Refining fertilizer recommendations for smallholder maize production systems in western Kenya

Muhati Stephen Ichami
A80/53327/2018

A thesis submitted for the fulfilment of the degree of
Doctor of Philosophy in Soil Science in the Department of Land Resource Management
and Agricultural Technology (LARMAT) of the University of Nairobi

2021

Declaration

This Ph.D. thesis is my original work and has not been submitted in part or entirely for award of degree in any other university

Signature.. Date 16/11/2021 ....

Muhati Stephen Ichami

A80/53327/2018

This proposal is submitted to the university with the approval of my supervisors as;

Signature.. Date 18/11/2021 ....

Professor George N. Karuku

Department of Land Resource Management and Agricultural Technology (LARMAT)

University of Nairobi

P.O BOX 29053 (00625), Kabete

Signature.. Date 23/Nov/2021 ....

Dr. Fredrick O. Ayuke

Department of Land Resource Management and Agricultural Technology (LARMAT)

University of Nairobi

P.O BOX 29053 (00625), Kabete

Signature .... Date 23 November 2021 ....

Dr. Keith D. Shepherd

World Agroforestry Centre (ICRAF)

P.O.BOX 30677 (00200), Nairobi

United Nation Avenue

Nairobi, Kenya.

Table of contents

List of Figures

List of Tables

List of Abbreviations and Acronyms

AEZ	Agro-Ecological Zones
ANOVA	Analysis of Variance
APNI	African Plant Nutrition Institute
BV	Plant Biomass or Plant BioVolume
CEC	Cation Exchange Capacity
CPI	Crop Performance Indicators
DRIS	Diagnosis Recommendation Intergrated System
DSM	Digital Soil Mapping
EV	Explained Variation
FR	Fertilizer Response
FURP	Fetrilizer Use Recommendation Project
GY	Grain Yield
ICRAF	World Agroforestry Centre
kg ha	kilogram per hectar
LDSF	Land Degradation Sampling Framework
LHS	Land Health Surveillence
ln	natural log
MANOVA	Multivariate Analysis of Varience
MFA	Multiple Factor Analysis
N-AE	Agronomic Use Efficiency of Nitrogen
PBFS	Population Based Farm Survey
PC	Principal Component
PCA	Princpal Component Analysis
PMM	Predictive Mean Matching
RE	Random Effects
SFG	Soil Fertility Gradient
SOC	Soil Organic Carbon
SSA	Sub-Sahara Africa
ST	Soil Type
TSBF	CIAT
VIP	Variable Importance Projection
WRB	World Reference Base

Acknowledgment

I thank God for his abundant grace, strength, good life, and health in this very, very long academic journey. This doctoral thesis would not have been possible without the support and encouragement of several people. Secondly, I wish to express my profound gratitude to my Ph.D. Supervisors; Dr. George N. Karuku, Dr. Fredrick O. Ayuke, Dr. Keith D. Shepherd, and Prof. Geoffrey Kironchi for their encouragement, support, and constructive comments, and guidance in the completion of my Ph.D. thesis.

My sincere gratitude goes to World Agroforestry Centre (ICRAF), Tropical Soil Biology and Fertility Institute (TSBF- CIAT), and Wageningen University for the financial support through the African Soil Information Service (AfSIS) project to pursue my Ph.D. studies. I wish to extend my appreciation Prof. Lijbert Brussaard, Prof. Ellis Hoffland and Dr. Jetse Stoorvogel of Wageningen University. The unlimited cooperation and financial support extended by the African Plant Nutrition Institute (APNI) through Dr. Shamie Zingore in carrying out the field research work is very much appreciated.

I acknowledge the support from staff and colleagues at ICRAF and TSBF- CIAT. Dr. Andrew M. Sila for assistance and guidance in Data Processing and Analysis, and your general encouragement in life. The entire ICRAF spectral family, for helping me during soil and plant samples characterization. Dr. Nteranya Sanginga, Dr. Jereon Huising, Dr. Peter Okoth, and staff of TSBF- CIAT for addressing all logistical issues, while conducting field activities in western Kenya.

I am highly indebted to the agricultural extension officers of the Ministry of Agriculture from Siaya and Kakamega Counties – Mr. Mark Opiyo, Mr. Seth Owido who were particularly helpful during field data collection. They not only helped in conducting farm surveys but in administering questionnaires and provided companionship during fieldwork. My field team include Mr. Boniface Pandi, Mzee Julius Okendo (now deceased), Ms. Maureen Achieng for their support in conducting field activities. To all the smallholder farmers from Siaya and Kakamega counties, who allowed me to conduct farm surveys in their maize fields. Since I cannot mention each individually, I wish to express my appreciation for your support. A big thank you, to all those who have contributed in diverse ways to the successful completion of this Ph.D. study, but whose names have not been mentioned.

I would like to express my heartfelt gratitude to my family. Words cannot express my gratitude to you all for your daily prayers, encouragement, patience, and love. Your daily inspiration and emotional support throughout this study have been tremendous. I thank God for his abundant grace, strength, good life and health in this very, very long academic journey.

Abstract

In western Kenya, soil nutrient depletion is one of the main problems that has led to declining crop yield. Agricultural intensification through the judicious application of fertilizers has been considered amongst mitigation options for these smallholder farming systems with an average land size of less than 3.0 ha. The blanket fertilizer recommendations used in this region, have led to poor response to the fertilizer applied and low nutrient use efficiency. These recommendations do not take into account the spatial variability occurring at the local level across the smallholder landscape. Furthermore, the methods used to diagnose soil nutrient constraints are inefficient, because they do not take into account the spatial extent to which the nutrient deficiencies occur. Digital Soil Mapping (DSM) technique and the Population-Based Farm Survey (PBFS) approach are promising strategies that can help address this problem though they have not been fully exploited for smallholder farming systems.

The main objective of this study was to develop and test nutrient management strategies that could be used to improve fertilizer recommendation using the DSM technique and the PBFS approach. The approach was tested to provide site-specific nutrient diagnostics and provide management recommendations in heterogeneous smallholder farming systems. First, evaluation of Fertilizer Response (FR) – a response ratio, and Agronomic Nutrient Use Efficiency (N-AE) was conducted using fertilizer trial data. Meta-analysis technique was employed to identify key factors that influence FR and N-AE in smallholder farming systems. The results indicated soil, climate, and management factors could explain only small amounts (< 30 %) of variation in FR and N-AE. Soil pH, phosphorus (P), texture, and rainfall had significant (P <0.001), but low levels of power in explaining variation in FR and N-AE. This implied that strategies for refining the blanket fertilizer recommendations should include soil-based information, but soil testing needs to be accompanied by nutrient response trials. Secondly, the utility of using the DSM technique was explored, to determine the optimum scale of using digital soil maps, relevant to nutrient management for maize farming systems. A farm survey was conducted and data on soil properties; soil pH, Soil Organic Carbon (SOC), Total Nitrogen (TN),

Potassium (K), Phosphorus (P), Cation Exchange Capacity (CEC), Calcium (Ca), and Magnesium (Mg), Grain Yield (GY) and Plant Biovolume (BV) were collected. Data on the soil properties and crop responses (GY and BV) were analyzed using Step-wise Multiple Linear Regression (SMLR) analysis and geostatistical techniques. The results showed high variability in GY, with 32 % of the observed variation being accounted for by the underlying soil properties. SOC was identified as the key driver of crop response to fertilizer application in the study area. Moderate spatial dependencies for SOC with an effective distance of 523 m were observed. The lower nugget value (0.0542) was indicative of short-distance spatial variability in soil properties. A threshold scale of 250 m was proposed, below which, local growing conditions within the study area were captured, implying that a soil nutrients map with a resolution < 250 m would capture the local variability. Lastly, a sampling approach on a population-based survey of smallholder maize fields was tested to diagnose soil nutrient constraints rather than the conventional agronomic trials. Soil test values were established using Cate-Nelson Analysis (1978) for NPK, which were used to define cases on nutrient constraints. In these study, three aspects are considered; evaluation of FR and N-AE to guide nutrient management strategies, the use of DSM techniques to provide fertilizer recommendations at a refined spatial scale, and utility of PBFS for diagnosis of nutrient limitations in smallholder farming systems. The main finding of the study includes: (*i*) FR and N-AE were highly variable in smallholder maize fields of western Kenya, (*ii*) SOC was the key soil factor that captured local spatial variability on farms. Thus, 250 m was the optimum soil sampling distance for nutrient management based on the spatial range of SOC. This study demonstrated that soil nutrient maps are useful tools, which can be implemented in strategies aimed at a refined fertilizer recommendation across SSA. The utility of DSM and the new PBFS approach has the potential for providing site-specific diagnostics to guide nutrient management decisions. Successfully developing such an integrated soil-based diagnostic system is warranted, and the wider application will be instrumental for refining fertilizer recommendation across maize smallholder agroecosystem systems.

Chapter One

Introduction

1.1 Background information

Smallholder farming systems play a pivotal role in crop production and are envisaged as the main actors in achieving food security and improving rural livelihoods (Tscharntke *et al.*, 2012). For this reason, several initiatives, which include the Abuja Conference (2006), African Green Revolution Alliance (2007), Global Food Security (2013), and Global Soil Security (2015), have sought to address the multifaceted problems facing these systems (Okoth *et al.*, 2011; McBratney *et al.*, 2014; Minasny *et al.*, 2017). In western Kenya, soil nutrient depletion is one of the main problems that has led to declining crop yields (Waswa *et al.*, 2013). Judicious nutrient management through the application of fertilizers is a strategy that could lead to increased production and productivity (Morris *et al.*, 2007; Bationo *et al.*, 2012). However, the production of adequate food for the ever-growing population in smallholder farming systems remains a major challenge (Drechsel *et al.*, 2001; UN, 2013), especially for staple cereal crops like maize.

Maize accounts for 30 to 50% of low-income household expenditures in sub-Sahara Africa (SSA), and nutrient depletion has exacerbated low productivity (Cairns *et al.*, 2013). Blanket fertilizer recommendations have led to poor response to fertilizer application and low nutrient use efficiency, which occur when applied fertilizer does not meet the maize nutrient requirements (Kihara *et al.*, 2016; Vanlauwe *et al.*, 2017). Blanket recommendations are general nutrient management guidelines, where a single fertilizer rate is applied uniformly across a broad geographic location regardless of the crop types. These recommendations do not take into account spatial variability occurring at the farm level. Understanding the magnitude and patterns of spatial variability in soil properties at this level is instrumental for fertilizer management, which leads to increased nutrient use efficiency (Bhatti *et al.*, 1998; Miller *et al.*, 2015). The spatial information may be useful for providing guidelines to refine fertilizer recommendations in smallholder farming systems. However, the refinement of fertilizer recommendations is hampered by the lack of detailed soil data (Sanchez *et al.*, 2009). Additionally, it is not clear, at what spatial resolution should the soil nutrient maps be provided, for them to be used effectively as decision support tools. In Kenya, the available soil maps are spatially coarse (1:250,000), and new more detailed soil surveys are rare.

The lack of rigorous methods for collecting spatial information on soil health at the farm level impedes the transfer of site-specific information needed by smallholder farmers for targeted nutrient management decisions (Shepherd & Walsh, 2007; Shepherd *et al.*, 2015). Conventional methods for nutrient diagnostics, do not take into account the spatial extent to which nutrient deficiencies occur (Huising *et al.*, 2011; Kihara *et al.*, 2016). Furthermore, there is a general lack of crop response data to calibrate soil information to soil test values for nutrient diagnosis in SSA (Vanlauwe *et al.*, 2017). Limited studies have been conducted to address these knowledge gaps, especially for smallholder farming systems in SSA. A need, therefore, arises, for developing novel methods that can be used for the diagnosis of soil nutrient constraints and mapping the spatial extent.

There is potential for a novel nutrient diagnostic approach, the population-based farm surveys (PBFS), which is anchored on the principles of Land Health Surveillance (LHS) (Vågen *et al.*, 2012; Beedy *et al.*, 2015; Shepherd *et al.*, 2015). Population-Based Survey surveillance has become popular in monitoring disease patterns (epidemiology) within human populations and designing targeted curative medical interventions (Lipscombe *et al.*, 2018; Frederiksen *et al.*, 2019). The PBFS approach uses a combination of tools such as rigorous ground sampling schemes, proximal techniques of infrared spectroscopy for rapid soil and plant testing, and statistical models, which provide population-based estimates from hierarchical data (Vågen *et al.*, 2012; Shepherd *et al.*, 2015). The PBFS approach can work in tandem with DSM, and provide an important synergy as a nutrient management tool. The DSM technique and PBFS approach have the potential of obtaining insight into the nutrient constraints, and studying patterns of spatial variation of soil properties at the farm level (Snoeck *et al.*, 2010; Hengl *et al.*, 2017). However, these tools have never been fully exploited for nutrient management in smallholder farming systems.

This study used the surveillance approach by surveying geographical variability in crop nutrition status and used the data to provide guidelines in deciding on nutrient management strategies to fit local fertility needs in western Kenya. The main objective of the study was to develop an integrated approach to assess nutrient constraints and provide guidelines for fertilizer recommendations for maize smallholder farming landscapes in western Kenya.

1.2 Problem Statement

The reduction in the average farm size dictates that most of the crop production to feed the increasing population would be achieved through intensification rather than expansion of farming land (Pradhan *et al.*, 2015; Wortmann *et al.*, 2017). Adequate and efficient use of fertilizer is, therefore, one of the main ingredients, since farmers apply small quantities (< 50 kg ha^{-1}) of fertilizer (Duflo *et al.*, 2008). There is a lack of evidence-based strategies, that would improve farmers' decisions concerning the proper application of nutrients to soils for optimal maize yields and sustainable utilization of scarce land resources.

Smallholder farmers face problems when deciding which type and rate of fertilizer to apply to their crops, given the local conditions and little available resources at the farm level. Typically, the farmers rely on their own experience from previous years, their expert knowledge, and blanket fertilizer recommendations that are spatially coarse (*e.g.* agro-ecological based) that result in poor fertilizer use efficiency (FURP, 1994; Mairura *et al.*, 2007; Bekunda *et al.*, 2010). The smallholder farmers do not allocate the scarce fertilizer resources efficiently, which often encourage unneeded nutrient applications and, in some cases, fail to include the limiting nutrients (Vanlauwe & Giller, 2006; Namonje-Kapembwa *et al.*, 2015). The lack of laboratory infrastructure and the high cost of soil analysis, limit the correct diagnosis of limiting soil nutrient deficiencies (Shepherd & Walsh, 2007; Vanlauwe *et al.*, 2017). Accurate diagnosis of the soil constraints can mitigate the challenges and improve the efficiency of the small amounts of fertilizer applied (Okalebo *et al.*, 2006; Muhati *et al.*, 2011). This study sought to develop a spatially-explicit approach for the diagnosis of soil nutrient constraints in smallholder farming systems.

The efficiency of blanket recommendation is often complicated by high soil spatial variability within smallholder landscapes (Jayne *et al.*, 2008; Ngome *et al.*, 2013).

Knowledge of the spatial patterns of smallholder farming landscapes may play a pivotal role in improving a farmer's decisions. These landscapes display tremendous spatial variability where soil properties vary markedly from field to field, farm to farm, and within the landscape level (Okeyo *et al.*, 2009; Tesfahunegn *et al.*, 2011; Tittonell *et al.*, 2015). The problem of variability can be addressed through targeted soil nutrient management strategies tailored to specific areas within the landscape where there is a kind of homogeneity. There is, however, a general lack of systematic studies identifying the key factors that influence FR and N-AE across smallholder farming systems to inform strategies for fertilizer recommendations.

1.3 Justification of the study

The global food demand is projected to increase by 60% in 2050 and is mainly attributed to the increase in the human population, which is estimated to be 9.5 billion by 2050. Improving soil nutrient status for food and income is a challenge to smallholder farmers, regardless of their landholdings and location. Several factors contribute to the decrease in crop productivity and include; lack of non-farm income sources, soil fertility depletion, and low levels of fertilizer use. Other challenges faced by farmers include diagnosis of soil nutrient constraint include; the high cost of chemical soil analysis, the coarse resolution of the available soil maps, and poor laboratory infrastructure across smallholder landscapes. Developing nutrient strategies for addressing challenges of soil nutrient depletion is necessary if the crop production increase is to be realised.

Current decisions on nutrient management for the smallholder farms in western Kenya are based on blanket fertilizer recommendations. The problem with these recommendations is that, farmers do not change or apply variable amounts of fertilizer, depending on their individual needs or what is really limiting on their farms. The high spatial variation, which characterizes the smallholder farms causes poor fertilizer response and low nutrient use efficiencies. There is a great need for incorporating information spatial variability in nutrient management decisions. This study sought to test and develop novel approaches, which may be required for decision support tools and could be included in strategies aimed at refining fertilizer recommendations. Additionally, scientific knowledge should form the basis for making informed management decisions. New techniques of DSM combined with PBFS could be incorporated and used in filtering spatial information for making informed nutrient management decisions for smallholder farms. Fine-scale soil properties mapping is not only needed for soil survey, but also for the management of soil nutrients. Adoption of DSM technique combined with proximal sensing techniques could help identify key soil and plant constraints in SSA in a cost-effective way and scale-up soil fertility replenishment programs. This study sought to test the utility of the DSM technique and PBFS approach for the diagnosis and management of soil constraints in smallholder farming systems of western Kenya.

1.4 Objectives

The main objective of this study was to develop an integrated approach to assess nutrient constraints and provide guidelines for refining fertilizer recommendations for smallholder farming landscapes in western Kenya.

1.4.1 Specific objectives

1.4.1.1. To determine variability in fertilizer response and nutrient use efficiency in smallholder landscapes in Western Kenya.

1.4.1.2. To determine the optimum sampling distance for developing digital nutrient maps for the provision of fertilizer recommendations.

1.4.1.3. To test the applicability of a population-based survey for diagnosis of yield-limiting soil nutrients in smallholder landscapes in Western Kenya,

1.4.1.4. To identify useful covariate data that can be utilized for a population-based survey in combination with DSM techniques for nutrient management in smallholder landscapes.

Chapter 2

Literature review

2.1 Smallholder farming systems

The smallholder farming systems are subsistence in nature with dominantly rain-fed crops and low fertilizer inputs (Waithaka *et al.*, 2007). The farm size is often less than three hectares and the mixed crop-livestock system includes maize (*Zea mays* L.) as the dominant staple crop, usually intercropped with common bean (*Phaseolus vulgaris* L.) and cattle (*Bos taurus*) and goats (*Capra aegagrus hircus*) for livestock (Diwani *et al.*, 2013). Approximately, 40-80% of the farming systems are mixed farming and are efficient for the production of maize, milk and meat (Herrero *et al.*, 2010). The smallholders are supported by government agricultural extension services to manage both their crops and livestock. The farming system is a kind of synergy as crops are used to provide feed for livestock, while animals produce manure to replenish soil nutrients to the farm (Giller *et al*., 2006). Several studies have reported average maize yield levels, under current local conditions, and using conventional farming practices that range from 400 to 3000 kg ha^{-1} for a long rain season (Vanlauwe *et al.*, 2014; Kihara *et al.*, 2016).

Removal of crop residue from the smallholder farms has been the main cause of nutrient mining, leading to soil fertility degradation in these smallholder farming systems (Smaling *et al.*, 1993; Tully *et al.*, 2015). High poverty levels, shortened fallow periods and low fertilizer inputs have exacerbated the decline in soil nutrients for these smallholder landscapes (Bekunda *et al.*, 2010; Bationo *et al.*, 2012; Namonje-Kapembwa *et al.*, 2015).

2.2 Soil fertility degradation in sub-Sahara Africa

Soil fertility degradation is a phenomenon characterized by the loss of productive capacity of land due to the decline in soil nutrient supply to crops (Stockdale *et al.*, 2013). In SSA, the decline in soil fertility continues to be a major concern to scientists and policymakers, due to its direct implication on food security and rural development (Bekunda *et al.*, 2010). Previous studies have reported an annual nutrient mining of 30 kg ha^{-1} per year of N, P, K in 85% of farmlands during the 2002 – 2004 cropping seasons

(Bationo *et al.*, 2003). Of this amount, 40 % of the farms had nutrient mining rates exceeding 60 kg ha^{-1} of N, P, K per year and were characterized as a severe rate (Bationo *et al.*, 2003).

In Kenya, the potential threat of soil fertility degradation to food security is well documented (Stoorvogel *et al.*, 1993; Tully *et al.*, 2015). For example, a study by Tully *et al*. (2015) indicates a decline in maize yield from 5.1 Mg ha^{-1} to 3.2 Mg ha^{-1} and N deficiency was the main cause. According to the study by Tully *et al.* (2015), the mean N balance was -88 kg N ha^{-1} in 2012 and -112 kg N ha^{-1} for 2013 N in the smallholder farms. Phosphorus, K, Sulphur (S) and other micro-nutrients also limit crop production in Western Kenya (Kihara *et al.*, 2016, 2017). Studies have shown unless concrete steps are taken, to curb the devastating consequences of soil fertility degradation in SSA, attaining food security in the near future will remain a myth (Vanlauwe & Giller, 2006).

2.3 Evolution of fertilizer recommendations in sub-Sahara Africa

Strategies of replenishing soil nutrients in SSA have evolved over long periods of time (Mafongoya *et al.*, 2006; Okalebo *et al.*, 2006). In the past decades, shift cultivation and fallowing were the traditional methods used of maintaining soil fertility and replenishing nutrients (Kumwenda *et al*., 1996; Mafongoya *et al*., 2006; Okalebo *et al*., 2004). However, the rapid increase in the population of SSA, projected to 1.2 billion by 2050 (UNPD, 2009; Ray *et al.*, 2013) has led to more demands for food and settlement on large areas of agricultural land, as the practise of shift cultivation has become a non-viable nutrient replenishment strategy (Kumwenda *et al.*, 1996). This has led to agricultural intensification, which involve fertilizer use to replenish soil nutrients (Bekunda *et al*., 2010; Cassman, 1999; Fischer *et al*., 2001), and gave rise to the development of fertilizer recommendation rates.

Earlier strategies of agricultural intensification in the 1980s led to the development of countrywide blanket recommendations and later followed by agro-ecologically based recommendations in the 1990s (FURP, 1994, 1987). The blanket recommendation was based on practical considerations for extension officers to communicate uniform messages rather than based on site-specific needs or the socio-economic circumstances of the poorly resourced smallholder farmers (Barreto & Bell, 1994; Kamanga *et al.*, 2014; Schnier *et al*., 1996). The effectiveness of this blanket recommendation was low because the soil variability was ignored and therefore the recommendations were incompatible with smallholders' resources and needs (Smaling *et al.*, 1992; Schnier *et al.*, 1996). The agro-ecologically based recommendations were aimed at improving the blanket recommendations for the whole country and were developed principally using 70 multi-locational fertilizer trial data collected nationwide, but delineated across agro-ecological zones (Smaling & Van De Weg, 1990; Smaling *et al.*, 1992; FURP, 1994).

The spatially explicit agro-ecologically based recommendation was therefore based on a decision tree considerate of the driving objectives for (*i*) production for home consumption (FURP, 1994) and (*ii*) production for the market (FURP, 1994; Schnier *et al.*, 1996). But still, these recommendations were spatially coarse and often resulted in a mismatch between nutrient application and actual requirements, translating into sub-optimal crop response, low fertilizer use efficiency hence leading to low crop yields (Vanlauwe & Giller, 2006). The main reason being, variability occurring in soil over short distances across the smallholder farming systems (Tittonell *et al.*, 2007a, 2013). The short distance variability led to proposals for site-specific recommendations by several workers (Moyer *et al*., 2012; Snoeck *et al*., 2010; Webb *et al*., 2011). Furthermore, some smallholder farmers were unable to afford fertilizers and instead applied animal manures (Zingore *et al*., 2007). Most studies propose a judicious mix of both organic and inorganic fertilizer within the framework of Integrated Soil Fertility Management (ISFM), thus balancing the chemical, physical and biological properties of the soil to improve the nutrients in the crop production environment (Vanlauwe *et al.*, 2010).

The Diagnosis and Recommendations Integrated System (DRIS) is a method used for interpreting plant tissue analyses so as to rank the importance of the various nutrients limiting optimal plant yields and to estimate the degree to which each of the limiting nutrients are deficient (Walworth & Sumner, 1987; Ramakrishna *et al.*, 2009). This information, in turn indicates remedial steps that can be taken to enhance the growth of the sampled or subsequent crops. The system has been used successfully to diagnose nutritional problems in many crops (Walworth & Sumner, 1987; Angeles *et al.*, 1990;

Bailey *et al.*, 1997; Ramakrishna *et al.*, 2009). Nziguheba *et al.* (2009) successfully applied the DRIS norms to diagnose limiting nutrients using fertilizer trials.

2.4 Spatial variability in smallholder farming systems.

Spatial variability occurs when measurements of a soil property, at different spatial locations exhibit values that differ across the locations. The spatial variability of soil properties can broadly be classified as; inherent and dynamic. The inherent soil spatial variability is due to natural soil-forming processes and its distribution is influenced by the soil-forming factors such as climate, vegetation, time, the geology of parent materials and topography (Persson *et al.*, 2005; Seibert *et al.*, 2007; Tesfahunegn *et al.*, 2011). Dynamic soil spatial variation, on the other hand, occurs as a result of management practices carried out across the smallholder farming systems (Zingore *et al*., 2007). Management choices affect the amount of soil organic matter, soil structure, soil depth, water and nutrient holding capacity, among others (Zingore *et al*., 2007). Soils, however, respond differently to management practices depending on their inherent properties.

Spatial variability of soil properties poses a great challenge to nutrient management strategies (Mzuku *et al.*, 2005). Previous studies have shown the effect of variability of soil properties on crop performance could be detrimental, especially when the fields are patchy (Haileslassie *et al.*, 2005; Kravchenko *et al.*, 2006; Chikuvire *et al.*, 2007; Diarisso *et al.*, 2015). Formulation of fertilizer recommendations for managing soils with highly variable properties, based on a few selected experimental sites, may thus lead to erroneous outcomes (Cassman, 1999; Tittonell & Giller, 2013). Alternative approaches, such as precision agriculture, which can adapt nutrient management practices to the location-specific fertility status have proved to be effective (Cassman, 1999; Tilman *et al.*, 2002).

Different studies have indicated that soil spatial variability is dependent on scale (Cambardella *et al.*, 1994; Tesfahunegn *et al.*, 2011). At the landscape scale, variation in soils can occur due to their relative positions within the landscape, as influenced by parent material and climatic factors. At the farm level, soil variability is due to the effects of management activities (Tittonell *et al*., 2013; Vanlauwe *et al*., 2007).

Attempts by smallholder farmers to address the adverse effects of spatial variation and the negative impact of soil nutrient depletion has led to soil fertility gradients across and within farms. The fertility gradients are attributed to the redistribution of organic matter across the smallholder farms within the landscapes (Rufino *et al*., 2007; Tittonell *et al*., 2013). Maiti *et al.* (2006) claimed that the significant difference of spatial variation (in terms of soils, climate and management) makes it impossible to extrapolate the results of fertilizer recommendations from few sites across farming systems.

2.5 Digital soil mapping for soil nutrient management in smallholder landscapes

Digital Soil Mapping (DSM) is the computer-assisted production of soil property maps or the creation of a geographically referenced soil database generated using field and laboratory observation methods coupled with environmental data through quantitative relationships (Lagacherie & Mcbratney, 2007). Digital soil mapping and the science of geostatistics has been developed for widespread practical application in mapping the variability of soil chemical and physical properties (Mora-Vallejo *et al*., 2008; Sanchez *et al.*, 2009; Snoeck *et al*., 2010).The DSM technique opens new opportunities to describe the spatial variation in soil conditions (McBratney *et al.*, 2003; Stoorvogel, 2014). Recent examples of DSM include the 100m resolution digital soil map of Machakos and Makueni district in Kenya (Mora-Vallejo *et al.*, 2008) and also the new global initiatives at a 30 arc sec resolution (Arrouays *et al.*, 2014; Stoorvogel, 2014). These maps can be used to provide information on the spatial variation of soil properties at a finer resolution (Hengl *et al.*, 2015). However, the mapping units delineated on these maps are essentially taxonomic and bear only limited direct relationship to soil fertility needs. By themselves, the digital soil maps have only minor utility for improving fertilizer recommendations to meet local conditions of smallholder farms.

The application of DSM in nutrient management requires geo-referenced point observations, which can provide additional information on soil fertility parameters. Soil and plant testing provide opportunities for geo-referenced data which can be used in ground-sampling schemes in combination with infrared spectroscopy (Shepherd *et al*., 2015). Ground-sampling schemes can be used to assess spatial variation at the farm level and if successful, DSM could predict soil nutrient conditions for smallholder farms and as a result replace conventional soil testing after the initial farm survey. Where relevant soil maps are unavailable, fertilizer recommendations can only consider the local conditions through the soil and/or plant testing.

2.6 Population-Based Survey Surveillance Approach

Principles of land health surveillance are used to develop this PBFS approach and would, therefore, help in providing guidelines to improve fertilizer recommendations (Shepherd *et al*., 2015). The objective of the surveillance is to make inferences on the sample to the population (Waters & Doyle, 2003). The population, in this case, maybe maize plots of equal sizes (measurements) within the smallholder farms, in which different soil and plant measurements are made to provide data at different scales (hierarchical data).

The applicability of soil testing for fertilizer use recommendation decision support systems is curtailed by high costs of conventional wet-analysis and lack of crop response data for calibrating soil tests to fertilizer use recommendation (Valkama *et al.*, 2009; Muyayabantu *et al.*, 2012). Lack of crop response data is attributed to the high cost of agronomic trials, which have to be conducted several times in space and time to establish reliable fertilizer use management recommendations. Under such situations, the population-based survey approach of small-scale farms has potential as an alternative method (Beedy *et al.*, 2015; Shepherd *et al.*, 2015). In this approach, surveys are conducted on small-scale farms for soil nutrient conditions and maize crop productivity to establish relationships. These relationships between soil test values and maize productivity are used to characterize and understand the variation among small-scale farms(Vågen *et al.*, 2012; Wang *et al.*, 2019). Diagnosis of soil nutrient constraints is made using reference values established from the relationships(Vågen *et al.*, 2012; Shepherd *et al.*, 2015). As a result, the variations across the small-scale farms and soil diagnostics are used as a basis for decision support for fertilizer use recommendations management.

The population-based survey approach becomes feasible due to the development of low-cost analytical techniques of Infrared Spectroscopy (IRS), a useful tool for high-density soil sampling. The ability to rapidly and non-destructively characterize soils using IRS permits high-density sampling and acquisition of information on nutrient status and variation within a target population (*e.g.* farms) (Stenberg *et al.*, 2010). IRS could also be used as an integrative measure of soil quality and employed as a screening tool of soil condition; hence its application to soil variability assessment and monitoring at a broad scale is a promising approach (Shepherd and Walsh, 2007). At the same time, the DSM technique is applied to map and predict soil nutrients based on point-observations from the population- surveys of small-scale farms across the landscape. These advances in the aforementioned techniques have not yet been tested and deployed into operational decision support for fertilizer use management recommendations in Kenya.

Chapter Three

Fertilizer response and agronomic nitrogen use efficiency in Africa smallholder farming systems[1]

Abstract

Improving fertilizer recommendations for farmers is essential to increase food security in smallholder landscapes. Currently, blanket recommendations are provided across agro-ecological zones, although Fertilizer Response (FR) and Nutrient Use Efficiency (N-AE) by maize crops are spatially variable. The objective of this study was to identify factors that could help to refine fertilizer recommendation by analysing the variability in FR and N-AE. Secondary data on on-farm fertilizer studies across Sub-Sahara Africa (SSA) yielded 71 publications. These studies were separated for Kenya and the rest of SSA. FR was expressed as a ratio between fertilized and control maize yield. The variability in FR was studied using a meta-analysis whereas key factors that influence FR and N-AE were studied with linear regression models. On average, the FR was 2 units higher, compared to the control, but it varied considerably from 1 to 28.5 9 (ratio). In the rest of SSA, 18% of the plots were non-responsive with an FR < 1. N-AE ranged from -27 kg dry weight per kg N to 165 kg dry weight kg per N. The main factors affecting N-AE for Kenya were P, silt content, soil pH, and rainfall, whereas only soil pH and texture were important for the rest of SSA. This study, however, indicates that available data on soil, climate and management factors could explain only a small part (< 30%) of the variation in FR and N-AE. Soil pH, P, silt content, and rainfall had significant ($p < 0.01$) but low levels of power in explaining variation in FR and N-AE. The findings indicate that strategies to refine fertilizer recommendation should include information on soil types and soil properties.

3.1 Introduction

The increasing food demand for the growing population in SSA, which is projected at 1.2 billion by 2050, require agricultural intensification with efficient fertilizer use. The current fertilizer recommendations in SSA are often only specified to the level of a region, for instance, an agro-ecological zone (AEZ) or administrative district (*e.g.*,

[1] This chapter is based on a published article in Nutrient Cycling and Agroecosystem Journal Vol. 113 (1): pages 1-19. https://doi.org/10.1007/s10705-018-9958-y

Mowo and Mlingano 1993; FURP 1994; Schnier *et al.* 1996; GoM-MoALD 1994). The fertilizer recommendations for these larger regions are commonly referred to as blanket fertilizer recommendations. However, environmental and management factors vary at short distances in the smallholder landscapes of SSA (Vanlauwe *et al*. 2011; Zingore *et al.* 2007; Tittonell *et al*. 2008). As a result, the blanket fertilizer recommendations are often considered to be of limited relevance to farmers (Tittonell *et al.*, 2013).

The blanket fertilizer recommendations can only be refined if the factors that influence the variation in FR are known. This is becoming increasingly urgent since an increasing number of farmers report decreasing FR for staple food such as maize. Giller *et al.* (2006) introduced the concept of non-responsive soils *i.e.,* soils on which crops do not respond to mineral fertilizer application. However, the causes behind non-responsive soils are not clearly understood. With a better understanding of factors that affect the variability in response to fertilizers, fertilizer recommendations can be improved. There is, however, lack of studies that systematically identify key factors that affect the fertilizer response across smallholder farming systems of SSA. A meta-analysis study pointed to the importance of secondary and micronutrient deficiencies in SSA in low fertilizer responses (Kihara *et al.*, 2017). Multiple studies have shown that the FR varies across smallholder landscapes due to environmental (soil-related and climatic) and management factors. Zingore *et al.* (2007) demonstrated that the low level of soil organic carbon in the maize fields of Zimbabwe led to a poor fertilizer response. Sileshi *et al*. (2008) attributed a high variability in rainfall amounts to low FR. Vanlauwe *et al.* (2016) observed that the poor fertilizer response in maize is a result of unbalanced soil fertilization. However, it remains unclear which are the key factors of variability in FR.

Fertilizer recommendations, both type and amount, can be evaluated using indicators such as FR and N-AE. The FR is defined as the incremental crop yield due to fertilization, independent of the quantity or the type of fertilizers applied. The FR is calculated as the ratio of fertilized crop yield and unfertilized crop yield of a control plot. The FR is a useful concept for identifying, for example, "non-responsive soils ", *i.e.* smallholder farmers' fields where no increase in crop yield is observed after sufficient amounts of fertilizers have been applied (Tittonell *et al*. 2007; Zingore *et al.* 2007; Njoroge *et al.* 2017). The FR can also be used to evaluate the overall effect of fertilizer use across farms in a region. The agronomic nutrient use efficiency is a measure of the crop yield increase for a given amount of nutrient added (Dobermann *et al.*, 2002) and can be used to evaluate the efficiency of a specific nutrient applied. For example, the agronomic nitrogen use efficiency (N-AE) is defined as the incremental crop yield per applied nitrogen – measured in kg dry weight (dw) per kg N.

Soil property maps of relevant variables to fertilizer management are becoming available (Hengl *et al*. 2015, 2017). These maps may help to give a better insight into the spatial variability of nutrient concentrations (Antwi *et al.*, 2016). However, these maps are generally at coarse spatial resolution and are only suitable for guiding recommendations at the regional scale but not at the farm level. Even though high throughput and cost-effective methods for soil analysis are also available (Shepherd & Walsh, 2007; Shepherd, 2010), most smallholder farmers do not have access to soil analysis services at the plot level. Knowledge of key factors that influence FR and N-AE is therefore critical for strategies aimed at improving nutrient management.

This study aimed in identifying key factors that influence FR and N-AE for refining fertilizer recommendations for smallholder farmers. The specific objectives of this study were to: (*i*) quantify the variation in FR and N-AE, and (*ii*) identify key environmental and management factors that influence variability in FR and N-AE. A meta-analysis approach was employed (Hedges *et al.*, 1999; Borenstein *et al.*, 2009) to analyse FR, and regression analysis to understand the driving factors for FR and N-AE.

3.2. Material and Methods

3.2.1 Data collection

Secondary data was collected on agronomic studies conducted between 1980 to 2018. These studies were searched on the internet using Google Scholar, Mendeley, and Web of Science databases. Criteria used for obtaining a set of comparable studies for the analysis were: (*i*) maize cultivated as a monoculture, (*ii*) the experiment conducted on a smallholder farm in SSA, and (*iii*) fertilizer treatments randomly allocated to the plots. The selected treatments included only inorganic N or combinations with inorganic P and or K fertilizer applications. Treatments in which additional organic fertilizers were applied were excluded. A systematic process for the selection of suitable fertilizer studies is presented in a flowchart (Figure 3.1).

3.3 Data extraction and treatment

Data on fertilizer treatments, crop yields, soils, climatic (agro-ecological zones) and management factors were extracted from the selected publications. A database was established with each record representing a treatment plot (Tables 3.1, 3.2 & 3.3).

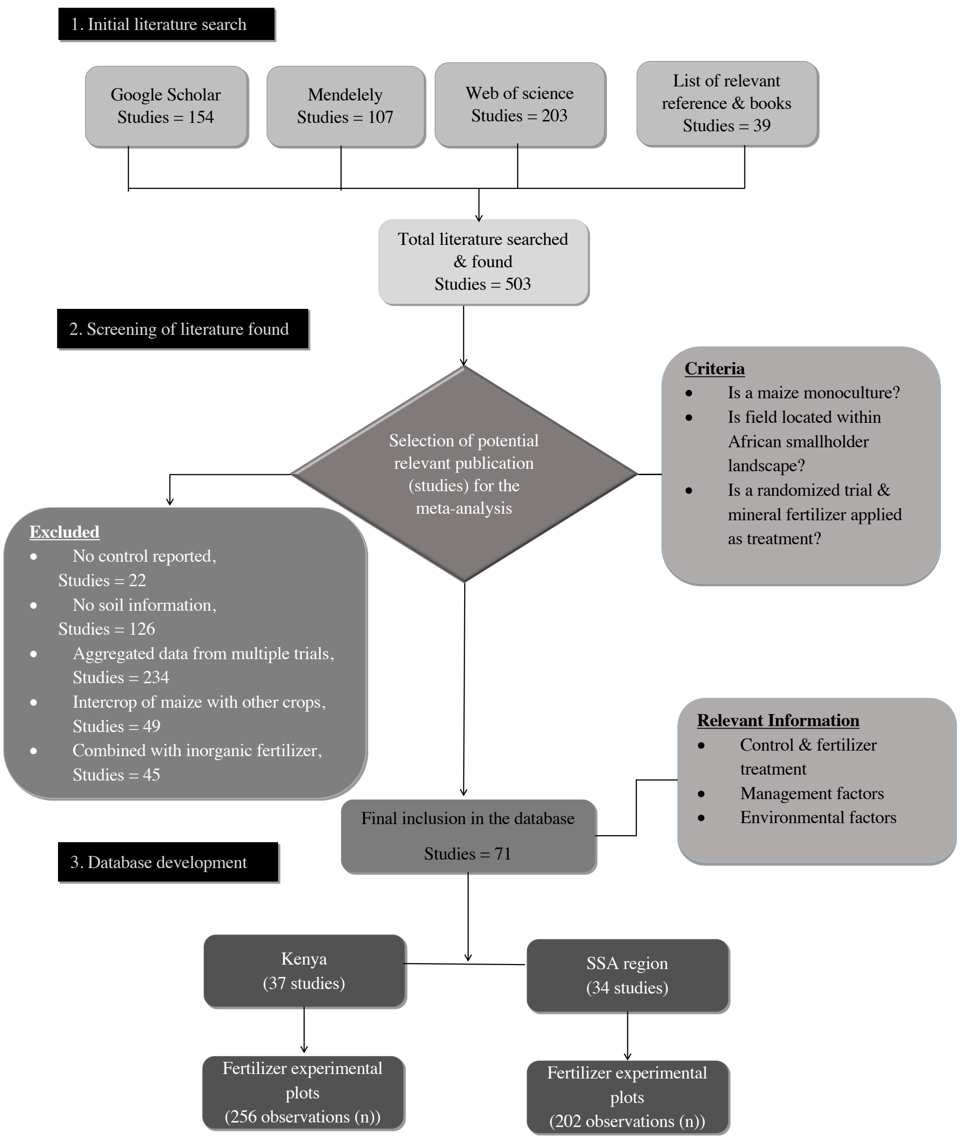

Figure 3.1 Meta-analysis flow chart for selection of studies and database development, n represents the total number of observations. Control was considered as plots where no fertilizer (organic or inorganic) had been applied.

Table 3.1: Crop yield and soil properties, altitude and climatic factors used in the analysis of the on-farm fertilizer experiments (n = number of observations, for the rest of SSA and Kenya) the numbers in   are the number of observations that data imputation was performed for the soil properties.

Variables	**Description**	**Units**	**n**	
			Kenya	**SSA**
	Yield (crop response)			
Control	Mean maize yield for control plot	kg ha^{-1}	202	256
Treatment	Mean maize yield for fertilized plot	kg ha-^{1}	202	256
	Environmental factors (continuous)			
Soil properties analysed prior to fertilizer trial	Soil pH,	-	194 (4)	230 (28)
	Total carbon (SOC)♣	g kg^{-1}	192(5)	181 (38)
	Total nitrogen (TN)♣	g kg^{-1},	155	216
	P-Olsen (P)♣	mg kg^{-1},	105 (19)	122(22)
	P-Bray 2	mg kg^{-1},	135	173
	Exchangeable K (K)♣	cmol kg^{-1}	104(34)	102(48)
	Exch. Ca	cmol kg^{-1}	78(37)	71(93)
	Exch. Mg	cmol kg^{-1}	67	198
	Clay	%	109(36)	152(77)
	Sand	%	86(38)	143(89)
	Silt	%	96(41)	161(67)
Rainfall	average per growing season	mm	198	220
Altitude	Height above sea level	m	186	85

Key: n=number of observations; N = Nitrogen, P = Phosphorus and K= Potassium, ♣ indicates variables that were imputed and the figures in brackets are the number of observations that data imputation was performed for soil properties

Table 3.2: Categorical environmental factors used in the analysis of the on-farm fertilizer experiments for Kenya and the rest of SSA and Kenya.

Variables	Description	Units	n	
			Kenya	SSA
	Soil properties and agro-climatic factors (categorical)			
Soil orders (World Resource Base Reference Soil Groups)	Cambisols		22	-
	Nitisols		68	-
	Vertisols		6	-
	Ferralsols		81	24
	Luvisols		4	-
	Lixisols		41	
	Arcrisols		37	22
	Alfisols		12	4
	Phaeozems		-	8
	Alisols		8	-
Soil textural classes (USDA)	Clay		28	14
	Clay loam		21	-
	Loamy sand		2	8
	Sand		-	64
	Sandy clay		4	6
	Sandy clay loam		6	10
	Sandy loamy		31	55
Agro-ecological zone for Africa (based on Dudal, 1980) and Kenya (based on Jätzold and Kutsch, 1982)	Sub-humid (SSA)		18	-
	Humid (SSA)		6	-
	Lowlands		-	4
	Lower midlands		-	6
	Upper midlands		-	38
	Lower highlands (sub-humid)		-	6
	Lower highlands humid		-	6

Table 3.3: Management factors used in the analysis of the on-farm fertilizer experiments for Kenya and the rest of SSA.

Variables	Description	Units	n	
			Kenya	SSA
	Management factors (continuous)			
Fertilization rate	Amount of N applied	kg N ha^{-1}	254	251
	Management factors (categorical)			
Nutrient applied	N only		149	235
	NPK		49	14
Manager	Farmer		159	234
	Researcher		81	63

Key:
n=number of observations, N = Nitrogen, kg N ha^{-1} = kilogram N per hectare

Typically, experiments included various fertilizer treatments and or multiple seasons and sites. The data contained multiple fertilizer treatments from different experiments. Formally, these observations cannot be considered independent. However, López-López *et al.* (2014) showed that multiple entries from a single experiment are valid and can help to increase the precision of the analysis when using literature data. Controls were considered as treatments with no application of fertilizers either organic or inorganic. The data variables were harmonized by (*i*) converting the reported soil measurement units that were reported different for a given soil, into similar units, (*ii*) reconverting major soil types classified differently to into the common World Resource Base Reference Soil Groups (IUSS Working Group WRB, 2014), and (*iii*) converting measured P-Olsen to P-Bray, because this variable were never measured in the same experiment and to convert to similar metrics. Therefore, published conversion factors were used to estimate P-Olsen from P-Bray 1: P-Olsen = 0.44 P-Bray 1 (Kleinman *et al.*, 2001) and from P-Bray 2 to P - Olsen = 0.79 P-Bray 2 (Wuenscher *et al.*, 2015).

The database still included many missing values because soil descriptions and analytical procedures differed. A flowchart deducing the different steps of estimating the missing data is presented in Figure 3.2. To handle the rest of the missing data on soil properties, the following approaches were used:

i) A pairwise correlation analysis was conducted to establish the correlation among paired soil properties. Paired soil properties with a Pearson correlation coefficient (r) > 0.8 were selected. From this pair, the property with the highest number of missing values was dropped. Prior to dropping the property out, a linear equation was established and used to estimate the missing value of the retained property (with fewer missing values). But, still there were missing soil test values. Thus, where the pair-wise could fill the missing, the next approach was employed.
ii) Predictive Mean Matching (PMM) approach was used to impute the remaining missing soil data using the "*mice*" R-package (Van Buuren & Groothuis-Oudshoorn, 2011). The PMM approach is based on regression analysis and estimates missing values by means of the nearest neighbour (Di Zio & Guarnera, 2009; Vink *et al.*, 2014). This approach was used so that the originality of the soil data and the underlying distribution are maintained (Little & Rubin, 2002; Vink *et al.*, 2014). Remaining missing values (18%) for soil pH, total C, Exch. K, silt and clay were estimated using the PMM approach.

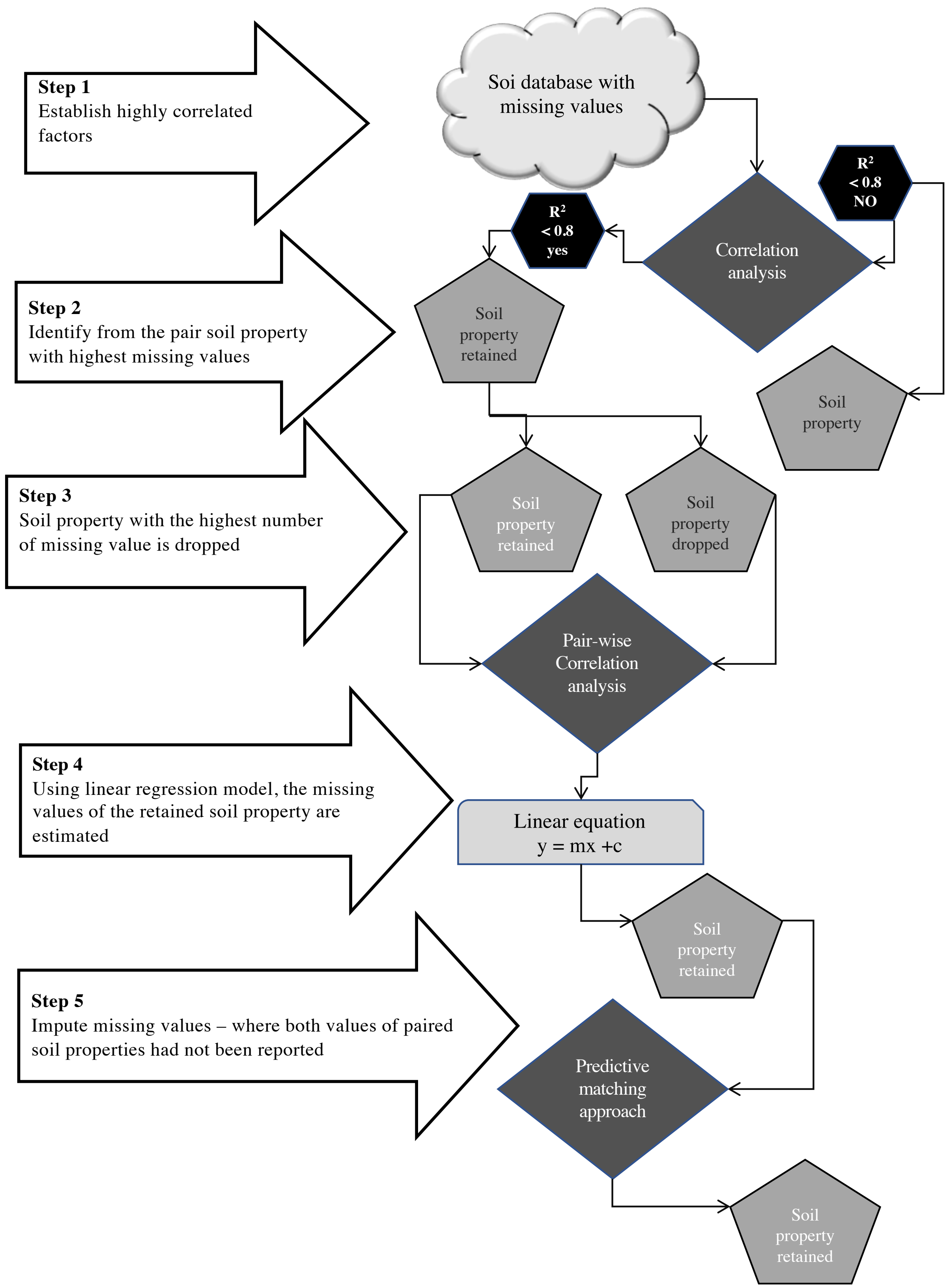

Figure 3.2: Flowchart showing the different steps to estimate the missing values for some of the soil properties in the database. The r is the Pearson correlation coefficient obtained from the correlation analysis. Soil property that was highly correlated ($r > 0.8$), the property with the highest number of missing values were removed from the analysis

Lastly, to calculate sampling variance for meta-analysis, the standard deviation (sd) was included in the database. For studies, where only the standard errors or coefficient of variation were reported, they were used to estimate the sd. In studies where no measures of variance were presented, a value of 1.5 times of the mean of all reported sd was assigned (Ishak *et al*. 2008; Ros *et al*. 2011).

3.4 Statistical analysis

A meta-analysis was conducted to: (i) quantify heterogeneity across fertilizer studies and, (ii) to evaluate causes of variation in FR and effect size across categorical variables. The effect size (response ratio *i.e.* FR) estimator was used to quantify the magnitude of the effect of fertilizer application on yield (Hedges *et al.*, 1999) and was considered a proxy index of soil responsiveness. Agronomic nitrogen use efficiency was taken to represent nutrient use efficiencies across the fertilizer studies. Regression analysis was done to discern the continuous independent variables that explain variability in FR and N-AE.

3.4.1 Fertilizer Response of Maize

Fertilizer response was taken as the ratio of the mean maize yield of the fertilized plot ($\bar{x}_t$ in kg ha^{-1}) and the mean maize yield of the control plot ($\bar{x}_c$ in kg ha^{-1}) (Hedges *et al*. 1999; Ros *et al*. 2011) and was computed as a natural log (*ln*) to normalize the data distribution (Johnson & Curtis, 2001). A normalized FR is required to develop random effect meta-regression models. The *ln* FR was computed as in Equation 3.1.

$$ln\,\text{FR} = \ln\left(\frac{\bar{x}_t}{\bar{x}_c}\right) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ 3.1$$

Soils with FR > 1 were categorized as responsive. Within the non-responsive soils (FR ≤ 1) the poor and fertile soils (less responsive) were distinguished, based on the maize yields in the control plots, and as described by Vanlauwe *et al*. (2014). The fertile soils category were soils where no significant (p < 0.005) increase in maize yield was realized after N fertilization or a combination of N with inorganic P or K addition, (Vanlauwe *et al.*, 2014), but would still have high maize yields (> 1,125 kg ha^{-1} for smallholder farm in SSA) as displayed in the control plots. A FR ≤ 1 meant that fertilization had no effect or negatively affected yield.

The sampling variance of the fertilizer response (FR_v) was used to compute the Heterogeneity (QT) between fertilizer studies and evaluate factors affecting FR. FR_v was calculated as equation 3.2.

$$ln\ \mathrm{FR}^{V} = \left(\frac{(sd_t)^2}{n_t(\bar{x}_t)^2} + \frac{(sd_c)^2}{n_c(\bar{x}_c)^2}\right) \text{............................ Equation 3.2}$$

where n is the sample size/number of replicates, sd_t is the standard deviation for the yields within the treatment and sd_c is the standard deviation for the yields within the control.

3.4.2 Agronomic Nitrogen Use Efficiency of maize

The agronomic nitrogen use efficiency was computed following Vanlauwe *et al*. (2011) as in equation 3.3.

$$\mathrm{N-AE} = \left(\frac{\bar{x}_t-\bar{x}_c}{\mathrm{FN}}\right) \text{........................Equation 3.3}$$

where FN (kg N ha^{-1}) is the amount of applied fertilizer N. The unit for N-AE is kg dry weight kg^{-1} N. The average N-AE was computed across the different groups of categorical factors (Table 3.1, 3.2, 3.3). The derived N-AE is not an effect size as defined in the meta-analysis. Therefore, sampling variance, as a requirement for computing an effect size was not computed for N-AE. Instead, N-AE statistical analysis involved conducting regression analysis.

3.4.3 Meta-analysis of fertilizer response for maize crop

A meta-analysis approach was used to evaluate FR following Hedges *et al.* (1999) and Luo *et al.* (2006). FR was used to evaluate soil responsiveness to N fertilization, or combinations with inorganic P and or K additions. To establish the different categories of soil responsiveness, the relationship between FR and maize yield of control plots was evaluated. The dataset was split into three categories of soil responsiveness to fertilizer application, similar to Njoroge *et al*. (2017). To further evaluate these categories, their corresponding soil properties were analysed.

To examine the heterogeneity (Q_T) of FR in fertilizer studies across Kenya and SSA, a random-effect (RE) meta-regression model was developed (Viechtbauer, 2010). The

RE model was fitted using the Restricted Maximum Likelihood method (Brown & Kempton, 1994). A test of Q_T was used to assess how comparable the studies were and to test the significance of Q_T of the FR (Hedges & Olkin, 1985). Significant QT of the FR indicates that the variation cannot only be attributed to the sampling error and other explanatory factors are playing a role as well (Huedo-Medina *et al.*, 2006). The latter situation would provide an option to identify explanatory factors of Q_T across fertilizer studies.

The potential effect of publication bias in the meta-analysis was tested using a regression test for the overall dataset (71 studies) (Viechtbauer 2010). The test is a quantitative representation of the importance of publication bias in the meta-analysis (Thornton & Lee, 2000). The publication bias was also evaluated through a "*funnel*" plot. The distribution of *ln* FR in the "*funnel*" plot was analysed, to check if indeed publication bias was likely to influence the meta-analysis results (Viechtbauer 2010). The trim and fill method were used to estimate the number of additional observations necessary to change the results of the analysis from significant to non-significant (Duval & Tweedie 2000; Viechtbauer, 2010).

To examine the influence of soil, climatic and management factors on FR, an analysis of the categorical variables was conducted, as a further step in meta-analysis (Table 3.1, 3.2, 3.3). The categorical variables included; soil types, soil textural classes, agro-ecological zones, type of management (farmer or researcher managed), range of N application rates and nutrient types (N, P and K). To compare the effect of fertilization across the groups, the weighted means ($ln\ \mathrm{FR}_w$) of FR and their corresponding 95% confidence intervals (CIs) were computed for each group, following Luo *et al*. (2006):

$$ln\ \mathrm{FR}_w = \left(\frac{\sum_{i=1}^{m} ln\ \mathrm{FR}\, w_i}{\sum_{i=1}^{n} w_i}\right) \text{.. Equation 3.4}$$

where 'i' is an observation, w_i is the weight of i, defined by the reciprocal of the *ln* FR_v ($w_i = 1/ln\ \mathrm{FR}^v$), and m is the number of observations within a group of that categorical variable. The effect of fertilization for each group was considered significantly different from 1 if the CI did not overlap the line of no effect (*ln* FR = 0, $p \leq 0.05$), and different from one another if their 95% CIs were non-overlapping (Hedges *et al.*, 1999). A back-transformed $ln\ \mathrm{FR}_w$ was reported in text and figures. The "*metafor*" R-package was used to conduct the meta-analysis (similar to Barto & Rillig (2010) and back transformed values (FR) were reported in the figures.

3.4.4 Regression analyses – influence of climatic and management factors on FR and N-AE

To further study how soil properties, management and climatic factors (the continuous factors) affect FR and N-AE, General Linear Regression (GLR) models were developed. In this analyses *ln* FR or N-AE was the dependent variable and independent variables were: N application rate (only for *ln* FR), total C, soil pH, P-Olsen, exchangeable K, clay, silt and rainfall. Variables were standardized by dividing each observation with the standard deviation of the variable, so that each factor had equal representation in the GLM. The relationship between dependent (*ln* FR or N-AE) and independent variables was assessed based on the level of significance ($p \leq 0.05$) and coefficient of determination (adjusted R^2).

Further evaluation of the GLMs were conducted by computing the variable importance projections (VIP) scores from each GLM (*ln* FR or N-AE), which primarily indicate the relative measure of the importance of each predictors in the model (Kuhn, 2008). These scores were considered robust, because they took into account the orthogonal variation between independent factors (Chong & Jun, 2005; Farrés *et al.*, 2015) and high variation in *ln* FR or N-AE. The VIP scores were used to discern the key factors, which also explain the underlying variation in FR or N-AE (Kuhn, 2008; Mehmood *et al*., 2012; Farrés *et al*., 2015). The scores were computed independently for each other (predictors) using the *t*-statistic (Kuhn, 2008). A criterion of VIP scores >1 was adopted for identifying the key factors, so that those with scores >1 were taken as the key ones (Chong & Jun, 2005). The "*pls*" R-package was used for regression analysis (Mevik & Wehrens, 2007). The "*caret*" R-package was used to compute VIP scores (Kuhn *et al.*, 2014).

3.5 Results

3.5.1 Fertilizer response for maize crop

The median FR was 1.8 for Kenya and 1.7 for the rest of SSA (excluding Kenya), which indicate maize yield almost doubled with N fertilization. There was a significant nonlinear, negative relationship between FR and the maize yields of control plots (Figure 3.3 a, b) with R^2 value of 0.47 for Kenya (p = 0.003) and 0.49 for SSA (p = 0.002). There was no obvious relationship between FR and N application rate (Figure 3.3 c, d) although the maximum attainable FR in Kenya tended to decrease with N application rate (Figure 3.3 d).

Responsive soils (FR >1) were common across sampled plots; 86% Kenya and 89% for SSA. The maize yields in control plots of these non-responsive soils varied from 100 to 7000 kg ha^{-1}. Of these soils, 72% were considered fertile non-responsive soils (control plots with maize yields >1125 kg ha^{-1}). At this point, the quadratic curve started to level, which was an indication of no significant effect (p < 0.001) of fertilization, and most observation (>20 %) were close to or below the line of no effect to fertilization (FR = 1, Figure 3.3). The mean FR was 2.2 for the responsive soils, 0.68 for poor, non-responsive and 0.89 for fertile, non-responsive soils in SSA including Kenya. The number of non-responsive soils for Kenya (51 plots) was too small for further analyses,

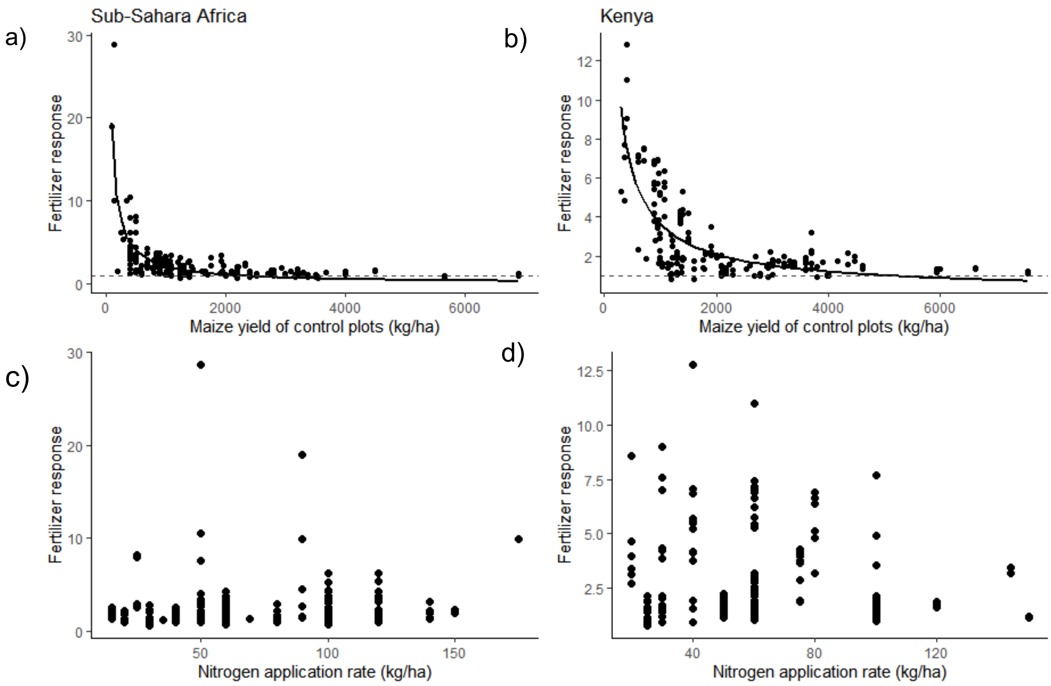

Figure 3.3: Fertilizer response (FR) as a function of maize yield in the unfertilized control plots (a, b) and or N application rate (c, d) for sub-Saharan Africa (a, c) and Kenya (b, d). The dashed line is the line of no response to the fertilizer (FR=1). The solid lines describe non-linear relationships function as: FR = 32244 (Control Yield) -0.7 (p = 0.003; R^2=0.47) for Kenya and FR =83(Control Yield)-0.5 (p=0.002; R^2=0.49) for Sub-Saharan Africa.

so the data of the non-responsive plots was pooled together for Kenya and the rest of SSA before being subjected for further analysis.

Soil characteristics varied within the three soil responsive categories (Figure 3.4). For example, average total C ranged from 2 - 27 g kg-1 for poor, non-responsive soils and from 1 – 56 g kg-1 for fertile, non-responsive soils. The average total C content for responsive plots was 63% higher than that of poor, non-responsive plots. The mean concentration of P-Olsen for the poor, non-responsive plots was higher (11 mg kg-1) than that of responsive soils (6 mg kg-1) and fertile, non-responsive plots (4 mg kg-1). Soil C and exchangeable K seemed to be the main separators between poor, non-responsive soils and the other two categories (Figure 3. 4 c, f). The mean N application rates were on average 22% lower for poor, non-responsive plots compared to the responsive soils.

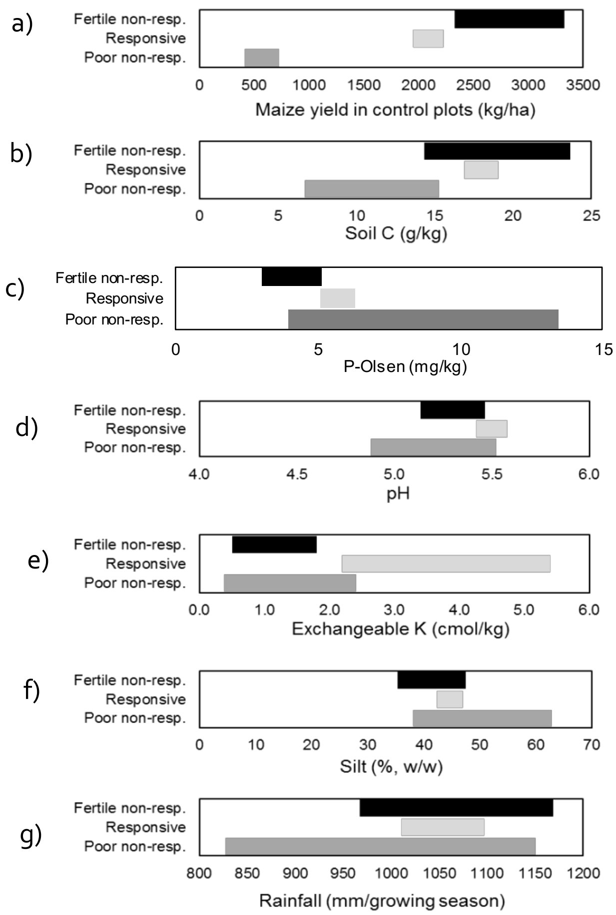

Figure 3.3: Means and standard errors of maize yields in control plots, soil variables and rainfall for poor, non-responsive soils responsive soils and fertile, non-responsive soils.

3.5.2. Heterogeneity in fertilizer studies and test of publication bias

Random effect (RE) meta-regression model results indicate significant variation in FR among the observations of the fertilizer studies for Kenya (Q_T = 15435, degree of freedom = 198, $p < 0.001$) and for SSA (Q_T = 1645, degrees of freedom = 245, $p < 0.001$). This implies independent variables explained a significant part of this variation other than the sampling error alone for all studies included in meta-analysis. Thus, the evaluation of factors that attribute to the variability in FR was necessary.

The regression test results (z value = 0.75, p = 0.39) suggests absence of publication bias across the 71 selected fertilizer studies. Although the distribution of *ln* FR observations (Figure 3.5) in the "*funnel*" plot was not symmetrical because of more relatively high values for *ln* FR, only 84 observation was missing and did not have any effect on the overall results of a meta-analysis. Additional observations would, however, have resulted in a more symmetrical "*funnel*" plot

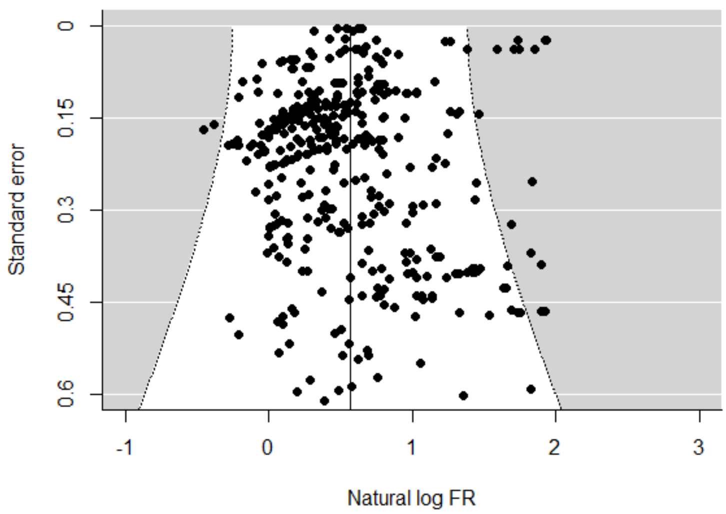

Figure 3.4: Funnel plot showing the distribution of the natural log fertilizer response (*ln* FR) in the funnel plot adjusted with the heterogeneity of the random effect model. The black vertical line is the regression line, which divides the funnel into equal portions to illustrate the asymmetry. Black dots are individual *ln* FR observations (n = 457); the grey area illustrates area outside the "*funnel*".

3.5.3. Variability in fertilizer response

Weighted mean of FR across categorical variables was used to assess variability between their sub-groups using CIs (Figure 3.6 & 3.7). The meta-analysis showed that the CIs around the FR of all soil orders except Cambisols (Kenya) and Areonsols (rest of SSA) overlapped with the line of no response. The FR was significantly ($p < 0.001$) higher than 1 for these two soil orders, which implied positive crop response to fertilizer application compared to other soils (Figure. 3.6 a).

Combined application of N, P and K led to a doubling of the mean FR ($p < 0.0001$) both SSA and Kenya (Figure 3.6 b) compared to application of N alone. For plots with N alone, the FR did not differ significantly from 1. Again, the variation in FR was large, indicating variability across the plots. For Kenya, sandy soils in general tended to show a higher FR than non-sandy soils (Figure. 3.6 c). This trend was, however, not confirmed for the rest of SSA. There, sandy loam soils were the only class of soils with a FR significantly higher than 1. For clay soils, the FR did not differ significantly ($p < 0.001$) from 1.

The FR was similar between the farmer and researcher-managed plots (Figure 3.7 d). The mean FR in SSA farmer-managed plots was significantly higher ($p < 0.001$) than 1. The response to fertilization did not vary significantly among agro-ecological zones (Figure 3.6 c). In Kenya, the FR was highest in the lower humid zone (4.8) and > 1 also in the upper midlands and lowlands. There was no significant response ($p > 0.01$) to N fertilization in the lower midlands and lower highlands. For the rest of SSA, the mean FR for the sub-humid zone was 2.9. The FR_w for sub-humid zone was significantly ($p < 0.001$) higher than 1.

The FR had wide CI range (0.8 to 1.5) across the N application rates <30 kg N ha^{-1} for Kenya (Figure 3.7 b). The average FR for 30-60 kg N ha^{-1} application ranges was 1.61 and was not significantly ($p > 0.01$) different from 1. For SSA, FR was not significantly ($p < 0.001$) different from 1 for N application rates range of 30-60 kg N ha^{-1} since the CI overlapped with the line of no effect.

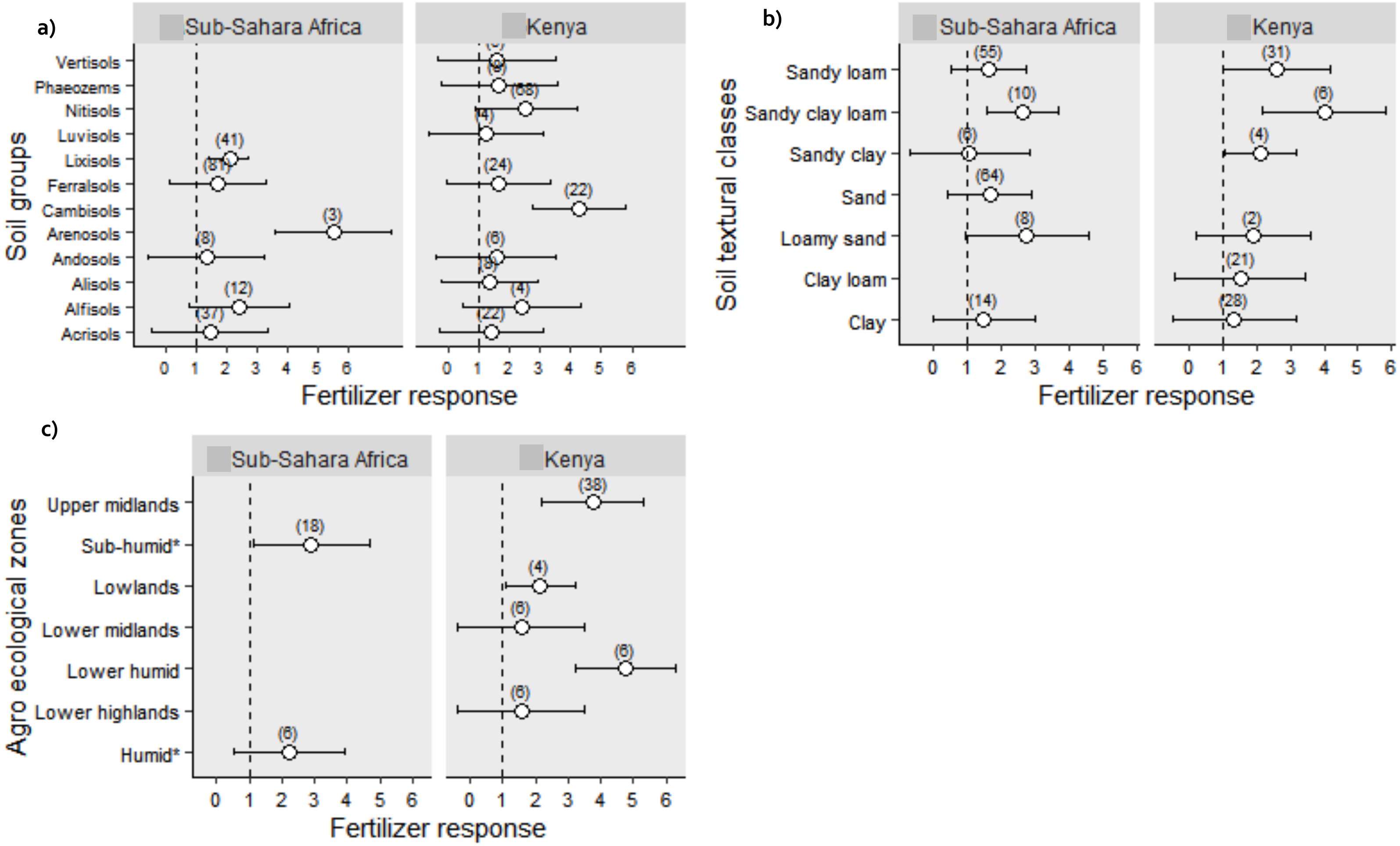

Figure 3.5: Means of the fertilizer response (FR) across (a) World reference soil groups, (b) soil textural classes and (c) agro-ecological zones. The dashed line is the line of no response to the fertilizer (FR=1). Error bars represent confidence intervals; numbers in brackets represent the number of observations per category.

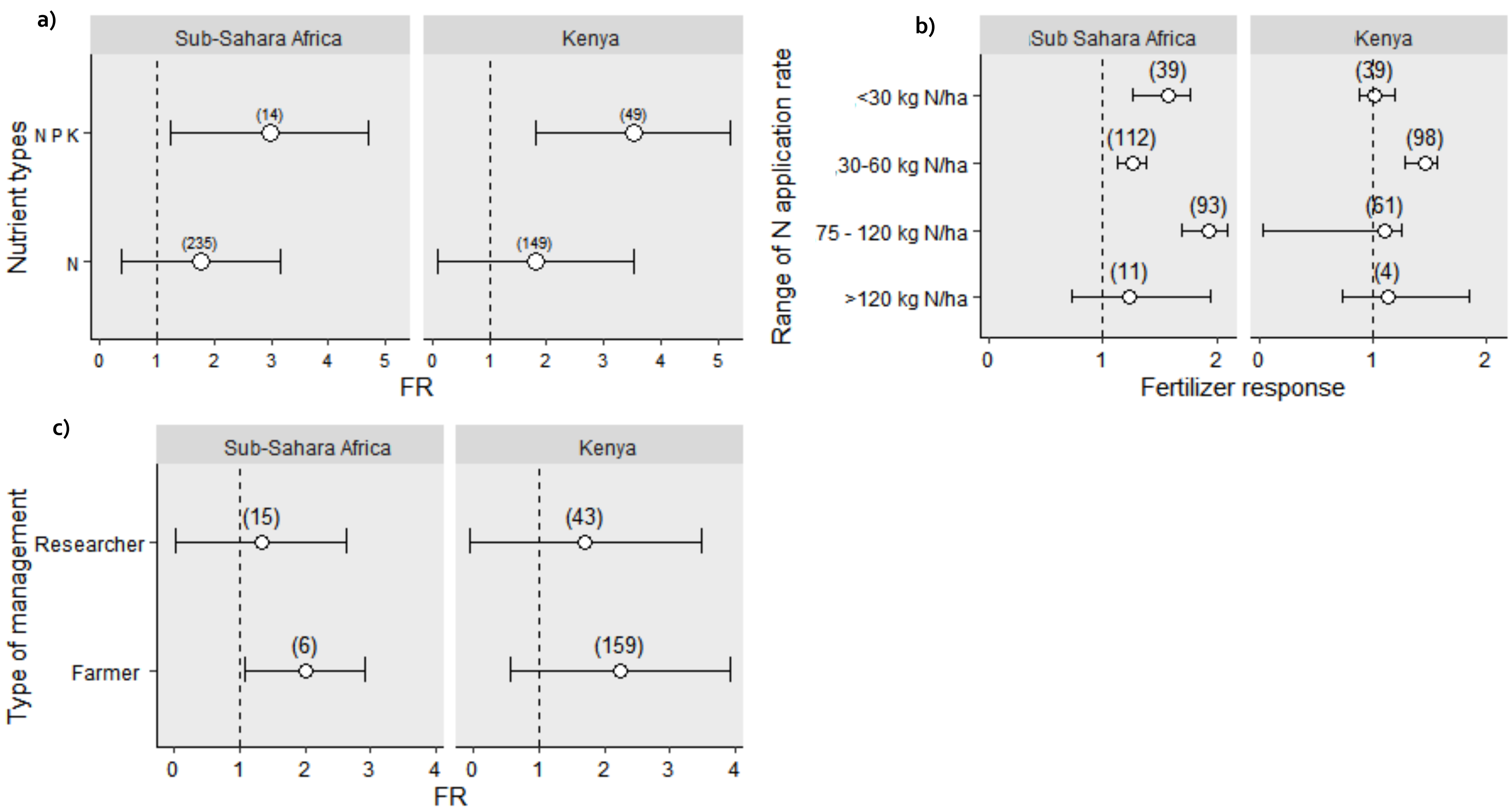

Figure 3.6: Means of the fertilizer response (FR) across categorical variables, (a) nutrient types, (b) range of nitrogen application rates and (c) type of management of the fertilizer experiments. The dashed line is the line of no response to the fertilizer (FR=1). Error bars represent confidence intervals; numbers in brackets represent the number of observations

3.5.4 Agronomic Nitrogen Use Efficiency

The average N-AE was 42 kg dw kg^{-1} N for Kenya and 18 kg dw kg^{-1} N for the rest of SSA. The N-AE varied from -27 to 165 kg dw kg^{-1} N across all observations (Figure 3.8 a, b). There were no significant (p > 0.01) relations between maize yield of the control plot (Figure 3.8 a, b) or N application rate (Figure 3.8 c, d) and N-AE, though the maximum attainable N-AE seemed to decline with increasing maize yields in the control plots and N application rates. Mean N-AE varied across the soil, climate and management factors (Table 3.4).

3.5.5 Key factors affecting Fertilizer response and Agronomic Nitrogen Use Efficiency

The regression (GLM) with the eight continuous predictors explained 31% of the variation in FR for Kenya and 9% for SSA (Table 3.5, 3.6). Fertilizer response decreased significantly (p < 0.0001) with increasing P-Olsen in Kenya, but not in the rest of SSA (see Appendix 2). Here, soil pH and average rainfall during a growing season were the significant predictors (p > 0.001) of variation in FR. They correlated positively with FR. Low values for soil pH and rainfall of a growing season displayed decreased FR (<1). Fertilizer response increased marginally significantly with soil total C in Kenya (p = 0.10), but not in the rest of SSA.

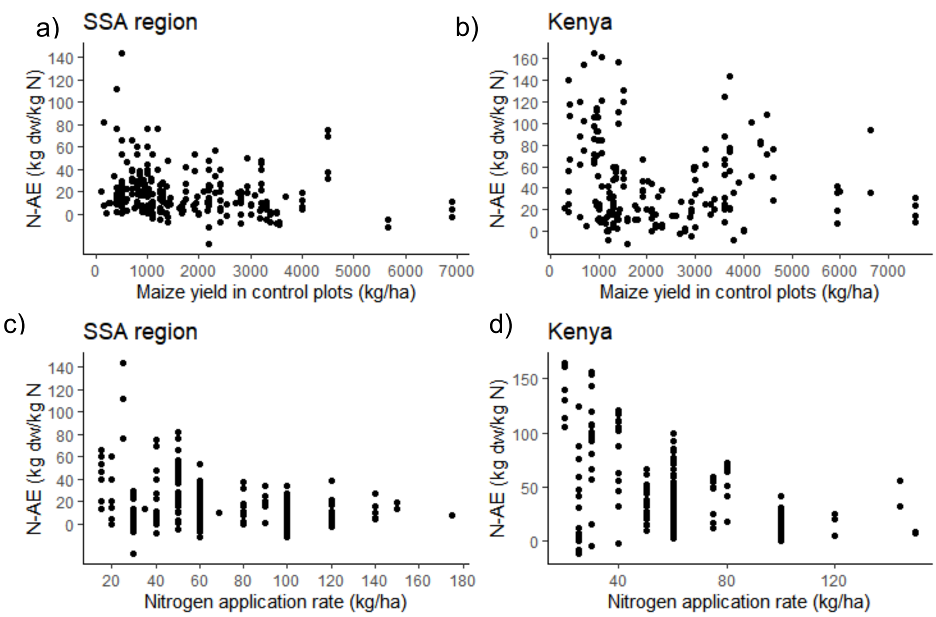

Figure 3.7 : Agronomic nitrogen use efficiency (N-AE) as a function of maize yield of the control plots (a, b) or nitrogen application rate (c, d) across fertilizer studies for sub-Saharan Africa (a, c) and Kenya (b, d)

Table 3.4 : Agronomic nitrogen use efficiency statistics for categorical variables of soil and climatic factors

		Kenya			*Sub-Saharan Africa*		
Factor	Level	Mean	SE	n	Mean	SE	n
		Environmental					
Soil order	Acrisols	17.3	4.9	22	20.0	2.5	37
	Alfisols	87.4	19.2	4	23.5	1.6	12
	Alisols	12.3	4.8	8	-	-	-
	Andosols	22.8	1.6	6	28.9	10.7	8
	Arenosols	-	-	-	15.3	2.2	3
	Cambisols	71.8	10.2	22			
	Ferralsols	42.0	8.0	24	16.7	2.1	81
	Lixisols	-	-	-	22.5	5.4	41
	Luvisols	10.5	3.0	4	-	-	-
	Nitisols	51.2	5.2	68	-	-	-
	Phaeozems	65.8	4.8	8	-	-	-
	Vertisols	22.1	4.6	6	-	-	-
Soil textural classes	Clay	26.7	6.1	28	19.5		14
	Clay loam	33.7	2.0	21			
	Loamy sand	20.8	-	2	22.5	4.7	8
	Sandy	-	-	-	10.8	0.7	64
	Sandy clay	20.4	2.2	4	3.5	9.6	6
	Sandy clay loam	54.6	21.6	6	17.2	5.9	10
	Sandy loam	49.8	7.0	31	16.8	3.1	55
Agro-ecological zone	Lower highlands 1	22.8	1.6	6	-	-	-
	Lower highlands 2	75.8	9.6	6	-	-	-
	Lower midlands	28.3	7.7	4	-	-	-
	Lowlands	20.4	3.5	4	-	-	-
	Upper midland	94.3	15.1	8			
	Humid	-	-	-	21.8	2.7	6
	Sub-humid	-	-	-	26.7	1.9	18
		Management					
Manager	Farmer	44.6	3.2	159	18.3	1.4	234
	Researcher	32.9	4.7	81	22.4	1.6	63
Nutrient type	N, P, K	80.0	5.7	49	23.3		14
	N	30.4	2.5	149	18.1	1.4	235

Key: - = missing statistic of the group; n=number of observations; SE = standard error

When tested whether addition of maize yields of control plots as predictor could improve the predictive ability of the model, the adjusted R^2 value increased from 31% to 59% for Kenya and 9% to 57% for SSA. In that case, P-Olsen no longer explained any variation in FR for Kenya, and for SSA rainfall dropped out of the model. The FR decreased with increasing maize yield in the control plots (Fig. 3.3).

Seven continuous predictors were used to develop the regression model for N-AE. The best model explained 29% of variability in N-AE for Kenya and 1.3% for SSA (Table 3). Similar to FR, rainfall and total C positively influenced variability in N-AE and P-Olsen did so negatively in Kenya; for SSA variation in N-AE silt was the best predictor

Table 3.5: Regression model estimates of soil and climatic factors on fertilizer response and agronomic nitrogen use efficiency.

Fertilizer response						
R^2	Adjusted R^2	Predictor	Estimate	Standard error	*p* value	*Significant Level*
		Kenya (n=202)				
0.29	0.26	Intercept	0.483	1.275	0.204	***d***
		Soil pH	0.166	2.679	0.008	***b***
		ln-Total C	-0.024	-3.353	0.001	***a***
		ln-P-Olsen	-0.020	-3.041	0.003	***a***
		ln-N rates	-0.002	-1.535	0.126	
		ln-Exch K	0.060	6.103	0.001	***a***
		Clay	0.001	0.425	0.671	
		ln-Silt	-0.007	-2.444	0.015	***d***
		ln-Rainfall	-0.0008	-0.771	0.004	***d***
		Sub-Saharan Africa *(n=255)*				
0.13	0.092	Intercept	-0.221	-0.734	0.464	
		Soil pH	0.154	2.714	0.007	***b***
		ln-Total C	-0.006	-1.649	0.101	
		ln-P-Olsen	0.021	3.049	0.003	***c***
		ln-N rates	0.002	2.043	0.042	***c***
		ln-Exch K	-0.022	-2.002	0.047	***c***
		Clay	-0.002	-0.591	0.555	
		ln-Silt	0.003	1.185	0.237	
		ln-Rainfall	-0.0001	-1.439	0.102	***b***

Key: Significant codes: ***a*** = 0.001, ***b*** = 0.01, ***c*** = 0.05, ***d*** = 0.1, *ln* = natural log

Table 3.6 Regression model estimates of soil and climatic factors on fertilizer response and agronomic nitrogen use efficiency

Agronomic Nitrogen Use Efficiency						
R^2	Adjusted R^2	Predictor	Estimate	Standard error	*p* value	*Significant Level*
0.32	0.29	Intercept	52.652	22.943	0.023	***d***
		Soil pH	7.019	3.780	0.065	***d***
		ln-Total C	-1.154	0.442	0.010	***c***
		ln-P-Olsen	-1.098	0.404	0.007	***b***
		ln-Exch.K	3.429	0.589	0.001	***b***
		Clay	-0.408	0.169	0.017	***c***
		ln-Silt	-0.634	0.185	0.001	***a***
		ln-Rainfall	-0.010	0.006	0.106	
Sub-Saharan Africa (n=255)						
		Intercept	11.183	11.791	0.344	
0.035	0.003	Soil pH	1.697	2.283	0.458	
		ln-Total C	0.056	0.138	0.687	
		ln-P-Olsen	0.544	0.273	0.048	***c***
		ln-Exch.K	-0.829	0.431	0.056	***d***
		Clay	-0.013	0.113	0.912	
		ln-Silt	-0.077	0.097	0.431	
		ln-Rainfall	-0.001	0.004	0.781	

Key: Significant codes: ***a*** = 0.001, ***b*** = 0.01, ***c*** = 0.05, ***d*** = 0.1, *ln* = natural log

Key factors explaining variation in FR using VIP scores were identified for Kenya and the rest of SSA with VIP scores > 1 (Figure 3.9). Exchangeable K, soil pH and rainfall (Figure 3.9 a, b) were the key factors for both Kenya and the rest of SSA. In addition, P-Olsen, total C and silt were relevant in Kenya, and N application rate in SSA. Clay was the least important factor for both Kenya and SSA. Results from the GLM indicate rainfall, as the significant ($p < 0.001$) factor influencing FR in Kenya and SSA. P-Olsen, total C and silt were additional key factors for Kenya while soil pH and exchangeable K the key ones for SSA. Nitrogen application rates was not significant ($p < 0.01$) but were important based on the VIP score that was 1.18 for SSA (Figure 3.9). For N-AE, results of relative importance of key explanatory factors based on VIP indicate P-Olsen, clay, silt, soil pH, and rainfall as key determinates for Kenya (Figure 3.9 c

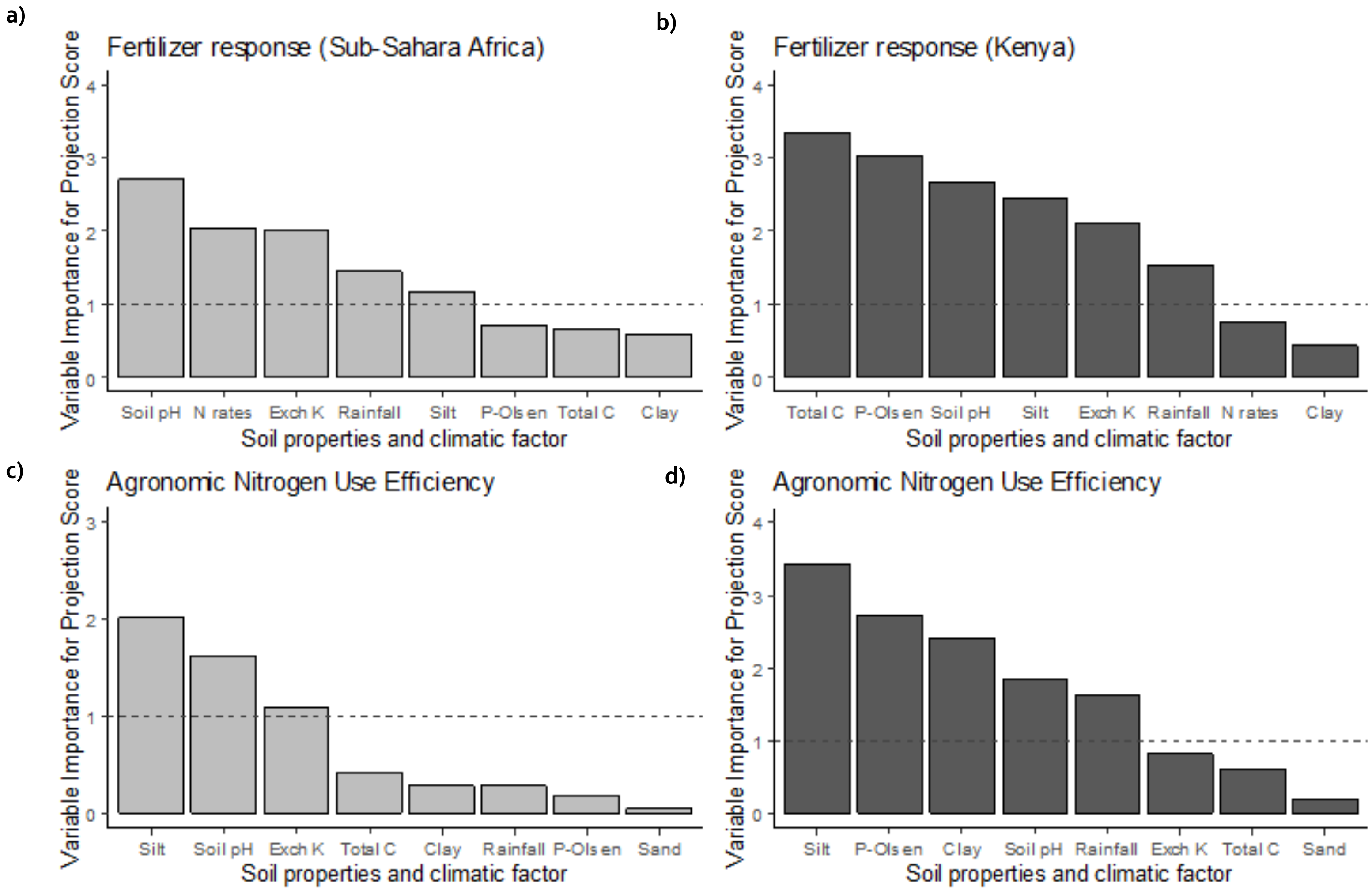

Figure 3.8: Relative importance of continuous management, soil and climate factors based on the variable importance projection (VIP) values computed from the general regression model explaining variation for in the fertilizer response (FR) (a, b) and agronomic nitrogen use efficiency (c, d). For sub-Sahara Africa and Kenya. The dotted line represents the threshold value for the VIP value (VIP=1) below which variables were considered not to be important predictors.

3.6 Discussion

3.6.1 Factors affecting variability of FR and N-AE

The results indicate that both FR and N-AE vary largely within Kenya, which justifies the need for refining fertilizer recommendations to a higher spatial resolution. The FR varied roughly from 1 – 28.8 and the N-AE from 0 – 160 kg dw kg^{-1} N. Fertilization, on average, nearly doubled the maize yield in both Kenya and the rest of SSA. The average response was statistically significant ($p < 0.001$) only when N was applied in combination with P and/or K (Figure 3.3 b). The average N-AE was 18 kg dw kg^{-1} N for SSA was similar to that of Vanlauwe *et al*. (2011) who reported 19 kg dw kg^{-1} N for farmer-managed experiments. The average N-AE for Kenya (42 kg dw N kg^{-1}) was substantially higher but not uncommon for East Africa due to high variability across smallholder farms (Vanlauwe *et al.*, 2011).

Reoccurring variables significantly ($p < 0.001$) explaining variation in both FR and N-AE in Kenya and SSA are total C, pH, P-Olsen, rainfall and silt (Figure 3.7, 3.6, 3.10 and Table 3.4). Soils with a lower pH < 5.2, rainfall < 1200 mm and silt < 10% tended to have lower FRs and N-AEs. In line with earlier studies (Kihara *et al.*, 2016), pH and FR and N-AE were positively related. This is probably because most soils in the study had a soil pH below the optimum of 5.5 - 6.5. At soil pH < 5.5, N mineralization rates decrease and P increasingly binds to the soil's solid phase. Dominant factors that explained variability in FR (based on the variable importance projection (VIP) score) varied to some extent between Kenya and SSA, which could be related to differences in agro-ecological zones, reference soil groups and soil textural classes between these two regions (Figure 3.5, 3.6). The results between factors that were significant from regression analysis and computed VIP scores also varied to some extent (Table 3.6, 3.7, Figure 3.9). For example, for Kenya, soil pH and exchangeable K were not significant ($p > 0.001$) based on the coefficient from the regression model (Table 3.6, 3.7), but were important based on computed VIP scores (> 1, Figure 3.9). The variation of factors can also be attributed to the difference in the statistical computation for regression analysis and calculation of VIP scores. Grömping (2009) explained such computation differences, which is caused by the non-unique decomposition of model sum of squares in the regression model, due to correlated predictors. However, the study used uncorrelated variables to develop the regression models, which is contrary to this observation. Unexpectedly, higher FR values (for N) were found in soils with higher P-

Olsen concentrations (>11 mg kg^{-1}), but below the critical level of 15 mg kg^{-1}. This may be attributed to the fact that in part of the entries, N fertilization was combined with P (and K) fertilization, and FR was >1 particularly in those cases (Figure 3.7 b).

The biophysical and management parameters available in the dataset for this study, appeared to be difficult to capture the relevant amount of variation in FR (Table 3.1). Nevertheless, the meta-analysis of factors affecting FR provided few points of departure for spatial refinement of fertilizer recommendations. The wide CIs for soil orders and texture were all overlapping each other, although some orders and textural classes with FRs significantly > 1 ($p < 0.001$) were identified (the more fertile Cambisols and Arenosols for Kenya and SSA, respectively, and sandy (clay) loam soils consistently for Kenya and SSA). The regression analyses showed that the set of continuous environmental characteristics used, explained a very limited proportion of the variation in FR and N-AE (Table 3.6, 3.7). The continuous variables (Table 3.6) explained only 31% of the total variation in FR in Kenya and as little as 9% in SSA. For the N-AE the respective percentages of explained variation were even lower.

Units aggregating several factors determining FR or N-AE would intuitively be most suitable to refine fertilizer recommendations for spatially relevant units. The spatial mapping unit AEZ potentially captures a combination of factors such as the length of the growing season, climate, landform and soils, all related to land use. As such, it aggregates some of the other individual variables tested and is currently used to refine fertilizer recommendations. However, the average FR in the AEZ distinguished for the rest of SSA did not differ significantly (Figure. 3.6 c), although the FR was significantly >1 ($p < 0.001$) in three of the seven zones of Kenya. This renders AEZs an unsuitable unit for refining recommendations based on these results. Extending them with soil information (pH, P-Olsen, texture, order) could be a promising strategy.

The maize yield in the non-fertilized control seemed to be the best predictor for the FR ($p = 0.0001$) (Figure 3.3 a, b). This variable can, similar to AEZ, can be regarded an integral proxy for environment (soil fertility, climate, weather), genetic (maize variety) and management factors (planting density, control of pests, weeds, diseases). Both for Kenya and for SSA, the adjusted r^2 increased substantially when these yields were added to the set of independent variables (from 31% to 59% for Kenya and from 9% to

57% for SSA). The FR was higher when the yield in the control plots was lower (Figure 3.3 a, b). However, this statistical relationship is probably a result of autocorrelation because the maize yield in the non-fertilized control (x_c) is in the denominator of FR (Equation 3.1). This suggestion of autocorrelation is supported by the absence of any relationship between control yields and N-AE (Equation 3.2, Figure 3.8 a, b).

3.6.2 Soil responsiveness to fertilizer application

The FR when considered as an index, can provide useful tool for diagnosing soil responsive to fertilizer application. Over 85% of the sites were responsive to fertilization for both SSA and Kenya. Farms that were responsive were 86 % for Kenya and 14% were non-responsive. To prevent complete failure of fertilizers, prior identification of non-responsive soils is of utmost importance. Non-responsiveness of poor soils is often related to low soil organic matter content (Tittonell & Giller, 2013), causing soil physical constraints (low water-holding capacity), low micronutrient availability (Kihara *et al.*, 2017) as well as low microbial activity leading to increased soil disease risk (Lal, 2016). The results confirm that although the variation was high, the average C content of poor, non-responsive soils (11 g kg^{-1}) was significantly lower ($p = 0.031$) than in responsive, (18 g kg^{-1}) and fertile non-responsive soils (19 g kg^{-1}; Figure 3.4 c). The soil responsiveness categories were clearly distinguished by total C and exchangeable K (Figure 3.4 c, f). Thus, total C and exchangeable K could act as useful indicators for discerning the different categories of soil responsiveness to N fertilization, which may be useful for nutrient management. The high variation indicates that non-responsiveness is a complex feature that is not easily operationalized using easily available environmental data. This is probably the reason that soil total C is not a powerful predictor of the FR (Table 3.6).

3.6.3 Challenges for meta-analysis in agronomic studies

This study adhered to standards recommended for meta-analysis in agronomy studies (Philibert *et al.*, 2012), by developing a criterion for data inclusion and establishment of a database (Figure 3.2). However, exclusion of publications (only 71 studies out of the total of 503 found were acceptable) is a clear indication of the challenges for merging and comparing data from past literature for meta-analysis, which may be attributed to differences in reporting across fertilizer studies. For example, all studies reported on fertilizer treatments, which allowed us to quantify the FR and N-AE.

Studies that did not report control treatments (5%) where no fertilization was done) were omitted from during inclusion, while developing the databases. For example, there was variation on the different set of soil properties used in characterizing each study area. As a result, missing soil properties (18 %) were imputed, since different analytical methods were used for soil characterization.

3.7 Conclusions

The basic premise of this study was to identify key factors that can be used to refine fertilizer recommendation across smallholder farms of SSA. The findings indicate that available data layers can explain only very small amounts (< 30 %) of variation in FR and N-AE and there is need for more systematic studies at high spatial resolution to identify yield-limiting factors. The data indicated that soil pH, P-Olsen, silt content and rainfall had significant but low levels of power in explaining variation in FR and N-AE. This finding implies that strategies for refining the current blanket fertilizer recommendation should include information on soil type, soil properties (texture, P-Olsen and total Carbon). Such information can be derived through soil testing, which should be accompanied by nutrient response trials and preferably plant nutrient testing to diagnose limiting factors. Due to the limitation of the dataset, this study did not comprehensively unravel the biophysical and managements factors that lead to soil non-responsiveness across smallholder farms. The complexity of soil responsiveness to fertilizer application requires further studies to fully understand other factors that led to non-responsive soils, besides total C, soil pH, exchangeable K and P-Olsen as indicated in this study. There is therefore a need of promoting standards of reporting the findings in agronomy, specifically in fertilizer-related studies for future meta-analytical inferences. There is need for developing standard to provide enough information for agronomic studies. For example, there should be a minimum list (set) of soil properties that should be included in future studies, and clear description of any other factors observed within the site under investigation. This may allow combination as well as comparability of datasets across agronomic studies. Supply of information describing the availed data (metadata) should be a requirement for all agronomic studies. However, developing guidelines, calls for detailed investigation that could avail a standard protocol of presenting additional information for agronomic studies similar to those developed for biochar and metabolic studies (Fiehn *et al.*, 2007; Jeffery *et al.*, 2011).

Chapter Four

Analysis of spatial variation to guide the development of fertilizer use recommendations for smallholder farms in western Kenya

Abstract

Refining fertilizer recommendations using digital property maps has been considered among options for increasing maize production for heterogeneous smallholder farms. However, it is not clear the suitable spatial scale, which is pivotal for strategies utilized in developing soil nutrient maps. The objective of this study is to determine a spatial scale, that captures and reflects local growing condition on smallholder farms. A farm survey was conducted within a 100 km^2 sampling block to collect data on the spatial variation in unfertilized maize bio-volume and grain yield in relation to soil organic carbon, total nitrogen, calcium (Ca), magnesium (Mg), potassium (K) and phosphorus (P). Key soil factors associated with crop performance were identified using Step-wise Multivariate Linear Regression (SMLR) modelling. Variation of key factors and crop performance indicators (CPIs) were described by soil types, sampling units, and administrative units through an analysis of variance. In this region, soil properties displayed high variability as exhibited by coefficient of variation of 60% for Ca and 89% for K. This result also showed high variability in grain yield, with 31% of the variation (cumulative) being accounted for by underlying soil properties. Soil Organic Carbon (SOC) was identified as key factor associated with variation in CPIs. SOC displayed moderate spatial dependency (65%) with a range of 523 m. This study provided insights of the association between key soil factors with CPIs that was utilized, to provide a framework for determining optimum sampling distance, appropriate scale for developing digital soil maps (< 250 m). Strategies aimed at refining fertilizer use recommendation can therefore be recommended to use this scale as a guideline.

4.1. Introduction

Agricultural production in sub-Saharan Africa (SSA) can be characterized by smallholder farming and low productivity. The latter is caused by low inherent soil fertility (Bekunda *et al.*, 2010), soil nutrient depletion (Stoorvogel *et al.*, 1993), limited

This chapter is based on a published article in the *Geoderma Regional* Journal, Volume 22, https://doi.org/10.1016/j.geodrs.2020.e00300.

nutrient inputs (Ade Freeman & Omiti, 2003), and poor germplasm (Vanlauwe *et al.*, 2011). Fertilizers are required to replenish soil nutrient stocks and provide nutrients to the crop to increase productivity. Smallholder farmers face the basic question of what type and how much fertilizer to apply given the local conditions on their farms and available resources. Farmers rely mostly on their own experience from previous years and fertilizer use recommendations provided by government agricultural extension institutions. Typically, the recommendations are spatially coarse and developed on the basis of soil surveys and agronomic experiments. The recommendations are valid for the area or a spatial unit, for which the experiment is considered representative, which could be an administrative unit such as country or an agro-ecological zone (AEZ) (Smaling *et al.*, 1992). This so-called 'blanket' fertilizer use recommendations are a single fertilizer use recommendation for a given area, and do not represent the variation in conditions within that area (Bationo *et al.*, 2012).

Currently, several countries use blanket fertilizer use recommendations to guide decisions on nutrient management (Rurinda *et al.*, 2020). In the past, many studies have focused on refining or improving fertilizer recommendations, with the aim of attaining higher crop yields (Benson, 1999; Mowo & Mlingano, 1993). In the 1980s, countries such as Kenya provided blanket fertilizer use recommendations for the entire country to guide decisions by smallholder farmers on their fertilizer management options (AIC, 1981; FURP, 1994). Later, in the 1990s, fertilizer use recommendations were refined on the basis of agro-ecological zones in an attempt to deal with within-country variation through the Fertilizer Use Recommendation Project (FURP) in Kenya (Smaling *et al.*, 1992; FURP, 1994). Agro-ecological zones were defined based on climate, soil and topography (Geurts & Van den Berg, 1998). However, the variability in growing conditions within AEZs can limit the use of fertilizer use recommendations developed at the AEZ-level (Giller *et al.*, 2006). For example, Diarisso *et al.* (2015) reviewed soil spatial heterogeneity in smallholder landscapes and soil responsiveness to interventions and concluded that a form of precision agriculture is required that recognizes fine scale spatial heterogeneity.

Many recent studies focus on improving fertilizer management and include the use of decision support tools such as the Quantitative Evaluation of Fertility in Tropical Soils (QUEFTS) model (Janssen *et al.*, 1990), the derived Nutrient Expert (Pampolino *et al.*, 2012; Sattari *et al.*, 2014), and crop growth simulation models like APSIM (Kisaka *et al.*, 2016). These tools have to be re-calibrated for every region of interest (Pampolino *et al.*, 2012; Xu *et al.*, 2013). The parameters in the QUEFTS model never consider local spatial variability which occurs on smallholder farms. Furthermore, Molefe *et al.* (2012) observed that such decision support tools fail to capture complexity within smallholder farms. Matthews *et al.* (2002) reported that poor data quality and the lack, therefore limit the application of such nutrient management tools in smallholder farms.

A further refinement of fertilizer recommendations is hampered by the lack of detailed soil data (Sanchez *et al.*, 2009). New, more detailed studies on soil surveys are rare, particularly in SSA. The available national soil survey maps are spatially course (1:250,000), and are produced using different methods, resulting to varying levels of accuracy (regional or national) and data incompleteness (Arnhold *et al.*, 2015; Baruck *et al.*, 2016). Two new developments in collection of soil data may create new options to refine fertilizer recommendations even further: *i)* Digital soil mapping (McBratney *et al.*, 2003) which has evolved into an operational tool that can provide detailed insight in soil variability in an efficient way. Examples include the 100 m resolution digital soil map of Machakos and Makueni districts in Kenya (Mora-Vallejo *et al.*, 2008) and also various continental to global initiatives (Arrouays *et al.*, 2014; Stoorvogel *et al.,* 2017; Hengl *et al.*, 2017). *ii.)* Fertilizer recommendations for a farm can be based on soil test values for that particular farm. Where traditional soil analysis is often too expensive and, therefore, out of reach for smallholder farmers, new proximal sensing techniques like infrared spectroscopy (Shepherd and Walsh, 2007) can be used to provide soil analysis at a low cost.

A better understanding of soil spatial variability may provide guidelines to improve decisions for refining fertilizer use recommendations to optimize crop productivity. Those guidelines should include the scale fertilizer use recommendations needed to be developed. Previous research conducted mainly focused on soil nutrient depletion (Lijzenga, 1998), within farm variability (Tittonell *et al.*, 2005, 2007a, 2008b, 2013) and, spatial and temporal variability in maize response (Njoroge *et al.*, 2017). In western Kenya, blanket fertilizer recommendations are still the norm.

This study aimed at developing an approach for assessing the optimal level of scale, that reflect the variability in local growing conditions in smallholder farms. The objectives of this study were to: *(i.)* describe the spatial variability of soil properties and crop performance, *(ii.)* identify the key soil factors of crop performance, and *(iii.)* identify a scale in which soil spatial variability can be sufficiently described. The hypothesis that the scale of variability of key soil drivers corresponds to that of fertilizer response across smallholder farming landscapes was tested. Smallholder farming systems in western Kenya region was the case study.

4.2 Materials and methods

4.2.1. Site description

The study area is a heterogeneous smallholder landscape in western Kenya (0°26′ - 0°18′ northern latitude; 33°58′ - 34°33′ eastern longitude) delimited by the administrative boundaries of Siaya and Kakamega counties (Figure 4.1). The area is characterized by sub-humid conditions and classified as the Lower Midland (LM_1) (Jaetzold and Kutsch, 2006). FURP provided a blanket fertilizer recommendation for the LM_1 AEZ of 60 kg N ha^{-1} and 30 kg P ha^{-1} for monocrop maize (FURP, 1994). The FURP experiment site is located 2 km away outside the current study area but within the LM_1 AEZ (Figure 4.1 a).

The smallholder landscape is characterized by a distinct bimodal rainfall pattern and a mean annual temperature of 20°C. Long rains (March–June) have a mean precipitation of 1350 mm whereas short rains (September-December) have a mean precipitation of 850 mm (Jaetzold and Kutsch, 2006). The mean potential evapotranspiration (ET_o) is estimated at 1287 mm per maize growing (four months) season (Ademba *et al*., 2015). The altitudes vary in the gently undulating landscape (slopes < 3 %) between 1400 and 1500 m above sea level. The main soil types (WRB, 2014) are presented in Figure 4.1 (b) and include Ferralsols (well drained, moderately to very deep, clay soils), Cambisols (well drained, moderately deep, loamy clay soils) on the hills, and Gleysol (poorly drained, shallow, sandy loam soils) in the plains (Waswa *et al*., 2013; Waveren, 1995). Most soils in these areas are considered moderately P fixing (Nziguheba, 2007). The fertility of the soils is limiting in N, P and K (Lijzenga, 1998).

Farming systems are subsistent with dominantly rain-fed crops and low fertilizer inputs. The mixed crop-livestock system includes maize (*Zea mays L.*) as the dominant staple crop, usually intercropped with common bean (*Phaseolus vulgaris L.*) (Diwani *et al.*, 2013). Using local maize varieties, the average yield levels achieved with the current local conditions using conventional farming practices range from 400 to 2000 kg ha^{-1} (Kwabiah *et al*., 2003; Vanlauwe *et al*., 2014) for both short and long rainy seasons. Other crops cultivated include bananas (*Musa paradisiaca L.*), sweet potatoes (*Ipomoea batata L.*) and groundnuts (*Arachis hypogaea L*). The smallholder farmers are supported by county governmental agricultural extension services

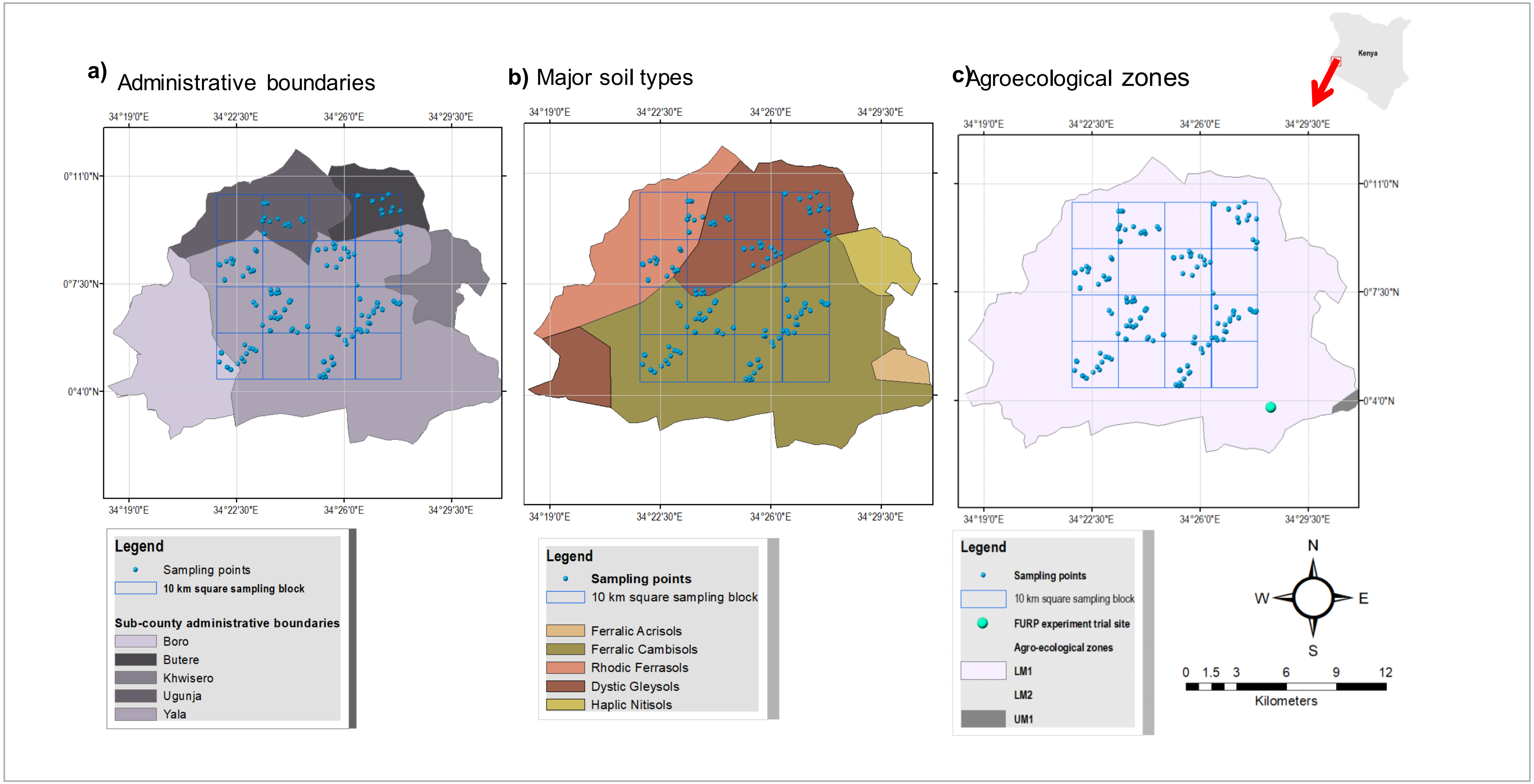

Figure 4.1: (**a**) Map showing the administrative divisional boundaries of the study area in western Kenya. The blue dots indicate the location of the sampling points and the site of the agronomic trial used to develop fertilizer recommendation for LM_1. (**b**), a map displays different the different soil types for the study area and, (**c**) The left panel display a map of the agro-ecological zone and the 100 km^2 block with the 6.25 km^2 tiles, which fell within one agro-ecological zone - LM_1. The top most panel displays location of study site within Kenya.

4.2.2. Overview of sequence of steps for determining a relevant scale

The five consecutive steps include; *(i.)* a farm survey, *(ii.)* determine variability in soil properties and crop performance indicators (CPIs), *(iii.)* relate soil properties to CPIs and identify main soil drivers of variation of CPIs, *(iv.)* characterize the scale of spatial variability of key soil drivers in order to describe a relevant scale of variability in inherent soil fertility, and *(v.)* validate if there is any correspondence in the scale of variability of key soil drivers with the scale of variability in fertilizer response. The steps are presented in the next sections.

4.2.3 Farm survey

A farm survey was carried out within the Land Degradation Sampling Framework (LDSF) (Vågen *et al*., 2010). The LDSF is a stratified hierarchical sampling design that captures variability at different scale levels: block, tiles, sub-tiles, fields and plots across a given landscape (Figure 4.2). A 100 km^2 block, which typifies a smallholder landscape, was first allocated within the study area. The block was sub-divided into 16 tiles and each tile further sub-divided into 10 sub-tiles. A total of 8 tiles and 32 sub-tiles (four from each tile) were randomly selected, within which three sites (maize fields), were randomly generated for sampling during the short and long rains maize cropping seasons.

Different maize fields were drawn for each of the two seasons. Unfertilized, well-managed (*e.g*., free of weeds, pests, not affected by drought and diseases), the mono-crop maize fields were selected for sampling that had a maize crop at the ear-leaf growth stage (silking stage *i.e*., 70 - 75 days after plant emergence) at the time of sampling. The ear-leaf growth stage is considered optimum for diagnosis of nutrient constraints in maize (Römheld, 2011). Per sub-tile, two different locations were selected for sampling in every two of the consecutive seasons, and a maize field according to the above criteria was searched. A third point was selected for cases where no appropriate maize field was found near two selected points. If the two selected points met the criteria, the third point was not sampled. During this farm survey, only unfertilized maize field were sampled where the yield reflects the inherent soil fertility. This ensured there was minimum variation introduced due to differences in fertilizer application across maize fields.

During sampling, GPS was used for navigation across the smallholder, and maize fields were identified and selected for measurements. In the selected fields, a Y frame layout was placed to locate four plots measuring 2.5 m^2 (Figure 4.2 d). The central plot was located first, by measuring 20 m from the main boundary, towards the centre of the maize field. The main boundary was defined as the boundary located from the direction of smallholder farmers' homestead, towards the maize field. Subsequently, three plots were located 12.2 m from the central plot and distributed uniformly around it. During the first visit, the exact coordinates, soils were sampled and plant density (count number of maize plants), and plant bio-volume (biomass) as a proxy of plant biomass were determined (Plate 4.1, 4.2). In a second visit, just before harvest by the farmer the Grain Yield (GY) was measured. The biomass was estimated using the basal diameter (BD), and height (H) of all the maize plant following Chomba *et al*., (2013):

$$BV(cm^3) = H(cm) \times \left(\frac{BD(cm)}{2}\right)^2 \pi \text{ ....Equation 4.1}$$

where BV is the Biovolume, BD is the basal diameter and H is the height. The BD was measured in duplicate, 2 cm above the soil surface for all maize plants in the plot (Plate 4.3). Plant biomass was used to test whether it could be an indicator of crop performance rather than GY. Maize grain yield (14% moisture content) was measured from dry maize that was hand-harvested and the kernels removed and weighed (kg) between 50-60 days after the silking stage. Yield and biomass, were the used as proxies of crop response, which reflect variability in the inherent soil fertility across the maize fields. In the rest of this Chapter 4, the term CPIs is used to refer to both GY or maize biomass and the term biomass and biovolume are used interchangeably, but have the same meaning.

To characterize soil properties, composite soil samples were taken per plot. Using a transect in a zig zag pattern, six topsoil (0-20 cm) samples were taken with an Edelman soil auger (600 cm^3) within the 2.5 m^2 plots to have a representative composite soil sample (Plate 4.4). Sub samples from the composite soil samples, obtained using coning and quartering, were analysed at Crop Nutrition Laboratory Services and World Agroforestry Centre laboratories.

The samples were air-dried, thoroughly mixed and ground to pass a 2 mm sieve prior to the analysis. Soil pH was measured with a pH meter with a 1:2.5 soil/water suspension (Okalebo *et al*., 2002). Soil Organic Carbon (SOC) and total N were analysed by dry combustion using a C/N analyser (Wright and Bailey, 2001), with an acidification pretreatment to remove carbonates for SOC detemination. Extractable Ca, Mg, K, Na, and P were determined using the Mehlich-3 extract (Mehlich, 1984) and an inductive coupled plasma optical emission spectrometer (Sikora *et al*., 2005).

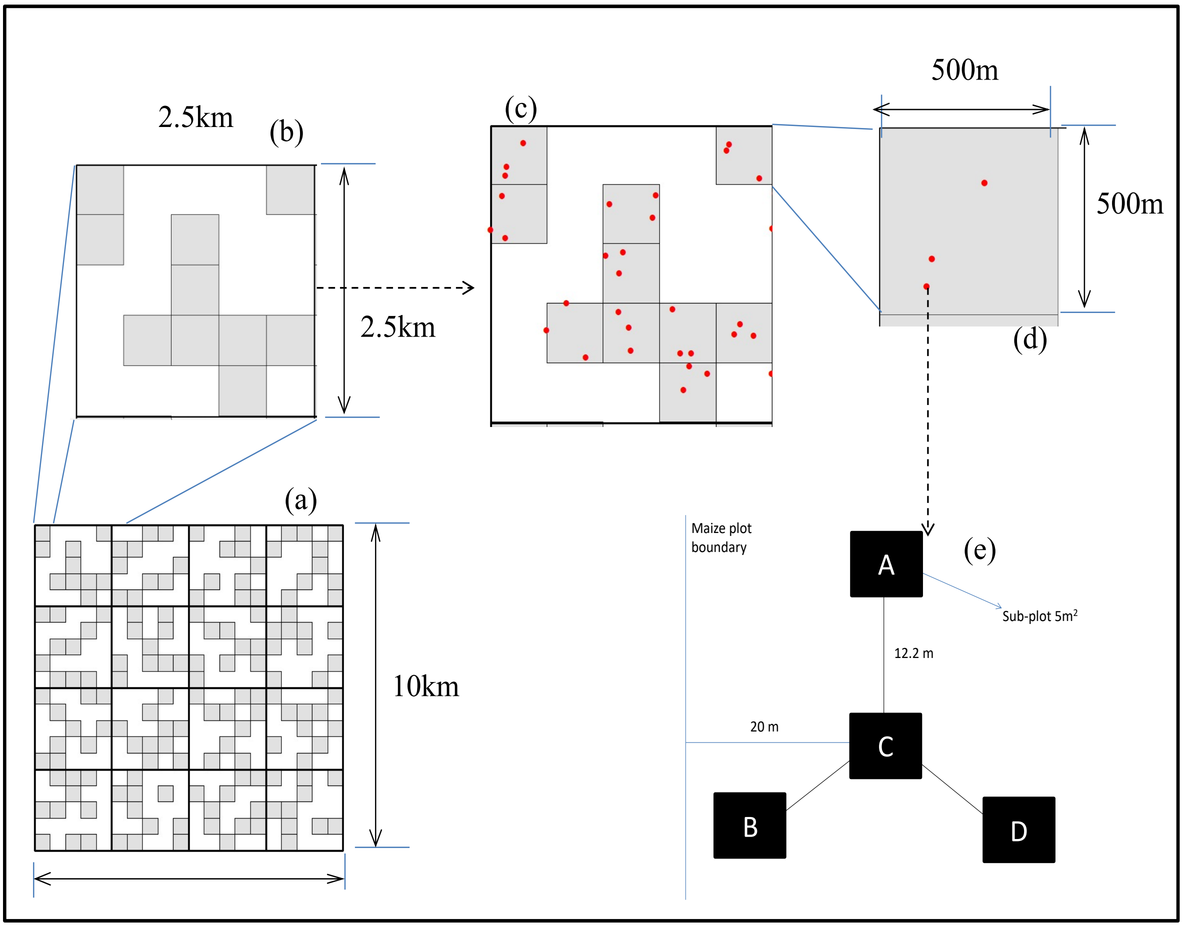

Figure 4.2: Schematic illustration of sampling strategy used for this study. (a) Block measuring 100km^2 with eight randomly selected tiles. (b) Tile measuring 6.25 km^2 with four randomly selected sub-tiles (c) Sub-tile measuring 0.25 km^2 with randomly selected maize fields, 2 for sampling and an alternate point. (d) Plots (A, B, C and D) within the maize field.

Plate 4.1: Global Positioning System (GPS) device was used to navigate and identify maize field being across the smallholder landscapes in western Kenya

Plate 4.2: Measurement of the height of a maize plant, was taken on the maize field for calculation of bio volume

Plate 4.3: Veiner calliper was used to make measurement of the diameter of maize plant taken on the smallholder farm for calculation of plant biovolume

Plate 4.4: Soil auger used to extract samples from the 2.5 m^2 plots for characterize soil properties for the farm survey.

4.2.4 Variability in soil properties and crop performance

To assess the variability in soil properties and CPIs, descriptive statistics; means, median, coefficient of variation (CV), minimum and maximum were used. Density plots were used to check whether soil properties and CPIs followed a near normal distribution as a requirement for the subsequent steps of multivariate statistical analysis of soil properties and CPIs. Variables with a skewed distribution were log-transformed to achieve the near normality. Relationships among soil properties, and between soil properties, and CPIs were evaluated using a pair-wise correlation analysis to derive the Pearson correlation coefficients (r). Challenges with multi-collinearity can occur when highly correlated variables are included in a regression model (Wold *et al*., 1984), Therefore, variables highly correlated to another variable ($r > 0.80$) were identified prior to further statistical processing (ter Braak & de Jong, 1998).

4.2.5 Key soil drivers of crop performance indicators

Key soil properties that can be attributed to the variation in CPIs were identified by Stepwise Multivariate Linear Regression (SMLR) modelling (Geladi *et al.*, 1999). The regression method was used to analyse the linear relationship between single dependent variables (CPIs) with the independent variables (soil properties) based on Equation 4.2.

$$y = a + \sum_{i=1}^{n} b_i \times x_i \pm \varepsilon \qquad \text{...4.2}$$

where "y" is the CPIs, "x_i" are the soil properties, "n" is the number of soil properties, "a" is the intercept, "b_i" are the regression coefficients and "ε" is the standard error of the estimate. The SMLR analysis was used because of; *(i.)* the few numbers of independent variables included in the analysis, (pH, SOC, TN, P, K, Ca, Mg, Na) and, *(ii.)* to avoid the problem of overfitting by adding or deleting variable in SMLR analysis (Guan *et al.*, 2013). The data was evaluated for outliers which were expunged from the analysis to minimize the problem of overfitting in the regression models.

The relative importance of the predictors in the regression model were calculated and used to discern the key soil factors. The significance probability ($p < 0.001$) and coefficient of determination (r^2) were used as basis of evaluation. The r^2 refer to the conventional coefficient of determination - the proportion of variance explained by soil factors. A bootstrap re-sampling strategy was used to assess the strength of evidence, that indeed the identified soil predictors were truly independent and reproducible. Hence, the mean confidence intervals for each soil predictors were estimated using 1000 iterations. The relative importance of the predictors was calculated using the "*relaimpo*" R package (Grömping, 2006).

To further discern which soil properties were key, the soil predictors were gradually eliminated using a stepwise criterion, a backward elimination and forward selection, to build the best regression equations describing CPIs as a function of the soil properties. The predictors with the lowest contribution ($p > 0.005$) to the regression model were eliminated and then tested until the remaining ones had a significant contribution. The key predictors of CPIs were then identified in the final regression equation. The Akaike's Information Criteria (AIC) value was computed and used it to evaluate whether the identified soil predictors were similar to those obtained using the aforementioned relative importance (r^2) statistic. The magnitude of AIC values formed the basis for interpretation, where the regression model with the lowest AIC was considered as the best, while the soil predictors with the highest AIC value were taken to be the key soil property. The regression modelling was done using the "*lme4*" R-package (Bates *et al.*, 2015).

4.2.6. Scale of variability

Discrete map units like AEZ, soil map units, and administrative boundaries form a logical basis for the development of blanket fertilizer recommendations (Smaling and Van De Weg, 1990). The study area fell within a single AEZ. However, different soil map units and administrative units were identified. These units are only useful for the refinement of the fertilizer use recommendations if they describe the variation in the key soil properties that describe the CPIs. Soil types are expected to describe the variation in soil properties, and could form a logical basis for fertilizer recommendations. Administrative units are instrumental to agricultural extension officers for logistical purposes of disseminating fertilizer recommendation. However, the administrative units may include considerable variation in soil properties, making them less useful for fertilizer recommendations. The hierarchy of scales across the LDSF sampling framework (Figure 4.2) was considered since nutrient variability occurs at different scale levels across smallholder landscape (Tittonell *et al.*, 2013; Zhu *et al.*, 2017).

ANOVA mixed-effect linear models were conducted to analyse the variation across the soil map units, administrative units, and the different scales of the LDSF sampling frame using the identified key soil predictors and CPIs as dependant variables. The "*nlme*" R-package was used to conduct the unbalanced ANOVA, where the plots (2.5 m^2) were fitted as random effects (Pinheiro *et al.*, 2019). The administrative units, soil types, tiles, sub-tiles and fields were considered as the fixed effects and provided a measure of explained variability (EV) by each of the mapping units. One mapping unit was modelled at a time with the response variable being the identified key soil properties or the CPIs.

To evaluate the proportion of EV between the mapping units, two pseudo R^2 summary statistics for mixed-effects models were estimated as described by Nakagawa & Schielzeth, (2013). These were "marginal" (R^2_m), which considers the variance of the fixed effects, and "conditional" (R^2_c) that takes the variance of both the fixed and random effects into account (Nakagawa *et al.*, 2017). These statistics were computed following Equations 4.3 and 4.4, respectively.

$$R^2_m = \frac{var_f}{var_f + var_r + var_e} \quad \text{...} 4.3$$

$$R^2_c = \frac{var_f + var_r}{var_f + var_r + var_e} \quad \text{... ....} 4.4$$

where var_f is the variance of the fixed effects, var_r is the variance of the random effect and var_e is the variance of the model residuals. High values of R^2 indicates the mapping units such as soil types/maize fields may be appropriate for a blanket fertilizer recommendation.

Often discrete mapping units do not properly describe the relevant variability, specifically the relevant variation, for example, where the administrative boundaries are used. An alternative approach to describe the spatial variation in the key soil properties was to carry out a geostatistical analysis to determine the spatial dependency of the key soil properties and CPIs. Semi-variograms for key soil drivers and CPIs were derived following Kerry *et al*. (2010) using the "*gstat*" *R*-package (Pebesma, 2004). The range was interpreted as the basis to define the relevant scale of variability following Kerry

& Oliver (2004). The interpretations depend on the strength of spatial dependencies that was determined by the nugget/sill ratio following Cambardella *et al*. (1994) with ratios < 0.25 depict strong, 0.25 – 0.75 moderate and > 0.75 weak spatial dependencies. Weak spatial dependence implied that there were considerable short distance variation and no logical patterns (Costa *et al*., 2015).

4.2.7 Validation

In steps 1- 4 (section 4.2.3 – 4.2.6) a farm survey methodology was developed and used to analyse the spatial variability in soils and crop performance. However, to properly derive fertilizer use recommendations, it is necessary to look at fertilizer response. However, under normal conditions, it would be very resource intensive to carry out a large number of fertilizer response trials. Therefore, the hypothesized notion that the scale of variability of key soil drivers of CPIs corresponds to the scale of variability in fertilizer response across smallholder farms was tested. In the study area a large number of fertilized maize trials were carried out by the African Soil Information Service (http://afsis-dt.ciat.cgiar.org; last accessed on 18th April 2019) and the International Institute of Plant Nutrition (IPNI) (Huising *et al*., 2011; Zingore *et al*., 2014). This allowed testing of the above hypothesis. Data of fertilized maize trials were obtained in the short and long rainy season of 2010 and 2013. These trials consisted of N, P and K fertilizer treatments. Maize yield of the fertilized and control plots was used to calculate the Fertilizer Response (FR), computed as a response ratio following Hedges *et al*. (1999) based on equation 4.5.

$$ln\,\mathrm{FR} = ln\left(\frac{y_c}{y_t}\right) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 4.5$$

where y_c is yield from the control plot and y_t is the yield from the treatment plots all reported in Mg ha^{-1}. The FR was transformed into the *ln*FR to achieve normality. Geographical coordinates, corresponding to each fertilized plot were used to determine the spatial dependency of *ln* FR as described in section 4.2.6. The spatial dependency in FR is compared to the spatial dependency in soil properties and CPI's.

4.3 Results

4.3.1 Farm survey

A total of 64 maize fields typifying smallholder farms were sampled within 32 sub-tiles that were randomly distributed within 8 tiles across the 100km^2 block. An average of 7 maize fields within each tile was sampled. Out of 256 plots sampled, 203 had complete observations on soil properties, GY and biomass. Yield was not observed in 20% of the fields as farmers harvested the field prior to the planned date. Seven observations were identified as outliers, as this would influence the overall results. The total number of maize plants at silking stage per plot ranged from 7 to 25. At harvest, a 10% reduction in total number of maize plants was observed for all fields sampled during biomass estimates.

4.3.2 Variability in soil properties and crop performance indicators

Table 4.1 presents the mean, median, maximum and minimum values for soil properties and CPIs. Soil properties in the topsoil indicate considerable variation in soil fertility. SOC concentrations ranged from 0.56 to 5.23%. The highest values of SOC were found in fields that had recently been converted to maize cultivation and those that displayed intensive soil management (14% of the observations). Low SOC values were observed in maize fields that were intensively cultivated. Soil pH varied from slight acidity (4.8) to near neutrality (7.4) and within the optimum range for maize growth. Mehlich-3 extractable P was below the critical concentration of 15 mg kg^{-1} in 55% of the sampled plots. Grain yield ranged widely from 0.8 to 11.8 Mg ha^{-1}.

Coefficients of variation indicated different degree of variation within soil properties and CPIs (Table 4.1). Mehlich-3 extractable P and K were highly variable with CVs of 74 and 89% respectively. SOC and total N were moderately variable with CVs of 32 and 26%, respectively. Yield and biomass exhibited a moderate variation as indicated by their CVs of 57 and 43%, respectively. Coefficient of variation was used to assess variability since it allowed comparison among variables with different units of measurement; soil properties and CPIs. However, the CV statistics could not allow the explicitly evaluation of spatial variation in soil properties (Haileslassie *et al*., 2005).

Density plots for Mehlich-3 extractable P, K, Ca, Mg and Na displayed a negatively skewed distribution, indicating presence of low values in the dataset, which signifies low nutrient levels in the study area. Soil pH, total N, GY and biomass displayed a near normal distribution.

Table 4.1: Soil properties and crop performance on 203 unfertilized maize plots across a smallholder landscape in Western Kenya.

Soil property	**Mean**	**Coefficient of Variation (%)**	**Minimum**	**Median**	**Maximum**
Soil pH (water)	5.70	8.84	4.78	5.61	7.36
SOC (%)	1.57	31.61	0.56	1.55	5.23
Total N (%)	0.15	26.21	0.06	0.14	0.32
Mehlich-3 P (mg kg^{-1})	22.97	74.55	3.65	17.60	89.90
Ca (cmol kg^{-1})	5.80	59.91	0.88	5.00	24.45
Mg (cmol kg^{-1})	1.99	52.10	0.29	1.76	6.61
K (cmol kg^{-1})	0.42	89.29	0.08	0.31	2.74
Bas (cmol kg^{-1})	8.39	54.91	1.42	7.48	31.92
Na (cmol kg^{-1})	0.18	62.01	0.01	0.16	0.77
Crop Performance Indicators (CPIs)	**Mean**	**Coefficient of Variation (%)**	**Minimum**	**Median**	**Maximum**
Grain yield (Mg ha^{-1})	3.62	56.83	0.08	163.60	11.28
Maize Biomass (cm^3)	170.40	43.02	31.00	3.20	392.90

Key : n= number of observations, pH =soil pH, total C= total carbon, total N= total nitrogen, Mehlich-3 P= phosphorus, Ex.K= Exchangeable bases, Ex.Ca= Exchangeable calcium, Ex.Mg= Exchangeable magnesium, Ex.Bas= Exchangeable bases, Ex.Na= Exchangeable sodium, ESR= Exchangeable sodium ratio, ESP= Exchangeable sodium percentage, Ca.Mg=Calcium magnesium ratio, Biomass = Maize biomass and Yield= Maize grain yield. Symbols in brackets are the units.

Significant correlations between soil properties and CPIs was evident, as shown by r^2 values (Table 4.2). *ln*SOC was positively correlated with yield ($r^2 = 0.55$, $p < 0.001$) and biomass (r=0.88, $p < 0.0001$). Relationship between *ln*P and yield ($r^2 = -0.01$, $p = 0.02$), *ln*Na and yield ($r^2 = -0.01$, $p = 0.101$) were weak and insignificant, and there was no relation between *ln*P and biomass.

Results of pairwise correlation also revealed high correlation among soil properties ($r^2 > 0.8$). *ln*SOC and total N were highly correlated ($r^2 = 0.95$, $p < 0.001$), and so were *ln*Ca and *ln*Mg ($r^2 = 0.89$, $p < 0.001$). One of the highly correlated soil properties was expunged, before the next step of regression analysis. For example, total N and *ln*Mg were removed, because they were highly correlated with *ln*SOC and *ln*Ca, respectively. Correlation coefficients between Soil pH, *ln*SOC, *ln*P, *ln*Ca, *ln*K and *ln*Na were relatively low and they were therefore all considered in identifying the key soil factors which influence underlying variation in CPIs for the study area.

Table 4.2: Pearson pair-wise correlation coefficients between soil properties, maize plant biomass and grain yield from maize fields across the smallholder landscape

Soil Property	**Grain yield (GY)**	**Maize biomass (BV)**
Soil pH	0.08**	0.16**
Total N	0.53***	0.87***
lnSOC	0.55***	0.89***
lnP	0.01	0.02
lnCa	0.32*	0.51**
lnMg	0.33	0.55
lnK	0.19	0.25
lnNa	0.001	0.08
lnBas	0.33**	0.53***
Maize biomass	0.55***	1
Grain Yield	1	

Key: Significant codes: **** = 0.001, ***= 0.01, **= 0.05, * = 0.1, *ln* = natural log

4.3.3 Key soil drivers of crop performance

Regression results indicate soil predictors explained 32 and 79% of the variability in maize yield and biovolume, respectively (Table 4.3). SOC was the main factor that significantly ($p < 0.001$) contributed to the variability in CPIs. The explained variance by the soil predictors indicate that each individual soil property played a role in influencing the underlying variation. But the contribution of pH, Total N, *ln*P, *In*Ca, *ln*Mg, *In*K and *In*Na was not statistically significant. The problem of multicollinearity could have had an influence on the performance of the model, since *ln*SOC and total N, as well as *ln*Mg and *ln*Ca were highly correlated, previously mentioned in Table 4.2. To test the aforementioned influence, total N and *ln*Mg were removed from the model. This reduced the explained variance to 31 and 78% for yield and biovolume, respectively, and the intercept become significant, meaning the model accuracy was not affected. Hence, the result indicates no influence of multicollinearity when all eight soil properties were included in the regression models. Multicollinearity creates high coefficient estimators that inflates variances and may lead to selecting the wrong soil predictors (Kroll & Song, 2013). The problem is magnified when the samples size is small contrary to this study, which had 196 observations (Kroll & Song, 2013). Thus, the multiple linear regression models predicted variability of CPIs fairly well as shown by the explained variance for the study area.

The relative importance results for the eight predictors of yield and biomass are shown Figure 4.3. The predictor with highest r^2 was *ln*SOC with values of 41% for yield and 43% for maize biomass. SOC was identified as the key factor that influence variation in CPIs for the study area. The lowest observed r^2 values were 0.12 (pH) for yield and 0.12 (*ln*Na) for maize biomass. The negative influence of soil and high sodicity observed by other workers explain why pH and sodium were the least important soil predictors (Mbakaya *et al.*, 2008). Contrary to obtained soil pH and sodium content that were within the optimum ranges (Table 4.1).

To test the robustness on the predictions obtained from the regression analysis, bootstrapping stimulations, using the pratt and last method were employed (Grömping, 2006). In the pratt method, the VIPs scores are calculated using weighting of sums of squares of the predictors, while last methods the coefficient of determination is used. The bootstrap results confirmed SOC as key factor (Figure 4.4.). Thus, the regression models represent part of variation in CPIs as explained by the soil properties, and was useful for differentiating the contribution of each soil factor. SMLR modelling as a strategy was used to reduce the number of soil predictors, that would be considered for evaluation of spatial structure. This also simplifies the proposed approach for determine a local scale of variability on smallholder landscapes for maize fields.

Table 4.3: Two regression models showing the explained variance r^2 values and the coefficient of determination of the maize yield and maize biovolume for the study area

Crop performance indicator	**Variable**	**Coefficient estimate**	**Standard error**	**t-value**	***p*-value**	
Yield	Intercept	2.429	3.763	0.646	0.5194	
	pH	-0.288	0.396	-0.727	0.4681	
	*ln*SOC	3.824	1.367	2.798	0.0057	**
	Total TN	-0.688	1.615	-0.426	0.6704	
	*ln*P	-0.319	0.257	-1.244	0.2150	
	*ln*ExK	0.272	0.264	1.031	0.3036	
	*ln*ExCa	0.586	0.525	1.116	0.2657	
	*ln*ExMg	-0.545	0.541	-1.008	0.3149	
	*ln*ExNa	-0.216	0.159	-1.354	0.1773	
	r^2 value	**0.324**				
	Adjusted r^2 value	**0.295**				
Biovolume	Intercept	76.671	69.785	1.0990	0.273	
	pH	8.641	7.741	1.1160	0.266	
	*ln*SOC	211.355	25.229	8.3780	0.001	***
	Total N	12.672	29.282	0.4330	0.666	
	*ln*P	-6.593	5.002	-1.3180	0.189	
	*ln*ExK	0.684	4.966	0.1380	0.891	
	*ln*ExCa	-1.174	10.022	-0.1170	0.907	
	*ln*ExMg	-8.432	10.043	-0.8400	0.402	
	*ln*ExNa	-5.602	3.107	-1.8030	0.073	.
	r^2 value	**0.789**				
	Adjusted r^2 value	**0.781**				

Key: level of significance *** = 0.001, ** = 0.01, *= 0.05 and • =0.1 pH = soil pH, N = Nitrogen, *ln*P = Natural log, P = phosphorus, ExK = Potassium, ExCa = Natural log of Calcium, ExMg = Magnesium and ExNa = Sodium

.

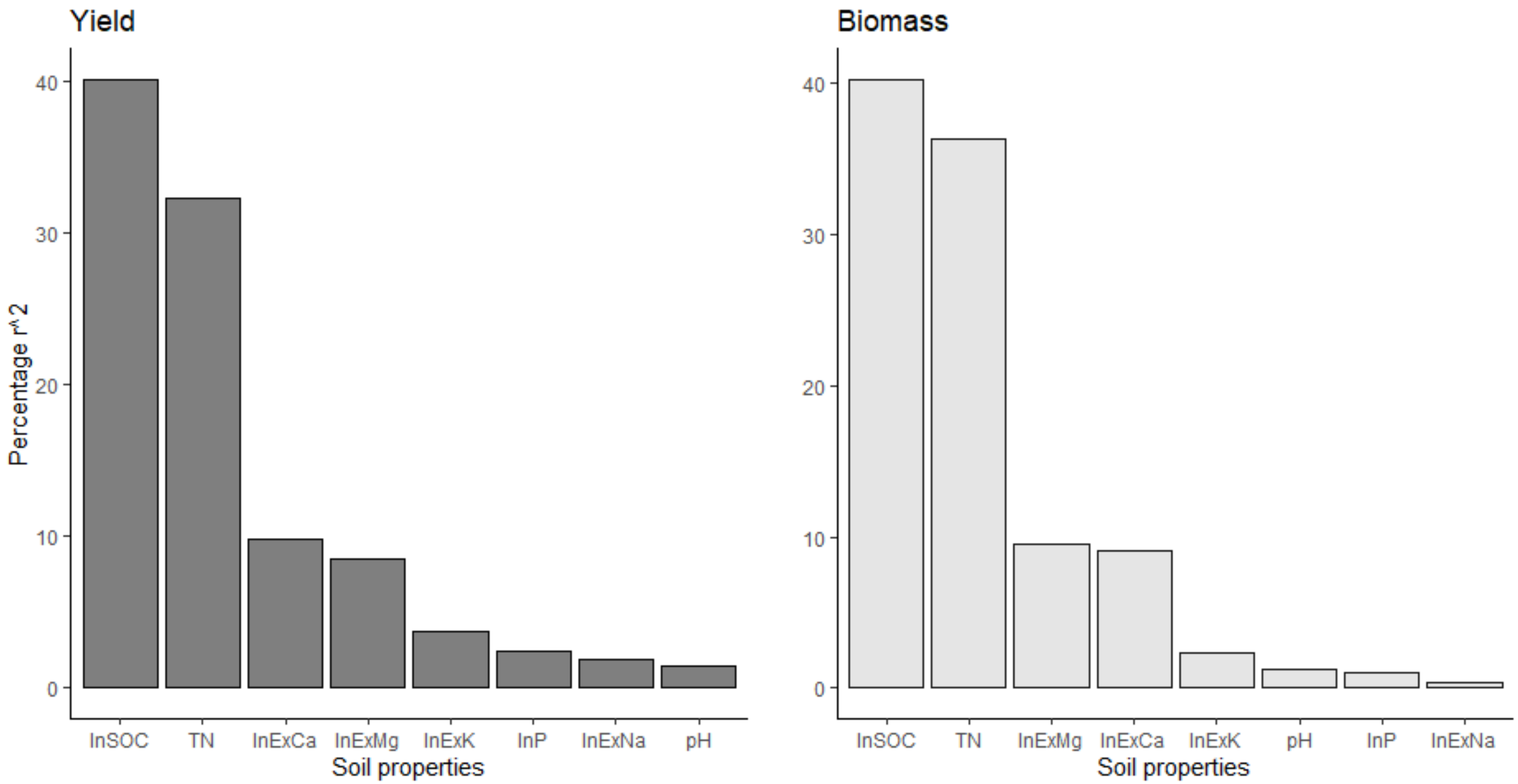

Figure 4.3: The relative importance (percentages) of the contribution of the eight soil predictors to the explained variance for (a) maize yield and (b) plant biomass from regression analysis across the study area. pH = soil pH, lnSOC = Natural log of SOC, soil pH, TN = Nitrogen, lnP = Natural log of phosphorus, InK = Natural log of potassium, InCa = Natural log of Calcium, lnMg Natural log of magnesium and InNa = Natural log of sodium

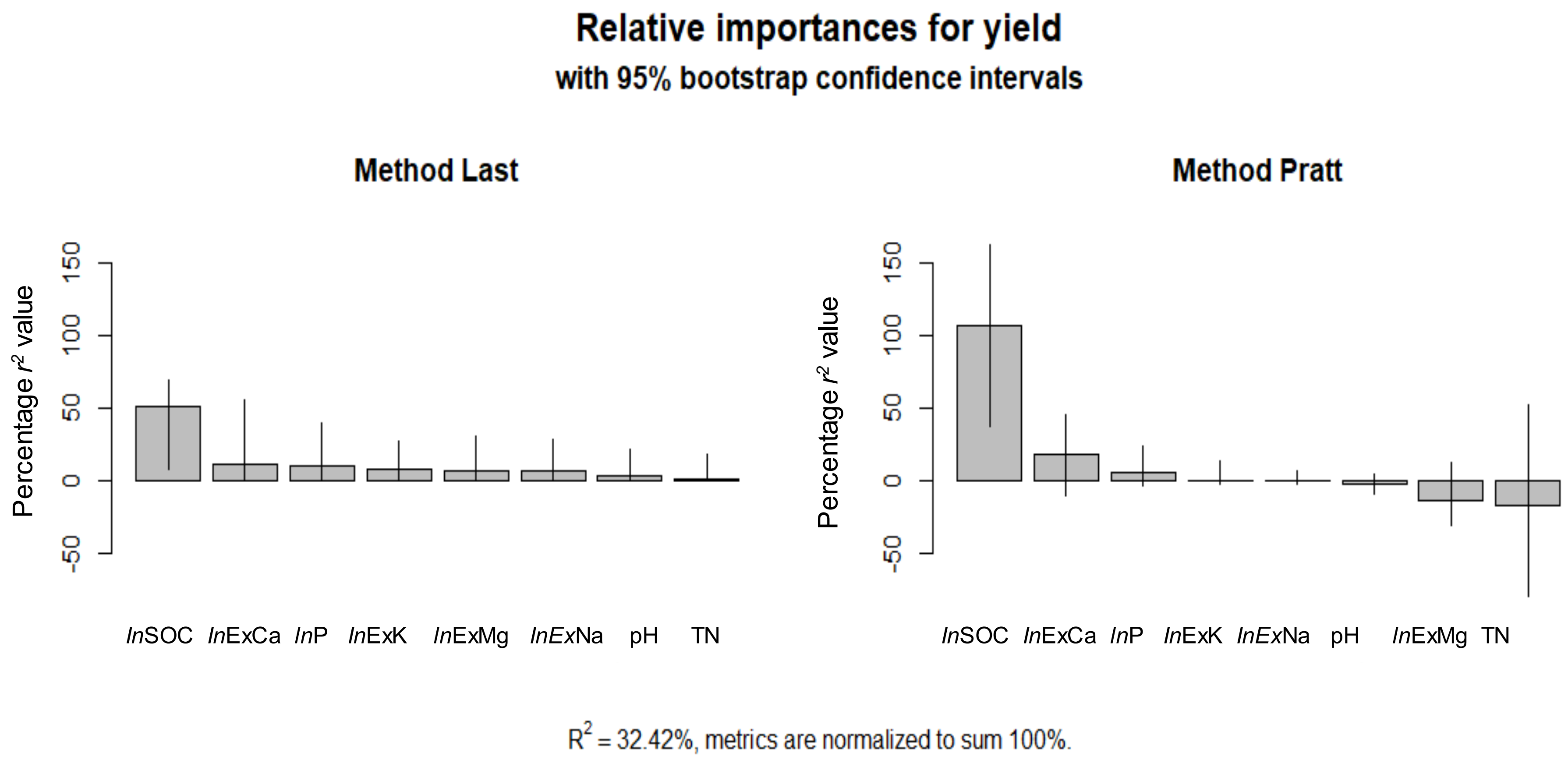

Figure 4.4: Bootstrap results of the confidence interval mean to test the strength relative importance selection, for identifying key soil properties that influence variation and crop performance indicators from the regression analysis. 1000 iteration were used in the stimulation with different methods as described by Grömping, (2006). pH = soil pH, lnSOC = Natural log of SOC, soil pH, TN = Nitrogen, lnP = Natural log of phosphorus, InK = Natural log of potassium, InCa = Natural log of Calcium, lnMg Natural log of magnesium and InNa = Natural log of sodium

To further discern which of these soil properties are key soil factors, AIC values were evaluated from the stepwise regression models (Figure. 4.5). The best regression model included lnSOC and InNa as the main soil predictors for CPIs. *In*SOC was the only significant ($p < 0.0001$) predictor in the models with the highest AIC value of 257 for yield and 2113 for biovolume. Even though InNa was included in this regression equation, its contribution was not significant ($p = 0.7635$) and the AIC values as a predictor for CPIs were the lowest, 187 and 1739, for yield and biomass, respectively. Thus, the results confirmed that indeed SOC was key soil factors that influence variation in crop performance of the study area

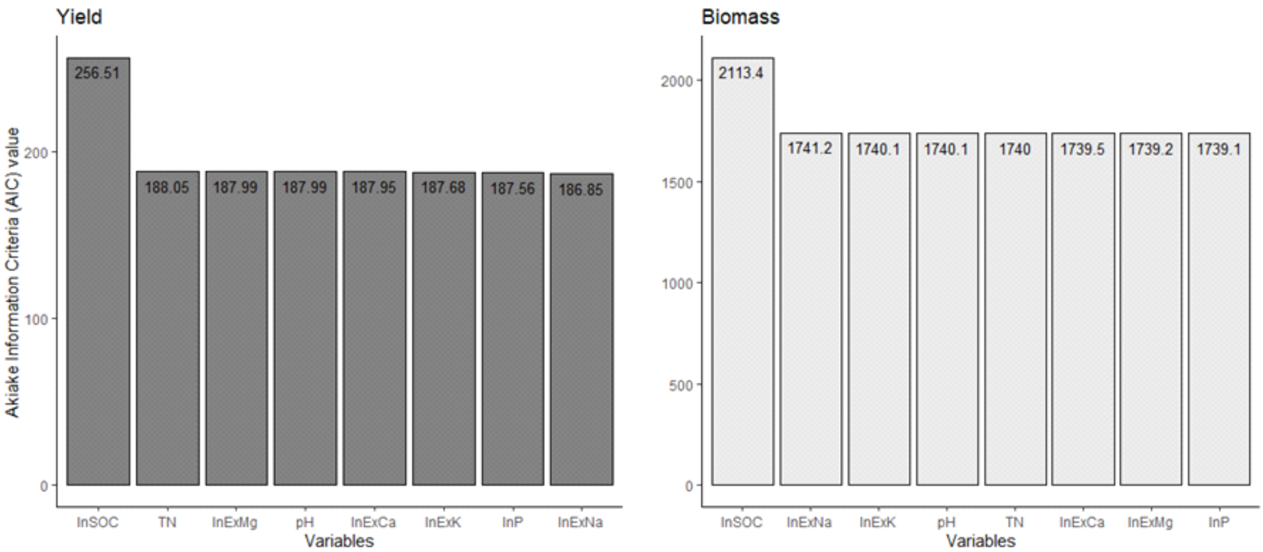

Figure 4.5: Graph with the ranking of the soil predictor based on the magnitude of the Akaike Information Criteria (AIC) values for the best regression model from the step-wise multivariate analysis. These values were used to identity the key soil driver that explain variability in maize grain yield or plant biomass. **Key**: pH = soil pH, lnSOC = Natural log of SOC, soil pH, TN = Nitrogen, *ln*P = Natural log of phosphorus, *In*K = Natural log of potassium, *In*Ca = Natural log of Calcium, *ln*Mg Natural log of magnesium and *In*Na = Natural log of sodium

4.3.4 Spatial variability of key soil drivers and crop performance indicators

The ANOVA models showed that the soil types and administrative boundaries describe less than 10% of the variation in the SOC (Figure 4.6). The mixed-effect models result indicated low marginal R^2 (Figure 4.6. a c) but high conditional R^2 (Figure 4.6 b d). This meant that the fixed effects (administrative boundaries, soil types, LDSF scales (tile, sub-tile and field)) explain low variability (< 5 %) in SOC. Most of the spatial variability in *ln*SOC was attributed to differences in fields, and between sampling the plots, indicated by high marginal R^2 values (Figure 4.6 a c). This implied that most of the local variability was capture at field level (within variation), which make them good basis for the development of a fertilizer use recommendation.

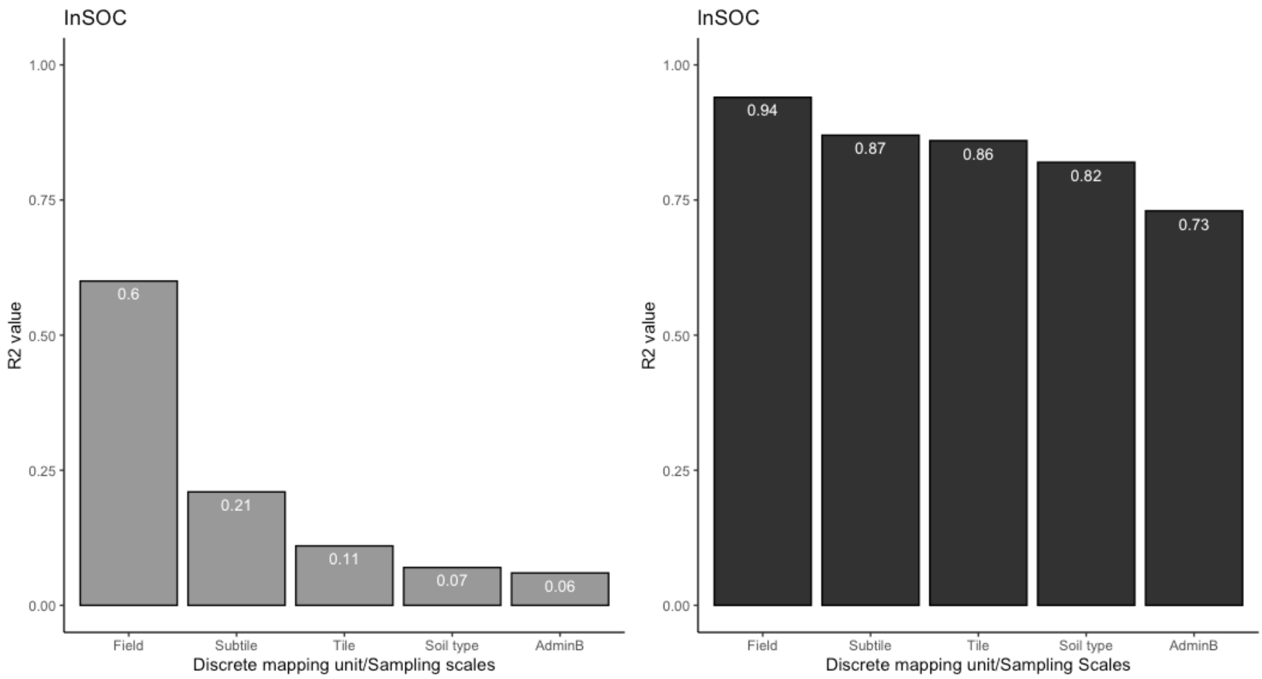

Figure 4.6: Explained variance for the fixed and mixed effects soil organic carbon (*ln*SOC) across the discrete mapping units. The units are soil types (SoilType), and administrative boundary (AdminBond.) and the scales of the LDSF sampling framework.

A similar trend was observed for the CPIs (Figure 4.7). Higher conditional R^2 for at field scale were observed across the CPIs. The high conditional R^2 is attributed to the inclusion of variance for both the larger mapping units (fixed effects) and plots (the random effects). Although the mean *ln*SOC, yield and biovolume were significantly ($p < 0.01$) different between the three soil types, there is considerable variation within these soil units. This is confirmed by the results of the stratification following the LDSF framework. Here, it became apparent that the smaller mapping units (fields) were describing considerable variation in *ln*SOC, yield and biovolume compared to all the other stratifications of the landscape that were applied, even despite the fact that they were just randomly located squared in the landscape.

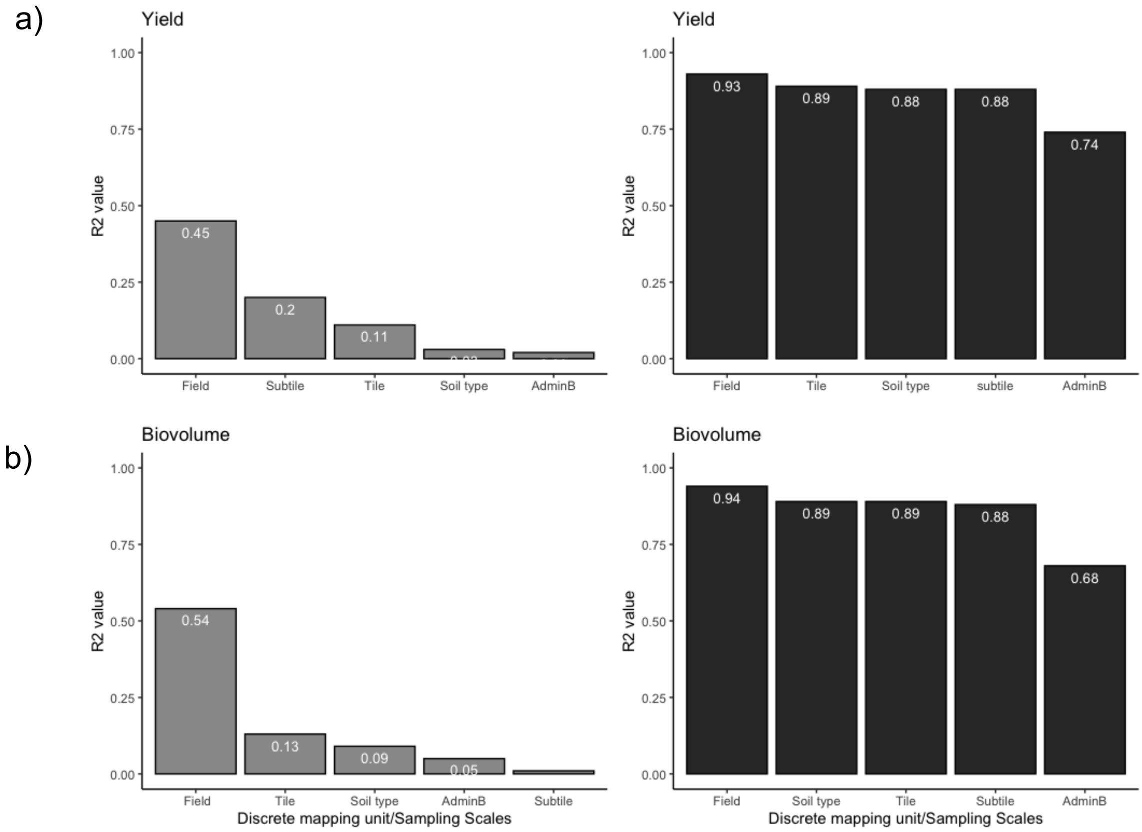

Figure 4.7: Explained variance for the fixed and mixed effects for crop performance indicators (CPIs), (a) grain yield and (b) biomass across the discrete mapping units. The units are soil types (SoilType), and administrative boundary (AdminBond.) and the scales of the LDSF sampling framework.

Table 4.4 shows the semi-variogram parameters for the key soil properties and CPIs. *ln*SOC and CPIs showed moderate to strong spatial dependencies. However, the semi-variograms showed considerable short distance variability, confirming the results in the literature that these systems present considerable short distance variability (Okeyo *et al.*, 2009). Although *ln*SOC showed considerable spatial dependency, the patterns were found to occur at relatively short distances with a range of 523 m. Despite the relations between the soil properties and the CPIs, the short distance spatial dependency shown for the soil properties (or a short range for *ln*SOC) is not found for the CPIs. The CPIs show a stronger spatial dependency and also a longer range. The results of the geostatistical analysis were in line with the analysis of variance. The short distance variability found for SOC and the CPIs indicated that the very general soil units and large tiles, which covered areas of > 5 km^2 did not describe the local variation on smallholder landscape. Smaller areas like the sub-tiles and the administrative units are roughly the size of the range of the semi-variograms and describe more variation for *ln*SOC and CPIs as shown by the analysis of variance.

The results confirmed that blanket fertilizer use recommendations on the basis of the soil map are not likely to be efficient in contributing to increase of food production. The development of fertilizer use recommendations would therefore require intensive sampling to describe the variation in soil conditions. Following Kerry & Oliver (2004), the results can be re-interpreted towards optimal sampling densities. The optimum sampling distance should be less than half the range of a fitted semi-variogram model. Given the relatively short range of 523 m for *ln*SOC, it is necessary to sample at distances of less than 250 m. Although this would describe the large trends, it should be recognized the results would still not be very effective due to short distance variability as indicated by the nugget. Thus, local variability within smallholder farms may be captured at 250 m resolution. This leaves two options. Either the interpolation can be enhanced by suitable co-variables derived from *e.g.,* detailed satellite imagery (digital elevation models) through regression kriging, or farmers rely on soil testing on their fields. The two CPIs showed different levels of spatial variation with different spatial dependencies. High range values were observed for biomass (3291 m) and yield (968 m). This is not surprising given the relatively low correlation coefficient between with *ln*SOC ($r = 0.56$). This can be attributed to the roles of other factors influencing crop response in the smallholder farming system (Waithaka *et al.*, 2007)

4.3.5 Validation

The results showed that there was considerable soil variation at local scale (< 543 m). The best semi-variogram model fitted for *ln*FR was spherical, which corresponded to the least root mean square error. The nugget/sill ratio suggest *ln*FR exhibited moderate of spatial dependencies across fertilized plots for the study area (Table 4.4). A high nugget effect (> 0.1) was observed (Table 4.5) suggesting that in fertilizer response there was small-scale variation cross fertilized maize plots. Moderate spatial dependencies (65%) implied that interpretation the range distance was necessary. The range of *ln*FR was 425 m and did not correspond to that of SOC (543 m) (Table 4.5). However, in practical terms the difference can be seen to insignificant ($p < 0.001$). Hence the hypothesis, that the scale of fertilizer response corresponds to the scale of variability of key soil properties is valid.

Table 4.4: Spatial dependency of key soil properties, crop performance indicators, and fertilizer response for maize fields in a smallholder western Kenya in terms of the semi-variogram. Strong < 25%, moderate 25 – 75 %, Weak > 75 %

Key soil divers	Fitted semi-variogram Model	Nugget (C_o)	Partial Sill (C_1)	(C_o)+(C_1)	Nugget: Sill (NS) ratio	Range (m)	Spatial dependency
			Farm survey				
lnSOC	Exponential	0.0394	0.0589	0.0983	0.60	543	Moderate
Crop response							
Grain yield	Linear	0.3094	0.0963	0.4057	0.24	3291	Strong
Maize Biovolume	Exponential	0.1078	0.1048	0.2126	0.49	968	Moderate
			Fertilizer trial				
Fertilizer response	Spherical	0.9026	0.9368	1.8394	0.50	425	Moderate

4.4. Discussion

Understanding the local spatial variability on smallholder maize fields can aid farmers and policy makers in making efficient nutrient management decisions. Empirical rules were derived, to describe local variability of key soil factors - SOC and CPIs, which were then used to determine a directional flow of decisions (Figure 4.8). The results provide evidence of existence of variability as indicated by soil properties, that displayed high variation, as shown by high CV values of the study area (Table 4.2). Anthropic influence affects spatial structure of soil properties in maize fields (Haileslassie *et al.*, 2005). Variability in soil properties can be attributed to natural intrinsic variation, parent material (Deckers, 2002) and difference in management across maize fields (Zingore *et al*., 2007). The high variability was observed in maize yield and were within ranges of maize yield reported by Kihara *et al*., (2016), and has been captured in other studies conducted in western Kenya at landscape level (Tittonell *et al.*, 2013; Burke & Lobell, 2017). Studies have shown the impact of high soil variability on nutrient requirement for maize crop in smallholder farms of Nigeria (Shehu *et al.*, 2018), that has consequently led to variable fertilizer use efficiency and poor fertilizer response (Tittonell *et al.*, 2007a; Njoroge *et al.*, 2017). Maps displaying spatial patterns of soil properties may capture the variability at specific locations and provide information for the local inherent soil variability (Antwi *et al.*, 2016). The use of auxiliary information in digital soil maps would further allow integration of information such as rainfall data or elevation (as raster maps) in creation of digital soil maps and improve on the accuracy of predicting soil properties. This study provided a sequential framework that can aid in capturing local variation in soil properties and crop responses (Figure 4.8)

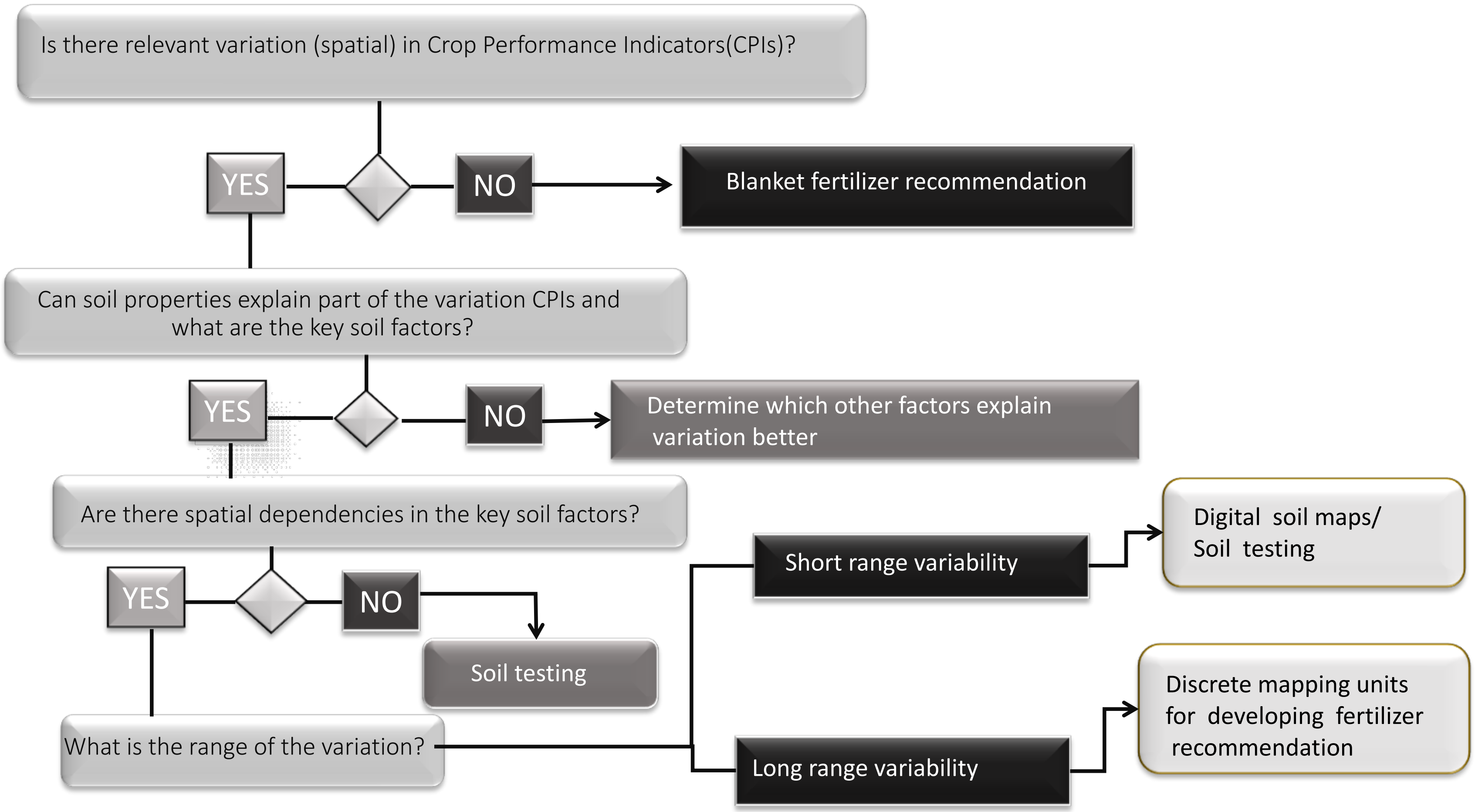

Figure 4.8: Proposed decision tree for determining which a logical way to determine options for providing fertilizer recommendation based on information on relevant spatial variability in crop response or key soil drivers.

The study showed robust relationships between CPIs and soil properties, which were evaluated further (Table 4.2). Pair-wise correlations results between soil properties and CPIs were statistically significant ($p < 0.001$) and provide evidence of the existence of a relation between inherent soil fertility and CPIs (Table 4.2). The most striking correlations were between SOC and biovolume ($r = 0.88$, $p < 0.0001$) and between SOC and yield ($r = 0.55$, $p < 0.001$), which confirm that SOC is an important soil factor, and proper management of SOC would result to high CPIs in the region. These findings also agreed with those reported by Chomba *et al*. (2013). It is difficult to give a reason why there were good correlation between SOC with biovolume and not yield (Table 4.2). However, can be attributed to fact that since measurement were taken when the maize crops were at optimum growth stage. A similar correlation between SOC and maize yield ($r = 0.59$, $p < 0.001$) was observed in other studies conducted in Uganda on Ferrosols (Musinguzi *et al*., 2016). Correlations between soil properties and CPIs, agree with those reported by (Mtangadura *et al*., 2017) who found positive relations between maize yield and Ca in Zimbabwe.

SMLR revealed SOC as key factors associated with the variation of CPIs for the study area (Table 4.3, Figure 4.5, 4.6). This could be explained by the fact that SOC has a dominant influence on N supply, nutrient retention and sulphur supply, soil structure and soil responsiveness to fertilizer application (Lal, 2016; Six *et al*., 2002; Zingore *et al*., 2007). These results were consistent with other studies that reported SOC as key soil factor that influence maize yield as well as fertilizer response (Musinguzi *et al*., 2016). The SMLR statistical methodology implemented in analysing soil data had wider applicability and could be applied to other similar sites and crops. However, 32% of the explained by the regression model for GY implied that other climatic (*e.g.,* rainfall) and management (*e.g.,* tillage) factors, not included in this analysis may influence crop performance (Tittonell *et al*., 2008; Waithaka *et al*., 2007), but would require a larger sample size to increase confidence level of the results (Maas & Hox, 2005). In this case the complex of the regression model would increase and the additional factors may not necessary explain more of the variation in CPIs than the eight predictors including SOC (Wheeler *et al.*, 2012).

Thirdly, an evaluation of the spatial structure variation of key soil factors and CPIs was done using ANOVA. Results of the model results displayed significant ($p < 0.001$) variation between administration boundaries, tile, and sub-tiles. Low explained for *ln*SOC suggests that not much variation (< 25%) was captured by the large discrete mapping units, *i.e.,* administration boundaries, tile and sub-tiles, for this region (Figure 4.5). Broad discrete mapping unit such as administrative boundaries are likely to be inappropriate for delineating fertilizer recommendations, as indicated by the fixed -effects, and marginal R^2 values (Figure 4.5 b, 4.6 b, d). The explained variance was considerable (> 25%) for the small spatial units (*e.g.,* field). This was also evident based on the high conditional R^2 values (> 50 %), which accounted for both the random (plots) and fixed effect (fields, soil types and administrative units). High explained variance for maize fields implied that fertilizer recommendations should be provided at field level for the study area. This may require each smallholder farmer to conduct soil testing of their fields.

Variation of key soil factors was described by very small spatial units and confirmed by low nugget values (0.0039) obtained from the semi-variogram models for *ln*SOC, indicating short distance variability (Table 4.5). This further confirmed that fertilizer use recommendations will need to rely on digital soil maps and/or local sampling at optimum sampling distance. Short distance variability has been reported by Diarisso *et al*. (2015) in the small villages of west Africa. However, short distance variability may limit application of digital soil mapping with coarse resolutions (Keskin & Grunwald, 2018). For such a scenario, soil testing would be an alternative. However, researchers have argued that, the use of soil testing could be more effective, when it is combined with plant tissue analysis (Webb *et al*., 2011) or when cheaper and rapid soil characterisation methods such as infrared spectroscopy (IR) are employed, especially as IR predicts SOC satisfactorily (Shepherd *et al*., 2015).

Accurate evaluation of local spatial variability relies on the scale of measurement or measurement unit. Small plots (2.5 m^2, Figure 4.2) were used to capture spatial variation on the maize fields since small measurement units lead to spatial variance close to the true value, while large units may introduce biases (Western & Blöschl, 1999). Therefore, it is important to consider the measurement unit, which captures the local spatial variability, since the measurement unit influence spatial variability. This approach demonstrates local spatial variability can be captured for the heterogenous smallholder landscape. However, it important to use smaller sampling units as they would increase the precision of estimating variance and the overall spatial variability.

Optimization of a relevant scale has been a major bottleneck for nutrient management in smallholder farms (Vasu *et al*., 2017). The spatial structure of SOC with a strong relation with the CPIs and moderate dependencies was a proper basis. Moderate spatial dependencies (60%) for SOC have previously been reported elsewhere in western Kenya (Okeyo *et al*., 2009). Occurrence of moderated spatial dependencies could be explained by influence of extrinsic management factors such as ploughing and other local management practices that weaken spatial dependencies after long history of cultivation (Mzuku *et al.,* 2005). For this study area, 250 m was proposed as the optimum resolution for digital soil maps, given the effective range of 543 m for SOC. This distance can serve as threshold scale below which maps would capture the local growing conditions of the study area. Other studies have proposed a similar distance of 323 m for rain-fed conditions (Vasu *et al*., 2017). The reliability of these digital nutrient maps would also depend on the sampling protocol and accuracy of the semi-variogram model (Liu *et al*., 2014). However, at 250 m sampling distance may have impact on the cost of soil analysis.

Finally, the hypothesis that the spatial range of SOC was similar to that of fertilizer response was tested. The mean range of SOC (523 m) and FR (423 m) were statistically insignificantly (p > 0.001), therefore the hypothesis was accepted (Table 4.4). Even though the ranges were not exact in their magnitude. The minor discrepancies in terms of lack of exact correspondence of range between SOC and fertilizer response for trial data can be attributed the difference in sampling density between the two approaches. The sampling density for farm survey across the landscape was higher at 2.3 samples per 100 km^2 compared to that of fertilizer trials at 0.42 samples per 100 km^2. Many studies reported that high density sampling is required for better results where soil pattern is complex due to the topography (Cobo *et al*., 2010; Tesfahunegn *et al*., 2011). Difference in factors such as weather patterns and germplasm can also explain the discrepancy, since data collection was conducted in the same study area, but during different seasons and, sampled different maize varieties among farms.

4.5 Conclusions

This study demonstrates an approach for establishing a relevant scale for making fertilizer recommendations that captures spatial variation of soil properties and CPIs based on local conditions. The following is a summary of the conclusions:

(i) SOC was the key soil factors that determined variability in unfertilized maize grain yield and plant biovolume of this region.

(ii) Discrete mapping units based on soil classification, administrative boundaries, or agro-ecological zones may not be suitable for delineating fertilizer recommendations for smallholder farms in the study area.

(iii) Only SOC show moderate spatial dependencies and was used for interpretation of a suitable scale that could provide the relevant spatial detail of maps for nutrient management for this study area.

(iv) Based on the spatial correlation distance of SOC, which displayed an effective range of 523 m, a resolution/distance of 250 m is proposed as the threshold scale for developing digital nutrient maps or optimum sampling distance for soil testing (Kerry & Oliver, 2004).

(v) This finding provide approximation of scale as a basis for guiding fertilizer recommendations for maps and future efforts should be directed at improving its accuracy of this rough estimate for smallholder farmer to use for crop prdoction.

Chapter Five

Spatially explicit approach for diagnosis of yield-limiting nutrients in smallholder agroecosystem in western Kenya

Abstract

Adept use of fertilizers is critical if sustainable development goal of zero hunger and agroecosystem resilience are to be achieved for African smallholder agroecosystems. These heterogenous systems are characterized by poor soil health attributed mainly to soil nutrient depletion. However, conventional methods do not take into account spatial patterns across geographies within landscapes, which poses great challenges for targeted interventions of nutrient management. This study aimed to develop a soil test value using a population-based farm survey method for diagnosing soil nutrient deficiencies. The approach embraces the principles of land health surveillance of problem definition and rigorous sampling scheme. A farm survey was conducted on 64 maize fields, to collect data on soil and plant tissue nutrient concentration and grain yield for maize crop. Correlation analysis was used to establish soil test values, by evaluating relations between grain yield and the tissue nutrient concentrations. Diagnosis Recommendation Integrated System (DRIS) indices for nitrogen, phosphorus and potassium (NPK) were used to rank and map prevalence of nutrient limitations. Weak but significant correlation existed between plant tissue nutrient concentration and grain yield with r of 0.089, 0.033 and 0.001 for N, P, K, respectively. Soil test cut-off values were 0.01 g kg^{-1}, 12 mg kg^{-1}, 4.5 $cmol_c$ kg^{-1} for N, P, K, respectively. Nitrogen and K were the most limiting nutrients for maize production in 67% of target population. The study demonstrated that, population-based farm survey of crop fields can be a useful tool in nutrient diagnosis and setting priorities for site-specific fertilizer recommendations. Therefore, larger scale application of the approach is warranted.

This chapter is based on an article accepted for publication in the *PLOS* One Journal.

5.1 Introduction

Smallholder agroecosystems supports livelihoods of 1.2 billion people and are the backbone of the rural economy (FAO 2015; Goswami *et al.,* 2017). The agroecosystems play a significant role in food production, poverty alleviation and mitigation against hunger for rural populations (Samberg *et al.,* 2016). The smallholder systems are characterized by soil fertility degradation (Bekunda *et al*., 2010) and poor quality germplasm (Vanlauwe *et al*., 2010), which constraint crop production (Ngome et al., 2013; Tittonell & Giller, 2013). Poor soil health, associated with nutrient limitation is pivotal, and a major consequences of low agricultural productivity (Lehmann & Kleber, 2015; Shepherd *et al.,* 2015). Soil health is the capacity of soil to respond to agricultural intervention, so that it continues to support both the agricultural production and provision of other ecosystem services (Kibblewhite *et al*., 2008; Vågen *et al.,* 2012). To mitigate the scourge of poor soil health, accurate and repeatable methods for determining nutrient deficiencies are a prerequisite, in order to reduce risks of fertilizer investments (Jordan-Meille *et al*., 2012; Shepherd *et al*., 2015).

It is now more apparent to policy makers and soil scientists that conventional methods used in nutrient diagnostic, may not be efficient, due to occurrence of soil heterogeneity. Several studies provide evidence of high soil heterogeneity within smallholder agroecosystem, which is among cardinal causes of low nutrient use efficiencies (Diarisso *et al.,* 2016; Tittonell *et al*., 2013, 2007; Zingore *et al*., 2008). However, few studies propose ways of dealing with soil heterogeneity, in relation to site-specific nutrient diagnostics, to inform decisions on fertilizer requirements (Tittonell *et al*., 2015). Conventional approaches lack rigours framework, which can help farmers make evidence-based decisions on nutrient management (Vågen *et al*., 2012; Shepherd *et al*., 2015). These diagnostics approaches can generally be summarized as follows.

First is visual symptoms observation (histology) which entails inspection of deficiency symptoms of nutrient that are most limiting to crop growth (Foster, 2001; Bekunda *et al*., 2010). The deficiency of individual nutrient produce characteristic effects on various organs of plants (Römheld, 2011), such as stunted growth and yellow greenish colour on leaves is normally associated with N limitation (Maynard, 1979). Ability to recognise these particular effects forms basis of histology, which is readily applied by smallholder farmers (Mairura *et al.*, 2007; Römheld, 2011). Farmers also use indigenous knowledge such as soil colour and presence of a particular weed within their farm to diagnose a limiting nutrient (Payton *et al.*, 2003; Mairura *et al.*, 2007). However, the symptoms observed, which can be attributed to nutrient limitations, could also be misinterpreted with other plant stress such as pest and diseases). This requires histology to be complemented with other methods such as soil testing, to ascertain actual cause of the observed symptom.

Secondly is the nutrient omission trials, where a single/few trials are established in a specified geographical location, to evaluate crop responses to fertilizer applications (Huising *et al*., 2011; Zingore *et al.*, 2014; Kihara *et al.*, 2015). The diagnostic results obtained are limited to that specific locality, and when extrapolated to other regions, it may lead to incorrect diagnostic conclusions (Brouwer *et al.*, 1993; Sileshi *et al.*, 2008). The assumption that few trials would represent soil heterogeneity, at regional or landscape level is rarely realistic. This limits the applicability of the omission trials, to accurately diagnose spatial pattern of limiting nutrients at regional level, within agroecosystem (Tittonell *et al.*, 2013; Diarisso *et al.*, 2016).

Lastly is the soil testing which may provide information on the limiting nutrients. Often, low soil test values signify a positive crop response to fertilizer application (Van Biljon *et al*., 2008; Petersen *et al.*, 2012). Soil test value need to be calibrated to crop response before they can be interpreted accurately (Havlin & Jacobsen, 1994), but lack of crop response data to calibrate soil tests is also major setback in many developing countries (Shepherd & Walsh, 2007; Webb *et al.*, 2011). Soil testing must be in tandem with plant tissue testing, which is a powerful tool for diagnosing micronutrient deficiencies that may prevent responses to macronutrients (Roth *et al.*, 1989; Njoroge *et al.*, 2018). High costs of wet chemical analysis, curtails the applicability of soil testing for large area assessments of limiting nutrients (Bekunda *et al*., 2010; Shepherd and Walsh, 2002).

Given the limitations of the conventional methods, alternative approaches for nutrient diagnosis for smallholder agroecosystems are necessary for the rationalization of fertilizer investments decisions. Population-Based Farm Survey (PBFS) for evaluation of disease prevalence has become popular in epidemiology, because it is a rapid and reliable way to assess a patient's health condition within a given population (Frederiksen *et al.,* 2019; Lipscombe *et al.,* 2018). This approach is particularly useful when monitoring disease patterns within human populations and designing targeted curative medical interventions (Kuper *et al.,* 2019; Rivera-Andrade *et al*., 2019). There is potential for developing a nutrient diagnostic approach, a PBFS, which can be anchored on the principles of Land Health Surveillance (LHS) that are borrowed from PBS (Shepherd *et al*., 2015). The LHS deploys a rigorous ground sampling scheme and uses proximal techniques such as infrared spectroscopy for rapid nutrient diagnosis (Shepherd *et al.,* 2015). Information on soil and plant nutrient relationships is collected, and statistical models employed to provide population-based estimates of means, DRIS indices and confidence intervals on nutrient limitations (Shepherd *et al.,* 2015; Vågen et al., 2012). The developed DRIS indices can be used to diagnose and rank limiting nutrients (Nziguheba *et al.,* 2009). This can be done in tandem with digital property mapping for evaluation of spatial patterns and prevalence of limiting nutrients. Consequently, spatial variability patterns of nutritional constraints are identified at landscape scale to guide nutrient management decisions within smallholder agroecosystems. The proposed approach, has never been tested in nutrient management for smallholder agroecosystems, more so in Kenya. Globally, NPK are considered as the major nutrients limiting plant growth (Ågren *et al*., 2012; Bekunda *et al.,* 2010). A hypothesis that the spatial pattern occurrence of N, P, K nutrient limitation is random within smallholder agroecosystems was tested. Previous studies conducted in western Kenya, have characterized the region with poor soil health (Bationo *et al.,* 2012; Bekunda *et al*., 2010; Okalebo *et al.,* 2006).

The study aimed at developing and testing the PBFS approach for nutrient diagnosis in the smallholder agroecosystems of western Kenya. The specific objective was to develop local soil test values for N, P, K, based on population distribution of smallholder maize fields. The region was deemed as a suitable testing site, because it is typified with heterogenous parcels of smallholder maize fields. The study evaluated NPK nutrients limitations using farm surveys data, and mapped their spatial distribution.

5.2. Material and methods

5.2.1 Study area

The study area encompassed the administrative sub-counties of Boro, Butere, Yala, Khwisero and Ugunja and has been described in section 4.2.1 (Figure 4.1).

5.2.2 Overview of the Population-Based Farm Survey approach

Population-Based Farm Survey (PBFS) approach involves conducting a survey on a number of maize fields. The term *"population-based"* is used to signify a population of smallholder maize field, within a smallholder agroecosystem. A target population was defined as a representative sample population, drawn from a population of maize fields, with defined characteristics to be being evaluated. The target population formed basis for making inferences about nutrient diagnostic for the whole population. The sequence of steps employed in the population-based farm approach are summarized in Figure 5.1, which provide guidelines for nutrient diagnostics, ranking and mapping.

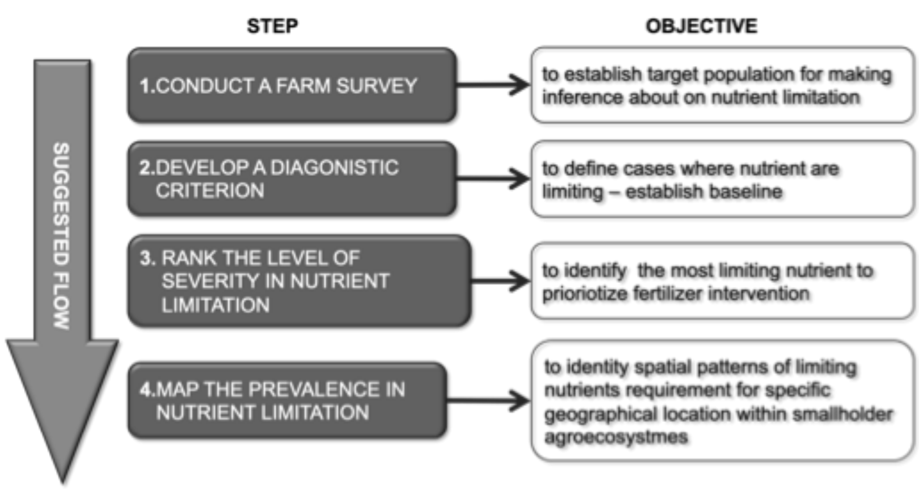

Figure 5.1: Sequence of step used for establishing a population-based farm survey approach for nutrient diagnostic.

5.2.3 Farm survey

Farm survey was conducted within a hierarchical sampling scheme, using Land Degradation Sampling Framework (LDSF) (Vågen *et al.*, 2010) similar to that in section 4.2.3. Purposive sampling strategy was also included, where a field considered to be either poor or good in terms of management, and fell within the proximity of a selected coordinate, were sampled (Plate 5.1). This ensured the target population was sufficiently characterized. Good fields belonged to resource endowed farmers, while poor fields were associated with those with resources (Zingore *et al*., 2007).

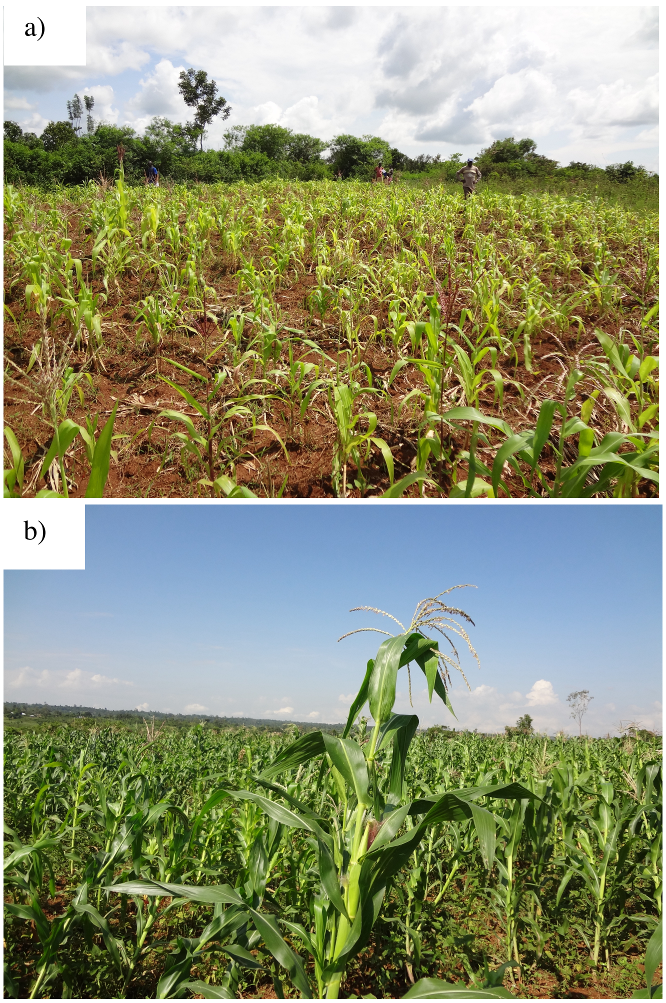

Plate 5.1 a) Poorly managed maize field, and **b)** good managed maize field (photo courtesy of Stephen Ichami during field measurement on 18th April 2018.

Soil sampling was achieved as described in section 4.2.2. In addition, plant sampling was done by extracting three maize ear leaves, for nutrient analysis at silking stage (60 to 75 days after emergence). Only three maize plants were randomly selected. The assumption was that at silking stage, maize tissue (ear leaf) would have the optimum concentration of nutrients (NPK). Maize biovolume and grain yield were measured, and represented crop responses of the study area. Maize biovolume (BV) was estimated as described in section 4.2.2.

Soil and plant samples were analysed at Crop Nutrition Laboratory Services and World Agroforestry Centre laboratories in Nairobi, for N, P, K as describe earlier in section 4.2.3. Preliminary preparation of soil samples involved air-drying and grounding, to pass through a 2 mm sieve to minimize variation due to moisture. Maize leaf samples were dried at 60° C, prior to grounding (< 1 mm). Conventional wet-chemical methods were used to obtain reference data with 25% of the collected sample, and used to develop calibration models for prediction of N, P, K using Infrared IR. In section 4.2.3, a description of the methods used for soil characterization is given.

All collected samples (256) were characterized for N, P, K nutrient concentrations using Infrared (IR) spectroscopy technique. Fourier-transform MIR spectrometer (FT-IR; Tensor 27, Bruker Optics, Karlsruhe, Germany) with a high throughput screening extension arm using a liquid Nitrogen cooled HgCdTe detector, was used to determine diffuse reflectance in the mid-infrared region (4000 – 600 cm^{-1}). Prior to IR determination, soil samples were fine ground using a sample mill. The fine samples (approximately 0.05 grams) were loaded and levelled into wells in aluminium micro-plates (A752-96, Bruker Optics, Karlsruhe) using a micro spatula, in four replicates (per sample), to enable spectral measurement. An empty well was used for reference readings, taken before each sample reading using an average of 32 scans. Absorbance was recorded at a spectral resolution of 4 cm^{-1} zero-filled to 2 cm^{-1} and single spectrum obtained for each sample. First derivative spectra with a smoothing gap of 3 points were used in all the spectral analysis.

Plant samples were analysed for total NPK tissue nutrient concentration following Okalebo *et al*. (2002). Total N was determined by sulphuric digestion followed by micro-Kjedhal distillation, while P was calorimetrically determined using vanadium molybdate after digestion with sulphuric acid (Zarcinas *et al.*, 1987; Okalebo *et al.*, 2002). Potassium determination was done using a flame emission spectroscopy, after digestion with sulphuric acid (Zarcinas *et al.*, 1987).

The NPK nutrient concentrations were predicted from spectral measurements for all soil and plant samples (Shepherd and Walsh, 2002; Sila *et al*., 2017). Reference data values were calibrated to the smoothed first derivative spectra using partial least square regression (PLSR), implemented in the "*soil. spec*" R package (Sila *et al.*, 2014). The reliability and robustness of the calibration models was evaluated by the hold-out cross-validation procedure, using the coefficient of determination (r^2) and the root mean square error of cross validation (*RMSECV*) calculated based on equations 5.1 and 5.2.

$$r^2 = \frac{SSR}{TSS} \quad \ldots\ldots\ldots\ldots.. \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.\ 5.1$$

$$RMSECV = \sqrt{\sum_{i=1}^{N_p} \frac{(\hat{y}cvi - y_i)}{N_p}} \quad \ldots\ldots\ldots\ldots\ldots\ldots.5.2$$

where *SSR* is the sum square of regression, and *TSS* is the total sum of squares, $\hat{y}_{cvi}$, and y_i are the predicted and measured reference values respectively and N_p is the number of samples tested. Models with highest r^2 and lowest *RMSECV* are considered to be the best and robust.

The farm survey was stratified into unfertilized maize fields, to minimize variation due to fertilizer management, and captured historical effects of management. 64 fields were sampled, which include a total of 258 plots that formed target population of the PBFS. This captured soil and maize information at different spatial scales (plot, field and landscape). The survey ended up with a target population database, containing geographical coordinates, spectral measurements on soil and maize nutrient concentrations, grain yield and plant biovolume.

Preliminary data preparation and analysis involved calculation of descriptive statistics for soil and maize nutrient concentrations. The mean, maximum, minimum, standard deviation, coefficient of variation (CV) and confidence intervals (CI) were computed for NPK nutrients, grain yield and plant biomass. Density plots, for testing normality in data distribution were developed. Skewed variables were normalized. The normalized data was required for correlation analysis and geostatistical modelling. Pearson correlation coefficient (*r*) formed basis for evaluation of the strength of this relation, with values > 0.50 considered as strong relations.

5.2.4 Developing nutrient diagnosis criterion

Diagnostic criterion was developed using established critical nutrient values for NPK maize tissue concentration (Campbell, 2000; Reuter and Robinson, 1997), (Table 5.1). Normally, soil values that occur below the critical levels were diagnosed as deficient (Reuter & Robinson, 1997; Campbell, 2000). Values were used to characterize our target population as "*deficient*" and "*sufficient*" of which the latter would imply serious nutrient limitations.

Table 5.1: Published critical nutrient concentration values for maize crop at 90% relative grain yield

Nutrient	Critical nutrient concentration (%) at deficiency level	Critical nutrient concentration (%) range sufficiency level	Source
Nitrogen	3.00	4.00 – 6.40	Campbell (2000)
Phosphorus	0.25	0.42 – 0.69	Campbell (2000)
Potassium	2.00	3.50 - 5.00	Campbell (2000)

Relationships between maize tissue nutrient concentration and crop response (grain yield and biovolume) were evaluated following Cate and Nelson (1971) method. The analysis enabled the calibration of nutrient tissues concentration to maize yield, and determined levels where addition of nutrients was likely to increase maize yield. Boundary fit analysis using "*drc*" R package were used to fit the response curve, and evaluate variation with respect to established critical nutrient values for each nutrient. The response curves were established using corresponding median in every quartile, the minimum and maximum values, for the NPK nutrient concentration as a function of maize yield (Kihara and Njoroge 2013).

To develop a soil test value (cut-off) for defining cases where of nutrient limitation, a frequency distribution plot (normal distribution) of "*deficient*" and "*sufficient*" sub-population were plotted as function of soil tests values for NPK nutrient concentrations. Soil cut-off value was identified based on overlaps of the normal distribution curves of the 2 sub-populations, at the upper and the lower 90% CI of "*deficient*" and "*sufficient*" sub-population, respectively.

5.2.2.3 Ranking level of nutrient limitation severity

The DRIS reference indices were used to measure the nutrient balance within a whole plant and ranked the order of limiting nutrients (Walworth *et al.*, 1986; Walworth & Sumner, 1987; Nziguheba *et al.*, 2009). The DRIS approach utilized indices and norms derived from maize tissues nutrient concentrations and corresponding yields representing variability encountered in maize fields. DRIS norms were established using the criterion of significant variance ratio between "*deficient*" and "*sufficient*" subpopulations (Walworth & Sumner, 1987). Means and variances of maize tissue nutrient concentration were calculated for two sub-populations for each nutrient. The N/P or N/K or P/K were computed only for the sufficient sub-populations and then divided by the number of observations of each expression: Equation 5.3, 5.4 and 5.5.

$$norm\ for\ nitrogen/phosphorus = \frac{N/P}{n} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots 5.3$$

$$norm\ for\ nitrogen/potassium = \frac{N/K}{n} \text{......................... } 5.4$$

$$norm\ for\ phosphorus/potassium = \frac{P/K}{n} \quad \text{...................... } 5.5$$

$3c$

where NPK refers to total nitrogen, phosphorus and potassium nutrient tissue concentration (%), respectively, and n is the number of sufficient populations.

DRIS indices are quantitative evaluations of the relative degree of imbalance among the nutrients under study and were as: Equation 5.6, 5.7 and 5.8

$$N\ Index = \left[\frac{f(N/P)+f(N/K)}{2}\right] \quad \text{...................................... } 5.6$$

$$P\ Index = \left[\frac{f(N/P)+f(P/K)}{2}\right] \quad \text{...................................... } 5.7$$

$$K\ Index = \left[\frac{f(P/K)+f(N/K)}{2}\right] \quad \text{...................................... } 5.8$$

where:

$$f(N/P) = \left[\frac{N/P}{n/p} - 1\right]\frac{1000}{CV} \quad \text{.............................. } 5.9$$

when the actual value of N/P > n/p or

$$f(N/P) = \left[1 - \frac{N/P}{n/p}\right]\frac{1000}{CV} \quad \text{.............................. } 5.10$$

when the actual value of N/P < n/p

n/p is the mean (norm) value for N/P, and CV is coefficient of variation for high-yielding populations. The other terms of *f*(N/K) and *f*(P/K) are derived in a similar way using the means, n/k for N/K and p/k for P/K, respectively in place of n/p.

Interpretation of DRIS indices was based on the positive and negative values, which sum to zero, and the more negative the value is, the more the nutrient is limiting.

5.2.5 Mapping prevalence of limiting nutrients

Geostatistical analysis was conducted to evaluate spatial variability of DRIS indices across the study area. First, the semi-variogram models for DRIS indices were developed using the "*gstat"* package (Pebesma, 2004), and the model parameters: range, nugget and sill, were used to evaluate spatial structure of the DRIS indices.

Interpretation of these parameters was done following Cambardella *et al*. (1994), with values < 25% were considered as weak, 26 - 75% moderate, and > 75% as strong spatial structure. Low nugget: sill ratio was indicative of short distance variability in the DRIS indices.

Secondly, the spatial pattern within the study area were analysed by developing maps of the DRIS indices using "*ggplot2*" R package (Wickham, 2009). The spatial pattern described levels of nutrient distribution within a geographic location, and were performed to identify the spatial trend of DRIS indices. This enabled identification of specific geographies where actual NPK nutrient limitation occurred.

Finally, the Moran's (I) Index was computed and used to identify spatial patterns in nutrient limitations based on DRIS indices, using the "*gstat*" package (Pebesma, 2004). The index depicted levels of spatial clustering, with positive values taken to indicate clustering of nutrient limitation, while negative values were indicated spatial dispersion.

5.3. Results

5.3.1. Farm survey

Table 5.2 shows calibration models within the mid infrared (MIR) region (400 – 4000 cm^{-1}) for predictions of soil and plant nutrient concentration for the study area. Calibration models for NPK gave good fits with cross-validated r^2 values of 0.94, 0.69 and 0.74, respectively. Total maize tissue N had the lowest *RMSECV* of 0.08 and r^2 value of 0.84. The predictions varied across NPK nutrients for soil and plant tissues, as exhibited by different fit of r^2 value. The results indicate prediction potential of nutrient concentration in soil and plant samples using MIR spectral signatures. The fundamental vibrations of molecules in soil and plant materials are normally found in MIR spectral region, where very distinct spectral signatures are displayed, because of strong absorption of overtones by hydroxyl ions (Brown *et al.,* 2006). Hence it was possible to quantify nutrient concentrations in soil and maize leaves tissues (concentrations) using IR spectral measurements.

Table 5.1: Mid-infrared calibration model statistics that predicted soil and plant nutrient concentrations of the study area.

Soil samples		
Nutrient concentration	r^2	**RMSECV**
Total N (%)	0.94	0.09
Extractable Potassium (cmol kg^{-1})	0.74	0.44
Extractable Phosphorus (mg kg^{-1})	0.69	0.40
Maize ear-leaf samples		
Nitrogen (%)	0.84	0.08
Phosphorus (%)	0.84	0.16
Potassium (%)	0.80	0.12

Key: r^2 = Coefficient of determination, RMSECV = Root Mean Square Error of Cross Validation

Total soil N varied from 0.06% to 0.36%, while extractable P and K had a median of 17.2 mg kg^{-1} and 4.6 cmol$_c$ kg^{-1}, respectively (Table 5.3). Low concentrations of nutrients characterized the soils of the study area, as exhibited by mean values of total soil N, extractable P and K, which were below established critical soil values of 0.2%, 10 mg kg^{-1} and 3 cmol$_c$ kg^{-1}, respectively (Okalebo *et al.*, 2002). According to these results, NPK were the principle limiting nutrients for maize production in the study area.

The study area was characterized by low variation in maize tissue nutrient concentration, but high variation in soil nutrient and crop responses (Table 5.3). High variation occurred in extractable P with a coefficient of variation (CV) value of 61% compared to a corresponding value of 25 % in maize tissue (Table 5.3). Total soil N exhibited the lowest variation (CV = 25%). Low variability for total P and K tissue nutrient concentration was also evident, with CV values of 25% for both. Low variability in maize tissue nutrient concentration can be attributed to homogeneity in the maize tissue samples. Grain yield and maize biovolume displayed high variability with CV values of 55% and 40%, respectively. These results provide evidence of variability in NPK nutrient limitations across the study area. Site-specific diagnosis of NPK nutrients may therefore be pivotal for amelioration as a nutrient management strategy. The CV statistics indicate that farm survey captured variability in crop response and inherent soil fertility, and did not describe the spatial variability (Haileslassie *et al.*, 2005).

Density plots for soil total N, grain yield and BV displayed a near normal distribution. Extractable P and K were negatively skewed, indicating presence of low values. These variables were transformed to natural log (*ln*) values, in order to attained approximate normal distribution, as required in the subsequent steps of the population-based farm survey approach.

5.3.2. Diagnosis of limiting nutrients

Observed maize yield response as a function of NPK maize tissue nutrient concentration is shown in Figure 5.2. The general trend was an increase in maize tissue nutrient concentration for total N that corresponded to a 2-fold increase in grain yield (Figure 5.2 a, d). This implied that improved uptake of N would double maize yields for the study area. Total N had a significant relationship (p value = 0.001, r = 0.0089) with maize yield, which tended to be positive, as would be expected (Figure 5.2 a). Relationship between grain yield and total P tissue nutrient contrasts results of total N, which were poor, as exhibited by low r value of 0.03 (Figure 5.2 b). Similar trend was observed for total K with r value of 0.009. Though the r value was low (0.089) for total N, this logarithmic relationship was the strongest amongst all the grain yield and plant tissue relationships.

Table 5.2: Soil properties, maize ear leaf total tissue nutrient concentration and crop response variables of unfertilized maize plots across smallholder agroecosystem in western Kenya.

Nutrient	Observations (n)	Minimum	Maximum	Median	Mean	Standard deviation	Coefficient of variation	Confidence Interval mean (95%)
Soil properties								
Total nitrogen (%)	256	0.06	0.36	0.14	0.15	0.04	25%	0.00
Extractable P (mg kg-^{1})	256	8.19	107.21	17.22	21.02	12.82	61%	1.58
Extractable K (cmol kg^{-1})	256	0.12	6.48	4.61	4.24	1.49	35%	0.18
Maize tissue (ear leaf samples)								
Nitrogen (%)	220	0.12	0.40	0.22	0.23	0.06	25%	0.01
Phosphorus (%)	220	0.57	2.76	1.84	1.80	0.45	25%	0.06
Potassium (%)	220	0.00	3.81	1.86	1.68	0.70	42%	0.11
Crop response (based on inherent fertility)								
Grain yield (Mg ha^{-1})	256	0.08	11.28	3.20	3.54	1.90	54%	0.23
Plant biomass (cm^{3})	256	31.00	392.91	161.00	170.38	73.33	43%	9.10

Crop response curves indicated maize yield increased with increasing tissue nutrient concentrations (Figure 5.2 d, e, f). Based on Cate and Nelson, (1971) recommendations, the crop response curve only fell in the deficiency zone (level), attributed to fact that, only unfertilized maize fields were sampled. However, there was a significant increase in maize yield beyond the established critical values. For example, 40% increase in yield was realised, when N tissue concentration was > 3.0 %. This result implies tissue nutrient assimilation by the maize crop, may lead to increased yields. Therefore, application N, P, K fertilizer would result into increased maize yield for the study area.

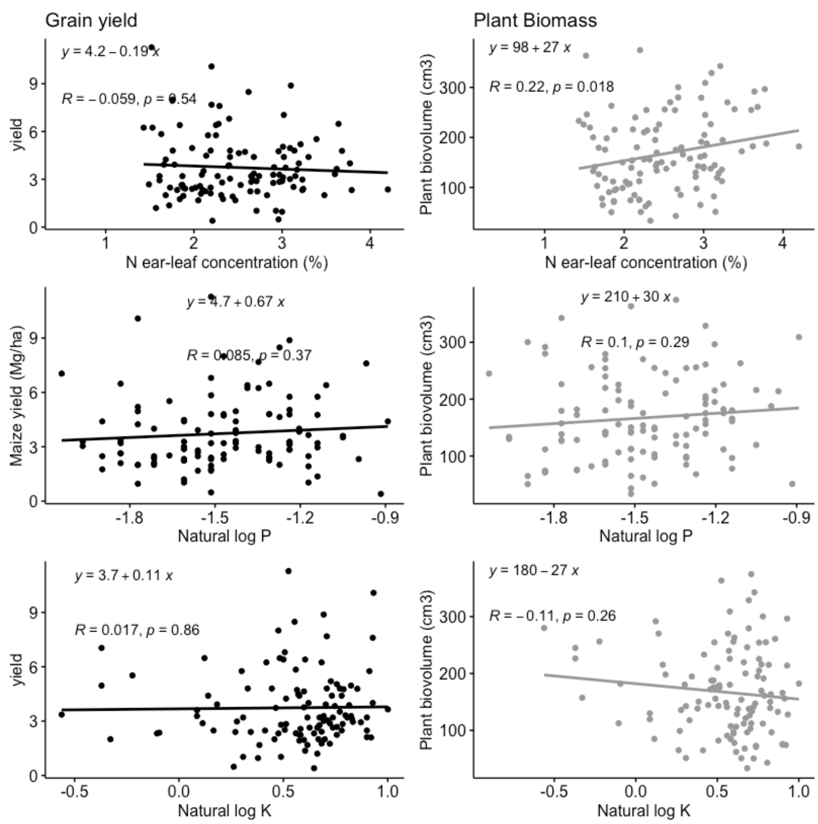

Figure 5.2: Scatter plot between grain yield (Mg ha^{-1}) and plant biovolume as function of maize ear-leaf tissue total nutrient concentration for; a, b) nitrogen (%), c, d) phosphorus (%), and e, f) potassium (%).

Significant relationship between grain yield and maize tissue concentration for total N,P,K were indicative that they could be inferred, and used to measure nutrient limitations. Therefore, established critical nutrient value for total N formed basis for dividing target population into a "*deficient*" and "*sufficient*" sub-population. The '*deficient'* sub-population constituted 46%, a representation of poor maize field in the study area. The two sub-population formed basis for establishing NPK soil cut-off values for the study area.

Frequency distribution plots of transformed soil tests values for total soil N, Extractable P and K are presented in Figure 5.3. Back transformed soil cut-off values were 0.074% for total soil N, 12.5 mg kg^{-1} for Extractable P and 4.5 $cmol_c$ kg^{-1} for Extractable K. There were no significant differences in the magnitude between developed farm survey soil cut-off values, compared to published critical soil test values for total soil N (0.2%), Extractable P (10 mg kg^{-1}) and Extractable K (3.0 $cmol_c$ kg^{-1}), as reported in literature Okalebo *et al*., (2002). For total N and Extractable K, there was no clear distinction of the '*deficient*' and '*sufficient*' sub-populations as would be expected (Figure 5.3 d f).

Developed soil cut off values were used to determine the prevalence of soil nutrient limitation for the study area. Only 67% of target population characterized showed deficiency in total soil N, 54% in phosphorus, while 37% had potassium deficiency. These soil test values were below developed soil cut-off values in target population from the farm survey dataset. The nitrogen status of the study area was limiting in most maize fields than phosphorus and potassium. About 15 % of the maize fields had N levels in the range of 0.1 - 0.25%. Therefore, it was clear that N was a major limiting soil factor for maize growth in most sampled fields.

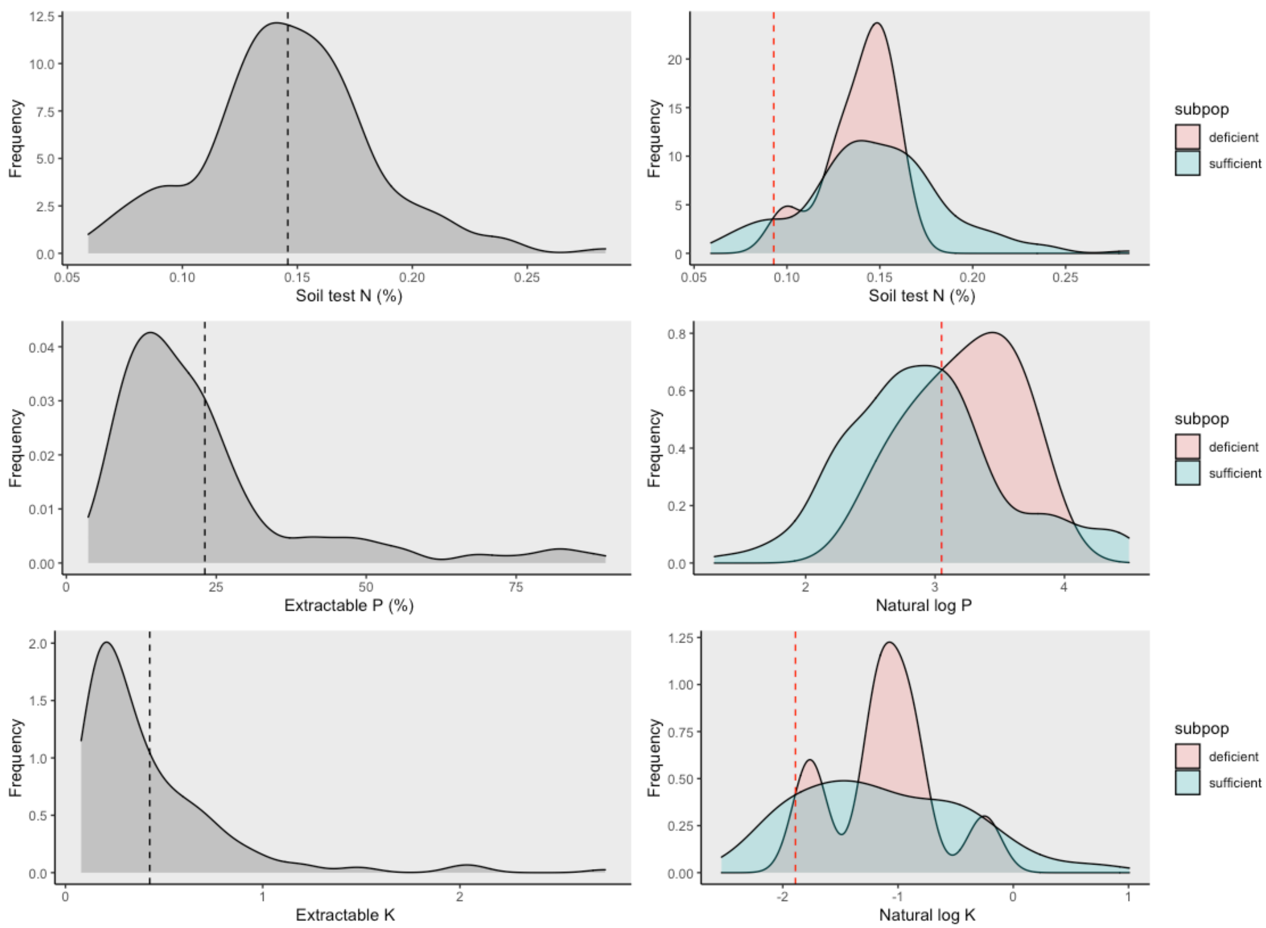

Figure 5.3: Relationship between grain yield (Mg ha^{-1}) as function of maize tissue total nutrient concentration for; a) nitrogen (%), b) phosphorus (%), and c) potassium (%) for establishing critical nutrient concentration. Maize response curves (d, e, f) for NPK nutrients. The dotted black line represents the critical nutrient concentration values established from literature.

5.3.3 Ranking limiting nutrients

The DRIS indices varied widely, from -35.86 to 36.87 for N, and were within ranges of the published international norms reported by Elwali *et al.* (1985). While applying the local DRIS norms derived from population-based farm survey data, the mean DRIS indices were -6.3 for N, -13.5 for K and -2.1 for K (Figure 5.4). Normally, the DRIS indices are not affected by differences in growth stages of maize crop (Walworth *et al.*, 1986). Phosphorus was ranked as the most limiting nutrient, followed by N, and K in a descending order of limiting nutrient. The relative ranking of the limiting nutrient, from the most to least limiting was phosphorus > nitrogen > potassium. The results imply that fertilization of the maize crop in the study area may prioritize fertilization of phosphorus since it is the most limiting nutrient.

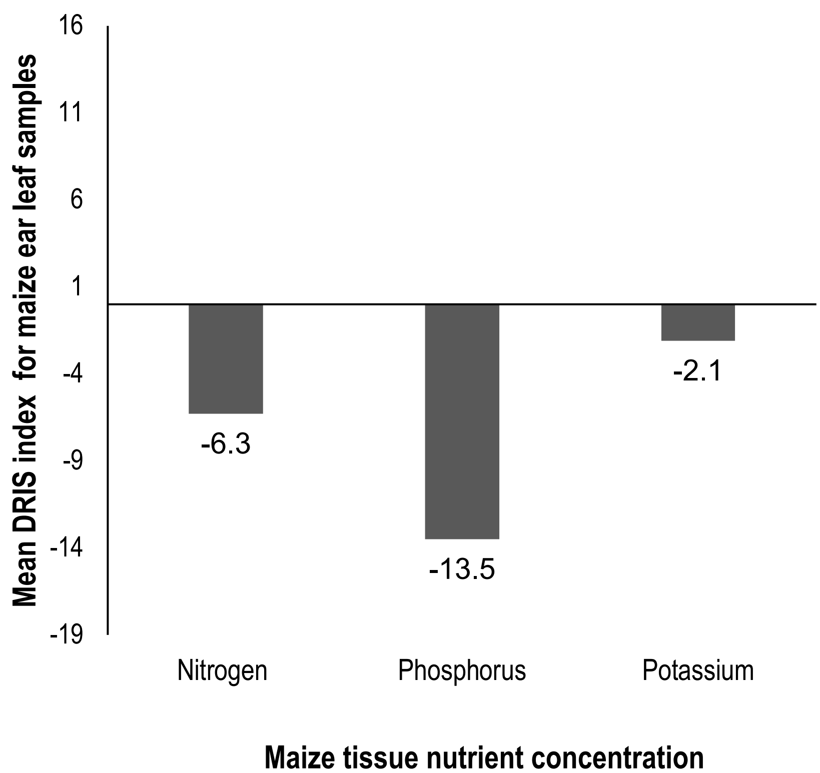

Figure 5.4: Bar graph showing the mean DRIS index for NPK maize ear leaf samples for the study area

Critical N concentration value separated the target population of 258 observations into '*deficient*' and '*sufficient*', which comprised of 116 and 141 observations, respectively. These sub-populations formed basis for computing DRIS indices. Calculated variance ratio of N/P, N/K and P/K were significant ($p < 0.01$) similar to those reported by Elwali *et al.,* (1985). The significance of variance may be considered as good evidence on the validity of the assumption used in separating the two sub-populations, which was set at critical N nutrient value of 2.9%, corresponding to grain yield of 2.4 Mg ha^{-1}. The '*deficient'* population showed high values of standard deviation and coefficient of variation as compared to '*sufficient'* population.

5.3.4 Mapping prevalence of limiting nutrients

Derived semi-variogram model parameters are presented in Table 5.4. The best semi-variogram model fitted for DRIS index for N and P was exponential and spherical, respectively. The variograms had the least root mean square error values of 0.00023 for N and 0.019 for P. The nugget/sill ratio ranged from 0.23 to 0.70, an indication of strong to moderate spatial dependencies for DRIS indices. The nugget effect of DRIS index for N suggest moderate spatial dependencies across unfertilized plots. A high nugget effect (0.0041) was observed for DRIS Index for P suggesting small-scale variation. Low nugget values were signified short distance variability for DRIS indices. Moderate spatial dependencies occurred at an effective range of 543 m for DRIS index for N, the lowest compared to DRIS index for P (3291 m). The spatial dependency of the DRIS index was strong for N and P. The existence of strong and moderate spatial dependencies meant that further evaluation of spatial patterns of nutrient deficiencies using maps of the study area was necessary.

Table 5.3: Spatial dependency of DRIS indices for maize fields in a smallholder Western Kenya in term of the semi-variogram. Strong < 25%, moderate 25 – 75 %, Weak > 75 %

	Semi-variogram model	**Nugget: Sill ratio**	**Range (m)**	**Spatial dependency**
DRIS Index N	Exponential	0.23	543	Moderate
DRIS Index P	Gaussian(normal)	0.61	302	Strong
DRIS Index K	Linear	0.34	3291	Strong

Spatial autocorrelation analysis was performed to test the significance in the pattern of geographical distribution in nutrient limitation (Table 5.5). A hypothesis that the spatial occurrence of NPK nutrient limitation is a random, within smallholder agroecosystem was rejected, since Moran Index for NPK was significantly positive ($p < 0.01$). This meant that NPK nutrients display a clustering pattern, where multiple NPK deficiencies occurred in the same geographical location, within the study area. Occurrence of multiple nutrient deficiencies can be attributed to the non-responsive characteristic of soil that is common in western Kenya.

Table 5.4:Moran Index for the for maize fields in a smallholder landscape in Western Kenya

	DRIS indices		
	Nitrogen	**Phosphorus**	**Potassium**
Moran Index	0.40	0. 23	0.42
p value	0.0003***	0.31·	0.01·

level of significance *** = 0.001, ** = 0.01, *= 0.05 and ·= 0.1

Simulated maps provided a display of the spatial distributions of NPK nutrient limitations for the 100 km^2 sampling block based on DRIS indices (Figure 5.5). There was a variation in nutrient limitation patterns across the study area. For example, there was a gradient of NPK deficiencies towards north western side of study area as indicated by the blue hue. Clustering of NPK nutrient limitation occurred on the south west part of the study area. The result meant that nutrient amelioration strategies require a holistic application of NPK, at varying rates for the maize crop.

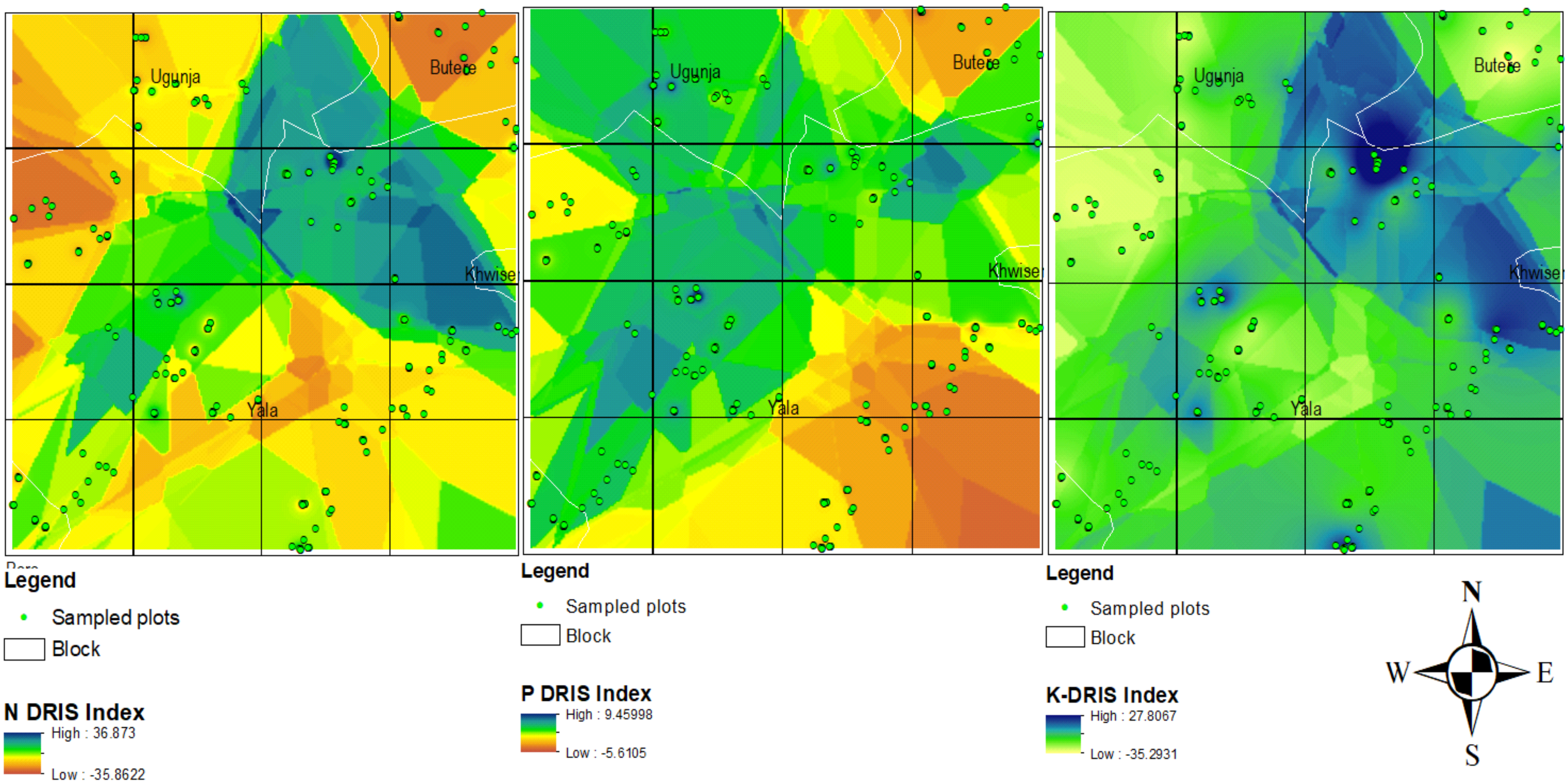

Figure 5.5: Map showing the spatial patterns of DRIS indices for the study area for: a) nitrogen (%), b) phosphorus (%), and c) potassium (%).

5.4 Discussion

This study has demonstrated development of NPK soil test values form farm survey, rather than using the crop response trials. The novelty of the population-based farm survey approach comes from the combination of technologies that generate synergy between them. For example, critical nutrient concentration concept established by Cate and Nelson (1979), DRIS index approach (Walworth *et al.*, 1986) and geostatistical tools (Pebesma, 2004). The study incorporated these tools in the population-based farm survey approach, following the principle of land health surveillance, to estimate nutrient limitation at specific geographical niches (Shepherd *et al.*, 2015).

5.4.1. Farm survey

Farm survey enabled the development of a geospatial spectral database, to map and diagnose N, P, K nutrient limitations for the study area. The calibration model showed prediction potential of MIR, which varied with the soil and maize tissue nutrient concentration for NPK (Table 5.2). The best prediction observed for total soil N, with r^2 of 0.88, was similar to that reported by Mouazen *et al*. (2010), followed by extractable P with r^2 = 0.74, an excellent value according to Conzen (2003). The findings were comparable to previous research findings from air-dried soil samples by Malley *et al.* (2004), but better compared to results reported by Janik *et al*. (2008) with r^2 value of 0.07 for extractable P. The coefficients of determination ranged from 0.60 to 0.90, and their *RMSECV* were satisfactory in diagnosis of nutrient deficiencies according to Cozzolino and Moron (2006) and Stevens *et al.* (2008). The results demonstrate potential of using MIR for assessment of nutrient deficiencies for large-scale evaluation, as a rapid and low-cost analytical tool, that can be embedded in the proposed approach (Towett *et al*., 2015).

Crop response, soil nutrient concentration and maize tissue samples displayed high degree of variability as shown by the CV values for the study area (Table 5.3). The observed variability is attributed to difference in the inherent soil fertility, which could be influenced by topography (Arnhold *et al.*, 2015), management aspects (Zingore *et al.*, 2007) and soil types (Raynaud & Leadley, 2004). However, the findings conform to those reported by Tittonell *et al.* (2008) who found high variability in maize fields in western Kenya. Variability in crop response demonstrates the need of spatially explicit nutrient diagnostics, which may be implemented as the first step, towards site-specific nutrient management strategy. The poor relations can be attributed to low P concentration in soil of the study area. These relationships are explained by a poor capacity of plant tissue testing to predict magnitude of yield response within and below the critical nutrient concentration range.

5.4.2. Diagnosis of limiting nutrients

Soil and plant relation were used to establish critical nutrient concentration and soil cut-off values useful in defining cases where nutrients were limiting (Figure 5.3). The developed soil cut-off values fell within the range reported in previous studies (Adeoye & Agboola, 1985; Okalebo *et al.*, 2002). Adeoye and Agboola (1985) reported values ranging from 10 to 16 mg kg^{-1} for P and 0.6 to 0.8 $cmol_c$ kg^{-1} for K in smallholder farms of Nigeria, which are comparable with finding by Okalebo *et al*. (2002) in western Kenya. Based on developed soil cut-off values, diagnosis of N, P, K limitation were established in 67, 54 and 37% of sampled maize fields studied, respectively and agree with those of Kihara *et al*. (2015) also in western Kenya.

The frequency plot of the deficient and sufficient sub-population as function of soil test values displayed a near normal distribution for total soil N and extractable K (Figure 5.3 a, c). The observed distribution can be explained by high soil heterogeneity (Ronner *et al.*, 2016), soil disturbances through ploughing (Wopereis *et al.*, 2006) and low concentration of phosphorus and potassium values for the study area (Webb *et al.*, 2011; Ngome *et al.*, 2013). These soil test values provided a criterion for defining cases of N, P, K limitations (Figure 5.3). However, soil test values present disadvantages for nutrient diagnostics, when used in isolation (Das *et al.*, 2009; Valkama *et al.*, 2011), hence the need of incorporating critical maize tissue nutrient concentration. The relationships between measured soil test values and nutrient availability was very tenuous (r = 0.02 for P), similar to observation made in previous studies (Roth *et al.*, 1989; Havlin & Jacobsen, 1994). This is attributed to the maize growth and its nutrient uptake rate in the field, which depend on many environmental factors such as soil nutrients concentration and their interactions, which varied (Table 5.3) across the maize field (Römheld, 2011). The growing plant integrates all these soil factors and is the best measure of true nutrient availability of the unfertilized maize fields (de Barros *et al.*, 2007). This study therefore demonstrated the use of critical nutrient concentration values as a criterion of establishing sub-population, which guide in the development of localized soil cut off values and is the novelty of this current approach.

Synergy of the current approach was first evident by evaluating nutrient concentration in both maize tissues and soil samples (Figure 5.3) where established critical nutrient concentration test values formed the basis for defining limiting nutrient. A strong correlation between maize tissue N concentration and grain yield were observed (Figure 5.2 a), which could be explained by high nitrogen uptake at the silking stage as expected (Setiyono *et al.*, 2010). Observations of strong and significant ($p < 0.01$) correlations between N concentration in maize leaves and grain yield are in line with those of Bak *et al.* (2015). Poor correlation between grain yield and P concentration (Figure 5.2 c) observed in the study could be attributed to low levels of P in the soil, below the critical soil nutrient concentration of 10 mg kg^{-1} (Okalebo *et al.,* 2002). Poor relation between extractable K and grain yield could be attributed to the poor prediction of the magnitude of grain yield (Figure 5.2 c). The results differ with those reported by Clover *et al.* (2007) who found good relations between potassium and grain yield, but in fertilized maize fields. The utility of critical nutrient concentration in maize tissue provides a synergy, since these values were used to establish the deficient and sufficient sub-populations, on whose basis soil test values for diagnosis of limiting nutrients were developed. This was important in the implementation of the farm survey approach.

5.4.3 Ranking limiting nutrients

The DRIS indices showed varying values for NPK nutrient limitation in the study, but were within ranges reported by Ewali *et al.* (1985) and Nziguheba *et al*. (2009) for maize crop (Figure 5.4). The observed DRIS index indicate K (- 2.1) as less deficient, compared to P (-6.3) and N (-13.5) for this diagnosis. Potassium was thus ranked as the least important nutrient limiting maize yields in the area and hence recommend application of N-based fertilizers with lower portions of P and K. This can be attributed to the occurrence of both responsive and non-responsive soils attributed to P fixation in the study area (Sanchez *et al.*, 2003a; Ichami *et al.*, 2018). However, the frequency distribution plot for extractable P performed better (Figure 5.3 d). The result affirms that indeed farm surveys can be an alternative option for conventional methods for developing critical soil test values for nutrient diagnostics. The result also suggests soil

cut-off values obtained from farm survey may be a representation of variability in maize field within agroecosystems, since the values are developed from a population of fields (258 plots within 64 maize fields). This also implied there is potential of capturing spatial variation in nutrient concentration for the study area through farm surveys.
The finding relate well to those by Kihara *et al*. (2015) who found NPK limiting maize production in western Kenya. However there is a difference in the sampling methodology, where 258 plots were analysed for the farm survey in this study, compared to 32 plots for conventional nutrient diagnostic trials by Kihara *et al*.(2015). In both studies, similar conclusions are made, *i.e.,* N, P, K are limiting. The farm survey may therefore be a better approach since it explicitly captured the within spatial variability in maize fields.

The DRIS indices demonstrate the incorporation of plant nutrient concentration in nutrient diagnostics, and bring the second synergy into the population-based farm survey approach. The DRIS technique was found to be advantageous for the farm approach, because it reflected the nutrient status of the whole maize plant (Ramakrishna *et al*., 2009). The DRIS indices can therefore be employed in identifying nutrient limitation and preference of mitigation criteria can be aligned to the most deficient nutrient (P) within in the study area.

5.4.4 Mapping prevalence of limiting nutrients

Diagnosis of limiting nutrients and ranking was established in the aforementioned sections (5.2 to 5.3). The results however do not explicitly explain, the spatial distribution of nutrient limitation of the study area. Thus, a geostatistical technique was employed to evaluate geographical distribution, which brings in the last synergy of the population-based farm survey approach.

Semi-variogram models' parameters provided good estimates of spatial structure of DRIS indices (Table 5.4). Different theoretical semi-variogram models were selected for the significant fit of DRIS indices for NPK (Pebesma, 2004). Exponential model provided the best fit to the semi-variogram of N, while Gaussian and Linear models were best fit for P and K. Several findings suggest that exponential model is the most suitable for assessing spatial variability in soil nutrients (Cobo *et al*., 2009, 2010; Snoeck *et al*., 2010), it explains the maximum variability in the spatial dataset

(Goovaerts, 2000). The DRIS indices for N displayed short distance variability as exhibited by the range (543 m) and low nugget values (0.00314), which implied that high resolution maps of DRIS indices may be appropriate for nutrient diagnostics. Few studies have reported the spatial structure of DRIS indies, however similar observation of short distance variability for total soil N have been reported by Okeyo *et al*. (2009) in smallholder farms of western Kenya.

A hypothesis that the spatial occurrence of N, P, K nutrient limitation is a random pattern within landscape was rejected, since the Moran Index (MI) for N, P, K were positive (Table 5.5). Test of significance for MI values returned by geostatistical analysis of N showed displayed significantly similar clustered distribution ($p < 0.001$, MI = 0.40), with low levels N observed in one location. The clustering pattern can be explained by differences in soil characteristic patterns, which are complex due to the topography of the area (Cobo *et al.,* 2010; Tesfahunegn *et al.,* 2011). Spatial pattern of P and K did not appear significantly *($p < 0.001$)* different from a random distribution for this region. The result for N clustering conforms to those by Panday *et al*. (2018) who found clustering for N in smallholder farms of Nepal. Clustering may be taken as an indication of occurrence of N, P, K limitations in one location, which require holistic approach for nutrient management for different geographies niches. The analysis of the spatial structure of DRIS indices provided synergy through evaluation of geographical pattern of nutrient limitations. In this way, nutrient management strategies could be implemented using the spatial distribution DRIS maps as a guide for identifying occurrence of nutrient limitation in specific geographical niches within smallholder landscapes.

The simulated maps depicted status of N, P, K contents in different geographical niches across the study site most of which displayed their deficiencies, as indicated by negative DRIS indices (Figure 5.5). The deficiency could be explained by different historical management practices (no fertilizers were applied) that have influenced inherent soil properties (Vanlauwe *et al*., 2010; Ngome *et al*., 2013; Wang *et al*., 2014). Sanchez *et al.,* (2003) attributed the deficiency of P to fixation by aluminium oxides. This finding implied that by making reference to the DRIS indices simulated maps, NPK fertility status could be assessed before recommending site-specific fertilizer inputs.

Management strategies to enhance soil nutrients status could be implemented in the site using these maps as a guide (Hengl *et al.*, 2015). Normally, low nutrient values require relatively higher amount of fertilizer application; therefore, these maps may lead to better understanding of existing nutrient limitation, allowing easier management and maintenance of sustainable maize productivity. This research thus sets a precedent for upscaling future digital property mapping of nutrient limitation in other parts of the country and implementation of site-specific fertilizer recommendations for smallholder agroecosystems.

5.5 Conclusion

The study developed a novel approach population-based farm survey for diagnosis of limiting nutrients for smallholder agroecosystems in western Kenya.

(i) Soil test values for N (0.01%), P 12.2 mg kg^{-1} and K (4.5 $cmol_c$ kg^{-1}) were developed from quantitative soil and plant relationships and then used to define cases of nutrient deficiencies.

(ii) Spatial maps for nutrient limitations were developed, which identified occurrence of nutrient limitations in specific geographies for the study area to guide nutrient management for agricultural value chains.

(iii) The N indicated clustering based the Moran Index, which show the multiple high N deficiency occur at several points within the study area. Simulated maps were utilized to show the exact location where nutrient deficiency occur.

(iv) This study demonstrated site-specific diagnosis of nutrients, using NPK soil test values and DRIS indices developed from farm survey.

(v) This information could lead to effectiveness and optimize fertilizer use recommendations in the region. Therefore population-based farm survey approach is an effective diagnostic approach for exploring the spatial variability of soil nutrients, and can be upscaled for future use in similar smallholder agroecosystems.

Chapter Six

Evaluating covariate information for diagnosis of yield-limiting nutrients in smallholder agroecosystem in western Kenya

Abstract

There is need to improve decision making for nutrient management, in order to minimize risks in fertilizer investments by farmers in smallholder agroecosystems. Understanding the influence of biophysical and management factors on soil nutrient and crop responses is a critical component of managing risks for site-specific management decisions. This study aimed to identify key co-variate factors that can be used to improve accuracy of the novel population-based farm survey approach for nutrient diagnostics. This study investigated relationships between soil test value, nutrient diagnostic indices with soil fertility gradient (SFG), soil type (ST) and landscape positioning (LP) factors. A total of 256 plots within 64 farmers' maize fields were surveyed in western Kenya using the Land Degradation Sampling Framework (LDSF). Data was collected on maize grain yield (GY) as a function of soil properties and plant nutrient content for NPK, which corresponded to the aforementioned factors. Diagnostic Recommendation Integrated System (DRIS) indices were calculated using GY and maize tissue data. This data was subjected Principal Component Analysis (PCA) to evaluate correlation between these factors with soil test values and DRIS indices. Multivariate Analysis of Variance (MANOVA) was used to identify key factors that influence variation in soil nutrients and crop response. SFG was a non-significant factor that influenced variation in crop response in the study area. PCA result indicated strong correlation between SFG, ST and LP with soil N and P. ST and LP significantly influenced ($p < 0.001$) soil test values, DRIS indices and GY, and explained 53% of the total variation for the study area. To improve on the accuracy of the population-based farm survey approach, stratification was recommended using ST and LP as the main stratum and key covariate information for this study site. The finding implied that implementation of the population-based farm survey approach should include important covariate information about the site, which may improve accuracy of site-specified nutrient diagnostics for heterogeneous smallholder agroecosystems.

This chapter is based on a manuscript to be submitted to Agricultural Systems Journal

6.1. Introduction

Unless the yield gaps for staple cereals are reduced, feeding the global population of 9.5 billion people projected for 2050, will be an uphill task (Ray *et al.*, 2013; Pradhan *et al.*, 2015). The past decade has seen agricultural intensification being considered amongst mitigation strategies for closing yield gaps, and has resulted to 20% increase of crop yield for smallholder agroecosystems (Van Ittersum *et al.*, 2013). Targeted agronomic intervention including judicious fertilizer application are amongst options for attaining higher crop yields (Andersson & Giller, 2019). Information on site-specific nutrient limitation is therefore crucial, and is required for these smallholder agroecosystems, which are highly heterogeneous in both management and biophysical aspects (Zingore *et al.*, 2007b; Njoroge *et al.*, 2017b). To fully operationalize site-specific nutrient diagnostics, accurate spatially-explicit information is needed to guide nutrient management decisions.

Digital Soil Mapping (DSM) combined with population-based farm survey approach, have the potential of providing accurate spatially-explicit information. DSM produces maps from point observation data using statistical modelling (Kempen *et al.*, 2012; Malone *et al.*, 2016). Often environmental covariates, defined as independent variables that explain variation in the dependent one, are used for statistical modelling (Hengl, 2007; Salkind, 2010). Studies have shown the use of covariates data derived from digital elevation model (*e.g.* relief) and vegetation maps, improve accuracy and reduces uncertainty of predicted soil maps (Mora-Vallejo *et al.*, 2008; Kempen *et al.*, 2012). Normally, DSM follows the Scorpan approach, where these covariates are selected to represent soil forming factors (McBratney *et al.*, 2006; Buol *et al.*, 2011). DSM has also demonstrated potential as a nutrient management tool. For instance, Hengl *et al.* (2015) successfully developed soil nutrient maps at 250 m resolution for maize crop across Africa while Snoeck *et al.* (2010) used DSM to provide fertilizer recommendation for cocoa plantations in Ghana. Burke & Lobell, (2017) used high resolution satellite imagery to map fertilizer response in western Kenya. However, information related to nutrient management is never fully included as covariate, and has not been studied to a great extent. A study by van Apeldoorn *et al.* (2014) revealed difficulty of DSM in capturing soil fertility gradient across smallholder farms of Zimbabwe. Another study conducted by Samuel-Rosa *et al.* (2015) revealed inclusion of many covariates does not increase accuracy of predicted soil properties. Instead, increasing soil point observations was considered the best option of improving the prediction accuracy (Samuel-Rosa *et al.*, 2015). Thus, population-based farm survey approach has potential, and provides a rigorous ground sampling scheme, and may be plausible with regard to increasing field observation and has potential for capturing spatial variation within smallholder agroecosystems (Shepherd *et al.*, 2015).

Several factors are known to influence spatial variability of nutrients in the smallholder agroecosystems. Management factors such as application of manure within smallholder farms result to soil fertility gradients (Zingore *et al.*, 2007b; Tittonell *et al.*, 2013). Landscape position of smallholder farm and soil type are also relevant information related to variability and nutrient management (Gómez-Plaza *et al.*, 2001; Costa *et al.*, 2015). However, heterogeneity within smallholder agroecosystems still present major challenges to understanding factors influencing spatial variation of nutrients.

Population-based farm survey approach offers an opportunity to include covariate information and can improve prediction accuracy of soil properties, especially if combined with DSM. These two approaches, have potential of providing a suitable framework for spatial diagnostic of limiting nutrients. The population-based survey approach employs the Land Health Surveillance (LHS), whose principles are derived from clinical medicine and have been successfully applied in epidemiological studies (Boulos, 2004; Shepherd *et al.*, 2015). In the epidemiological approach, disease test value is used as a criterion to establish diagnostic norms from a study population, which are adjusted to covariates such as sex and age, and then used to assess prevalence of diseases, which later help in designing appropriate targeted interventions (Krall *et al.*, 2014). The LHS encompass a wide array of intervention, including implementation of guidelines, monitoring soil health patterns and advancement of evidence-based management practices (Shepherd *et al.*, 2015). Its concepts were put into operation, and field implementation achieved through the LDSF (Vågen *et al.*, 2010, 2012). The LDSF provided a monitoring and evaluation framework for assessing processes of soil fertility degradation and the effectiveness of intervention measures (Vågen *et al.*, 2012). However, few studies have tested the use of LHS for environmental management, particularly for nutrient management within smallholder agroecosystems (Shepherd *et al.,* 2015). Beedy *et al.* (2015) identified areas with high risk of land degradation in four agro-ecologies of Malawi and proposed targeted management interventions. Smith *et al.* (2008) used surveillance to monitor and manage tree health in Australian forests.

Borrowing LHS principles, this study used a population distribution derived from a farm survey of soil and plant data to derive soil test values and develop DRIS indices. These variables were used to assess prevalence of yield nutrient constraints for maize crop, while including covariate information (factors). The objective of this study was therefore to identify covariates (biophysical and management factors), which could be used to stratify a target population of smallholder farms. It was postulated, biophysical and management factors, which account for the highest variability, and were highly correlated to DRIS indices and crop responses, would be suitable for stratification. The main assumption here was, by including covariate information that accounts for most variation in a population of smallholder maize fields, the overall accuracy of the approach improved. These covariates depend on available information, and were identified based on soil fertility gradient, soil type and landscape positioning of a maize fields within a smallholder agroecosystem in western Kenya. This approach mimics the one used to establish population cohorts in epidemiological population-based surveys (Krall *et al.*, 2014).

6.2. Material and methods

6.2.1. Study area

The study area, a heterogeneous smallholder landscape in western Kenya (0°26′ - 0°18′ northern latitude; 33°58′ - 34°33′ eastern longitude) is delimited by the administrative sub-counties of Boro, Butere, Yala, Khwisero and Ugunja. Ugunja and has been described in section 3.2.1 (Figure 4.1).

6.2.2 Farm survey – Selection of sampling farms

A farm survey was carried out within the LDSF scheme (Vågen *et al.,* 2010). The LDSF is a stratified hierarchical sampling design that captures variability at different scale levels: block, tiles, sub-tiles and fields across a given landscape. The farm survey has been described in section 4.2.3.

The conducted farm survey also captured the biophysical and management factors, which influenced soil properties, DRIS indices and crop response. Data collection involved recording information on biophysical and management information on randomly selected and sampled maize fields. These factors, summarized in Table 6.1 and included: SFG, ST and LP. Selection was guided by information from previous studies (Akponikpè *et al.*, 2011; Ngome *et al.*, 2013; Tittonell *et al.*, 2013; Zingore *et al.*, 2007), and were linked to with corresponding soil test values, DRIS indices, grain yield and maize biomass measured at plot level. The influence of the aforementioned factors on soil test values, DRIS indices, GY and BV were evaluated for different categories across of these factors (Table 6.1).

Table 6.1: Biophysical and management factors recorded during the population-based farm survey of maize fields in the study area.

Factor	Description based on this study	Categories	Author(s)
Soil fertility gradient	Variation of soil nutrients within smallholder farms with respect to management of soil organic matter and location of farmer's homestead	Outfield Mid-field In- field	Tittonell *et al.*, 2007; Zingore et al., 2007a
Reference Soil Groups	Based on distinct characteristics that provide growing benefits and limitations to crop yields.	Nitisols Acrisols	WRB, 2014
Landscape position	Location of the smallholder farm with respect to landscape elevation	Foot slope Mid Ridge Slope	(Arnhold et al., 2015)

6.2.3 Soil and plant sampling

Soil sampling was conducted once during the development stage, the maize silking stage, procedure in section 5.2.2.1. Collected soil and plant samples were analysed for N, P, K nutrient concentrations using Infrared (IR) spectroscopy technique as described in section 5.2.2.1.

A database was used to establish DRIS diagnostic indices for N, P, K. Data on plant samples for N, P, K and maize yield of the target population were used to calculate the DRIS norms in section 5.2.2.3. The functions for N, P, K and DRIS indices were then determined according to the methodology of Beaufils, (1973).

6.2.3 Statistical analysis

Statistical analyses were implemented using R statistical packages (R Core Team, 2018). Descriptive statistics were computed for soil test, plant NPK nutrient concentration, DRIS indices, GY and BV using '*ddplyr'* package (Wickham, 2009). Density plots were developed to check the normal distribution of data using *ggplot2* package (Wickham, 2009). Skewed variables were normalized, prior to further statistical processing. Pearson correlation coefficient (*r*) formed basis for evaluation of strength of the relationships between variables, with $r > 0.50$ considered as strong.

Principle Component Analysis (PCA) was conducted to determine influence of landscape position, soil fertility gradient and soil type on soil test values, DRIS indices, GY and BV using *FactorMineR* (Lê *et al.*, 2008). PCA was used to identify the key factors explaining variance in the dataset without losing important defining information. The PCs are optimal linear combinations of initial variables explaining the variance in descending order. Correlation between soil test values, DRIS indices, GV and BV were also established from the PCA analysis using *factoextra* packages (Lê *et al*., 2008). Multiple Factor Analysis (MFA) were conducted to evaluate association of categories of categorical factors (soil fertility gradient, soil types and landscape position) as function of explained variability in soil test values, DRIS indices, GY and BV (Lê *et al*., 2008).

To identify covariate factors, that may be suitable for stratification of a population of smallholder maize fields, Multivariate Analysis of Variance (MANOVA) was conducted (Huberty & Olejnik, 2005). Wald-Type test statistic were used to assess the influence of the aforementioned categorical factors on soil test values, DRIS indices and crop responses, with $p < 0.005$ taken to be statistically significant. Prior to PCA, MFA and MANOVA each variable was checked to meet assumptions of normality and homoscedasticity. For PCA and MFA, each variable was standardized due to difference in the units of measurement following equation 6.1 (Hengl, 2007) and for equal weighting in the analysis.

$$Z_i \frac{x_i - x_{min}}{x_{max} - x_{min}} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 6.1$$

where z_i is the standardized value, x_i the observed value, while x_{min} and x_{max} are minimum and maximum values, respectively.

6.3 Results

6.3.1 Distribution N, P, K nutrient concentrations

Density plots were used to evaluate the distribution N, P, K concentrations in soil and maize tissue samples. Soil P and K nutrient concentration displayed a skewed distribution (Figure 6.1 d, f). N, P, K maize tissue nutrient concentration displayed a near normal distribution compares to N, P, K in soil. Soil P values were skewed to the left indicating lower P concentration in the soil for this site. Thus, strategic P fertilization may mitigate its deficiency. Soil P and K were log transformed to attain near normal distribution, prior to PCA and MFA analysis.

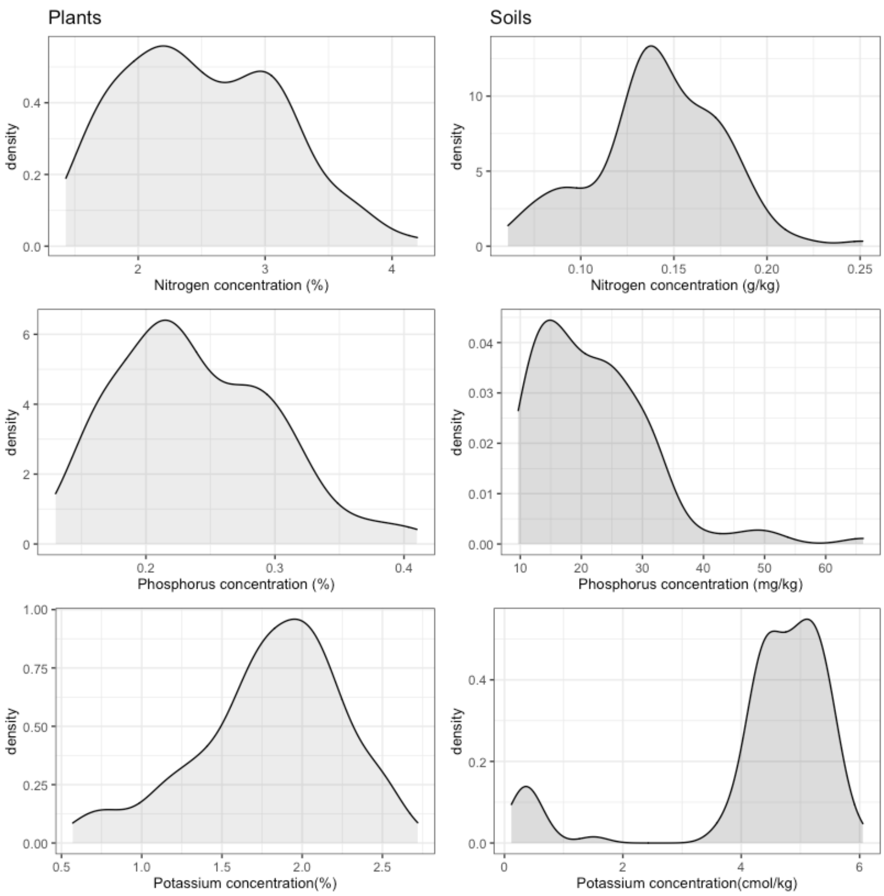

Figure 6.1: Density plots showing distribution of soil and plant tissue NPK nutrient

Descriptive statistics for soil properties, plant samples, yield and plant biovolume, shown Table 6.1 displayed median values of 3.2 Mg ha^{-1} for GY, 0.14 g kg^{-1} for soil N and 17.22 g kg^{-1} for soil P. The median soil N value (0.14 g kg^{-1}) was below 0.2 g kg^{-1}, the critical level for N (Okalebo *et al.,* 2004).

Table 6.1: Descriptive statistics for maize yield soil test values and diagnostic recommendation integrated system (DRIS) indices for NPK nutrients for the target population of smallholder maize fields.

Variable	Observations (n)	Minimum	Maximum	Median	Standard Error	Confidence Interval	Coefficient of Variation
Soil N (g/kg)	203	0.06	0.36	0.14	0.01	0.01	0.25
Soil P (mg/kg)	203	8.19	107.20	17.22	0.80	1.58	0.61
Soil K (cmol/kg)	203	0.12	6.48	4.61	0.09	0.18	0.35
Maize tissue N (%)	237	1.40	4.70	2.52	0.04	0.08	0.25
Maize tissue P (%)	237	0.12	0.49	0.22	0.01	0.01	0.27
Maize tissue K (%)	237	0.57	2.78	1.84	0.03	0.06	0.26
DRIS N index	219	-37.38	58.86	-31.78	1.00	1.96	_
DRIS P index	219	-22.93	21.74	-1.11	0.33	0.65	_
DRIS K index	219	-36.50	52.52	-32.78	0.96	1.90	_
Maize biomass (BV) (cm^3)	255	31.00	392.91	160.86	4.57	9.00	0.43
Grain yield (Mg/ha	257	0.08	11.28	3.20	0.12	0.23	0.54

DRIS – Diagnostic Recommendation Integrated System, N = Nitrogen, P = Phosphorus, K = Potassium, - negative CV values were not report

Figure 6.2 present a correlation matrix of soil test values, maize tissue nutrient concentration and DRIS indices for N, P, K. Significant correlation between soil N and BV (r = 0.67, $p < 0.05$), plant P and soil P (r =0.52, p = 0.00035) and plant K and soil K (r = 0.27, p = 0.000067) were displayed for this region. DRIS N index was significant positive relation ($p < 0.005$) with BV, maize tissue N, soil K, soil P with r values of 0.59, 0.45, 0.34 and 0.38, respectively. Similar positive trend for DRIS K index were observed between maize biomass (r = 0.64), soil N (r = 0.33), soil K (r =0.33) and plant N (r = 0.29). This result support the premise that soil test values and DRIS indices were suitable in determining nutrient status for maize production for this region.

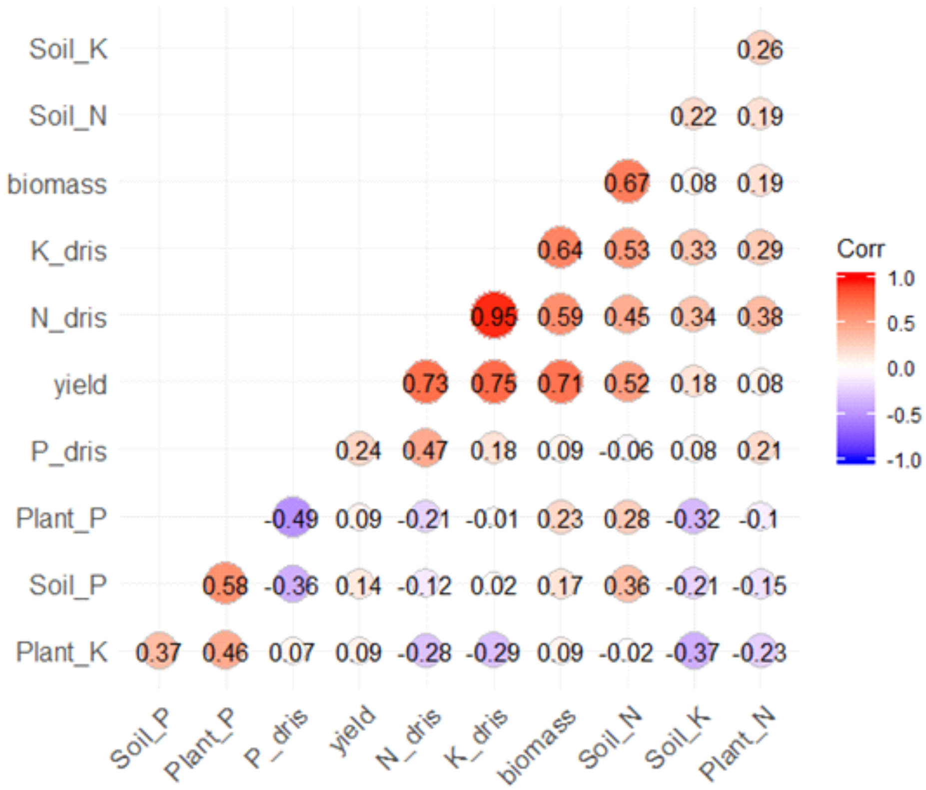

Figure 6.2: Correlation matrix of soil and maize tissue NPK nutrient concentration, DRIS diagnostics norms, grain yield and plant biomass

6.3.2 Effect of soil fertility gradient, landscape position and soil type on crop response.

Means and standard deviation of GY and BV across categories of soil fertility gradient, soil types and landscape position of smallholder maize fields in the study area are presented in Table 6.2. The composition of the sampled maize field, based on relative distance to the homestead constitute SFGs categories; 41% mid-field, 43% infields and the rest were out fields. The infields were located close to homesteads. Infields displayed significant (p = 0.00035) high GY compared to midfields (3.53 Mg ha^{-1}) and outfields (3.41 Mg ha^{-1}). Nitisols and Acrisols were the major reference soil groups for the study area. Nitisols had higher GY (3.76 Mg ha^{-1}) compared to Arcisols (3.44 Mgha^{-1}) and were not statistically different (p = 0.46). Foot slope, mid slope, upslope and ridge were taken to represent landscape position and evaluate their effect crop response. Maize fields located in ridges constitutes 16% of target population, with significantly (p < 0.001) higher mean of GY (4.83 Mg ha^{-1}) compared to those on mid slope (3.39 Mg ha^{-1}). Mean GY of fields located on foot slopes (3.76 Mg ha^{-1}) were higher compared to those on up slope (3.51 Mg ha^{-1}), and were not statistically different (p = 0.67).

Table 6.2: Means and standard deviation for grain yield and plant biomass of the sampled population of smallholder maize fields (p < 0.05)

Covariates (Biophysical and management factors)	Categories	n	Grain yield Mean	Grain yield Standard deviation	Maize biomass Mean	Maize biomass Standard deviation
Soil fertility gradient	Infield	47	4.01	1.98	197.97	69.93
	Midfield	45	3.53	1.93	168.28	66.22
	Outfield	19	2.41	1.05	107.81	55.2
Reference Soil Groups	Acrisols	91	3.44	1.89	159.71	72.96
	Nitisols	21	3.76	1.93	196.34	66.31
Landscape position	Foot slope	36	3.76	2.22	157.55	73.36
	Mid slope	53	3.39	2.19	153.47	79.13
	Ridge	23	4.83	1.47	215.4	56.56
	Upslope	17	3.51	1.65	171.25	62.78

6.3.3 Prevalence of N, P, K nutrient limitations

Soil test values and DRIS indices were used to evaluate prevalence of N, P, K limitation and nutrient imbalance, respectively. Means soil test and confidence interval (CI) values indicated significant difference between outfield and infield for N and K (Figure 6.3 a, e). Conversely, there were no significant difference in soil fertility gradient categories for soil P, although outfield displayed high variability as shown by wide range of CI values (Figure 6.3 d, c, d). Infields displayed high soil N and K test values compared to outfields (Figure 6.3 a, e), which is in agreement to finding reported by Tittonell *et al.* (2013), who observed decreasing soil P concentration with increasing distance from homestead. Outfields recorded the lowest soil test N.

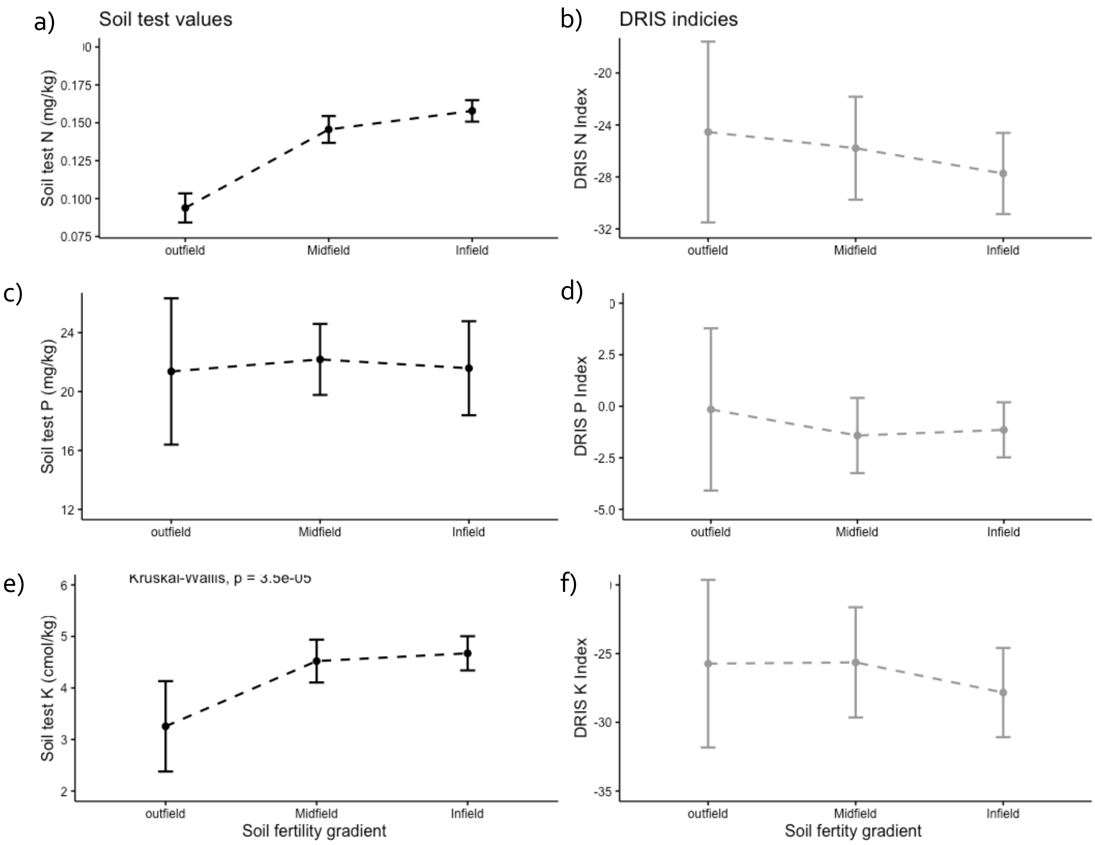

Figure 6.3: Means for grain yield and DRIS indices for NPK across categories of soil fertility gradient of the target population. The error bars represent the confidence interval at 0.05 significant level.

Low soil test values indicated N,P, K nutrient limitations in study area (Römheld, 2011). In the target population of maize fields, prevalence of N nutrient limitations was observed in 45% of infields, 67% of midfield and 74% of outfield, based on critical soil value of

0.2 g kg^{-1}. While phosphorus had 35% of infields, 54% of midfield and 58% of outfield diagnosed as P deficient, based on critical P value of 15 mg kg^{-1}.

Figure 6.4 presents mean plots for soil test values for N, P, K between Nitisols and Acrisols. Significant difference in mean soil P and K (p = 0.00002) were evident in the region. Similar trends were observed for DRIS indices for P and K (Figure 6.4 d, f). High variation was observed in both P concentration in soil and maize tissue, as displayed by the 95% confidence intervals. In this region, 75% of target population farms were diagnosed as N deficient on Arcisols, while Nitisol had 46% and 34% with P and K limitations, respectively.

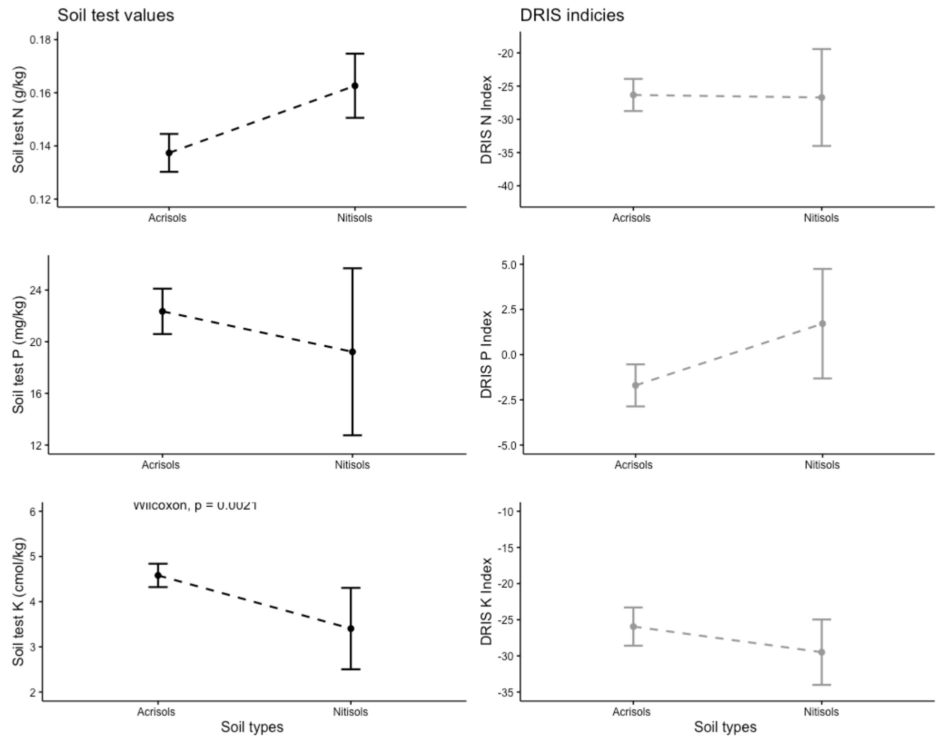

Figure 6.4: Means for soil test value on the left panel, and DRIS indices on the right panel for NPK across categories of soil types of the study area. The error bars represent the confidence interval at 0.05 significant level.

Four categories of landscape position of maize field indicated no significant differences across soil test values (Figure 6.5 a, c, e). High soil N and P values were observed for maize fields located on foot slope and ridges. The higher soil N and K values on the foot slopes can be explained by the possibility low leaching in the smallholder landscape (Waswa *et al*., 2013). Contrary to higher slopes, where soil materials from farms located on mid slope and up slope are transported downwards, leading to accumulation of nutrients (Waswa *et al*., 2013). Prevalence of N limitations was found in 64% of target population of maize fields located in upslope, and 35% of fields on ridges of the study site.

Maize fields located in up slope fields recorded lowest N index value (-28.2), while ridges recorded DRIS N mean of -5.3 an indication of N nutrient imbalance (Nziguheba *et al.*, 2009; Vanlauwe *et al.*, 2014). The DRIS P Index was significantly higher for fields on mid slopes relative to those located on foot slopes and ridges (Figure 6.5 b, d, f). The DRIS N indices were within ranges reported by Nziguheba *et al*. (2009) in western Kenya.

Evaluation of means of soil test values and DRIS indices across categories of SFGs, ST and LPs the target population of smallholder maize farms, provided evidence of variation within the study area (Figure 6.3. 6.4 and 6.5). This variability indicated the importance of adjusting these values, based on covariates such as soil fertility gradient, soil types and landscape positioning of farms. Reliance on one traditionally established critical soil test values, for a wider region without taking into account the variation, can lead to inaccurate diagnosis of limiting nutrients (Olfs *et al.*, 2005).

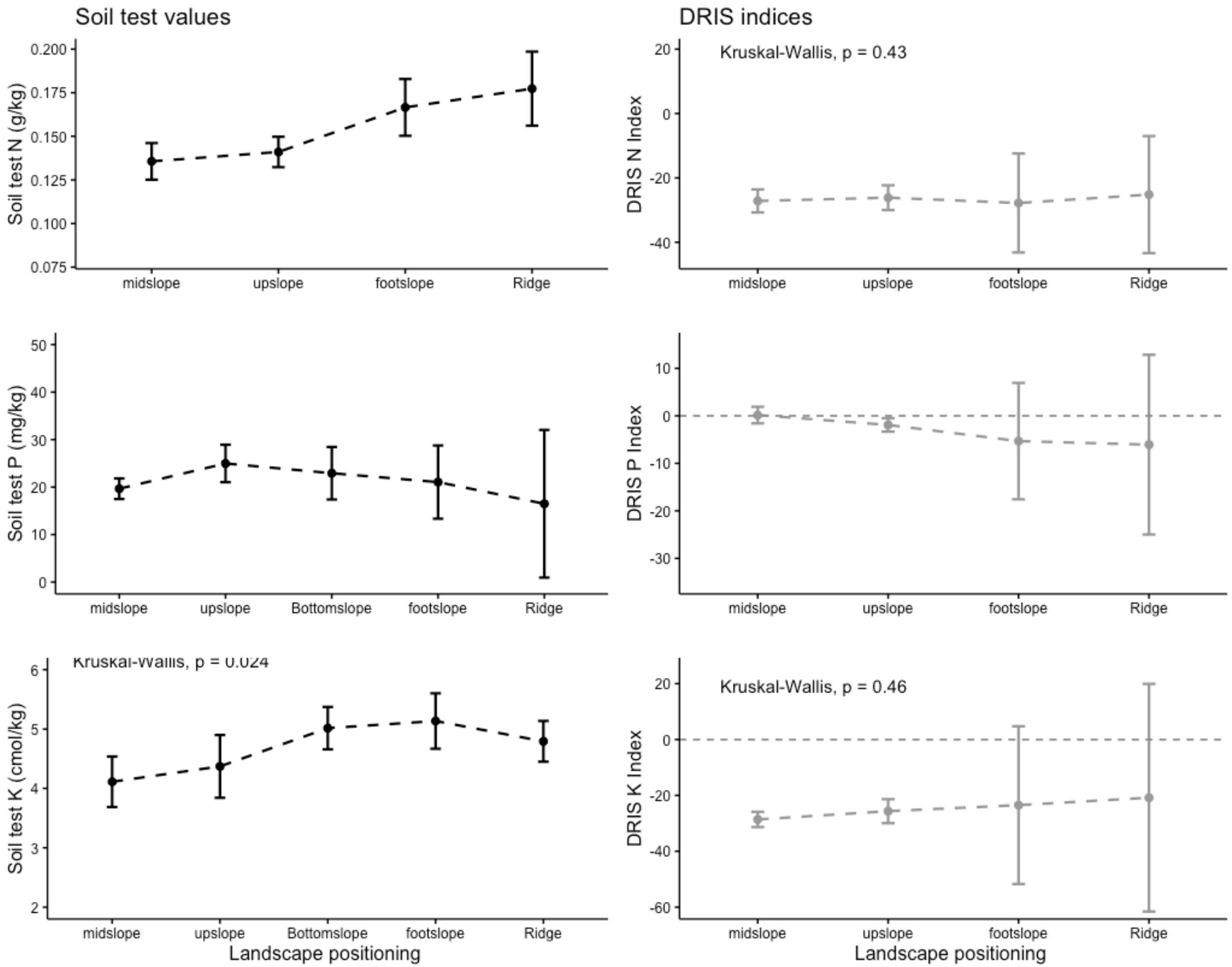

Figure 6.5: Means for grain yield and DRIS indices for NPK across categories of topographical position of the location of smallholder maize field target population. The error bars represent the confidence interval at 0.05 significant level.

6.3.4 Key covariate information - biophysical &management factors

Five principal components (PCs) explained 92% of total variability in soil test values, DRIS indices, GY and BV (Table 6.3). The first component (PC1) explained 33% of total variation, while only 20% was accounted for with the second PC2 (Figure 6.6). PC1 had the highest eigen value (2.64) and was the most important explanatory factor for this study area. PC1 and PC2 sufficiently accounted for 53% of total variation in soil test values, DRIS indices and crop response, which indicated high communality (de Winter & Dodou, 2016). Importance of communality between these variables suggested the practical significance in capturing spatial variability of nutrients for the study area.

Table 6.3: Eigen values and explain variance by five principal components from PCA of soil test values, DRIS indices, grain yield and plant biomass. PC = principal component.

Principal component	Eigen values	Explained variance (%)
PC1	2.64	33.03
PC2	1.67	20.86
PC3	1.35	16.92
PC4	0.91	11.43
PC5	0.82	10.23

Major contributors to variation on PC1 were soil N, BV and GY with strong ($p < 0.005$) negative correlation, and r values of -0.88, -0.77 and -0.75, respectively (Figure 6.6). High factor loading of soil N, GY and BV on PC1 meant they accounted for 33% of total explained variation by PC1. Soil N was highly correlated with GY and BV, an indication that variability in soil N affected GY and BV. DRIS P index had the least contribution, and poorly correlated with PC1 (r = - 0.00008). High loadings were observed on PC2 with soil P and plant N identified as its main contributors to total explained variance of 20%. Previous studies have shown nitrogen is a major contributor on PC1 (Muhati *et al.*, 2011; Mavunganidze *et al.*, 2016; Junqueira *et al.*, 2016). The PCA findings demonstrated importance of utility soil test values and DRIS indices captured, variability in crop responses (GY and BV), since 53% was explained. Thus, a soil-based fertilizer recommendation tool, that would include DRIS indices as indicator of nutrient imbalances may be appropriate for nutrient management strategies for this smallholder agroecosystem (Olfs *et al*., 2005).

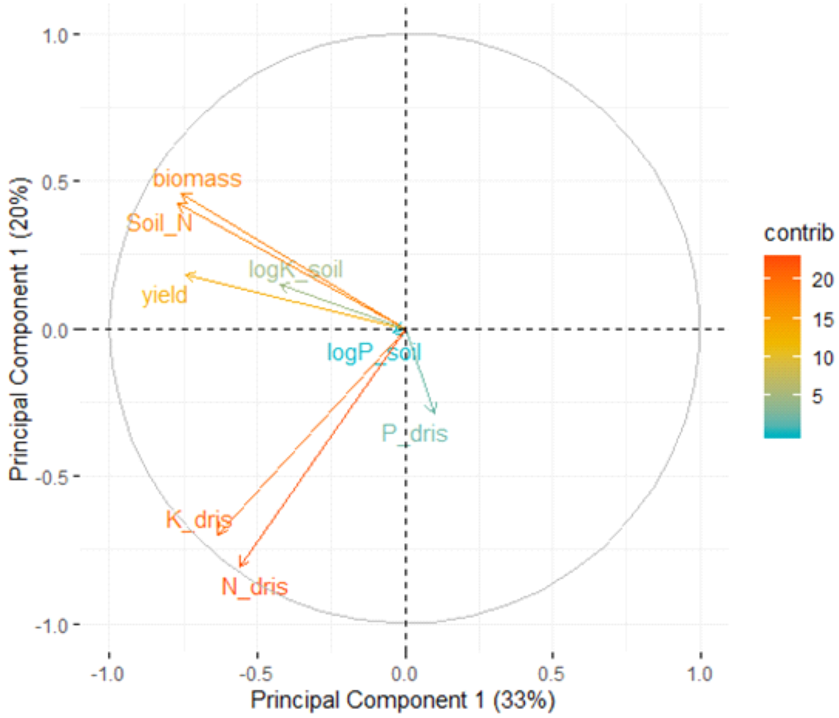

Figure 6.6: Characterization of soil test values, DRIS indices, maize yield and plant biomass through principal components analysis showing observed variance by principal component 1 and 2 projected in biplot of 203 sampled plots. The contrib is the contribution intensity, with blue color representing lowest contribution and red/orange the major contributors.

Figure 6.7 presents multiple factor analysis (MFA) results for categories of factors projected in two-dimensional space (DC1 and DC2). The results revealed 34% of total explained variance was influenced by soil fertility gradient, soil types and landscape positioning (categories of these factors). The DC1 accounted for 17.5 % of total explained variation and was closely related to outfields, infields and mid slopes (Figure 6.7 a). DC2 was closely related to nitisols and midfield, and it explained 16.6% of total variance. Key categorical factors that contributed to the variation on both DC1 and DC2 were outfields nitisols and up slope (Figure 6.7 b). Acrisols and bottom slopes were the least contributors to DC1 and DC2.

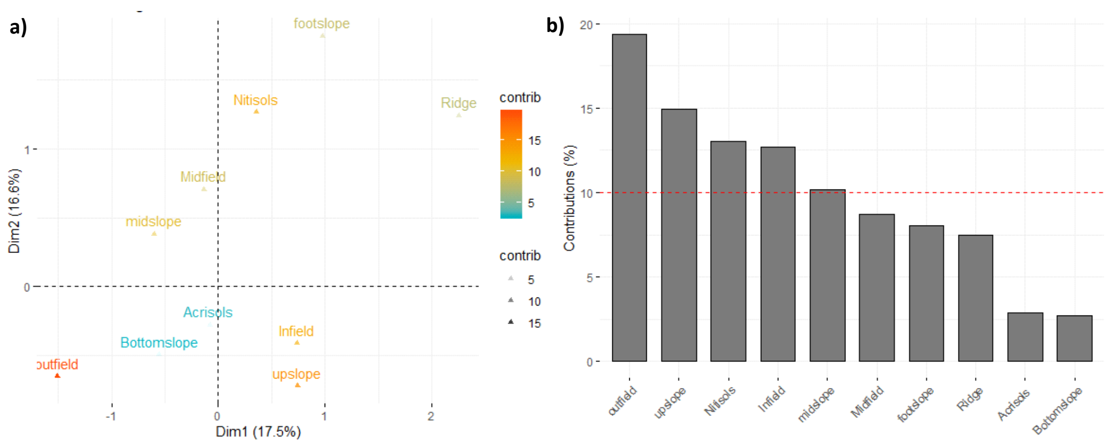

Figure 6.7: Characterization of (a) categories of soil fertility gradient, soil type and landscape positioning factors on two dimensional components (DCs) for multiple factor analysis for 203 plots, b) Relative contribution of factors on their influence of total explained variance for DC1 and DC for the study area, the red dotted line is the threshold line.

MANOVA results indicated soil fertility gradient and soil type significantly influenced ($p < 0.005$) BV, soil nitrogen and potassium (Table 6.4). Only soil fertility gradient significantly influenced GY ($p = 0.004$) contrary observations made by van Apeldoorn et al. (2014), who found no relationships between maize yields and soil fertility gradient on smallholder farms of Zimbabwe. Landscape positioning had a significant influence on soil nitrogen ($p = 0.019$). There were no significant interactions between soil fertility gradient, soil type and landscape positioning on all soil test values and DRIS indices for this region (Table 6.4). However, significant interactions ($p < 0.005$) were observed between soil type and landscape positioning on soil phosphorus and potassium levels. The results have shown a significant influence of soil fertility gradient on GY. This study depicted interaction between soil type and landscape position a key biophysical factor ($p < 0.005$), implying as suitable covariate information for stratification of the population based-farm survey.

Table 6.4: F probability values (*P* value) of biophysical and management factors obtained from multivariate analysis of variance of soil test values, DRIS indices and crop response for the study area based on Wald-type test.

Factor	Degrees of freedom	BV	GY	Soil N	lnP	lnK	N-dris	P-dris	K-dris
SFG	2	**1.345e-10** ***	**0.004247** **	**2.789e-15** ***	0.921	**2.450e-05** ***	0.583	0.728	0.645
landscape	4	0.197	0.831	0.019 *	0.059.	0.084.	0.898	0.097.	0.411
Soil	1	**1.385e-08** ***	0.373	**0.001** **	0.271	**8.449e-06** ***	0.996	0.015 *	0.287
SFG: landscape	6	0.229	0.701	0.197	0.489	0.900	0.665	0.449	0.362
SFG: Soil	2	0.515	0.071	0.575	0.417	0.426	0.266	0.212	0.641
landscape: Soil	4	0.672	0.762	0.237	**0.005** **	**3.428e-07** ***	0.129	0.791	0.327
SFG: landscape: Soil	2	0.302	0.755	0.985	0.075	0.789	0.059	0.086	0.380

Significance codes: *** = 0.001 ** = 0.01 * = 0.05. Bold numbers significant influence ($p < 0.005$)
SFG = soil fertility gradient, landscape = landscape positioning, Soil = soil types, log = natural log, e = exponential

6.4 Discussion.

6.4.1 NPK nutrient concentration and prevalence in limitation

The mean DRIS indices for N, P, K (-37.8, -22.91, and -36.55) were negative, implying an occurrence of nutrient imbalance in the study area (Table 1). Evaluation of means of soil test values and DRIS indices, provided evidence of prevalence in N, P, K nutrient imbalance. Soil test values of 0.10 g kg^{-1} for N 17.29 g kg^{-1} for P and 4.6 cmol kg^{-2} for K, which were below the established critical indicate prevalence in nutrient limitation in the study area (Table 6.1, 6.2). The findings corroborate past studies which report N, P, K deficiency in western Kenya, that lead negative nutrient balances, reflecting crop nutrient mining by farmers (Smaling *et al.,* 1993; Tully *et al.,* 2015). Studies report N deficits in excess of -100 kg ha^{-1} yr^{-1} for maize systems (Montanarella *et al*., 2015). In this region, soil fertility recommendations are provided on the basis of broad soil type and agroecological zones (Bekunda *et al*., 2010; Smaling *et al*., 1992). The conventional practice involve a single critical soil value, established from crop response trials, and is used to diagnose soil nutrient limitations (Olfs *et al.*, 2005; Bekunda *et al.*, 2010; Römheld, 2011). Reliance on one traditionally established critical soil test values, for a wider region without taking into account the variation, can lead to inaccurate diagnosis of limiting nutrients (Olfs *et al.*, 2005). Farmers in this region do not apply balanced doses of fertilizer, which resulted to nutrient imbalance as shown with the DRIS values (Vanlauwe *et al*., 2014), which can be rectified through balance application of NPK fertilizer at optimum rates (Kihara *et al.*, 2016). The inability to produce higher maize yield is a critical problem that is related to nutrient deficiencies (Table 6.1). Farmer require information on nutrient diagnostics as well as imbalance to improve their decision on fertilizer investments. Thus, to mitigate the problem of N, P, K nutrient deficiency, planners need to consider these issues when developing an integrated nutrient management plan for refining blanket fertilizer recommendations for this region.

Variation in means of GY across SFGs categories was attributed to application of more manure as nutrient source to infields compared to outfield, which led to accumulation of soil organic matter (Vanlauwe *et al.*, 2016). These findings corroborate those observed by Tittonell *et al.* (2013), where maize yields varied as a function of distance with respect to location of homestead. There was evidence in variation of soil test values across STs, SFGs and LPs (Table 6.2, Figure 6.3, 6.4, 6.5), which implied that a single soil test value may not be appropriate to guide decisions on nutrient management for the region. Therefore, it is necessary to adjusting soil test values, based on covariates such as SFGs, STs and LPs.

6.4.2 Effect of soil type, soil fertility gradient and landscape position of crop response

Soil types and landscape positioning were key factors influencing variation in soil test values, DRIS norms and crop responses. This is attributed to the explained variance (53%) from PCA, which displayed significant association of nitisol and farms located on up slope landscape position in the study area (Table 6.3, Figure 6.6).

Nitisols and Acrisols were the major reference soil groups for the study area. Nitisols had higher GY (3.76 Mg ha^{-1}) compared to Arcisols (3.44 $Mgha^{-1}$) (Table 6.1), attribute to differences in soil characteristic between these two RSGs. Soil characteristic such as nutrient availability and moisture content influence maize yields in smallholder agroecosystems (Ngome *et al.,* 2013; Vanlauwe *et al*., 2012). Acrisols (80% of observations) were sandy with low water holding capacity, compared to Nitisol, that are clayey (Brady & Weil, 2007; Elias, 2017). Hence higher mean of GY observed on Nitisols were attributed to the good inherent characteristics compare to those of Arcisols, which were shallow with potential of P fixation, due to high aluminium content as reported elsewhere (Sanchez *et al.*, 2003). The observed mean grain yields were in the ranges reported by Ngome *et al*. (2013) on Nitisols and Acrisols in western Kenya, but below potential yield of 15 Mg ha^{-1} (Kihara *et al.,* 2016).

Results show importance of location of maize field relative to position on landscape on GY, which decreased on fields located in on upper slope position compared to those in lower slopes Higher yields on farms located on foot slope (37% of target population) can be explained by nutrient accumulation, that receive alluvial deposits from farms located in up slope and foot slope positions (Afyuni *et al.*, 2010). Variation in available water across the landscape positions also affected maize yield (Afyuni *et al.,* 2010). This can be attributed to the critical role played by landscape on soil formation (Buol *et* *al*., 2011), nutrient availability and water storage, that influenced levels of GY in the study area (Arnhold *et al*., 2015). Studies have also shown soil P and K significantly decreased with elevation in smallholder farms of mount Elgon in western Kenya (De Bauw *et al.*, 2016). The result show soil fertility gradient cause variation in soil test values and crop responses, which agree to early findings where soil fertility gradient was reported as the main factor that influence nutrient on smallholder farms (Zingore *et al.*, 2007b; Diarisso *et al.*, 2016). Contrary the findings of this study, where soil type and land scape position were had higher influence compared to SFGs based on factor loading on PCA (figure 6.7). the variation. Thus, while refining fertilizer recommendation for this region STs, SFGs and LPs may be considered.

6.4.3 Covariate information – biophysical and management factors.

There were no significant interactions between soil fertility gradient, soil type and landscape positioning on all soil test values and DRIS indices for this region (Table 6.4). However, significant interactions ($p < 0.005$) were observed between soil type and landscape positioning on soil phosphorus and potassium levels. The results have shown a significant influence of soil fertility gradient on GY. This study depicted interaction between soil type and landscape position a key biophysical factor ($p < 0.005$), implying as suitable covariate information for stratification of the population based-farm survey. The finding is echoed by those reported by Hermann & Táth, (2011) in smallholder fields in Hungary. The significant influence of soil type and landscape positioning can be explained by the effect on soil nutrient levels, that vary depending on their minerology and soil characteristics across different types (Raynaud & Leadley, 2004; Shehu *et al.*, 2018). This also explains why the current blanket fertilizer recommendations take into account variation in soil types and agroecological zones (Mowo & Mlingano, 1993; FURP, 1994; Wopereis *et al.*, 2006) even though they are spatially coarse.

The objective of this study was to identify key factors that can be used for stratification of a target population in farm surveys for nutrient diagnostics. Stratification reduces sampling variance and improve overall accuracy of predicted nutrient values (Wheeler *et al.*, 2012). Selection of sampling locations across the population distribution stratified along major ST, and LP, can potentially improve nutrient diagnostics of the study area (Sun *et al.*, 2012). Furthermore, through DSM a separate opportunity to incorporate covariate information (e.g. aspect and vegetation) exists (Odeh *et al.*, 1994; Samuel-Rosa *et al.*, 2015). This can be complemented with current approach through rigorous sampling and field monitoring of nutrients in smallholder agroecosystems. Stratification run the risk of failing to capture actual spatial variation in nutrients, especially for smallholder agroecosystems, when used insolation, thereby decreasing the efficiency of sampling (Wheeler *et al*., 2012). This may be especially the true where prior observations within the area of interest are not available and relationships are not evaluated (Odeh *et al.,* 1994). Therefore, combination of population-based survey and DSM provides appropriate synergy. The study has demonstrated the potential of population-based survey to provide prior observation and evaluation of soil and plant relationships for the study area.

The utility of DSM using geostatistical techniques, may allow interpolation of spatial patterns of soil test values and DRIS indices (Hengl *et al*., 2007; Kempen *et al*., 2012). DSM provide an advantage of estimating soil test values in areas not sampled; hence it can be used to complement this current approach (Vašát *et al.*, 2010; Viscarra Rossel *et al.*, 2016). The parting shot here is the reinforcement of the need for accurate spatially explicit nutrient diagnostics in smallholder agroecosystems, which can be implemented in through a stratified population-based farm survey complemented with DSM across smallholder landscapes. The study has contributed new knowledge by providing an understanding of site-specific nutrient diagnostic information that is needed for targeted fertilizer recommendations for this region. This approach can be utilized for analysis of systems anywhere in the world and is necessary because findings from one scale (*e.g*., landscape and farm) can sometimes be counterintuitive when applied at a different scale.

6.5 Conclusion

The study has demonstrated potential of population-based farm survey approach for nutrient diagnostics for smallholder agroecosystems. The main conclusion are as follow;

(i) The DRIS diagnostic norms are essential for evaluating the nutrient status of the soil and plant under smallholder farming system for the study areas.

(ii) These diagnostic norms varied and were influenced by soil fertility gradient, soil type, topographical position and slope characteristics of the study area.

(iii) Soil types were the important factors; therefore, stratification of the population approach should be done using soil stratum for this study area.

(iv) To improve on the accuracy of the population-based farm survey approach, incorporation of covariate information (soil type and landscape position) is necessary.

(v) Inclusion of environmental covariates information improves estimation of soil nutrient deficiencies at smallholder landscape scales. As such, there is potential of using large-scale DSM approaches combine with population-based farm survey may improve prediction nutrient limitations in smallholder farming systems.

(vi) These results show geostatistical analysis using kriging is an effective prediction tool for exploring the spatial variability of soil nutrients, and is recommend for future soil sampling campaigns for smallholder farming systems in Kenya.

(vii) Smallholder farmers need simple empirical soil and DRIS maps that can used to guide them, showing regions of nutrient imbalance and N,P,K deficiencies.

Chapter Seven

General Conclusions and Recommendations

This thesis developed series of approaches for diagnosis and mapping soil nutrient deficiencies for smallholder farming landscapes, capturing the local variability on smallholder farms. The results presented show that; *(i.)* Phosphorus, silt and potassium were the key factors that influence variation and in fertilizer response and nutrient use efficiency in Chapter three, *(ii.)* A spatial resolution of 250 m was proposed as threshold for developing digital soil maps for nutrient management in smallholder landscape in Chapter four. Such maps captured the local spatial variability occurring at farm level. Thus, maps at this spatial resolution would be effective decision support tools and would be instrumental for strategies aimed at refining fertilizer recommendations. *(iv.)* A population-based farm survey approach was used to develop soil test values for NPK, rather than the conventional crop response trials. Soil test values of 0.01 g kg^{-1}, 12 mg kg^{-1}, 4.5 $cmol_c$ kg^{-1} for NPK, respectively, were established and used to diagnose their limitations across the landscape in Chapter five. *(iv.)* Covariate information was analysed to identity information that could be used stratifying target population of maize fields in the novel approach – population-based farm survey, in Chapter 5. Soil types and Landscape position were the important covariate information that was to be incorporated the farm survey in order to improve its accuracy. These findings suggest a soil-based fertilizer recommendation, with potential of using farm survey data and digital soil mapping in strategies aimed at refining fertilizer recommendations in smallholder farming systems. The main finding and implication of the thesis are summarized as follows.

7. 1 Variability in fertilizer response and nutrient use efficiency.

Using meta-analysis approach on secondary data, this chapter three established high variability in fertilizer response and nutrient use efficiency for the study area. High variability in fertilizer response and nutrient use efficiency implied that the current blanket fertilizer recommendations may not be efficient. Fertilizer response was proposed as a proxy indicator for evaluating soil responsiveness to mineral fertilizer application. A criterion for identifying the occurrence of non-responsive soils for smallholder agroecosystem was also established.

The findings of chapter three have major implications for strategies of refining fertilizer recommendation. The main finding indicates only a small part of variation was explain by soil properties; phosphorus, silt and potassium. These soil factors influence fertilizer response and nutrient use efficiencies, and may be targeted as entry points in strategies aimed at refining fertilizer recommendations. Thus, the use of digital nutrient maps would be a good basis as a starting point for directing strategies aimed at refining fertilizer recommendations, for obtaining spatial explicit fertilizer recommendations. A soil-based fertilizer recommendation system may be appropriate for this study area.

7.2 Analysis of scale for provision of fertilizer recommendations

The objective of chapter four was to determine a relevant scale for provision of fertilizer recommendation or optimum sampling distance for developing digital nutrient maps. Provision of appropriate fertilizer recommendations for smallholder farmers requires that they are provided at a scale that is suitable, reflecting the local variability on smallholder farm. The main finding was the occurrence of short distance variability across smallholder agroecosystems of the study area. The short distance variability implies that nutrient diagnostics should be based on the local farm conditions. Using a farm survey approach, data on soil properties and crop responses, a scale of < 273 m for developing digital soil nutrient maps was established. The distance of 250 was considered optimum for soil testing for the study area.

Based on the LDSF sampling strategy, this study established fertilizer recommendation should be provided at field scale. Hence, soil testing should be conducted on every field in the study area. However, conducting soil testing for every field may be very expensive due to the high cost in wet chemical laboratory analysis (Bekunda *et al.*, 2010). Soil testing requires that the analytical methods used are efficiency and cost-effective. Infrared spectral methods constitute rapid diagnostic screening tools that can be used to diagnose fields with limiting nutrients (Shepherd and Walsh, 2007) and are relatively cheap for large scale assessment. The utility of infrared spectroscopy may provide adequate soil spatial information on nutrients and would help provide fertilizer recommendations at field scale.

The sequence of main steps used to establish the scale for providing fertilizer recommendations in the population-based farm survey approach have been outlined. This chapter studied the use of the farm survey data to establish a scale that can be used to develop digital nutrient maps.

7.3 Population-based survey approach for diagnosis of nutrient limitations

In Chapters five and six, a novel population-based farm survey approach was established. This approach may be used to guide nutrient diagnostic for targeted soil fertility replenishment programs at fine spatial resolution. Critical soil nutrient values for defining cases of constraints were developed. Digital soil mapping combined with infrared spectroscopy technique were used to map and characterize spatial variation of soil nutrient constraints across the study area. Rather than conventional agronomic trial used for nutrient diagnostics in smallholder agroecosystems, the developed population-based farm survey approach may be an alternative since it captures the spatial variation of the study area.

Deficiency of NPK contents in the study area as displayed by the DRIS indices could be improved with the use of site-specified fertilizer to reduce the wide variation within the contents. However, the application of the fertilizer material should be done based on the nutrient contents already in the soil as depicted by the spatial distribution maps and other resources available to smallholder farmers of the study area.

The spatial distribution maps generated through this study showed locations of low, moderate and high nutrient contents. It suggests that management zones could be easily targeted without going through any tedious and laborious means to identify areas of low or adequate NPK nutrient contents for decision making purposes. The NPK nutrient concentrations displayed a moderate strength of spatial dependencies within each of them. The spatial dependencies of the nutrient contents in the study area confirmed that the variation in the spread of their distributions were influenced soil fertility gradient, soil type, topographical and slope characteristics. Determining within-field nutrient levels allows the variable-rate application of fertilizers. When considerable variability is present, immediate economic returns are possible, provided the variability is on a portion of the yield/nutrient curve which allows increased yield or quality if application rates are varied.

Biophysical and management factors such organic matter management, topography and slope of the study area had weak association with the distribution pattern of diagnostic norms (DRIS indices). Soil type was the only factor that has a significant relationship with of diagnostic norms and crop responses. Thus, to improve on the accuracy of the population-based farm survey, stratification should be done across soil types for the study area, in order to reduce soil sampling variance.

The present study has demonstrated the use of population-based farm survey in combination with spectroscopic technique and digital soil mapping as potential tools that for providing spatially explicit information on nutrient limitations. Combining these techniques provides new opportunities for characterizing variation in nutrient limitations across smallholder agroecosystems. Thus, enhances the potential of developing a soil-based fertilizer use decision support for smallholder agroecosystems

7.4 Recommendations for priorities in future research

The broad objective of this study was to develop a diagnostic system for deploying rapid spectral analysis techniques of soil and plant samples within a spatial sampling framework to guide refining fertilizer recommendations for smallholder agroecosystems.

Despite the valuable insights provided on nutrient diagnostics, there remain a number of unanswered questions that require further research. Further research is needed to refine the population-based farm survey nutrient assessment, to enhance its wider applicability. For instance, research could explore the possibility of estimating the actual fertilizer rates for each of the geographical niche within the agroecosystem. Research findings may provide options of nutrient management that might result in better yields than those of obtained by farmers in the study area. But the economic benefit, considering fertilizer investment as a proposed intervention, may not be beneficial to improve farmers' livelihood if not significant. This is because not all yield increment is profitable, and challenges of the occurrence of non-responsive soils need to addressed. Therefore, there need to prioritise research on how smallholder farmers will economically benefit from site-specific recommendations. Therefore, refining fertilizer recommendations should further investigate variability in their profitable that the smallholder farmer may obtain from the current practice, before refining fertilizer recommendations. This will be important in managing risks of fertilizer investments by the smallholder farmers.

8. Reference

Ade Freeman, H. & Omiti, J. 2003. Fertilizer use in semi-arid areas of Kenya: analysis of smallholder farmers' adoption behavior under liberalized markets. *Nutrient Cycling in Agroecosystems*, **66**, 23–31, (At: http://dx.doi.org/10.1023/A:1023355011400.).

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

Afyuni, M.M., Cassel, D.K. & Robarge, W.P. 2010. Effect of Landscape Position on Soil Water and Corn Silage Yield. *Soil Science Society of America Journal*.

AIC. 1981. Agricultural Information Centre (AIC) Major Crops Technical Handbook. .

Akponikpè, P.B.I., Minet, J., Gérard, B., Defourny, P. & Bielders, C.L. 2011. Spatial fields' dispersion as a farmer strategy to reduce agro-climatic risk at the household level in pearl millet-based systems in the Sahel: A modeling perspective. *Agricultural and Forest Meteorology*, **151**, 215–227, (At: http://www.sciencedirect.com/science/article/pii/S0168192310002820. Accessed: 5/1/2015).

Andersson, J. & Giller, K. 2019. *Doing development-oriented agronomy : Rethinking methods, concepts and direction*.

Angeles, D.E., Sumner, M.E. & Barbour, N.W. 1990. Preliminary nitrogen, phosphorus, and potassium DRIS norms for pineapple. *HortScience*, **25**, 652–655.

Antwi, M., Duker, A.A., Fosu, M. & Abaidoo, R.C. 2016. Geospatial approach to study the spatial distribution of major soil nutrients in the Northern region of Ghana. *Cogent Geoscience*, **2**, (At: http://doi.org/10.1080/23312041.2016.1201906.).

Arnhold, S., Otieno, D., Onyango, J., Koellner, T., Huwe, B. & Tenhunen, J. 2015. Soil properties along a gradient from hillslopes to the savanna plains in the Lambwe Valley, Kenya. *Soil & Tillage Research*, **154**, 75–83, (At: http://10.0.3.248/j.still.2015.06.021.).

Arrouays, D., Grundy, M.G., Hartemink, A.E., Hempel, J.W., Heuvelink, G.B.M., Hong, S.Y., Lagacherie, P., Lelyk, G., McBratney, A.B., McKenzie, N.J., Mendonca-Santos, M. d L., Minasny, B., Montanarella, L., Odeh, I.O.A., Sanchez, P.A., Thompson, J.A. & Zhang, G.-L. 2014. Chapter Three - GlobalSoilMap: Toward a Fine-Resolution Global Grid of Soil Properties. In: *Advances in Agronomy* (ed. Donald, L.S.), pp. 93–134. Academic Press.

Bailey, J.S., Cushnahan, A. & Beattie, J.A.M. 1997. The diagnosis and recommendation integrated system (DRIS) for diagnosing the nutrient status of grassland swards: II. Model calibration and validation. *Plant and Soil*, **197**, 137–147.

Barreto, H.J. & Bell, M.A. 1994. Assessing risk associated with N fertilizer recommendations in the absence of soil tests. *Nutrient Cycling in Agroecosystems*, **40**, 175–183, (At: http://dx.doi.org/10.1007/BF00750463.).

de Barros, I., Gaiser, T., Lange, F.M. & Römheld, V. 2007. Mineral nutrition and water use patterns of a maize/cowpea intercrop on a highly acidic soil of the tropic semiarid. *Field Crops Research*, **101**, 26–36.

Barto, E.K. & Rillig, M.C. 2010. Does herbivory really suppress mycorrhiza? A meta-analysis. *Journal of Ecology*, **98**, 745–753.

Baruck, J., Nestroy, O., Sartori, G., Baize, D., Traidl, R., Vrščaj, B., Bräm, E., Gruber, F.E., Heinrich, K. & Geitner, C. 2016. Soil classification and mapping in the Alps: The current state and future challenges. *Geoderma*.

Bates, D., Mächler, M., Bolker, B.M. & Walker, S.C. 2015. Fitting linear mixed-effects models using lme4.

Journal of Statistical Software.

Bationo, A., Hartemink, A., Lungu, O., Naimi, M., Okoth, P., Smaling, E., Thiombiano, L. & Waswa, B. 2012. Knowing the African Soils to Improve Fertilizer Recommendations. In: *Improving Soil Fertility Recommendations in Africa using the Decision Support System for Agrotechnology Transfer (DSSAT) SE - 3* (eds. Kihara, J., Fatondji, D., Jones, J.W., Hoogenboom, G., Tabo, R. & Bationo, A.), pp. 19–42. Springer Netherlands.

Bationo, A., Mokwunye, U., Vlek, P.L.G., Koala, S., Shapiro, B.I. & Yamoah, C. 2003. Soil fertility management for sustainable land use in the West African Sudano-Sahelian zone. In: Gichuru MP, Bationo A, Bekunda MA, Goma PC, Mafongoya PL, Mugendi DN, Murwira HM, Nandwa SM, Nyathi P and Swift MJ (Eds.). Soil fertility management in Africa.

De Bauw, P., Van Asten, P., Jassogne, L. & Merckx, R. 2016. Soil fertility gradients and production constraints for coffee and banana on volcanic mountain slopes in the East African Rift: A case study of Mt. Elgon. *Agriculture, Ecosystems and Environment*.

Beaufils, E. 1973. Diagnosis and recommendation integrated system (DRIS). *Soil Science Bulletin-University of Natal (Sudafrica).(*.

Beedy, T., Njoloma, J., Aynekulu, E., Coe, R., Takoutsing, B. & Shepherd, K. 2015. Land health surveillance in four agroecologies in Malawi. *ICRAF Working Paper - World Agroforestry Centre*.

Bekunda, M., Sanginga, N. & Woomer, P.L. 2010. Restoring soil fertility in sub-sahara Africa. **108**, 183–236, (At: http://www.scopus.com/inward/record.url?eid=2-s2.0-77956584538&partnerID=40&md5=353bb2fec63f0b58e7c89ba655a5962a.).

Benson, T. 1999. Area-specific fertilizer recommendations for hybrid maize grown by Malawian smallholders.

Bhatti, A.U., Hussain, F., Ullah, F. & Khan, M.J. 1998. Use of spatial patterns of soil properties and wheat yield in geostatistics for determination of fertilizer rates. *Communications in Soil Science and Plant Analysis*.

Van Biljon, J.J., Fouche, D.S. & Botha, A.D.P. 2008. The lower and upper threshold values, biological optimum and mineralization of nitrogen in the main maize producing soils of South Africa. *South African Journal of Plant and Soil*, **25**, 8–13, (At: http://www.scopus.com/inward/record.url?eid=2-s2.0-44349150429&partnerID=40&md5=dae02b29d07270c8c6fea7ed71048fc5.).

Borenstein, M., Hedges, L. V., Higgins, J.P.T. & Rothstein, H.R. 2009. *Introduction to Meta-Analysis*.

Boulos, M. 2004. Towards evidence-based, GIS-driven national spatial health information infrastructure and surveillance services in the United Kingdom. *International Journal of Health Geographics*, **3**, 1, (At: http://www.ij-healthgeographics.com/content/3/1/1.).

Brady, N.C. & Weil, R.R. 2007. *The nature and properties of soils*. (At: http://www.amazon.com/Nature-Properties-Soils-14th-Edition/dp/013227938X/ref=sr_1_1?ie=UTF8&qid=1342639425&sr=8-1&keywords=the+nature+and+properties+of+soils+14th+edition.).

Brouwer, J., Fussell, L.K. & Herrmann, L. 1993. Soil and crop growth micro-variability in the West African semi-arid tropics: a possible risk-reducing factor for subsistence farmers. *Agriculture, Ecosystems and Environment*, **45**, 229–238.

Brown, H.K. & Kempton, R.A. 1994. The application of REML in clinical trials. *Statistics in medicine*, **13**, 1601–1617.

Buol, S.W., Southard, R.J., Graham, R.C. & McDaniel, P.A. 2011. Soil-Forming Factors: Soil as a Component of Ecosystems. *Soil Genesis and Classification*, 89–140, (At: http://dx.doi.org/10.1002/9780470960622.ch3.).

Burke, M. & Lobell, D.B. 2017. Satellite-based assessment of yield variation and its determinants in smallholder African systems. *Proceedings of the National Academy of Sciences*.

Van Buuren, S. & Groothuis-Oudshoorn, K. 2011. Multivariate Imputation by Chained Equations. *Journal Of Statistical Software*, **45**, 1–67, (At: http://igitur-archive.library.uu.nl/fss/2010-0608-200146/UUindex.html.).

Cairns, J.E., Hellin, J., Sonder, K., Araus, J.L., MacRobert, J.F., Thierfelder, C. & Prasanna, B.M. 2013. Adapting maize production to climate change in sub-Saharan Africa. *Food Security*, **5**, 345–360, (At: http://dx.doi.org/10.1007/s12571-013-0256-x.).

Cambardella, C.A. & Karlen, D.L. 1999. Spatial Analysis of Soil Fertility Parameters. *Precision Agriculture*, **1**, 5–14, (At: http://dx.doi.org/10.1023/A:1009925919134.).

Cambardella, C.A., Moorman, T.B., Parkin, T.B., Karlen, D.L., Novak, J.M., Turco, R.F. & Konopka, A.E. 1994. Field-Scale Variability of Soil Properties in Central Iowa Soils. *Soil Sci. Soc. Am. J.*, **58**, 1501–1511, (At: https://dl.sciencesocieties.org/publications/sssaj/abstracts/58/5/1501.).

Campbell, C.R. 2000. *Reference sufficiency ranges for plant analysis in the southern region of the United States Southern Cooperative Series Bulletin #394. Southern Cooperative Series Bulletin (Vol. 394)*.

Cassman, K.G. 1999. Ecological intensification of cereal production systems: yield potential, soil quality, and precision agriculture. *Proceedings of the National Academy of Sciences*, **96**, 5952–5959.

Chikuvire, T.J., Mpepereki, S. & Foti, R. 2007. Soil Fertility Variability in Sandy Soils and Implications for Nutrient Management by Smallholder Farmers in Zimbabwe. *Journal of Sustainable Agriculture*, **30**, 69–87.

Chong, I.G. & Jun, C.H. 2005. Performance of some variable selection methods when multicollinearity is present. *Chemometrics and Intelligent Laboratory Systems*, **78**, 103–112.

Cobo, J.G., Dercon, G. & Cadisch, G. 2009. Nutrient balances in African land use systems across different spatial scales: A review of approaches, challenges and progress. *Agriculture, Ecosystems & Environment*, **136**, 1–15, (At: http://www.sciencedirect.com/science/article/B6T3Y-4XY4S9H-1/2/d24374fac82adc1fa7868393f06a5ceb.).

Cobo, J.G., Dercon, G., Yekeye, T., Chapungu, L., Kadzere, C., Murwira, A., Delve, R. & Cadisch, G. 2010. Integration of mid-infrared spectroscopy and geostatistics in the assessment of soil spatial variability at landscape level. *Geoderma*, **158**, 398–411.

Costa, C., Papatheodorou, E.M., Monokrousos, N. & Stamou, G.P. 2015. Spatial Variability of Soil Organic C, Inorganic N and Extractable P in a Mediterranean Grazed Area. *Land Degradation and Development*.

Das, D.K., Maiti, D. & Pathak, H. 2009. Site-specific nutrient management in rice in Eastern India using a modeling approach. *Nutrient Cycling in Agroecosystems*, **83**, 85–94.

Deckers, J. 2002. A systems approach to target balanced nutrient management in soilscapes of sub-Saharan Africa. *Integrated plant nutrient management in Sub-Saharan Ấfrica. From concepts to practice. CAB International, Wallingford, Oxon, UK*, 47–62.

Diarisso, T., Corbeels, M., Andrieu, N., Djamen, P., Douzet, J.-M. & Tittonell, P. 2015. Soil variability and crop yield gaps in two village landscapes of Burkina Faso. *Nutrient Cycling in Agroecosystems*, 1–18, (At: http://dx.doi.org/10.1007/s10705-015-9705-6.).

Diarisso, T., Corbeels, M., Andrieu, N., Djamen, P., Douzet, J.-M. & Tittonell, P. 2016. Soil variability and crop yield gaps in two village landscapes of Burkina Faso. *Nutrient Cycling in Agroecosystems*, **105**, 199–216, (At: https://doi.org/10.1007/s10705-015-9705-6.).

Diwani, T.N., Asch, F., Becker, M. & Mussgnug, F. 2013. Characterizing farming systems around Kakamega Forest, Western Kenya, for targeting soil fertility–enhancing technologies. *Journal of Plant Nutrition and Soil Science*, **176**, 585–594, (At: http://dx.doi.org/10.1002/jpln.201200036.).

Dobermann, A., Witt, C., Dawe, D., Abdulrachman, S., Gines, H.C., Nagarajan, R., Satawathananont, S., Son, T.T., Tan, P.S., Wang, G.H., Chien, N. V, Thoa, V.T.K., Phung, C. V, Stalin, P., Muthukrishnan, P., Ravi, V., Babu, M., Chatuporn, S., Sookthongsa, J., Sun, Q., Fu, R., Simbahan, G.C. & Adviento, M.A.A. 2002. Site-specific nutrient management for intensive rice cropping systems in Asia. *Field Crops Research*, **74**, 37–66, (At: http://www.scopus.com/inward/record.url?eid=2-s2.0-18244393225&partnerID=40&md5=569a76220ff14fa3be521d53f2ae5d61.).

Drechsel, P., Gyiele, L., Kunze, D. & Cofie, O. 2001. Population density, soil nutrient depletion, and economic growth in sub-Saharan Africa. *Ecological Economics*, **38**, 251–258, (At: http://www.sciencedirect.com/science/article/pii/S0921800901001677. Accessed: 13/8/2015).

Dudal, R. 1980. Soil-related constraints to agricultural development in the tropics. *Priorities for alleviating soil-related constraints to food production in the tropics*, 23–37.

Duflo, E., Kremer, M. & Robinson, J. 2008. How High Are Rates of Return to Fertilizer? Evidence from Field Experiments in Kenya. *The American Economic Review*, **98**, 482–488, (At: http://www.jstor.org/stable/29730068.).

Duval, S. & Tweedie, R. 2000. Trim and Fill: A Simple Funnel-Plot–Based Method of Testing and Adjusting for Publication Bias in Meta-Analysis. *Biometrics*, **56**, 455–463, (At: http://dx.doi.org/10.1111/j.0006-341X.2000.00455.x.).

Elias, E. 2017. Characteristics of Nitisol profiles as affected by land use type and slope class in some Ethiopian highlands. *Environmental Systems Research*, **6**, 20, (At: https://doi.org/10.1186/s40068-017-0097-2.).

FAO, F. and A.O. of the U.N., IFAD, I.F. for A.D. & WFP, W.F.P. 2015. The State of Food Insecurity in the World. In: *International Hunger Targets: taking stock of uneven progress*.

Farrés, M., Platikanov, S., Tsakovski, S. & Tauler, R. 2015. Comparison of the variable importance in projection (VIP) and of the selectivity ratio (SR) methods for variable selection and interpretation. *Journal of Chemometrics*, **29**, 528–536, (At: http://doi.wiley.com/10.1002/cem.2736. Accessed: 15/8/2017).

Fiehn, O., Sumner, L.W., Rhee, S.Y., Ward, J., Dickerson, J., Lange, B.M., Lane, G., Roessner, U., Last, R. & Nikolau, B. 2007. Minimum reporting standards for plant biology context information in metabolomic studies. *Metabolomics*, **3**, 195–201, (At: https://doi.org/10.1007/s11306-007-0068-0.).

Fischer, G., Shah, M., van Velthuizen, H. & Nachtergaele, F.O. 2001. *Global agro-ecological assessment for agriculture in the 21st century*. IIASA. (At: http://www.iiasa.ac.at/Research/LUC/Papers/gaea.pdf.).

Foster, E.F. 2001. Nutrient Deficiencies and Toxicities of Plants CD-ROM. *Crop Science*, **41**, 917.

Frederiksen, M.E., Vázquez-Prada Baillet, M., Jensen, P.T., Rygaard, C., Hallas, J. & Lynge, E. 2019. Conization and healthcare use: A population-based register study. *European Journal of Cancer Prevention*.

FURP. 1994. The Fertilizer Use Recommendation Project , Kenya Agricultural Research Institute, National Agricultural Research Laboratories, Nairobi. *Final Report. Annex I: Fertilizer trial documentation (FERDOC). Ministry of Agriculture, Nairobi*.

Geladi, P., Hadjiiski, L. & Hopke, P. 1999. Multiple regression for environmental data: Nonlinearities and prediction bias. In: *Chemometrics and Intelligent Laboratory Systems*.

Geurts, P.M.H. & Van den Berg, M. 1998. A simple agro-ecological zonation for fertilizer recommendations in Mozambique. *Soil Use and Management*, **14**, 136–141, (At: http://dx.doi.org/10.1111/j.1475-2743.1998.tb00138.x.).

Giller, K.E., Rowe, E.C., de Ridder, N. & van Keulen, H. 2006a. Resource use dynamics and interactions in the tropics: Scaling up in space and time. *Agricultural Systems*, **88**, 8–27, (At: http://www.sciencedirect.com/science/article/pii/S0308521X05000867.).

Giller, K.E., Rowe, E.C., De Ridder, N. & Van Keulen, H. 2006b. Resource use dynamics and interactions in the tropics: Scaling up in space and time. *Agricultural Systems*, **88**, 8–27.

Gómez-Plaza, A., Martínez-Mena, M., Albaladejo, J. & Castillo, V.M. 2001. Factors regulating spatial distribution of soil water content in small semiarid catchments. *Journal of Hydrology*, **253**, 211–226.

Goovaerts, P. 2000. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *Journal of Hydrology*, **228**, 113–129, (At: http://www.sciencedirect.com/science/article/pii/S002216940000144X.).

Goswami, R., Saha, S. & Dasgupta, P. 2017. Sustainability assessment of smallholder farms in developing countries. *Agroecology and Sustainable Food Systems*, **41**, 546–569.

Grömping, U. 2006. Relative importance for linear regression in R: The package relaimpo. *Journal of Statistical Software*.

Grömping, U. 2009. Variable importance assessment in regression: Linear regression versus random forest. *American Statistician*, **63**, 308–319.

Guan, Y., Wei, J., Zhang, D., Zu, M. & Zhang, L. 2013. To identify the important soil properties affecting dinoseb adsorption with statistical analysis. *The Scientific World Journal*.

Haileslassie, A., Priess, J., Veldkamp, E., Teketay, D. & Lesschen, J.P. 2005. Assessment of soil nutrient depletion and its spatial variability on smallholders' mixed farming systems in Ethiopia using partial versus full nutrient balances. *Agriculture, Ecosystems and Environment*, **108**, 1–16.

Havlin, J.L. & Jacobsen, J.S. 1994. *Soil testing: prospects for improving nutrient recommendations*. Soil Science Society of America/American Society of Agronomy; SSSA Special Publication, 40. (At: http://www.scopus.com/inward/record.url?eid=2-s2.0-0028682824&partnerID=tZOtx3y1.).

Hedges, L. V., Gurevitch, J. & Curtis, P.S. 1999. The meta-analysis of response ratios in experimental ecology. *Ecology*, **80**, 1150–1156.

Hedges, L. V & Olkin, I. 1985. Statistical methods for meta-analysis. Orlando.

Hengl, T. 2007. *A Practical Guide to Geostatistical Mapping of Environmental Variables*.

Hengl, T., Heuvelink, G.B.M., Kempen, B., Leenaars, J.G.B., Walsh, M.G., Shepherd, K.D., Sila, A., MacMillan, R.A., Mendes de Jesus, J., Tamene, L. & Tondoh, J.E. 2015. Mapping Soil Properties of Africa at 250 m Resolution: Random Forests Significantly Improve Current Predictions. *PloS one*, **10**, e0125814, (At: http://dx.doi.org/10.1371/journal.pone.0125814. Accessed: 26/10/2015).

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., Wheeler, I. & Kwabena, N.A. 2017. Soil nutrient maps of Sub-Saharan Africa: assessment of soil nutrient content at 250 m spatial resolution using machine learning. *Nutrient Cycling in Agroecosystems*, **109**, 77–102, (At: https://doi.org/10.1007/s10705-017-9870-x.).

Hermann, T. & Táth, G. 2011. Evaluating the effect of nutrient levels of major soil types on the productivity of wheatlands in Hungary. *Communications in Soil Science and Plant Analysis*.

Herrero, M., Thornton, P.K., Notenbaert, A.M., Wood, S., Msangi, S., Freeman, H.A., Bossio, D., Dixon, J., Peters, M., Van De Steeg, J., Lynam, J., Rao, P., MacMillan, S., Gerard, B., McDermott, J., Seré, C. & Rosegrant, M. 2010. Smart investments in sustainable food production: Revisiting mixed crop-livestock systems. *Science*, **327**, 822–825.

Huberty, C.J. & Olejnik, S. 2005. *Applied MANOVA and Discriminant Analysis*.

Huedo-Medina, T.B., Sánchez-Meca, J., Marín-Martínez, F. & Botella, J. 2006. Assessing heterogeneity in meta-analysis: Q statistic or I2 index? *Psychological methods*, **11**, 193–206.

Huising, E.J., Zingore, S., Kihara, J., Nziguheba, G. & Palm, C.A. 2011. Diagnostic trials: a field guide.

Ichami, S.M., Shepherd, K.D., Sila, A.M., Stoorvogel, J.J. & Hoffland, E. 2018. Fertilizer response and nitrogen use efficiency in African smallholder maize farms. *Nutrient Cycling in Agroecosystems*, (At: https://doi.org/10.1007/s10705-018-9958-y.).

Ishak, K.J., Platt, R.W., Joseph, L. & Hanley, J.A. 2008. Impact of approximating or ignoring within-study covariances in multivariate meta-analyses. *Statistics in Medicine*, **27**, 670–686, (At: http://dx.doi.org/10.1002/sim.2913.).

IUSS Working Group WRB. 2014. *World reference base for soil resources 2014. International soil classification system for naming soils and creating legends for soil maps*.

Janssen, B.H.H., Guiking, F.C.T.C.T., van der Eijk, D., Smaling, E.M.A.M.A., Wolf, J. & van Reuler, H. 1990. A system for quantitative evaluation of the fertility of tropical soils (QUEFTS). *Geoderma*, **46**, 299–318, (At: http://www.sciencedirect.com/science/article/pii/001670619090021Z. Accessed: 20/12/2014).

Jayne, T.S., Myers, R.J. & Nyoro, J. 2008. The effects of NCPB marketing policies on maize market prices in Kenya. *Agricultural Economics*, **38**, 313–325, (At: http://doi.wiley.com/10.1111/j.1574-0862.2008.00302.x. Accessed: 26/6/2017).

Jeffery, S., Verheijen, F.G. a., van der Velde, M. & Bastos, a. C. 2011. A quantitative review of the effects of biochar application to soils on crop productivity using meta-analysis. *Agriculture, Ecosystems & Environment*, **144**, 175–187, (At: http://linkinghub.elsevier.com/retrieve/pii/S0167880911003197. Accessed: 19/7/2012).

Johnson, D.W. & Curtis, P.S. 2001. Effects of forest management on soil C and N storage: Meta analysis. *Forest Ecology and Management*, **140**, 227–238.

Jordan-Meille, L., Rubæk, G.H., Ehlert, P.A.I., Genot, V., Hofman, G., Goulding, K., Recknagel, J., Provolo, G. & Barraclough, P. 2012. An overview of fertilizer-P recommendations in Europe: Soil testing, calibration and fertilizer recommendations. *Soil Use and Management*.

Junqueira, A.B., Stomph, T.J., Clement, C.R. & Struik, P.C. 2016. Variation in soil fertility influences cycle dynamics and crop diversity in shifting cultivation systems. *Agriculture, Ecosystems & Environment*, **215**, 122–132, (At: https://www.sciencedirect.com/science/article/pii/S0167880915300815. Accessed: 24/7/2019).

KAmanga, b.c.g., waddington, s.r., whitbread, a.m., almekinders, c.j.m. & giller, k.e. 2014. improving the efficiency of use of small amounts of nitrogen and phosphorus fertiliser on smallholder maize in central malawI. *Experimental Agriculture*, **50**, 229–249, (At: http://journals.cambridge.org/article_S0014479713000513.).

Kempen, B., Brus, D.J., Stoorvogel, J.J., Heuvelink, G.B.M. & de Vries, F. 2012. Efficiency Comparison of Conventional and Digital Soil Mapping for Updating Soil Maps. *Soil Science Society of America Journal*.

Kerry, R. & Oliver, M.A. 2004. Average variograms to guide soil sampling. *International Journal of Applied Earth Observation and Geoinformation*, **5**, 307–325.

Keskin, H. & Grunwald, S. 2018. Regression kriging as a workhorse in the digital soil mapper's toolbox. *Geoderma*.

Kibblewhite, M.G., Ritz, K. & Swift, M.J. 2008. Soil health in agricultural systems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **363**, 685 LP – 701, (At: http://rstb.royalsocietypublishing.org/content/363/1492/685.abstract.).

Kihara, J., Huising, J., Nziguheba, G., Waswa, B.S., Njoroge, S., Kabambe, V., Iwuafor, E., Kibunja, C., Esilaba, A.O. & Coulibaly, A. 2015. Maize response to macronutrients and potential for profitability in sub-Saharan Africa. *Nutrient Cycling in Agroecosystems*, 1–11, (At: http://dx.doi.org/10.1007/s10705-015-9717-2.).

Kihara, J., Nziguheba, G., Zingore, S., Coulibaly, A., Esilaba, A., Kabambe, V., Njoroge, S., Palm, C. & Huising, J. 2016. Understanding variability in crop response to fertilizer and amendments in sub-Saharan Africa.

Agriculture, Ecosystems & Environment, **229**, 1–12.

Kihara, J., Sileshi, G.W., Nziguheba, G., Kinyua, M., Zingore, S. & Sommer, R. 2017. Application of secondary nutrients and micronutrients increases crop yields in sub-Saharan Africa. *Agronomy for Sustainable Development*, **37**.

Kleinman, P.J.A., Sharpley, A.N., Gartley, K., Jarrell, W.M., Kuo, S., Menon, R.G., Myers, R., Reddy, K.R. & Skogley, E.O. 2001. Interlaboratory comparison of soil phosphorus extracted by various soil test methods. *Communications in Soil Science and Plant Analysis*, **32**, 2325–2345, (At: http://www.tandfonline.com/doi/abs/10.1081/CSS-120000376.).

Krall, M.A., Gundlapalli, A. V. & Samore, M.H. 2014. Big Data and Population-Based Decision Support. *Clinical Decision Support*, 363–381, (At: https://www.sciencedirect.com/science/article/pii/B9780123984760000130. Accessed: 3/7/2019).

Kravchenko, a. N., Robertson, G.P., Snap, S.S. & Smucker, a. J.M. 2006. Using information about spatial variability to improve estimates of total soil carbon. *Agronomy Journal*, **98**, 823–829.

Kroll, C.N. & Song, P. 2013. Impact of multicollinearity on small sample hydrologic regression models. *Water Resources Research*.

Kuhn, M. 2008a. caret Package. *Journal Of Statistical Software*, **28**, 1–26, (At: http://www.jstatsoft.org/v28/i05/paper.).

Kuhn, M. 2008b. Building Predictive Models in R Using the caret Package. *Journal Of Statistical Software*, **28**, 1–26, (At: http://www.jstatsoft.org/v28/i05/.).

Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T. & Mayer, Z. 2014. caret: Classification and regression training. R package version 6.0-21.

Kumwenda, J.D.T., Waddington, S.R., Snapp, S.S., Jones, R.B. & Blackie, M.J. 1996. Soil fertility management research for the maize cropping systems of smallholders in Southern Africa: A review. *Natural Resources Group Paper*, 157–172, (At: http://repository.cimmyt.org/xmlui/bitstream/handle/10883/531/61970.pdf.).

Lagacherie, P. & Mcbratney, A.B. 2007. Spatial Soil Information Systems and Spatial Soil Inference Systems : Perspectives for Digital Soil Mapping. *Development in Soil Science*, **31**, 3–22.

Lal, R. 2016. Soil health and carbon management. *Food and Energy Security*, **5**, 212–222.

Lê, S., Josse, J., Husson, F. & others. 2008. FactoMineR: an R package for multivariate analysis. *Journal of statistical software*, **25**, 1–18.

Lehmann, J. & Kleber, M. 2015. The contentious nature of soil organic matter. *Nature*.

Lipscombe, L.L., Hwee, J., Webster, L., Shah, B.R., Booth, G.L. & Tu, K. 2018. Identifying diabetes cases from administrative data: A population-based validation study. *BMC Health Services Research*.

Little, R.J. a & Rubin, D.B. 2002. *Statistical Analysis with Missing Data*.

López-López, J.A., Marín-Martínez, F., Sánchez-Meca, J., Van den Noortgate, W. & Viechtbauer, W. 2014. Estimation of the predictive power of the model in mixed-effects meta-regression: A simulation study. *British Journal of Mathematical and Statistical Psychology*, **67**, 30–48.

Luo, Y., Hui, D. & Zhang, D. 2006. Elevated CO2 stimulates net accumulations of carbon and nitrogen in land ecosystems: A meta-analysis. *Ecology*, **87**, 53–63, (At: http://dx.doi.org/10.1890/04-1724.).

Mafongoya, P.L., Bationo, A., Kihara, J. & Waswa, B.S. 2006. Appropriate technologies to replenish soil fertility in southern Africa. *Nutrient Cycling in Agroecosystems*, **76**, 137–151.

Mairura, F.S., Mugendi, D.N., Mwanje, J.I., Ramisch, J.J. & Mbugua, P.K. 2007a. Assessment of farmers' perceptions of soil quality indicators within smallholder farms in the central highlands of Kenya. *Advances* *in Integrated Soil Fertility Management in Sub-Saharan Africa: Challenges and Opportunities*, 1035–1046.

Mairura, F.S., Mugendi, D.N., Mwanje, J.I., Ramisch, J.J., Mbugua, P.K. & Chianu, J.N. 2007b. Integrating scientific and farmers' evaluation of soil quality indicators in Central Kenya. *Geoderma*, **139**, 134–143, (At: http://www.sciencedirect.com/science/article/B6V67-4NC5T4N-1/2/ead91e87f85a6961e37c71fae87e751d.).

Maiti, D., Das, D.K. & Pathak, H. 2006. Simulation of fertilizer requirement for irrigated wheat in eastern India using the QUEFTS model. *TheScientificWorldJournal*, **6**, 231–245.

Malone, B.P., Jha, S.K., Minasny, B. & McBratney, A.B. 2016. Comparing regression-based digital soil mapping and multiple-point geostatistics for the spatial extrapolation of soil data. *Geoderma*, **262**, 243–253.

Mavunganidze, Z., Madakadze, I.C., Nyamangara, J. & Mafongoya, P. 2016. Influence of selected soil properties, soil management practices and socio-economic variables on relative weed density in a hand hoe-based conservation agriculture system. *Soil Use and Management*, **32**, 433–445, (At: https://onlinelibrary.wiley.com/doi/abs/10.1111/sum.12287.).

Maynard, D.N. 1979. Nutritional disorders of vegetable crops: A review. *Journal of Plant Nutrition*, **1**, 1–23, (At: http://www.tandfonline.com/doi/abs/10.1080/01904167909362696.).

Mbakaya, D.S., Okalebo, J.R., Muricho, M. & Lumasayi, S. 2008. EFFECTS OF LIMING AND INORGANIC FERTILIZERS ON MAIZE YIELD IN KAKAMEGA NORTH AND UGUNJA DISTRICTS , WESTERN KENYA. *Sustainable Development*, 123–129.

McBratney, A., Field, D.J. & Koch, A. 2014. The dimensions of soil security. *Geoderma*, **213**, 203–213, (At: https://www.sciencedirect.com/science/article/pii/S0016706113002954. Accessed: 21/11/2019).

McBratney, A.B., Mendonça Santos, M.L. & Minasny, B. 2003. On digital soil mapping. *Geoderma*, **117**, 3–52, (At: http://www.sciencedirect.com/science/article/pii/S0016706103002234.).

McBratney, A.B., Minasny, B. & Viscarra Rossel, R. 2006. Spectral soil analysis and inference systems: A powerful combination for solving the soil data crisis. *Geoderma*, **136**, 272–278, (At: http://www.sciencedirect.com/science/article/pii/S0016706106001066.).

Mehmood, T., Liland, K.H., Snipen, L. & Sæbø, S. 2012. A review of variable selection methods in Partial Least Squares Regression. *Chemometrics and Intelligent Laboratory Systems*, **118**, 62–69.

Mevik, B.-H. & Wehrens, R. 2007. The pls Package: Principal Component and Partial Least Squares Regression in R. *Journal of Statistical Software*, **18**, 1–24.

Miller, B.A., Koszinski, S., Wehrhan, M. & Sommer, M. 2015. Impact of multi-scale predictor selection for modeling soil properties. *Geoderma*, **239**, 97–106.

Minasny, B., Malone, B.P., McBratney, A.B., Angers, D.A., Arrouays, D., Chambers, A., Chaplot, V., Chen, Z.S., Cheng, K., Das, B.S., Field, D.J., Gimona, A., Hedley, C.B., Hong, S.Y., Mandal, B., Marchant, B.P., Martin, M., McConkey, B.G., Mulder, V.L., O'Rourke, S., Richer-de-Forges, A.C., Odeh, I., Padarian, J., Paustian, K., Pan, G., Poggio, L., Savin, I., Stolbovoy, V., Stockmann, U., Sulaeman, Y., Tsui, C.C., Vågen, T.G., van Wesemael, B. & Winowiecki, L. 2017. Soil carbon 4 per mille. *Geoderma*.

Montanarella, L., Badraoui, M., Chude, V., Baptista, I., Mamo, T., Yemefack, M., Aulakh, M., Yagi, K., Hong, S.-Y., Vijarnsorn, P., Zhang, G., Arrouays, D., Black, H., Krasilnikov, P., Sobocká, J., Alegre, J., Henríquez, C., Henriquez, R., Mendonça Santos, M. & Charzynski, P. 2015. *Status of the World's Soil Resources (SWSR) – Main Report*.

Mora-Vallejo, A., Claessens, L., Stoorvogel, J. & Heuvelink, G.B.M. 2008. Small scale digital soil mapping in Southeastern Kenya. *Catena*, **76**, 44–53.

Morris, M., Kelly, V. a, Kopicki, R.J. & Byerlee, D. 2007. *Fertilizer Use in African Agriculture*. (At:

http://www.journals.cambridge.org/abstract_S0014479707005777.).

Mowo, J.G. & Mlingano, A.R.I. 1993. *Review of fertilizer recommendations in Tanzania*. National Soil Service, ARI Mlingano, Ministry of Agriculture.

Moyer, G.R., Sweka, J.A. & Peterson, D.L. 2012. Past and Present Processes Influencing Genetic Diversity and Effective Population Size in a Natural Population of Atlantic Sturgeon. *Transactions of the American Fisheries Society*, **141**, 56–67, (At: http://dx.doi.org/10.1080/00028487.2011.651073.).

Muhati, S.I., Shepherd, K.D., Gachene, C.K.K., Mburu, M.W.K., Jones, R.B., Kironchi, G.O. & Sila, A. 2011. Diagnosis of Soil Nutrient Constraints in Small-scale Groundnut (Arachis hyopaea L .) Production Systems of Western Kenya Using Infrared Spectroscopy. *Journal Of Agricultural Sciencies and Technology*.

Muyayabantu, G.M., Kadiata, B.D. & Nkongolo, K.K. 2012. Response of maize to different organic and inorganic fertilization regimes in monocrop and intercrop systems in a Sub-Saharan Africa Region. *Journal of Soil Science and Environmental Management*, **3**, 42–48, (At: http://www.google.co.ke/url?sa=t&rct=j&q=&esrc=s&source=web&cd=8&sqi=2&ved=0CFcQFjAH&url=http://www.academicjournals.org/jssem/PDF/pdf2012/February/Muyayabantu et al.pdf&ei=N42OUJaCMo-yhAf2qYHwCw&usg=AFQjCNHm91GjIffmWzxmnwmjPyVRpo5M-A.).

Mzuku, M., Khosla, R., Reich, R., Inman, D., Smith, F. & MacDonald, L. 2005. Spatial Variability of Measured Soil Properties across Site-Specific Management Zones. *Soil Science Society of America Journal*, **69**, 1572.

Nakagawa, S., Johnson, P.C.D. & Schielzeth, H. 2017. The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society Interface*.

Nakagawa, S. & Schielzeth, H. 2013. A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*.

Namonje-Kapembwa, T., Black, R. & Jayne, T.S. 2015. Does Late Delivery of Subsidized Fertilizer Affect Smallholder Maize Productivity and Production ? *Paper prepared for presentation at the 2015 Agricultural& Applied Economics Association and Western Agricultural Economics Association Annual Meeting, San Francisco, CA, July 26-28, 2015*.

Ngome, A.F.E., Becker, M. & Mtei, K.M. 2011. Leguminous cover crops differentially affect maize yields in three contrasting soil types of Kakamega, Western Kenya. *Journal of Agriculture and Rural Development in the Tropics and Subtropics*, **112**, 1–10.

Ngome, A.F., Becker, M., Mtei, M.K. & Mussgnug, F. 2013. Maize productivity and nutrient use efficiency in Western Kenya as affected by soil type and crop management. *International Journal of Plant Production*, **7**, 517–536, (At: http://www.scopus.com/inward/record.url?eid=2-s2.0-84876958669&partnerID=40&md5=6d7493e7c250389ebd6d17083cc1ffa0.).

Njoroge, R., Otinga, A.N., Okalebo, J.R., Pepela, M. & Merckx, R. 2017a. Occurrence of poorly responsive soils in western Kenya and associated nutrient imbalances in maize (Zea mays L.). *Field Crops Research*, **210**, 162–174, (At: http://www.sciencedirect.com/science/article/pii/S0378429017302538.).

Njoroge, R., Otinga, A.N., Okalebo, J.R., Pepela, M. & Merckx, R. 2018. Maize (Zea mays L.) Response to Secondary and Micronutrients for Profitable N, P and K Fertilizer Use in Poorly Responsive Soils. *Agronomy*.

Njoroge, S., Schut, A.G.T., Giller, K.E. & Zingore, S. 2017b. Strong spatial-temporal patterns in maize yield response to nutrient additions in African smallholder farms. *Field Crops Research*, **214**, 321–330.

Nziguheba, G., Tossah, B.K., Diels, J., Franke, A.C., Aihou, K., Iwuafor, E.N.O., Nwoke, C. & Merckx, R. 2009. Assessment of nutrient deficiencies in maize in nutrient omission trials and long-term field experiments in the West African Savanna. *Plant and Soil*, **314**, 143–157, (At: http://dx.doi.org/10.1007/s11104-008-9714-1.

).

Odeh, I.O.A., Mcbratney, A.B. & Chittleborough, D.J. 1994. Spatial prediction of soil properties from landform attributes derived from a digital elevation model. *Geoderma*, **63**, 197–214.

Okalebo, R.J., Gathua, W.K. & Woomer, P.L. 2002. Laboratory Methods of Soil and Plant Analysis: A Working Manual. TSBF-CIAT and Sacred Africa, Nairobi, Kenya.

Okalebo, J.R., Othieno, C.O., Woomer, P.L., Karanja, N.K., Semoka, J.R.M., Bekunda, M.A., Mugendi, D.N., Muasya, R.M., Bationo, A. & Mukhwana, E.J. 2006. Available technologies to replenish soil fertility in East Africa. *Nutrient Cycling in Agroecosystems*, **76**, 153–170, (At: http://dx.doi.org/10.1007/s10705-005-7126-7.).

Okalebo, J.., Palm, C.., Lekasi, J.., Nandwa, S.., Othieno, C.. & Waigwa, M. 2004. Use of organic and inorganic resources to increase maize yields in some Kenyan infertile soils: A five year experience. In: Bationo A (Ed), Managing Nutrient Cycles to Sustain Soil Fertility in Sub-Saharan Africa, Academy of Science Publishers, Nairobi, K.

Okeyo, J.M., Shepherd, K.D., Wamicha, W. & Shisanya, C. 2009. Spatial Variation in Soil Organic Carbon within Smallholder Farms in Western Kenya: A geospatial Approach.

Okoth, P.F., Murua, E., Sanginga, N., Chianu, J., Mungatu, J.M., Kimani, P.K. & Ng'ang'a, J.K. 2011. Some Facts About Fertilizer Use in Africa: The Case of Smallholder and Large-Scale Farmers in KenyaInnovations as Key to the Green Revolution in Africa. In: (eds. Bationo, A., Waswa, B., Okeyo, J.M.M., Maina, F. & Kihara, J.M.), pp. 869–878. Springer Netherlands.

Olfs, H.-W., Blankenau, K., Brentrup, F., Jasper, J., Link, A. & Lammel, J. 2005. Soil- and plant-based nitrogen-fertilizer recommendations in arable farming. *Journal of Plant Nutrition and Soil Science*, **168**, 414–431, (At: http://dx.doi.org/10.1002/jpln.200520526.).

Pampolino, M.F., Witt, C., Pasuquin, J.M., Johnston, A. & Fisher, M.J. 2012. Development approach and evaluation of the Nutrient Expert software for nutrient management in cereal crops. *Computers and Electronics in Agriculture*, **88**, 103–110, (At: http://www.sciencedirect.com/science/article/pii/S0168169912001925.).

Payton, R.W., Barr, J.J.F., Martin, A., Sillitoe, P., Deckers, J.F., Gowing, J.W., Hatibu, N., Naseem, S.B., Tenywa, M. & Zuberi, M.I. 2003. Contrasting approaches to integrating indigenous knowledge about soils and scientific soil survey in East Africa and Bangladesh. *Geoderma*.

Pebesma, E.J. 2004. Multivariable geostatistics in S: the gstat package. *Computers & Geosciences*, **30**, 683–691, (At: http://www.sciencedirect.com/science/article/pii/S0098300404000676.).

Persson, A., Pilesjö, P. & Eklundh, L. 2005. Spatial Influence of Topographical Factors on Yield of Potato (Solanum tuberosum L.) in Central Sweden. *Precision Agriculture*, **6**, 341–357, (At: http://dx.doi.org/10.1007/s11119-005-2323-6.).

Petersen, J., Thomsen, I., Mattsson, L., Hansen, E. & Christensen, B. 2012. Estimating the crop response to fertilizer nitrogen residues in long-continued field experiments. *Nutrient Cycling in Agroecosystems*, 1–12, (At: http://dx.doi.org/10.1007/s10705-012-9482-4.).

Philibert, A., Loyce, C. & Makowski, D. 2012. Assessment of the quality of meta-analysis in agronomy. *Agriculture, Ecosystems and Environment*, **148**, 72–82, (At: http://www.sciencedirect.com/science/article/pii/S016788091100404X. Accessed: 24/12/2014).

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & R Core Team. 2019. {nlme}: Linear and Nonlinear Mixed Effects Models. (At: https://cran.r-project.org/package=nlme.).

Pradhan, P., Fischer, G., van Velthuizen, H., Reusser, D.E., Kropp, J.P. & Nelson, A. 2015. Closing Yield Gaps: How Sustainable Can We Be? (JL Gonzalez-Andujar, Ed.). *PLOS ONE*, **10**, e0129487, (At:

http://dx.plos.org/10.1371/journal.pone.0129487. Accessed: 14/3/2017).

Quemada, M., Baranski, M., Nobel-de Lange, M.N.J., Vallejo, A. & Cooper, J.M. 2013. Meta-analysis of strategies to control nitrate leaching in irrigated agricultural systems and their effects on crop yield. *Agriculture, Ecosystems & Environment*, **174**, 1–10, (At: http://www.sciencedirect.com/science/article/pii/S016788091300128X. Accessed: 13/3/2015).

R Core Team. 2018. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2018. R Foundation for Statistical Computing. *R Foundation for Statistical Computing, Vienna, Austria.*

Ramakrishna, A., Bailey, J. & Kirchhof, G. 2009. A preliminary diagnosis and recommendation integrated system (DRIS) model for diagnosing the nutrient status of sweet potato (<i>Ipomoea batatas</i>). *Plant and Soil*, **316**, 107–116, (At: http://dx.doi.org/10.1007/s11104-008-9763-5.).

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**.

Raynaud, X. & Leadley, P.W. 2004. Soil characteristics play a key role in modeling nutrient competition in plant communities. *Ecology*, **85**, 2200, (At: http://www.scopus.com/scopus/inward/record.url?eid=2-s2.0-8144219856&partner=40&rel=R4.5.0.).

Reuter, D. & Robinson, J. Ben. 1997. *Plant analysis: an interpretation manual*. CSIRO publishing.

Römheld, V. 2011. Diagnosis of Deficiency and Toxicity of Nutrients. In: *Marschner's Mineral Nutrition of Higher Plants: Third Edition*.

Ronner, E., Franke, A.C., Vanlauwe, B., Dianda, M., Edeh, E., Ukem, B., Bala, A., van Heerwaarden, J. & Giller, K.E. 2016. Understanding variability in soybean yield and response to P-fertilizer and rhizobium inoculants on farmers' fields in northern Nigeria. *Field Crops Research*, **186**.

Ros, G.H., Temminghoff, E.J.M. & Hoffland, E. 2011. Nitrogen mineralization: a review and meta-analysis of the predictive value of soil tests. *European Journal of Soil Science*, **62**, 162–173, (At: http://dx.doi.org/10.1111/j.1365-2389.2010.01318.x.).

Roth, G.W., Fox, R.H. & Marshall, H.G. 1989. Plant Tissue Tests for Predicting Nitrogen Fertilizer Requirements of Winter Wheat. *Agronomy Journal*, **81**, 502.

Rufino, M.C., Tittonell, P., van Wijk, M.T., Castellanos-Navarrete, A., Delve, R.J., de Ridder, N. & Giller, K.E. 2007. Manure as a key resource within smallholder farming systems: Analysing farm-scale nutrient cycling efficiencies with the NUANCES framework. *Livestock Science*, **112**, 273–287.

Rurinda, J., Zingore, S., Jibrin, J.M., Balemi, T., Masuki, K., Andersson, J.A., Pampolino, M.F., Mohammed, I., Mutegi, J., Kamara, A.Y., Vanlauwe, B. & Craufurd, P.Q. 2020. Science-based decision support for formulating crop fertilizer recommendations in sub-Saharan Africa. *Agricultural Systems*.

Salkind, N. 2010. Encyclopedia of Research Design. (At: https://methods.sagepub.com/reference/encyc-of-research-design.).

Samberg, L.H., Gerber, J.S., Ramankutty, N., Herrero, M. & West, P.C. 2016. Subnational distribution of average farm size and smallholder contributions to global food production. *Environmental Research Letters*.

Samuel-Rosa, A., Heuvelink, G.B.M., Vasques, G.M. & Anjos, L.H.C. 2015. Do more detailed environmental covariates deliver more accurate soil maps? *Geoderma*.

Sanchez, P. a, Ahamed, S., Carré, F., Hartemink, A.E., Hempel, J., Huising, J., Lagacherie, P., Mcbratney, A.B., Mckenzie, N.J., De, M., Mendonça-santos, L., Minasny, B., Montanarella, L., Okoth, P., Palm, C. a, Sachs, J.D., Shepherd, K.D., Vågen, T., Vanlauwe, B., Walsh, M.G., Winowiecki, L. a & Zhang, G. 2009a. Digital Soil Map of the World. *Science*, **325**, 6–7, (At: http://www.sciencemag.org/content/325/5941/680.short.).

Sanchez, P., Denning, G. & Nziguheba, G. 2009b. The African Green Revolution moves forward. *Food Security*, **1**, 37–44, (At: http://dx.doi.org/10.1007/s12571-009-0011-5.).

Sanchez, P.A., Palm, C.A. & Buol, S.W. 2003a. Fertility capability soil classification: A tool to help assess soil quality in the tropics. *Geoderma*.

Sanchez, P.A., Palm, C.A. & Buol, S.W. 2003b. Fertility capability soil classification: A tool to help assess soil quality in the tropics. *Geoderma*, **114**, 157–185.

Sattari, S.Z., van Ittersum, M.K., Bouwman, A.F., Smit, A.L. & Janssen, 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.

Schnier, H.F., Recke, H., Muchena, F.N. & Muriuki, A.W. 1996. Towards a practical approach to fertilizer recommendations for food crop production in smallholder farms in Kenya. *Nutrient Cycling in Agroecosystems*, **47**, 213–226, (At: http://dx.doi.org/10.1007/BF01986276.).

Seibert, J., Stendahl, J. & Sørensen, R. 2007. Topographical influences on soil properties in boreal forests. *Geoderma*, **141**, 139–148.

Setiyono, T.D., Walters, D.T., Cassman, K.G., Witt, C. & Dobermann, A. 2010. Estimating maize nutrient uptake requirements. *Field Crops Research*, **118**, 158–168.

Shehu, B., Merckx, R., Jibrin, J., Kamara, A. & Rurinda, J. 2018. Quantifying Variability in Maize Yield Response to Nutrient Applications in the Northern Nigerian Savanna. *Agronomy*.

Shepherd, K.D. 2010. Soil spectral diagnostics–infrared, x-ray and laser diffraction spectroscopy for rapid soil characterization in the Africa Soil Information Service. In: *19th World Congress of Soil Science*.

Shepherd, K.D., Shepherd, G. & Walsh, M.G. 2015. Land health surveillance and response: A framework for evidence-informed land management. *Agricultural Systems*, **132**, 93–106, (At: http://www.sciencedirect.com/science/article/pii/S0308521X14001206. Accessed: 15/6/2015).

Shepherd, K.D. & Walsh, M.G. 2002. Development of Reflectance Spectral Libraries for Characterization of Soil Properties. *Soil Science Society of America Journal*, **66**, 988, (At: https://www.soils.org/publications/sssaj/abstracts/66/3/988.).

Shepherd, K.D. & Walsh, M.G. 2007a. Review: Infrared spectroscopy - Enabling an evidence-based diagnostic surveillance approach to agricultural and environmental management in developing countries. *Journal of Near Infrared Spectroscopy*.

Shepherd, K.D. & Walsh, M.G. 2007b. Infrared spectroscopy-enabling an evidence-based diagnostic surveillance approach to agricultural and environmental management in developing countries. Journal of Near Infrared Spectroscopy. *Journal of Near Infrared Spectroscopy*, **1**, 1–20.

Sila, A., Hengl, T. & Terhoeven-Urselmans, T. 2014. soil. spec: Soil Spectroscopy Tools and Reference models. *R package version*, **2**.

Sila, A., Pokhariyal, G. & Shepherd, K. 2017. Evaluating regression-kriging for mid-infrared spectroscopy prediction of soil properties in western Kenya. *Geoderma Regional*.

Sileshi, G., Akinnifesi, F.K., Ajayi, O.C. & Place, F. 2008. Meta-analysis of maize yield response to woody and herbaceous legumes in sub-Saharan Africa. *Plant and Soil*, **307**, 1–19, (At: http://www.springerlink.com/index/10.1007/s11104-008-9547-y. Accessed: 19/7/2012).

Smaling, E.M.A., Nandwa, S.M., Prestele, H., Roetter, R. & Muchena, F.N. 1992. Yield response of maize to fertilizers and manure under different agro-ecological conditions in Kenya. *Agriculture, Ecosystems & Environment*, **41**, 241–252, (At: http://www.sciencedirect.com/science/article/pii/016788099290113P.).

Smaling, E.M.A., Stoorvogel, J.J., Windmeijer, P.N., Smaling, E.M.A. & Janssen, B.H. 1993. Calculating soil nutrient balances in Africa at different scales. *Fertilizer Research*, **35**, 237–250, (At: http://dx.doi.org/10.1007/BF00750641.).

Smaling, E.M.A. & Van De Weg, R.F. 1990. Using soil and climate maps and associated data sets to select sites for fertilizer trials in Kenya. *Agriculture, Ecosystems & Environment*, **31**, 263–274, (At: http://www.sciencedirect.com/science/article/pii/0167880990902253.).

Smith, D., Smith, I., Collett, N. & Elms, S. 2008. Forest health surveillance in Victoria. *Australian Forestry*.

Snoeck, D., Afrifa, A.A., Ofori-Frimpong, K., Boateng, E. & Abekoe, M.K. 2010. Mapping fertilizer recommendations for cocoa Production in Ghana using soil siagnostic and GIS tools. *West African Journal of Applied Ecology*, **17**, 97–107, (At: http://www.scopus.com/inward/record.url?eid=2-s2.0-80855131338&partnerID=40&md5=4c9f99866a244577170612c796f0df87.).

Stenberg, B., Rossel, R. a V, Mouazen, a M. & Wetterlind, J. 2010. Visible and Near Infrared Spectroscopy in Soil Science. *Advances in Agronomy, Vol 107*, **107**, 163–215.

Stockdale, E.A., Goulding, K.W.T., George, T.S. & Murphy, D. V. 2013. Soil fertility. In: *Soil Conditions and Plant Growth*, pp. 49–85.

Stoorvogel, J.J. 2014. S-world: A global map of soil properties for modeling. *In: Proceedings of the 1st GlobalSoilMap Conference on GlobalSoilMap: Basis of the global spatial soil information system, Orleans, France, 7 - 9 October 2013. - London, UK : Taylor & Francis Group, 2014 - ISBN 9781138001190 - p. 227 - 231*.

Stoorvogel, J.J. & Smaling, E.M.A. 1998. Research on soil fertility decline in tropical environments : integration of spatial scales. *Nutrient Cycling in Agroecosystems*, **50**, 151–158.

Stoorvogel, J.J., Smaling, E.M.A. & Janssen, B.H. 1993. CALCULATING SOIL NUTRIENT BALANCES IN AFRICA AT DIFFERENT SCALES .1. SUPRA-NATIONAL SCALE. *Fert Res*, **35**, 227–235.

Sun, W., Minasny, B. & McBratney, A. 2012. Analysis and prediction of soil properties using local regression-kriging. *Geoderma*, **171–172**, 16–23.

Tesfahunegn, G.B., Tamene, L. & Vlek, P.L.G. 2011. Catchment-scale spatial variability of soil properties and implications on site-specific soil management in northern Ethiopia. *Soil and Tillage Research*, **117**, 124–139.

Thornton, a & Lee, P. 2000. Publication bias in meta-analysis: its causes and consequences. *Journal of clinical epidemiology*, **53**, 207–16, (At: http://www.ncbi.nlm.nih.gov/pubmed/10729693.).

Tilman, D., Cassman, K.G., Matson, P.A., Naylor, R. & Polasky, S. 2002. Agricultural sustainability and intensive production practices. *Nature*, **418**, 671–7, (At: http://dx.doi.org/10.1038/nature01014.).

Tittonell, P., van Dis, R., Vanlauwe, B. & Shepherd, K.D. 2015. Managing Soil Heterogeneity in Smallholder African Landscapes Requires Form of Precision Agriculture.

Tittonell, P. & Giller, K.E. 2013. When yield gaps are poverty traps: The paradigm of ecological intensification in African smallholder agriculture. *Field Crops Research*, **143**, 76–90.

Tittonell, P., Muriuki, A., Klapwijk, C.J., Shepherd, K.D., Coe, R. & Vanlauwe, B. 2013. Soil Heterogeneity and Soil Fertility Gradients in Smallholder Farms of the East African Highlands. *Soil Science Society of America Journal*, **77**, 525–538, (At: http://www.scopus.com/inward/record.url?eid=2-s2.0-84877279021&partnerID=40&md5=de27a2ace1d1e915bb253ea6bf4f2dd0.).

Tittonell, P., Shepherd, K.D., Vanlauwe, B. & Giller, K.E. 2008a. 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. *Agriculture Ecosystems & Environment*, **123**, 137–150.

Tittonell, P., Vanlauwe, B., Corbeels, M. & Giller, K.E. 2008b. Yield gaps, nutrient use efficiencies and response to fertilisers by maize across heterogeneous smallholder farms of western Kenya. *Plant and Soil*, **313**, 19–37.

Tittonell, P., Vanlauwe, B., Leffelaar, P.A., Rowe, E.C. & Giller, K.E. 2005. Exploring diversity in soil fertility management of smallholder farms in western Kenya: I. Heterogeneity at region and farm scale. *Agriculture, Ecosystems and Environment*, **110**, 149–165.

Tittonell, P., Vanlauwe, B., de Ridder, N. & Giller, K.E. 2007a. Heterogeneity of crop productivity and resource use efficiency within smallholder Kenyan farms: Soil fertility gradients or management intensity gradients? *Agricultural Systems*, **94**, 376–390.

Tittonell, P., Zingore, S., van Wijk, M.T., Corbeels, M. & Giller, K.E. 2007b. Nutrient use efficiencies and crop responses to N, P and manure applications in Zimbabwean soils: Exploring management strategies across soil fertility gradients. *Field Crops Research*, **100**, 348–368, (At: http://www.sciencedirect.com/science/article/pii/S0378429006001857.).

Tscharntke, T., Clough, Y., Wanger, T.C., Jackson, L., Motzke, I., Perfecto, I., Vandermeer, J. & Whitbread, A. 2012. Global food security, biodiversity conservation and the future of agricultural intensification. *Biological Conservation*.

Tully, K.L., Wood, S.A., Almaraz, M., Neill, C. & Palm, C. 2015. The effect of mineral and organic nutrient input on yields and nitrogen balances in western Kenya. *Agriculture, Ecosystems and Environment*, **214**, 10–20.

UN. 2013. World population projected to reach 9.6 billion by 2050. *World Population Prospects: The 2012 Revision, Press Release (13 June 2013): 'World Population to reach 9.6 billion by 2050 with most growth in developing regions, especially Africa'*, 4, (At: http://esa.un.org/unpd/wpp/Documentation/publications.htm.).

United Nations Population Division. 2009. *2009 World Population to Exceed 9 Billion by 2050: Developing Countries to Add 2.3 Billion Inhabitants with 1.1 Billion Aged Over 60 and 1.2 Billion of Working Age*. (At: http://www.un.org/esa/population/publications/wpp2008/pressrelease.pdf.).

Vågen, T.-G., Davey, F.A. & Shepherd, K.D. 2012. Land Health Surveillance: Mapping Soil Carbon in Kenyan Rangelands. In: *Agroforestry - The Future of Global Land Use* (eds. Nair, P.K.R. & Garrity, D.), pp. 455–462. Springer Netherlands, Dordrecht.

Vågen, T.-G., Winowiecki, L., Tamene, T. & Tondoh, J.E. 2010. The Land Degradation Surveillance Framework (LDSF): Field Guide. World Agroforestry Centre (ICRAF).

Valkama, E., Uusitalo, R. & Turtola, E. 2011. Yield response models to phosphorus application: a research synthesis of Finnish field trials to optimize fertilizer P use of cereals. *Nutrient Cycling in Agroecosystems*, **91**, 1–15, (At: http://dx.doi.org/10.1007/s10705-011-9434-4.).

Valkama, E., Uusitalo, R., Ylivainio, K., Virkajärvi, P. & Turtola, E. 2009. Phosphorus fertilization: A meta-analysis of 80 years of research in Finland. *Agriculture, Ecosystems & Environment*, **130**, 75–85, (At: http://www.sciencedirect.com/science/article/pii/S016788090800340X.).

van Apeldoorn, D.F., Kempen, B., Bartholomeus, H.M., Rusinamhodzi, L., Zingore, S., Sonneveld, M.P.W., Kok, K. & Giller, K.E. 2014. Analysing soil organic C gradients in a smallholder farming village of East Zimbabwe. *Geoderma Regional*, **2–3**, 32–40, (At: http://www.sciencedirect.com/science/article/pii/S2352009414000145. Accessed: 12/1/2015).

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*.

Vanlauwe, B., AbdelGadir, A.H., Adewopo, J., Adjei-Nsiah, S., Ampadu-Boakye, T., Asare, R., Baijukya, F., Baars, E., Bekunda, M., Coyne, D., Dianda, M., Dontsop-Nguezet, P.M., Ebanyat, P., Hauser, S., Huising, J., Jalloh, A., Jassogne, L., Kamai, N., Kamara, A., Kanampiu, F., Kehbila, A., Kintche, K., Kreye, C., Larbi, A., Masso, C., Matungulu, P., Mohammed, I., Nabahungu, L., Nielsen, F., Nziguheba, G., Pypers, P., Roobroeck, D., Schut, M., Taulya, G., Thuita, M., Uzokwe, V.N.E., van Asten, P., Wairegi, L., Yemefack, M. & Mutsaers, H.J.W. 2017. Looking back and moving forward: 50 years of soil and soil fertility management research in sub-Saharan Africa. *International Journal of Agricultural Sustainability*.

Vanlauwe, B., Bationo, A., Chianu, J., Giller, K.E., Merckx, R., Mokwunye, U., Ohiokpehai, O., Pypers, P., Tabo, R., Shepherd, K.D., Smaling, E.M.A., Woomer, P.L. & Sanginga, N. 2010. Integrated soil fertility management: Operational definition and consequences for implementation and dissemination. *Outlook on Agriculture*, **39**, 17–24.

VANLAUWE, B., COE, R. & GILLER, K.E. 2016. BEYOND AVERAGES: NEW APPROACHES TO UNDERSTAND HETEROGENEITY AND RISK OF TECHNOLOGY SUCCESS OR FAILURE IN SMALLHOLDER FARMING. *Experimental Agriculture*, 1–23.

Vanlauwe, B., Descheemaeker, K., Giller, K.E., Huising, J., Merckx, R., Nziguheba, G., Wendt, J. & Zingore, S. 2014. Integrated soil fertility management in sub-Saharan Africa: unravelling local adaptation. *SOIL Discussions*, **1**, 1239–1286, (At: http://www.soil-discuss.net/1/1239/2014/.).

Vanlauwe, B. & Giller, K.E. 2006. Popular myths around soil fertility management in sub-Saharan Africa. *Agriculture, Ecosystems and Environment*, **116**, 34–46.

Vanlauwe, B., Kihara, J., Chivenge, P., Pypers, P., Coe, R. & Six, J. 2011. Agronomic use efficiency of N fertilizer in maize-based systems in sub-Saharan Africa within the context of integrated soil fertility management. *Plant and Soil*, **339**, 35–50.

Vanlauwe, B., Tittonell, P. & Mukalama, J. 2007. Within-farm soil fertility gradients affect response of maize to fertiliser application in western Kenya (A Bationo, B Waswa, J Kihara, and J Kimetu, Eds.). *Advances in Integrated Soil Fertility Management in sub-Saharan Africa: Challenges and Opportunities*, **76**, 171–182, (At: http://dx.doi.org/10.1007/978-1-4020-5760-1_10.).

Vašát, R., Heuvelink, G.B.M. & Borůvka, L. 2010. Sampling design optimization for multivariate soil mapping. *Geoderma*.

Viechtbauer, W. 2010. Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software*, **36**, 1–48, (At: http://www.jstatsoft.org/v36/i0.).

Vink, G., Frank, L.E., Pannekoek, J. & van Buuren, S. 2014. Predictive mean matching imputation of semicontinuous variables. *Statistica Neerlandica*, **68**, 61–90.

Viscarra Rossel, R.A., Brus, D.J., Lobsey, C., Shi, Z. & McLachlan, G. 2016. Baseline estimates of soil organic carbon by proximal sensing: Comparing design-based, model-assisted and model-based inference. *Geoderma*.

Waithaka, M., Thornton, P., Shepherd, K. & Ndiwa, N. 2007. Factors affecting the use of fertilizers and manure by smallholders: the case of Vihiga, western Kenya. *Nutrient Cycling in Agroecosystems*, **78**, 211–224, (At: http://dx.doi.org/10.1007/s10705-006-9087-x.).

Walworth, J.L., Letzsch, W.S. & Sumner, M.E. 1986. Use of Boundary Lines in Establishing Diagnostic Norms. *Soil Science Society of America Journal*, **50**, NP.

Walworth, J.L. & Sumner, M.E. 1987. The Diagnosis and Recommendation Integrated System (DRIS). In: *Advances in Soil Science SE - 4* (ed. Stewart, B.A.), pp. 149–188. Springer New York.

Wang, E., Bell, M., Luo, Z., Moody, P. & Probert, M.E. 2014. Modelling crop response to phosphorus inputs and phosphorus use efficiency in a crop rotation. *Field Crops Research*, **155**.

Wang, H., Snapp, S.S., Fisher, M. & Viens, F. 2019. A Bayesian analysis of longitudinal farm surveys in Central Malawi reveals yield determinants and site-specific management strategies. *PLOS ONE*, **14**, e0219296, (At: https://doi.org/10.1371/journal.pone.0219296.).

Waswa, B.S., Vlek, P.L.G., Tamene, L.D., Okoth, P., Mbakaya, D. & Zingore, S. 2013. Evaluating indicators of land degradation in smallholder farming systems of western Kenya. *Geoderma*, **195–196**, 192–200, (At: http://www.sciencedirect.com/science/article/pii/S0016706112004028.).

Waters, E. & Doyle, J. 2003. Evidence-based public health: Cochrane update. *J Public Health Med*, **25**, 72–5.

Webb, M.J., Nelson, P.N., Rogers, L.G. & Curry, G.N. 2011. Site-specific fertilizer recommendations for oil palm smallholders using information from large plantations. *Journal of Plant Nutrition and Soil Science*, **174**, 311–320.

Wheeler, I., McBratney, A., Minasny, B. & de Gruijter, J. 2012. Digital Soil Mapping to inform design-based sampling strategies for estimating total organic carbon stocks at the farm scale. In: *Digital Soil Assessments and Beyond*.

Wickham, H. 2009. ggplot2: elegant graphics for data analysis.

de Winter, J.C.F. & Dodou, D. 2016. Common Factor Analysis versus Principal Component Analysis: A Comparison of Loadings by Means of Simulations. *Communications in Statistics - Simulation and Computation*, **45**, 299–321, (At: https://doi.org/10.1080/03610918.2013.862274.).

Wopereis, M.C.S., TamÃ©lokpo, A., Ezui, K., GnakpÃ©nou, D., Fofana, B. & Breman, H. 2006. Mineral fertilizer management of maize on farmer fields differing in organic inputs in the West African savanna. *Field Crops Research*, **96**, 355–362, (At: http://www.scopus.com/inward/record.url?eid=2-s2.0-32644465601&partnerID=40&md5=e28e6d618638d256c568a13b3bf65c75.).

Wortmann, C.S., Milner, M., Kaizzi, K.C., Nouri, M., Cyamweshi, A.R., Dicko, M.K., Kibunja, C.N., Macharia, M., Maria, R., Nalivata, P.C., Demissie, N., Nkonde, D., Ouattara, K., Senkoro, C.J., Tarfa, B.D. & Tetteh, F.M. 2017. Maize-nutrient response information applied across Sub-Saharan Africa. *Nutrient Cycling in Agroecosystems*, **107**, 175–186.

Wuenscher, R., Unterfrauner, H., Peticzka, R. & Zehetner, F. 2015. A comparison of 14 soil phosphorus extraction methods applied to 50 agricultural soils from Central Europe. *PLANT SOIL AND ENVIRONMENT*.

Zarcinas, B., Cartwright, B. & Spouncer, L.R. 1987. Nitric acid digestion and multi-element analysis of plant material by inductively coupled plasma spectrometry. *Communications in Soil Science and Plant Analysis*, **18**, 131–146, (At: http://www.tandfonline.com/doi/abs/10.1080/00103628709367806.).

Zingore, S., Murwira, H.K., Delve, R.J. & Giller, K.E. 2007a. Soil type, management history and current resource allocation: Three dimensions regulating variability in crop productivity on African smallholder farms. *Field Crops Research*, **101**, 296–305.

Zingore, S., Murwira, H.K., Delve, R.J. & Giller, K.E. 2007b. Influence of nutrient management strategies on variability of soil fertility, crop yields and nutrient balances on smallholder farms in Zimbabwe. *Agriculture, Ecosystems and Environment*, **119**, 112–126.

Zingore, S., Murwira, H., Delve, R. & Giller, K. 2008. Variable grain legume yields, responses to phosphorus and rotational effects on maize across soil fertility gradients on African smallholder farms. *Nutrient Cycling in Agroecosystems*, **80**, 1–18, (At: http://dx.doi.org/10.1007/s10705-007-9117-3.).

Zingore, S., Njoroge, S., Chikowo, R., Kihara, J., Nziguheba, G. & Nyamangara, J. 2014. 4R plant nutrient management in African agriculture: an extension handbook for fertilizer management in smallholder farming systems.

Di Zio, M. & Guarnera, U. 2009. Semiparametric predictive mean matching. *AStA Advances in Statistical Analysis*, **93**, 175–186, (At: http://dx.doi.org/10.1007/s10182-008-0081-2.).

Appendix

Appendix 1: List of studies used develop the database for meta-analysis.

No.	**Reference article**	**Journal name**	**Publication year**	**Volume, (Issue),Page No.**
1	Kayuki and Wortman	Agronomy Journal	2001	93 , 929-935
2	Ayoola and Adeniyan	African Journal of Biotechnology	2006	5, 1336-1392
3	Jama and Kiwia	Experimental Agriculture	2009	45, 241-260
4	Phiri et al.	African Journal of Agricultural Research	2010	5, 1235-1242
5	Baijuka et al.	Plant and Soil	2006	279, 77-93
6	Ikerra et al.	Nutrient Cycling in Agroecosystems	2007	76 , 333-344
7	Opala et a.l	Nutrient Cycling in Agroecosystems	2010	86, 317-329
8	Mucheru-muna et al.	Agroforestry Systems	2007	69 , 189-197
9	Onyango et al.	Proceedings of the Seventh Eastern and Southern Africa Regional Maize Conference, 11–15 February.	2001	330-334
10	Mwangi	World Journal of Agricultural Sciences	2010	3, 313-321
11	Abunyewa et al.	Journal of Agronomy	2007	6, (2), 302-309
12	Olasantan et al.	Nutrient Cycling in Agroecosystems	1997	46, 215-223
13	Shisanya et al.	Soil & Tillage Research	2009	103, 239-246
14	Kaizzi et al.	Agricultural Systems	2006	88, 44-60
15	Jenssen et al.	Agricultural Water management	2003	59,(3),217-237

No.	Reference article	Journal name	Publication year	Volume, (Issue),Page No.
16	Gitari and Friesen	Proceedings of the Seventh Eastern and Southern Africa Regional Maize Conference, 11–15 February.	2001	3, (9), 234-238
17	Obaga et al.	KARI report	-	Unpublished
19	Fening et al.	African Journal of Environmental Science and Technology	2009	59,(3),217-237
20	Mucheru	Bationo et al (eds) Managing Nutrient Cycles to Sustain soil fertility in SSA	2002	1,(4), 592003
21	Mukuralinda et al.	Agroforestry Systems	2009	80(211-221)
22	Mtambanengwe et al.	Nutrient Cycling in Agroecosystems	2006	76,271-284
23	Achieng et al.	Agriculture and Biology Journal of North America	2010	3, 234-238
24	Kimani et al.	Bationo et al (eds) Advances in Soil Fertility Management	2007	15,111-126
25	Smaling et al.	Agriculture, Ecosystems and Environment	1992	80,211-221
26	Macharia et al.	Journal of Animal & Plant Sciences	2005	76,271-284
27	Sakala et al.	Bationo et al (eds) Managing Nutrient Cycles to Sustain soil fertility in SSA	2004	1(4), 430-439
28	Okalebo et al .	Bationo et al (eds) Managing Nutrient Cycles to Sustain soil fertility in SSA	2004	1(4), 360-372
29	Nyongesa et al .	African Crop Science Conference Proceedings	2009	41, 241-252
30	Onyango et al.	KARI report	-	Unpublished
31	Macharia et al.	KARI report	-	Unpublished
32	Mathuva et al.	Field Crop Research	1998	55, 57-72
33	Jeranyama et al.	Agronomy Journal	2000	92,239-244

No.	Reference article	Journal name	Publication year	Volume, (Issue),Page No.

34	Kimetu et al.	Nutrient Cycling in Agroecosystems	2004	68,127-135
35	Amusan et al.	Nutrient Cycling in Agroecosystems	2011	90, (3), 321-330
36	Ssali	Fertilizer Research	1990	23, (2), 63-72
37	Akinnifesi et al.	Plant and Soil	2007	294 (2), 203-217
38	Kang et al .	Fertilizer Research	1980	1, (2), 87-93
39	Kaizzi et al.	Nutrient Cycling in Agroecosystems	2006	88, 44-60
40	Sangiga et al .	Plant and Soil	1996	179,119-129
40	Sangiga et al .	Biological Agriculture and Horticulture	1986	3, 347–352
41	Osiname et al.	Nutrient Cycling in Agroecosystems	2000	56,209-217
42	Gacheru and Rao	International Journal of Pest Management	2001	47,(3), 233-239
43	Casky et al.	Nutrient Cycling in Agroecosystems	2002	243,1-10
44	Nziguheba et al.	Plant and Soil	2001	198,159-168
45	Nyamangara and Nyagumbo	Nutrient Cycling in Agroecosystems	2010	88, 103-109
46	Adjei-Nsaih	Field Crop Research	2007	103, 87-9)
47	Anyanzwa et al .	Nutrient Cycling in Agroecosystems	2010	88, 39-47
48	Fofana et al.	Nutrient Cycling in Agroecosystems	2004	68, 213-222
50	Esilaba et al.	Agricultural Systems	2005	86,144-165
51	Kihara et al.	Nutrient Cycling in Agroecosystems	2011	90, 213–225

No.	**Reference article**	**Journal name**	**Publication year**	**Volume, (Issue),Page No.**
52	Usiri et al.	Communication in Soil Science and Plant Analysis	1998	27, (17&18), 2815-2828
53	Mtambanengwe and Mapfumo	Plant and Soil	2006	281, 173–191
54	Sigunga et al.	Nutrient Cycling in Agroecosystems	2002	62, (3), 263-275

55	Mureithi et al.	Agroforestry Systems	1994	27, (1), 31-51
56	Nguu	Fertilizer Research	1987	14, (2), 135-142
57	Titonell et al.	Plant and Soil	2008	313, (1), 19-37
58	Msolla et al.	Nutrient Cycling in Agroecosystems	2005	72, (3), 299-308
59	Ayuke et al.	Bationo et al (eds) Managing Nutrient Cycles to Sustain soil fertility in SSA	2004	1,(4), 65-77
60	Nabuhungu et al.	Bationo et al (eds) Advances in soil fertility management	2011	1, 325-335
61	Nekesa et al.	Bationo et al (eds) Innovations as key to green revolution	2011	2, 335-342
62	Ssila et al.	Nutrient Cycling in Agroecosystems	1990	23, (2), 63-72
63	Vanlauwe et al.	Agronomy Journal	2001	93,1191-1199
64	Saidou	Agriculture, Ecosystems and Environment	2003	100, (3),265-273
65	Otinga et al.	Field Crop Research	2013	140, 32-43
66	Githinji et al .	Bationo et al (eds) Innovation as key to Green Revolution in Africa	2011	2, 281-288
67	Kathuku et al.	Bationo et al (eds) Innovation as key to Green Revolution in Africa	2011	2, 265-270
68	Kaizzi et al.	Agronomy Journal	2012	104, 73–82
69	Kamanga et al.	Experimental Agriculture	2014	50, (2), 229-249
70	Rusinamhodzi et al.	Field Crop Research	2012	136, 12-22

Appendix 2: Papers used in analysis, showing the country where the experiment was conducted and nitrogen application rates

Reference	Country	N application rate (kg ha^{-1})

Kimani et al. (2007)	Kenya	20,40, 60, 80, 100
Kathuku et al. (2011)	Kenya	20, 80
Gitari and Friesen (2001), Smaling et al. (1992)	Kenya	25, 50
Szali (1990)	Kenya	25,50, 100, 150
Mwangi (2010)	Kenya	30, 40, 50, 60
Achieng et al (2010)	Kenya	30
Okalebo et al (2004), Onyango et al. (Unpublished)	Kenya	30, 60
Mathuva et al. (1998)	Kenya	40
Onyango et al. (2001), Shisanya et al. (2009), Obaga et al .(Unpublished), Mucheru (2002), Achieng et al. (2010), Macharia et al. (2005), Kimetu et al. (2004), Anyanzwa et al. (2010), Kihara et al. (2011), Githinji et al. (2011), Kathuku et al (2011)	Kenya	60
Macharia et al. (Unpublished)	Kenya	60, 120
Mureithi et al. (1994)	Kenya	75, 100
Nekesa et al. (2011)	Kenya	75
Sigunga et al. (2002), Titonell et al. (2000), Ngome et al. (2011)	Kenya	100
Ayuke et al. (2003)	Kenya	120
Achieng et al. (2010)	Kenya	144
Saidou (2003)	Benin	60
Nguu (1987)	Cameroon	30, 60, 120
Abunyewa et al. (2007), Adjei-Nsaih (2007)	Ghana	60
Fening et al. (2009)	Ghana	90
Sakala et al. (2004)	Malawi	35, 69

Reference	**Country**	**N application rate (kg ha^{-1})**
Casky et al. (2002)	Nigeria	40

Amusan et al. (2011)	Nigeria	50, 100
Ayoola & Adeniyan (2006), Olasantan et al. (1997)	Nigeria	60
Kang et al. (1980)	Nigeria	80
Vanlauwe et al. (2001)	Nigeria	90
Nabuhungu et al. (2007)	Rwanda	50, 175
Baijuka et al. (2006)	Tanzania	50
Usiri et al. (1991)	Tanzania	60
Jensen et al. (2003)	Tanzania	140
Fofana et al. (2004)	Togo	20, 40, 50, 100
Kaizzi et al. (2004)	Uganda	40,80
Kaizzi et al. (2012)	Uganda	50, 80
Kayuki & Wortman (2001), Esilaba et al .(2005),	Uganda	80
Jeranyama et al. (2000)	Zimbabwe	60, 120
Nezomba et al. (2010)	Zimbabwe	90
Nyamangara and Nyagumbo (2010)	Zimbabwe	100
Mtambanengwe et al. (2006), Mtambanengwe & Mapfumo (2006), Kurwakumire et al. (2014)	Zimbabwe	120