggests that soil factors were more important in influencing maize yield gaps than management-related variables. Both sites showed some nutrient concentrations below the critical value for maize growth (Table 3), suggesting that nutrient supply was inadequate to support maize development (NAAIAP, 2014). Factors leading to low soil nutrition status were a high percentage of sand, high concentrations of extractable Fe and Al oxides, and considerable

TABLE 4 Coefficient value*, *R*-value, and *p*-value for factors influencing maize yield gaps in Mukuyu

Soil and management-related factors	Coefficient value	*R*-value	*p*-value
Intercept	2.311	.999	.0006
Weed coverage in stage 1	0.033	.529	.0000
Extractable (Zn)	−0.144	−.173	.0001
Weed height in stage 1	0.058	.333	.0016
SPAD readings in stage 3	−0.061	−.897	.0015
Available (P)	−0.023	−.200	.010
Silt %	−0.0029	−.386	.051

*Level of increase or decrease in maize yield gap with a one unit increase or decrease in the factor.

erosion, especially in Shikomoli (Table 3). Low soil nutrition status is also an indication of other underlying factors occurring within smallholder farming systems, such as insufficient use of inorganic and organic fertilizers and crop residues (Achieng et al., 2010; Oseko & Dienya, 2015). Since maize is a staple crop in Kenya and in most parts across the globe, policy measures aimed at improving general soil fertility, such as leaving crop residues in situ, applying organic and inorganic fertilizers, and growing cover crops, are important. Low-cost technologies to chop crop residues for easier incorporation into the soil could also improve soil fertility. Making soil testing services available to smallholder farmers and having a supply chain for suitable fertilizers are also important measures.

The CART and GLMM analyses showed that maize yield gaps were influenced by both soil and within-season management-related factors that were consistent across sites. These were maize density, SPAD values, maize height, depth to compact layer, and silt content (Figures 4 and 5; Tables 4 and 6). Previous studies have found that low plant density exacerbates the maize yield gap on smallholder farms (Tittonell, Shepherd, Vanlauwe, & Giller, 2008; Keating et al., 2010; Delmotte, Tittonell, Mouret, Hammond, & Lopez-ridaura, 2011). Maize density decreased between stages 1 and 3, indicating reduced yield potential, and decreased more in Shikomoli than in Mukuyu, as shown by CART analysis (Figure 5). This was due to lodging of maize resulting from strong winds in June–July, exacerbated by the steep terrain in Shikomoli (Table 3). Stem lodging in maize plants could also have been caused by morphological traits such as tall plants, short internode length, low basal strength, and low soil nutrition (Mi et al., 2011; Shah, Tanveer, Rehman, & Anjum, 2017). Low maize density at harvest means that a smaller amount of crop residues is available to improve soil organic matter and fertility, contributing to a vicious low soil fertility cycle. The strong influence of maize density on maize yield gaps at both sites indicates a need for farmers and government authorities to introduce measures aimed at achieving and maintaining high plant density throughout the production period. At farm level, measures such as timely planting to escape adverse wind effects and ensuring adequate soil nutrition status would be helpful. At government level, measures could include breeding for maize varieties with high lodging resistance and carrying out on-farm research to establish the optimal plant population for different agroecological zones.

Maize yield gaps were also influenced by SPAD values and maize height, confirming findings by Ghimire and Timsina (2015) and Boomsma et al. (2010). The SPAD values observed in maize development stages 1 and 3 (27–43) were below the value reported to prevent yield loss (52) (Lindquist, Evans, Shapiro, & Knezevic, 2010). The SPAD values decreased as maize developed, indicating decreasing N concentration in the leaves, a sign of reduced assimilation and remobilization of N to yield components at the silking and tasseling stage (Yan, Chen, Dadouma, Tao, & Sui, 2017; Han, Okamoto, Beatty, Rothstein, & Good, 2015). Insufficient and untimely fertilizer use at both study sites most likely contributed to the low SPAD values and maize

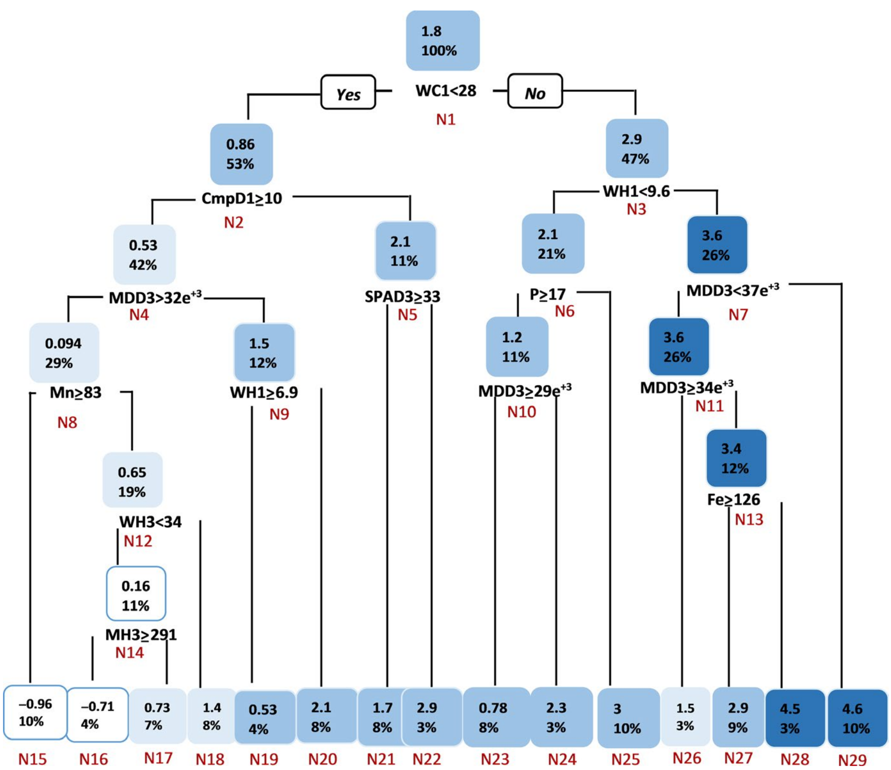

FIGURE 4 Classification and Regression Tree (CART) showing factors resulting in small or large maize yield gaps in Mukuyu. Boxes show average maize yield in t/ha (e.g., 1.8, 0.86, 2.8…0.4.6) and percentage of farms achieving that yield. N1, N2, N3……N29 are nodes. Intensity of box coloration increases with increasing yield gap. (WC1 = weed coverage in stage 1, SPAD3 = chlorophyll readings in stage 3, MH3 = maize height in stage 3, WH1 = weed height in stage 1, WH3 = weed height in stage 3, CmpD1 = depth of compaction at 500 pressure units, Mn = manganese, Fe = iron, P = phosphorus)

TABLE 5 Factor regrouping of soil (S) and management-related (M) variables influencing maize yield gaps in Mukuyu

Variable	Factor 1 (S)	Factor 2 (S)	Factor 3 (M)	Factor 4 (S)	Factor 5 (M)
Weed height	0.53*				
Depth to compaction layer	0.79*				
Iron (Fe)	0.75*				
Zinc (Zn)		0.96*			
Phosphorus (P)		0.66*			
Weed coverage			0.94*		
Silt	0.58*			0.81*	
Manganese (Mn)	0.43			0.57*	
Chlorophyll content (SPAD)					0.82*
Maize height					0.42
	Factor 1	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**
Sum of squared loadings	2.04	1.44	1.06	1.05	1.02
Proportion variance	0.19	0.13	0.10	0.10	0.09
Cumulative variance	0.19	0.32	0.41	0.51	0.60**
Eigen values	2.46	2.03	1.37	1.24	1.03

Note: Test of hypothesis that five factors are sufficient (chi-square statistic = 10.11 on 10 degrees of freedom, p-value = .431). The p-value is the probability that the source data perfectly fit the number of factors specified, five in this case, so larger values are better.

*Factors with higher variance (>0.5).

**Cumulative variance.

TABLE 6 Coefficient value*, R-value, and p-value for soil and management-related factors influencing maize yield gaps in Shikomoli

Soil and management-related factors	Coefficient value	R-value	p-value
Intercept	8.35	.999	.0000
Maize density at harvest	−0.007	−.63	.0000
Maize height at stage 3	−0.008	−.450	.0061
Cation exchange capacity (CEC)	−0.2	−.335	.0076
Depth of compaction at 500 psi	−0.06	−.732	.0069
Magnesium concentration (Mg)	−0.010	−.440	.0078
S (%)	−0.117	−.397	.0073

*Level of increase or reduction in maize yield gap with a one unit increase or decrease in the factor.

height, as previously described (Oseko & Dienya, 2015). This reflects the limited ability of smallholder farmers to access adequate fertilizer and their limited awareness of the importance of supplying adequate nutrition to maize during the critical yield determination period. Hence, increasing the accessibility of fertilizers and educating smallholder farmers on the need to time fertilizer applications to critical crop nutrient requirement stages are important. The lower than average precipitation experienced at the silking and tasseling stages (Uhe et al., 2018) could also have contributed to plant stress, leading to low nutrient use efficiency, low SPAD values, and shorter maize in Mukuyu. To mitigate such challenges, managing soil moisture by growing cover crops and legume intercrops for enhanced water use efficiency will become increasingly important under high climate variability. National and county governments could invest in irrigation facilities.

Dense soil at shallow depth (short distance to compact layer) (Table 5) indicated restricted maize rooting depth, limiting uptake and assimilation of nutrients and moisture (FAO, 2007). The effect of root resistance on maize yield gaps has been documented previously (Chen & Weil, 2011; Głąb, 2011). Soil compaction at shallow depth could have been caused by ploughing when the soil was wet (Elaoud & Chehaibi, 2011). This is more likely to have occurred in Mukuyu, where farm size is larger. In Shikomoli, shallow soil could be due to the rockiness and stoniness of the terrain (Jaetzold et al., 2010). Farmers at both study sites graze animals in the maize fields, due to inadequate pasturage. Trampling by animals could have resulted in hard layers in

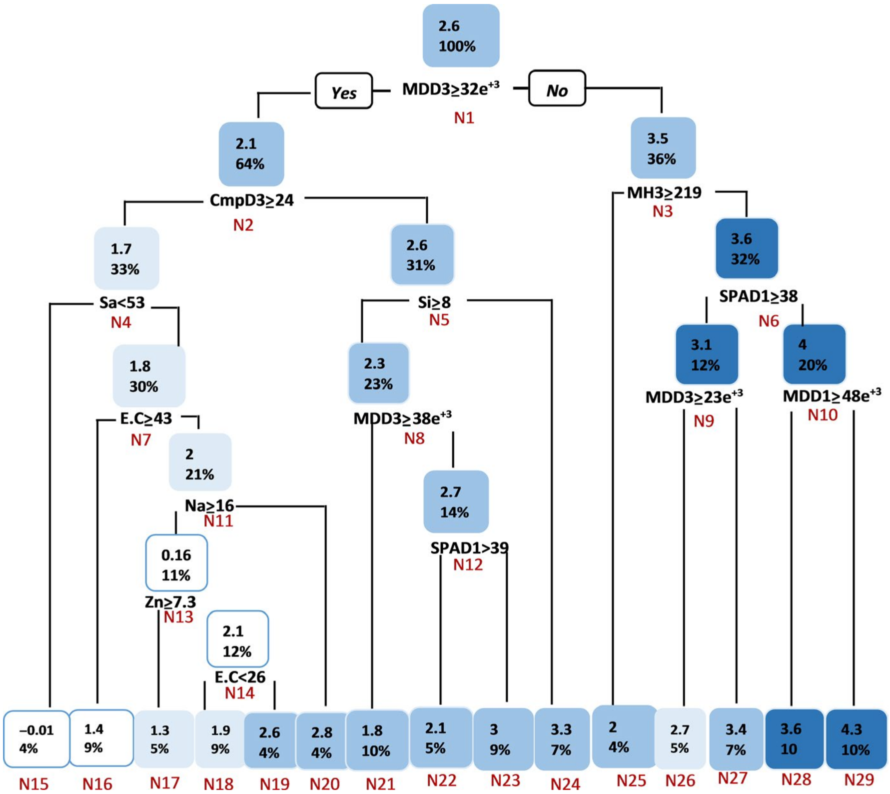

FIGURE 5 Classification and Regression Tree (CART) showing factors resulting in small or large maize yield gaps in Shikomoli. Boxes show average maize yield in t/ha (e.g., 2.6, 2.1, 3.6…4.3) and percentage of plots achieving that yield, N1, N2, N3……N29 are nodes. Intensity of box coloration increases with increasing yield gap. (MDD3 = maize density at harvest, CmpD3 = depth of compaction at 500 pressure units, MH3 = maize height in stage 3, SPAD1 = chlorophyll content in stage 1, Na = sodium, WH3 = weed height in stage 3, MDD1 = maize density at stage 1, Si = silt content, Sa = sand content, EC = electrical conductivity)

TABLE 7 Factor regrouping of soil (S) and management-related (M) variables influencing maize yield gaps in Shikomoli

Variables	Factor 1 (S)	Factor 2 (M)	Factor 3 (S)	Factor 4 (S/M)	Factor 5 (M)
Zinc (Zn)	0.73*				
Cation exchange capacity (CEC)	0.61*				
Magnesium (Mg)	0.80*		0.38		
Maize density at stage 3		0.99*			
Maize density at stage 1		0.68*			
Silt percentage			0.59*		
Sand percentage			0.94*		
Chlorophyll content (SPAD)					0.75*
Compaction depth				0.54*	
Maize height				0.31	
	Factor 1	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**
SS loadings	1.77	1.56	1.54	1.11	0.82
Proportion variance	0.17	0.15	0.15	0.06	0.09
Cumulative variance	0.17	0.32	0.47	0.53	0.62**
Eigen values	2.9	1.7	1.6	1.10	1.02

Note: Test of hypothesis that five factors are sufficient (chi-square statistic = 10.11 on 10 degrees of freedom, p-value = .844). The p-value is the probability that the source data perfectly fit the number of factors specified, five in this case, so larger values are better.

*Factors with higher variance (>0.5).

**Cumulative variance.

the soil. Given the rocky and stony terrain that characterizes most smallholder farms, untimely use of tractor/animal-drawn ploughs, and use of crop residues in situ as animal feed (FAO, 2012), the negative effect of restricted rooting depth on maize yields could be aggravated. Hence, there is a need for government measures to increase water and nutrient use efficiency, such as breeding for varieties that grow well on soils with high penetration resistance and have high N and P acquisition efficiency. There is also a need to inform smallholder farmers about the importance of avoiding ploughing wet soil in order to minimize compaction.

4.2 | Agroecology-specific factors influencing maize yield gaps

The GLMM and CART analysis identified a number of agroecology-specific factors influencing maize yield gaps. The findings are consistent with previous recommendations for site-specific extension services on soil and crop management strategies to reduce yield gaps (Banerjee et al., 2014; Krupnik et al., 2015; Tamene, Mponela, Ndengu, & Kihara, 2016; Yengoh, 2012).

Weed cover during early maize stages was the most yield-limiting factor in Mukuyu (Figure 4). High weed cover in stages 1 and 3 increased competition for soil nutrients and moisture, resulting in lower soil N availability for maize growth, as shown by low SPAD values (Table 2). High weed cover in stage 1 has been previously shown to reduce dry matter accumulation in earlier maize stages, which is essential for grain formation (Page et al., 2012). High weed cover in stage 3 has been shown to lengthen anthesis and silking stages, resulting in lower grain number (Reid et al., 2014). High weed cover in both stages 1 and 3 is therefore a predictor of low maize yield. Imoloame and Omolaiye (2017) also report negative effects of weed cover on maize performance between the 3rd and 6th weeks of development, resulting in low yield. High weed cover may be the result of unavailability of labor for timely weed control, because of financial constraint (Sims et al., 2018). Low returns from agriculture have also been shown to decrease investment in hired labor for weeding (Usman, 2017). Interviews with farmers in Mukuyu confirmed that low returns from maize sales had prevented them from hiring labor for weeding. High weed cover during early stages indicates an effect of other factors related to field management, such as delayed land preparation and past farmer weed management measures, which need to be investigated. Ongoing migration to cities is reducing the labor force in rural areas, with impacts on weed control (Sims et al., 2018). The high engagement of young people in other income-generating activities at the study sites seemingly has a similar effect. Given the strong influence of weeds on maize yield, farmers need to invest in early weed management measures such as labor-saving technologies (low-cost tillage equipment), and extension services need to create awareness among farmers of the significance of early land preparation in controlling weeds.

Previous work in western Kenya has demonstrated the negative effects of low soil nutrient concentrations on crop yield (Kihara et al., 2016). Low nutrient availability in soils contributed to yield gaps at both sites in the present study, with the most important factors being low Zn and P in Mukuyu, and low Mg and CEC in Shikomoli (Tables 3 and 4). Masood et al. (2011) and Tariq, Anjum, Randhawa, and Ullah (2014) have previously reported effects of P and Zn deficiency on maize production. Low P values result in reduced root development and can lead to decreased uptake of moisture and nutrients (Fageria, 2016). Low Zn impairs protein metabolism and has been shown to affect yield (Cakmak, 2000). Low Mg is an indication of reduced biomass formation and increased susceptibility of maize crops to environmental stress (Senbayram, Gransee, Wahle, & Thiel, 2015). Low CEC resulted from the dominant soil types (Acrisols, Ferrasols, Nitisols) at the study sites and indicates low ability of the soils to hold important nutrients in a plant-available pool. This contributed to low crop performance, as shown by low SPAD values (Table 2) and high maize yield gaps. Low P, Zn, and Mg also reflect lack of access by smallholders to fertilizers containing Zn, Mg, and potentially other nutrients, or unaffordability of sufficient doses. Measures that raise the availability of P, Zn, and Mg in soils on smallholder farms are therefore needed. Such measures at national and county government level might include increasing the accessibility and affordability of Zn, S, and P fertilizers to smallholder farmers. Measures at farm level might include improving soil organic matter content by applying organic manure to increase nutrient availability and CEC.

4.3 | Methodological applicability and relevance in yield gap analysis

This study examined causes of maize yield gaps using an integrated approach and identified relevant consistent and agroecology-specific factors. The study found support of the set hypothesis that using an integrated approach resulted in both consistent and agroecological specific factors influencing maize yield gaps. This improved the applicability of the findings on smallholder farms and gave direction of management options for enhancing yields. The CART, GLMM, and FA all revealed some factors that influenced maize yield gaps at both sites, and some that were specific for each site. Factor analysis showed that, when evaluated together, soil factors exerted more influence on maize yield gaps than management-related factors. Classification and Regression Tree analysis revealed more consistent factors influencing maize yield gaps at the two sites with contrasting agroecological potential than GLMM analysis (Section 3.6), while GLMM analysis revealed more agroecology-specific factors. Thus, despite the complexity of the smallholder farming systems, it proved possible to identify consistent factors and agroecology-specific soil and management-related factors limiting crop yield. Knowledge of consistent factors across agroecological zones can assist regional and national authorities in policy development, while knowledge of agroecology-specific factors can assist county and local authorities in prescribing soil and crop management measures to improve yields. In devising measures, it is important to consider the relative occurrence of the factors limiting yield, in order to enhance resource utilization efficiency.

Smallholder farms require a suite of management options they can select and adapt to improve yields, since they operate under resource constraints (Ronner, Descheemaeker, Almekinders, Ebanyat, & Giller, 2018). The CART method showed variability in occurrence of maize yield gaps and identified a combination of interacting factors that led to small (Figure 4; Node 15) and large (Figure 4; Node 29) yield gaps. The method thus revealed interacting factors not identified by GLMM and factor analysis. The CART analysis also showed the weight of soil and management-related factors on yield gaps (Figures 4 and 5). The study thus improved on findings from past studies that have used classical regression methods and only shown soil and management-related factors without considering the weight and combination of the factors influencing maize yield gaps Affholder et al., 2013; Fermont et al., 2009; Okumu et al., 2011). This will aid in devising a suite of soil and crop management measures for use on smallholder farms, based on extension work, on-farm research, and agronomic trials. More specifically, previous studies have either shown only soil factors such as P (Umeri, Moseri, & Onyemekonwu, 2015), Zn (Tariq et al., 2014), or soil texture (Tremblay et al., 2012), or management-related factors, that is, weed pressure (Reid et al., 2014), SPAD values (Ghimire & Timsina, 2015), maize height (Ashraf et al., 2016), or maize density (Shafi et al., 2012), limiting yields and causing yield gaps. This study showed the effect of a combination of site-specific (soil) and management-related factors on maize yield gaps. The combined approach using CART analysis unravels interaction between soil and management-related factors to influence maize yield gaps.

The results demonstrate the significance of an integrated approach in providing complementary findings, which have relevance at different levels of authority (regional, national, local) and usability on smallholder farms. This indicates that simultaneous and concerted efforts at different levels are needed to close maize yield gaps on smallholder farms. Future work could employ on-farm trial research and remote-sensing technologies as complementary methods to further unravel causes of low yields on smallholder farms.

5 | CONCLUSIONS AND RECOMMENDATIONS

This study makes a novel contribution to the existing literature on yield gaps by demonstrating the usefulness of combining different multivariate methods (CART, GLMM, FA) in revealing consistent and site-specific factors limiting yields in different agroecological regions. Maize yield gaps were found to be consistently influenced by maize density, chlorophyll values (SPAD), maize height, depth to compact layer, and soil texture. In an area with high agroecological potential (Mukuyu), maize yield gaps were also increased by high weed pressure. Low soil fertility contributed to yield gaps at both sites, but the deficient nutrients differed between the sites. The study also provides new knowledge on the variability of soil properties, including depth to compact layer (rooting depth), and the effect on maize yield gaps. The results indicate great potential to increase maize yields on smallholder farms through simultaneous measures that address multiple constraints affecting yields. To achieve high crop performance, measures aimed at improving soil fertility, sustaining an optimal plant population, and managing weeds need to be introduced by national, regional, and local authorities.

ACKNOWLEDGMENT

The authors gratefully acknowledge the Swedish Research Council Formas (Grant # 220-2014-646) and VR (Grant # 04440) for financial support. Special thanks to smallholder farmers in Mukuyu and Shikomoli, who willingly permitted field activities on their land, and to the field guides and local authority at both study sites, for providing guidance during the field activities.

ORCID

Sussy Munialo https://orcid.org/0000-0001-6684-9021
A. Sigrun Dahlin https://orcid.org/0000-0003-2831-9163
Ingrid Öborn https://orcid.org/0000-0003-3847-6082

REFERENCES

Affholder, F., Poeydebat, C., Corbeels, M., Scopel, E., & Tittonell, P. (2013). The yield gap of major food crops in family agriculture in the tropics: Assessment and analysis through field surveys and modelling. *Field Crops Research*, *143*, 106–118. https://doi.org/10.1016/j.fcr.2012.10.021

Banerjee, H., Goswami, R., Chakraborty, S., Dutta, S., Majumdar, K., Satyanarayana, T., … Zingore, S. (2014). Understanding biophysical and socio-economic determinants of maize (*Zea mays* L.) yield variability in eastern India. *NJAS - Wageningen Journal of Life Sciences*, *70–71*, 79–93. https://doi.org/10.1016/j.njas.2014.08.001

Beaujean, A. (2014). *Factor Analysis using R; Practical Assessment Research and Evaluation*, 18(4). Retrieved from http://pareonline.net/getvn.asp?v=18&n=4

Beza, E., Silva, J. V., Kooistra, L., & Reidsma, P. (2016). Review of yield gap explaining factors and opportunities for alternative data collection approaches. *European Journal of Agronomy*, *82*, 206–222. https://doi.org/10.1016/j.eja.2016.06.016

Berk, R. A. (2008). *Statistical learning from a regression perspective* (Vol. 14). New York:Springer.

Boomsma, C. R., Santini, J. B., West, T. D., Brewer, J. C., Mcintyre, L. M., & Vyn, T. J. (2010). Soil & Tillage Research Maize grain yield responses to plant height variability resulting from crop rotation and tillage system in a long-term experiment. *Soil and Tillage Research*, *106*(2), 227–240. https://doi.org/10.1016/j.still.2009.12.006

Cakmak, I. (2000). Tansley Review No. 111 Possible roles of zinc in protecting plant cells from damage by reactive oxygen species. *New Phytologist*, *146*, 185–205. https://doi.org/10.1046/j.1469-8137.2000.00630.x

Cassman, K. G., Dobermann, A., Walters, D. T., & Yang, H. (2003). Meeting cereals demand while protecting natural resources and improving environmental quality. *Annual Review of Environment and Resources*, *28*, 315–358. https://doi.org/10.1146/annurev.energy.28.040202.122858

Chen, G., & Weil, R. R. (2011). Root growth and yield of maize as affected by soil compaction and cover crops. *Soil and Tillage Research*, *117*, 17–27. https://doi.org/10.1016/j.still.2011.08.001

Delmotte, S., Tittonell, P., Mouret, J., Hammond, R., & Lopez-ridaura, S. (2011). On farm assessment of rice yield variability and productivity gaps between organic and conventional cropping systems under Mediterranean climate. *Eur. J. Agron.*, *35*, 223–236. https://doi.org/10.1016/j.eja.2011.06.006

Djurfeldt, A. A., & Wambugu, S. K. (2011). In-kind transfers of maize, commercialization and household consumption in Kenya. *Journal of Eastern African Studies*, *5*, 447–464. https://doi.org/10.1080/17531055.2011.611671

Djurfeldt, G., Andesron, A., Holmen, H., & Jirstrom, M. (2011). *The millennium development goals and the African Food Crisis – Report from the Afrint II project*. Stockholm, Sweden: SIDA.

Duncan, J. M. A., Dash, J., & Atkinson, P. M. (2015). The potential of satellite-observed crop phenology to enhance yield gap assessments in smallholder landscapes. *Frontiers in Environmental Science*, *3*, 1–16. https://doi.org/10.3389/fenvs.2015.00056

Elaoud, A., & Chehaibi, S. (2011). Soil compaction due to tractor traffic. *Journal of Failure Analysis and Prevention*, *11*, 539–545. https://doi.org/10.1007/s11668-011-9479-3

Fageria, N. K. (2016). *The use of nutrients in crop plants*. CRC press.

FAO (2006). *Guidelines for soil description*. Rome, Italy: FAO.

FAO (2007). *Mapping biophysical factors that influence agricultural production and rural vulnerability, by Velthuizen, Van Harrij Huddleston, Barbara Fischer, Gunther Salvatore, Mirella Ergin, Ataman Nachtergaele, O. Freddy Marina, Zanetti Mario, Bloise. Rome, Italy*.

FAO, (2012). *Why has Africa become a net food importer? by Rakotoarisoa M. A, Iafrate M, Paschali M., Trade and Market Division, FAO*

Fermont, A. M., Asten, P. J. A. V., Tittonell, P., Wijk, M. T. V., & Giller, K. E. (2009). *Closing the cassava yield gap: an analysis from smallholder farms in East Africa Field Crops Research Closing the cassava yield gap: An analysis from smallholder farms in East Africa*.

Fischer, T., Byerlee, D., & Edmeades, G. (2014). *Crop yields and global food security: Will the yield increase continue to feed the world?* ACIAR Monograph No. 158. Canberra, Australia: Australian Centre for International Agricultural Research. xxii + 634 pp.

Gathala, M. K., Ladha, J. K., Saharawat, Y. S., Kumar, V., Kumar, V., & Sharma, P. K. (2011). Effect of Tillage and crop establishment methods on physical properties of a medium-textured soil under a seven-year rice − Wheat rotation. *Soil Science Society of America Journal*, *75*, 1851–1862. https://doi.org/10.2136/sssaj2010.0362

Ghimire, B., & Timsina, D. (2015). Analysis of chlorophyll content and its correlation with yield attributing traits on early varieties of maize (*Zea mays* L.). *Journal of Maize Research and Development*, *1*, 134–145. https://doi.org/10.5281/zenodo.34263

Głąb, T. (2011). Effect of soil compaction on root system morphology and productivity of alfalfa (*Medicago sativa* L.). *Polish Journal of Environmental Studies*, *20*, 1473–1480.

Gordon, L. (2013). *Using classification and regression trees (CART) in SAS Enterprise Miner for applications in public health (Paper No. 089–2013)*. Lexington, KY: University of Kentucky. Retrieved from http://support.sas.com/resources/papers/proceedings13 /089–2013.pdf. Accessed May 21, 2018.

Han, M., Okamoto, M., Beatty, P. H., Rothstein, S. J., & Good, A. G. (2015). The genetics of nitrogen use efficiency in crop plants. *Annual Review of Genetics*, *49*, 269–289. https://doi.org/10.1146/annurev-genet-112414-055037

Hengl, T. (2007). *A practical guide to geostatistical mapping of environmental variables*. Luxembourg: European Commission.

Hui, K. C. F., Mueller, S., & Welsh, A. H. (2016). *Regularized PQL for Joint Selection in GLMMs*.

Imoloame, E., & Omolaiye, J. (2017). Weed infestation, growth and yield of maize (*Zea mays* L.) as influenced by periods of weed interference. *Advances in Crop Science and Technology*, *05*(02), 1–14. https://doi.org/10.4172/2329-8863.1000267

Jaetzold, R., Schmidt, H., & Hornetz Berthhold, S. C. (2010). *Farm Management Handbook: Natural Conditions and Farm Management; Atlas of Agro-ecological Zones, Soils and Fertilizing, Kakamega & Vihiga County*. Kenya/Nairobi, Kenya: Ministry of Agriculture/Germany Agency for Technical Cooperation

Karugia, J. T. (2003). *A Micro level analysis of agricultural intensification in Kenya: The case of food staples*. Nairobi, Kenya: University of Nairobi.

Keating, B. A., Carberry, P. S., Bindraban, P. S., Asseng, S., Meinke, H., & Dixon, J. (2010). Eco-effi cient Agriculture : Concepts, Challenges, and Opportunities. *Crop Sci.*, *50*, 109–119. https://doi.org/10.2135/cropsci2009.10.0594

Kihara, J., Nziguheba, G., Zingore, S., Coulibaly, A., Esilaba, A., Kabambe, V., … Huising, J. (2016). Understanding variability in crop response to fertilizer and amendments in sub-Saharan Africa. *Agriculture, Ecosystems and Environment*, *229*, 1–12. https://doi.org/10.1016/j.agee.2016.05.012

KNBS (2010). *The 2009 Kenya Population and Housing Census - Population Distribution by Age, Sex and Administrative Units*. [WWW Document]. http://www.knbs.or.ke/index.php?option=com_phocadownload&view=category&download=584:volume-1c-population-distribution-by-age-sex-and-administrative-units&xml:id=109:population-and-housing-census-2009&Itemxml:id=599

Krupnik, T. J., Uddin, Z., Timsina, J., Yasmin, S., Hossain, F., Al, A., … Mcdonald, A. J. (2015). Untangling crop management and environmental influences on wheat yield variability in Bangladesh: An application of non-parametric approaches. *AGSY*, *139*, 166–179. https://doi.org/10.1016/j.agsy.2015.05.007

Lindquist, J. L., Evans, S. P., Shapiro, C. A., & Knezevic, S. Z. (2010). Effect of nitrogen addition and weed interference on soil nitrogen and corn nitrogen nutrition. *Weed Technology*, *24*, 50–58. https://doi.org/10.1614/WT-09-070.1

Liu, Z., Yang, X., Hubbard, K. G., & Lin, X. (2012). Maize potential yields and yield gaps in the changing climate of northeast China. *Global Change Biology*, *18*(11), 3441–3454. https://doi.org/10.1111/j.1365-2486.2012.02774.x

Lobell, D. B., Cassman, K. G., & Field, C. B. (2009). Crop yield gaps: Their importance, magnitudes and causes. *Annual review of environment and resources*, *34*, 179-204.

Mackay, I., Horwell, A., Garner, J., White, J., McKee, J., & Philpott, H. (2011). Reanalyses of the historical series of UK variety trials to quantify the contributions of genetic and environmental factors to trends and variability in yield over time. *Theoretical and Applied Genetics.*, *122*, 225–238. https://doi.org/10.1007/s00122-010-1438-y

Masood, T., Gul, R., Munsif, F., Jalal, F., Hussain, Z., Noreen, N., & Khan, H. H. (2011). Effect of different phosphorus levels on the yield and yield components of maize. *Sarhad Journal of Agriculture*, *27*, 167–170.

Mavuthu, A. K. (2017). *Effect of the national accelerated agricultural inputs access subsidy program on fertilizer usage and food production in Kakamega County, Western Kenya* (Doctoral dissertation). Minnesota, MN : Walden University.

MEMR (2013). *Vihiga district environment action plan 2009–2013*. Nairobi, Kenya: Ministry of Environment and Mineral Resources.

Mi, C., Zhang, X., Li, S., Yang, J., Zhu, D., & Yang, Y. (2011). Assessment of environment lodging stress for maize using fuzzy synthetic evaluation. *Math. Comput. Model.*, *54*, 1053–1060. https://doi.org/10.1016/j.mcm.2010.11.035

NAAIAP (2014). *Soil suitability evaluation for maize production in Kenya*. Nairobi, Kenya: Ministry of Agriculture, Livestock and Fisheries.

Neumann, K., Verburg, P. H., Stehfest, E., & Muller, C. (2010). The yield gap of global grain production: A spatial analysis. *Agricultural Systems*, *103*, 316–326. https://doi.org/10.1016/j.agsy.2010.02.004

O'Keeffe, K. (2009). *Maize Growth and Development*. NSW: NSW Dept. of Primary Industries.

Oh, J. H., Thor, M., Olsson, C., Skokic, V., Jörnsten, R., Alsadius, D., … Deasy, J. O. (2016). A factor analysis approach for clustering patient reported outcomes. *Methods of Information in Medicine*, *55*, 431–439. https://doi.org/10.3414/ME16-01-0035

Ohshiro, A., & Ueda, S. (2018). Comparative study of explanatory factor analysis for construction of Clinical Research education model. *International Journal of Computer Science and Network Security*, *18*, 27–30.

Okumu, M. O., Asten, P. J. A. V., Kahangi, E., Okech, S. H., Jefwa, J., & Vanlauwe, B. (2011). Production gradients in smallholder banana (cv. Giant Cavendish) farms in Central Kenya. *Scientia Horticulturae*, *127*(4), 475–481. https://doi.org/10.1016/j.scienta.2010.11.005

Oloo, T. O. K., Ranabhat, N. B., & Gemenet, D. C. (2013). Potentials of hybrid maize varieties for small-holder farmers in Kenya: A review based on Swot analysis. *African Journal of Food, Agriculture, Nutrition and Development*, *13*, 1–26.

One Acre Fund (2016). *Optimizing maize variety adoption and performance 2015 trial report*. Nairobi, Kenya: ONE ACRE FUND.

Oseko, E., & Dienya, T. (2015). *Fertilizer consumption and Fertilizer Use by Crop (FUBC) in Kenya*. Nairobi, Kenya: Study Conducted For Africafertilizer.Org.

Page, E. R., Cerrudo, D., Westra, P., Loux, M., Smith, K., Foresman, C., … Swanton, C. J. (2012). Why early season weed control is important in maize. *Weed Science*, *60*(3), 423. https://doi.org/10.1614/WS-D-11-00183.1

Pansu, M., & Gautheyrou, J. (2006). *Handbook for Soil Analysis; Mineralogical, Organic and Inorganic Methods, New York*.

Ray, D. K., Mueller, N. D., West, P. C., & Foley, J. A. (2013). Yield trends are insufficient to double global crop production by 2050. *PLoS ONE*, *8*(6), 1–7. https://doi.org/10.1371/journal.pone.0066428

Reid, A., Gonzalez, V., Sikkema, P. H., Lee, E. A., Lukens, L., & Swanton, C. J. (2014). Delaying weed control lengthens the anthesis-silking interval in maize delaying weed control lengthens the anthesis-silking interval in maize. *Weed Science*, *62*, 326–337. https://doi.org/10.1614/WS-D-13-00099.1

Ronner, E., Descheemaeker, K., Almekinders, C. J. M., Ebanyat, P., & Giller, K. E. (2018). Farmers' use and adaptation of improved climbing bean production practices in the highlands of Uganda. *Agric. Ecosyst. Environ.*, *261*, 186–200. https://doi.org/10.1016/j.agee.2017.09.004

Roel, A., Firpo, H., & Plant, R. E. (2007). Why do some farmers get higher yields? Multivariate analysis of a group of Uruguayan rice farmers. *Computers and Electronics in Agriculture*, *58*, 78–92. https://doi.org/10.1016/j.compag.2006.10.001

Sawasawa, H. (2003). *Crop yield estimation: Integrating RS, GIS and management factors: A Case Study of Birkoor and Kortgiri Mandals – Nizamabad District, India*. Enschede, The Netherlands: International Institute for Geo-information Science and Earth Observation Enschede.

Senbayram, M., Gransee, A., Wahle, V., & Thiel, H. (2015). Role of magnesium fertilisers in agriculture: Plant – Soil continuum. *Crop and Pasture Science*, *66*, 1219–1229. https://doi.org/10.1071/CP15104

Shah, A. N., Tanveer, M., Rehman, A., & Anjum, S. A. (2017). Lodging stress in cereal — effects and management : an overview. *Env. Sci Pollut Res*, *24*, 5222–5237. https://doi.org/10.1007/s11356-016-8237-1

Shafi, M., Bakht, J., Ali, S., Khan, H., Khan, M. A., & Sharif, M. (2012). Effect of planting density on phenology, growth and yield of maize (Zea mays L.). *Pakistan J. Bot.*, *44*, 691–696.

Sheahan, M., Black, R., & Jayne, T. S. (2012). Are Farmers Under-Utilizing Fertilizer? Evidence from Kenya (No. 1007-2016-79766).

Sherpherd, G. T. (2010). *Visual soil assessment-field guide for Maize*. Rome, Italy.

Sims, B., Corsi, S., Gbehounou, G., Kienzle, J., Taguchi, M., & Friedrich, T. (2018). Sustainable weed management for conservation agriculture: Options for smallholder farmers. *Agriculture*, *8*, 1–20. https://doi.org/10.3390/agriculture8080000

Sultan, B., Baron, C., Dingkuhn, M., Sarr, B., & Janicot, S. (2005). Agricultural impacts of large-scale variability of the West African monsoon. *Agricultural and Forest Meteorology*, *128*, 93–110. https://doi.org/10.1016/j.agrformet.2004.08.005

Tamene, L., Mponela, P., Ndengu, G., & Kihara, J. (2016). Assessment of maize yield gap and major determinant factors between smallholder farmers in the Dedza district of Malawi. *Nutrient Cycling in Agroecosystems*, *105*(3), 291–308. https://doi.org/10.1007/s10705-015-9692-7

Tariq, A., Anjum, S. A., Randhawa, M. A., & Ullah, E. (2014). Influence of zinc nutrition on growth and yield behaviour of maize (*Zea mays* L.) Hybrids. *American Journal of Plant Sciences*, *5*, 2646–2654.

Therneau, T. M., & Atkinson, E. J. (2015). *An introduction to recursive partitioning using the RPART routines*.

Tittonell, P., Shepherd, K. D., Vanlauwe, B., & Giller, K. E. (2008). Unravelling the effects of soil and crop management on maize productivity in smallholder agricultural systems of western Kenya — An application of classification and regression tree analysis. *Agric. Ecosyst. Environ.*, *123*, 137–150. https://doi.org/10.1016/j.agee.2007.05.005

Tobergte, D. R., & Curtis, S. (2013). Yield and yield components: A practical guide for comparing crop management practices. *Journal of Chemical Information and Modeling*, *53*, 1689–1699. https://doi.org/10.1017/CBO9781107415324.004

Tremblay, N., Bouroubi, Y. M., Bélec, C., Mullen, R. W., Kitchen, N. R., Thomason, W. E., … Ortiz-monasterio, I. (2012). Corn Response to Nitrogen is Influenced by Soil Texture and Weather. *Agron. J.*, *104*, 1658–1671. https://doi.org/10.2134/agronj2012.0184

Uhe, P., Philip, S., Kew, S., Shah, K., Kimutai, J., Mwangi, E., … Otto, F. (2018). Attributing drivers of the 2016 Kenyan drought. *International Journal of Climatology*, *38*, e554–e568. https://doi.org/10.1002/joc.5389

Umeri, C., Moseri, H., & Onyemekonwu, R. (2015). Effects of Nitrogen and Phosphorus on the Growth Performance of Maize (Zea mays) in Selected Soils of Delta State. *Nigeria. Adv. Crop Sci. Technol.*, *04*, 1–3. https://doi.org/10.4172/2329-8863.1000207

Usman, J. (2017). Technical efficiency in rain-fed maize production in Adamawa state Nigeria: Stochastic approach. *International Journal of Environmental and Agriculture Research*, *3*, 67–73.

van Bussel, L. G. J., Grassini, P., Van Wart, J., Wolf, J., Claessens, L., Yang, H., … van Ittersum, M. K. (2015). From field to atlas: Upscaling of location-specific yield gap estimates. *Field Crops Research*, *177*, 98–108. https://doi.org/10.1016/j.fcr.2015.03.005

van Ittersum, M. K., Cassman, K. G., Grassini, P., Wolf, J., Tittonell, P., & Hochman, Z. (2013). Yield gap analysis with local to global relevance—A review. *Field Crops Research*, *143*, 4–17. https://doi.org/10.1016/j.fcr.2012.09.009

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

Yan, P., Chen, Y., Dadouma, A., Tao, Z., & Sui, P. (2017). Effect of nitrogen regimes on narrowing the magnitude of maize yield penalty caused by high temperature stress in North China Plain. *Plant, Soil Environ.*, *63*,131–138–138. https://doi.org/10.17221/6/2017-PSE

Yengoh, G. T. (2012). Determinants of yield differences in small-scale food crop farming systems in Cameroon. *Agriculture and Food Security*, *1*, 19. https://doi.org/10.1186/2048-7010-1-19

Yong, A. G., & Pearce, S. (2013). A Beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology*, *9*, 79–94. https://doi.org/10.20982/tqmp.09.2p079

How to cite this article: Munialo S, Dahlin AS, Onyango M. C, Oluoch-Kosura W, Marstorp H, Öborn I. Soil and management-related factors contributing to maize yield gaps in western Kenya. *Food Energy Secur*. 2019;00:e189. https://doi.org/10.1002/fes3.189