Zentrum für Entwicklungsforschung (ZEF)

Options For Sustainable Agricultural Intensification In Maize Mixed Farming Systems:

Explorative Ex-Ante Assessment Using Multi-Agent System Simulation

Inaugural Dissertation

zur Erlangung des Grades
Doktor der Agrarwissenschaften (Dr. agr.)

der Landwirtschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

von

Powell Mponela

Aus Karonga, Malawi

Bonn, 2021

Referent : *Prof. Dr. Christian Borgemeister*
Korreferentin : *Prof. Dr. Siegelinde Snapp*

Tag der mündlichen Prüfung: 14-12-2020

Angefertigt mit Genehmigung der Landwirtschaftlichen Fakultät der Universität Bonn

ABSTRACT

Nutrient depletion is a major limiting factor to agricultural sustainability in cereal dominated smallholder farming systems in Africa where over 80% of arable land is unsuitable to support primary productivity. This constrains food and nutritional security of rural communities. For appropriate design of interventions, there is need for empirical evidence on drivers of change.

A common sampling frame is used to integrate social-ecological data from farm surveys of soil, biomass and crop yield, nutrient inputs and outputs, and their determinants. The nutrient distributions are predicted using randomForest machine learning algorithm in R with remotely sensed reflectance for topography (30 m STRM-DEM), vegetation and soil (10 m Sentinel2 imagery) as co-variates. We use behavioural economics to unravel farm-type specific drivers of human induced nutrient inputs and a mixed model for crop yield function for outputs. Further, existing nutrient stoichiometry and transfer functions based on NUTMON, FarmDESIGN models with parameters from the study region are used to capture dynamic stocks and flows. Lastly, we build a multi-agent system for simulating sustainable agricultural intensification (MASSAI) in NetLogo and piloted to explore, *ex ante*, the agentic behaviours of farmers when faced with ambiguity in fertilizer subsidy regimes and its implications on nutrient budgets, human decision making and land productivity.

Though soil management in smallholder farming systems aims at addressing the most critical nutrient(s), the results from this study show that the soils are deficient in all three major nutrients (NPK) and structurally unstable due to low soil organic carbon (SOC). Farmers strive to utilise the commonly available soil fertility management: nine in every ten households used inorganic fertilizers, a third integrated legumes and almost half applied manures of various forms. From the empirical and simulated results, it is indicative that the maize mixed smallholder farming system in Malawi has become inelastic to changes in input policies.

Much as improvement in contribution of women in decision-making widens the scope for legume cropping, it negatively affects manuring. Therefore, addressing challenges that women face in manuring could offer greater opportunities for integrated soil fertility management.

After 15 years of fertilizer subsidy program, farmers have internalized it in their expenditure plan: some exclusively relying on subsidy while others source increasing amounts from the market and are becoming self-reliant. Those that rely on limited fertilizer acquired through subsidy proactively reduce the nutrient gap by increasing manuring. These behaviors have implications on nutrient management and sustainability of the farming systems. Although subsidy alone might not significantly shift the nutrient and productivity trajectories for the next 20 simulated years, increased subsidy could relatively accelerate nitrogen and phosphorus losses.

ZUSAMMENFASSUNG

OPTIONEN FÜR EINE NACHHALTIGE LANDWIRTSCHAFTLICHE INTENSIVIERUNG IN GEMISCHTEN MAISANBAUSYSTEMEN: EXPLORATIVE EX-ANTE-BEWERTUNG MIT EINER MULTI-AGENTEN-SYSTEMSIMULATION

Nährstoffverarmung ist ein wesentlicher limitierender Faktor für die landwirtschaftliche Nachhaltigkeit in Getreide-dominierten kleinbäuerlichen Anbausystemen in Afrika, wo mehr als 80% des Ackerlandes für die Primärproduktion ungeeignet ist. Dies schränkt die Nahrungsmittel- und Ernährungssicherheit der ländlichen Gemeinden ein. Um dem gezielt durch Interventionen entgegenwirken zu können, sind empirische Studien über Einflussfaktoren, die zu Veränderung beitragen können, nötig.

Eine integrierte Stichprobenstrategie wird verwendet, um sozio-ökologische Daten aus Betriebserhebungen zu Boden, Biomasse und Ernteerträgen, Nährstoffein- und -austrägen und deren Bestimmungsfaktoren zu integrieren. Nährstoffverteilungen werden unter Verwendung des ‚randomForest- machine learning algorithm' in R, mit aus Fernerkundung stammenden Daten, für die Co-Variablen Topographie (30 m STRM-DEM), Vegetation und Boden (10 m Sentinel2-Bilder) simuliert. Verhaltensökonomische Ansätze geben Aufschluss über die betriebstypspezifischen Faktoren, die eine Rolle spielen, wenn es um gezielte Nährstoffeinträge durch den Menschen geht. Ein gemischtes Modell wird für die Ermittlung einer Ernteertragsfunktion verwendet, um die Erträge zu ermitteln. Darüber hinaus werden vorhandene Nährstoffstöchiometrie- und Transferfunktionen basierend auf den NUTMON und FarmDESIGN Modellen mit Parametern aus der Untersuchungsregion verwendet, um dynamische Nährstoffbestände und -Flüsse zu erfassen. Schließlich bauen wir in NetLogo ein Multi-Agenten-System zur Simulation einer nachhaltigen Intensivierung der Landwirtschaft (MASSAI) auf und untersuchen, *ex ante*, das Verhalten von Landwirten, die aufgrund eines Düngemittelsubventionssystems mit Unklarheiten konfrontiert werden und dessen Auswirkungen auf Entscheidungsprozesse, Nährstoffbudgets und Landproduktivität.

Wenngleich Bodenbewirtschaftung in kleinbäuerlichen Anbausystemen darauf abzielt, Defizite in den kritischsten Nährstoffen auszugleichen, zeigen die Ergebnisse dieser Studie, dass die Böden in allen drei Hauptnährstoffen (NPK) mangelhaft und aufgrund niedriger organischer Kohlenstoffgehalte im Boden strukturell instabil sind. Landwirte streben an, allgemein verfügbare Methoden für Bodenfruchtbarkeitsmanagement zu nutzen: neun von zehn Haushalten verwendeten anorganische Düngemittel, ein Drittel integrierte Leguminosen und fast die Hälfte brachten Mist verschiedenster Art aus. Die empirischen und simulierten Ergebnisse deuten darauf hin, dass das gemischte kleinbäuerliche Maisanbausystem in Malawi gegenüber Veränderungen in politischen Richtlinien, die den Einsatz von Betriebsmitteln betreffen, unelastisch geworden ist.

Die verbesserte Einbindung von Frauen in Entscheidungsprozesse macht zwar den Anbau von Leguminosen wahrscheinlicher, wirkt sich jedoch negativ auf den Einsatz von Mist aus. Daher könnte sich die Bewältigung von Herausforderungen, denen Frauen beim Düngen ausgesetzt sind, positiv auf ein integriertes Bodenfruchtbarkeitsmanagement auswirken.

Nach 15 Jahren Düngersubventionen haben die Landwirte diese fest in ihre Budgetplanung integriert, wobei sich einige ausschließlich auf Subventionen verlassen, während andere mit eigenen Einkäufen mischen und nicht-Subventionsempfänger zunehmend Dünger über den Markt beziehen. Diejenigen die auf begrenzten, durch Subventionen erworbenen Dünger angewiesen sind, verringern proaktiv die Nährstofflücke durch Mistgaben. Diese Verhaltensweisen haben Auswirkungen auf das Nährstoffmanagement und die Nachhaltigkeit der landwirtschaftlichen Systeme. Obwohl Subventionen allein die Nährstoff- und Produktivitätsverläufe für die simulierten nächsten 20 Jahre möglicherweise nicht wesentlich verändern, könnten erhöhte Subventionen die Stickstoff- und Phosphorverluste beschleunigen.

THESIS STRUCTURE

The thesis is written in 5 chapters. Chapter 1 highlights the need for addressing soil fertility as primer for sustainable agricultural development in rural landscapes. As the reality at the doorstep of smallholder farmers unravel, the complexity of the system challenges the research and development practitioners.

In chapter 2, we review existing literature on farming systems and nutrient modelling.

Chapter 3 is a methodological one. First, it establishes principles and methodological structures for integrated analysis of a farming system, using concepts and methods from both livelihood and ecological theories. The study site is described by highlighting the human and ecological states and drivers. Based on the empirical parameterisation, the analytical approaches for human actions and ecological processes are developed. Particularly, the transfer functions for nutrient input and output flows are updated using data from the study region. Lastly to integrate human and ecological processes, multi-agent simulation (MAS) is proposed and used to build feedback mechanisms for the nutrient budgets, human decision and land productivity.

Chapter 4 presents the results and their discussion. In the first subsection, farming households are grouped into distinct farm types using input levels and resource endowments. Second, the distribution of major nutrients (NPK) and SOC and their stoichiometric thresholds is presented (Manuscript 1, which is under review). Third, the household soil fertility management choices within the gender imbalance and high dependency are explored (Manuscript 2, also under review). The fourth subsection reports the plot level input choices which are essential and used to parametrise spatially explicit the probability of input choice and intensification levels. Fifth, the maize productivity for smallholder farms under varying subsidy intensities, the input levels, plot fertility and household resource endowments are explored. The sixth subsection presents the human induced and ecological nutrient flows and balances. Lastly, the MAS results are presented that focuses on the implications of alternative subsidy regimes on fertilisation, manuring, maize yield, nutrient balance and profitability as indicators of sustainability.

Chapter 5 is a concluding one with a subchapter on extension of current research and areas for further research.

ACKNOWLEDGEMENTS

First, I would like to thank Prof. Christian Borgemesiter for his enduring academic, moral and financial support. Prof. Siegelinde Snapp (Michigan State University) helped me to put the science into local perspective emanating from her extensive knowledge of soil research in smallholder farming systems of Malawi. Many thanks also to Dr Bao Quang Le (International Center for Agricultural Research in the Dry Areas (ICARDA)) who introduced me to the multi-agent simulation, shared his LUDAS code and took time to teach me how to implement it in NetLogo. I am also thankful to my tutor Dr Grace Villamor for her relentless backstopping of the model framework. Dr Lulseged Tamene Desta (International Center for Tropical Agriculture (CIAT)) encouraged me to apply for the PhD and supported my field study. When administrative arrangements and costs seemed out of reach, he extended his hand, devoted his time and offered financial and material support. Dr. Guido Lüchters, Dr. Chiara Cofol and Mr. Henry Kankwamba have been very instrumental in parameterisation, coding and refining the statistical models in Stata and Dr. Tomislav Hengl (OpenGeoHub/ EnvirometriX) helped me to refine the randomForest model for mapping the nutrient distributions in R. Several ZEF staff members contributed to my work through colloquiums, many thanks for your constructive criticisms.

This study has been possible because of the scholarship from the German Academic Exchange service (DAAD). Dr. Lulseged Desta and staff of the International Centre for Tropical Agriculture (CIAT) in Malawi and Kenya provided administrative, material and financial support during the field work. Dr. Wilson Makumba and Dr. Moses Munthali of the Department of Agricultural Research Services provided soil augering equipment, staff and vehicle support. Special thanks go to farmers for allowing me to study their farms. I also would like to extend my thanks to Frank Babka who laid the foundation by supporting me during my BSc, MSc and also cushioning me with further financial support during my studies.

I also extend my appreciation to Dr. Guenther Manske, Maike Retat-Amin and the entire ZEF doctoral program, Parisa Asemi-Salot and Birgitt Skailes of DAAD, and Sabine Aengernndt-Baer of the ZEFc Secretariate for administrative support during my studies. I also thank my ZEF friends Shova, Nani, Henry, Bisi, Evelyn and Doro and office mates Dela and Stephane for their moral support. Although, I acquired knowledge and grew a little older, two things have not changed: am still a husband and a father. Many thanks to my wife Silimyake and children Powell, Lusako and Mtendeko for understanding the essence of my absence.

DEDICATION

You believed in me. I always miss your love mama Nyamjimila and your care dada Mgongo.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS AND ABBREVIATIONS

AfSIS	Africa Soil Information Service
ASWAp	Agricultural Sector Wide Approach program
BI	Brightness index
BNF	Biological nitrogen fixation
CEC	Cation exchange capacity
CI	Coloration index
CIAT	International Centre for Tropical Agriculture
CLHS	Conditioned latin hypercube sampling
DEM	digital elevation model
DSS	Decision support system
ES	Ecosystem services
FGD	Focus group discussion
FISP	Farm input subsidy program
FYR	Farmer reported yields
GIS	Geographic information system
GLM	Generalised linear model
GoM	Government of Malawi
GPS	Global positioning system
GRVI	Green-Red vegetation index
GSI	Grain Size Index
HES	Human-Environmental System
ISFM	Integrated soil fertility management
K-CA	K-mean cluster analysis
LLSF	Liebig's linear scoring function
LSD	Least significant differences
LU	Livestock units
LUDAS	Land Use Dynamics Simulator
MAS	Multi-agent system
MASSAI	Multi-Agent System for Sustainable Agricultural Intensification
MIR	Mid infrared
MODIS	Moderate resolution imaging spectroradiometer
MSAVI2	Modified Secondary Soil-Adjusted Vegetation Index
NDVI	Normalised Difference vegetation index
NIR	Near infrared
NNIR	Narrow NIR
NPK	Nitrogen, Phosphorus, Potassium
ODD + D	Overview, design concepts, and details + decision
OK	Ordinary kriging
OOB	Out of the bag error
PCA	Principal components analysis
Ri	Roughness index
RI	Redness index
RUSLE	Revised universal soil loss equation
SAI	Sustainable agricultural intensification
SDR	Sediment delivery ratio
SFM	Soil fertility management
SI	Saturation index
SLF	Sustainable Livelihood Framework
SLM	Sustainable land management
SOC	Soil organic carbon
SOM	Soil organic matter
SPI	Stream power index
SSE	Sum of the squared error
StI	Structural stability index
STRM	Shuttle Radar Topography Mission
USGS	United States Geological Survey
WEAI	Women empowerment in agriculture index

1 SUSTAINABLE AGRICULTURAL INTENSIFICATION OF SMALLHOLDER FARMING SYSTEMS

1.1 Sustainable Agricultural Intensification premised on soil fertility management

In Africa, sustaining farming systems has been the major aim for researchers, practitioners and land managers for more than five decades (Vanlauwe et al., 2017). Although water is still one of the major constraints due to low level of irrigation, soil fertility decline is the major biophysical limitation that needs to be addressed if sustainability of the farming systems is to be attained (Vanlauwe et al., 2015). Of all the systems, the maize mixed farming system that spans across the east and southern Africa (covering Ethiopia, Kenya, Uganda, Tanzania, Zambia, Malawi, Mozambique, Zimbabwe and South Africa) is greatly affected (Dixon et al., 2001). According to the global change diagnosis by Petschel-Held et al. (1999), the region is characterised by overuse of marginal areas, overexploitation of natural ecosystems, abandonment of unproductive land, degradation due to in-appropriate farming methods. Expanding agricultural land followed by continuous cropping with minimal inputs are the most critical forces altering the environment. In case of Malawi, estimates show that the associated land degradation increases the probability of households to become poor by 30% (Kirui, 2016). Consequently, the inherently low soil fertile landscapes, such as the marginal escarpments of the country where the rural poor live and cultivate, reinforces the chronic poverty (Barrett & Bevis, 2015; Jean et al., 2016). Yet, for most rural populations in Malawi, agricultural productivity remains the determinant of welfare and is one of the key strategies for livelihood improvement in the short to medium-term (Sachs et al., 2004).

If the current trends continue unabated, in the long-term, arresting land degradation in sub-Saharan Africa (SSA) through sustainable land management technologies could cost around US$ 3 billion which is far lower than the annual cumulative losses in total economic value of ecosystem services estimated at US$ 15 billion (Nkonya et al., 2016). To avert environmental degradation and support economic development, the concept of Sustainable Agricultural Intensification (SAI) has been mainstreamed in research and development programs (Pretty, 1997; United Nations, 1987, 2000, 2015b). Although sustainability goals make political sense at global and national levels, they are more relevant to most of the rural farmers who depend on environmental resources (United Nations, 2015a). To them, knowing the impact of their interventions on land productivity and ecological sustainability underpins their livelihoods.

Reports show both negative and positive historical trends. The 2016 Food and Agriculture Organisation (FAO) report on forests and agriculture showed that the rate of deforestation was reduced due to sustainable land management (SLM) (FAO, 2016b), but high levels of land degradation have been estimated (Holt-Giménez et al., 2006; Le et al., 2016; van Ittersum et al., 2013; Vlek et al., 2008) affecting agricultural productivity in many parts of Africa (FAO, 2016c). The 2019 IPCC report shows that SLM practices increase soil organic carbon (SOC) and reduce erosion and can potentially help communities to cope with climate change (Arneth et al., 2019).

However, since the inception of SAI in 1997, a decade of no substantial interest from the scientific community past but the concept has reported gained momentum between 2011 to 2016 (Weltin et al., 2018). The increased interest in SAI has been attributed to the need to intensify production while maintaining the resource base. SAI is defined as: "*increased production, income, nutrition and other returns on the same amount of or less, land and water with efficient and prudent use of inputs, minimising greenhouse gas emissions while increasing natural capital and flow of environmental services, strengthening resilience and reducing environmental impact through innovative technologies and processes*" (The Montpellier Panel, 2013). From this conceptual and operational definition, it is apparent that a broad range of farming operations and farming styles fall under the umbrella of SAI.

Notably, SAI has many facets and the interactions among targets and indicators is overwhelmingly complex to be holistically implemented (Smith et al., 2015). To increase production, incomes, nutrition and other returns, there are several factors of which soil and climate are overarching (Bouma et al., 2019). In east and southern Africa, increasing production is currently the main development and research focus because, despite increases in total food production, the net average per-capita agricultural production has declined by a quarter in reference to 1960 levels (Pretty et al., 2011). Although food sovereignty is not a prerequisite for national food sufficiency, the current food demand in Africa can only be met by substantial imports. Without other economic opportunities, in the short to medium term, small-scale farmers ought to find ways of increasing their own production (FAO, 2015; van Ittersum et al., 2016). At their doorsteps, managing the soils - that exist right beneath their feet and for which they have higher control- precedes all other options (Vanlauwe et al., 2015). In Malawi, cereal production can increase by 14.6 and 9.6% following improvements in nitrogen use efficiency and sustainable land management, respectively (FAO, 2016c; Rosegrant et al., 2014), with the highest potential to reduce poverty (Sachs et al., 2004).

Moreover, among natural resources, soil is one non-renewable resource that significantly house and contribute to ecosystem services and human wellbeing (Bouma et al., 2019). However, the IPCC report shows that as of 2018, almost 30% of the world's soils were degraded and around 12 million hectares of land are still lost each year (Arneth et al., 2019). The report further shows that in addition to harming the wellbeing of at least 3.2 billion people, land degradation costs more than 10% of annual global gross domestic product (GDP) in lost ecosystem services like preventing harmful nutrient run-off into streams or decreasing the effects of floods. Halting and reversing current trends of land degradation could generate up to USD 1.4 trillion per year of economic benefits and go a long way in helping to achieving the Sustainable Development Goals (SDGs).

1.2 Relevant soil and land management interventions

In Africa, the intricate nexus of meeting the polarised goals of increased food and biomass production, generating economic opportunities and sustaining land health in the region has resulted in an influx of projects supported by national governments, non-governmental organisations (NGOs) and some self-started by the communities. Quite often, the scalability and sustainability of the externally driven projects by governments and NGOs have been variable and non-defined (Snapp et al., 2010). In Malawi, due to downstream problems associated with land degradation, incentive mechanisms for rewarding individuals who sustainably manage their farmlands through locally supported payment for ecosystem services (ES) are being established within the Shire River Basin (Fleskens & Chilima, 2013). In this regard, cross-sectoral policies that benefit downstream ES, including hydro power generation, urban water supply and irrigation and upstream activities on farmlands and (de)forested hill slopes, need to be developed (GoM, 2013). Revegetation, avoiding deforestation of steep slopes coupled with soil and water conservation practices on farmland are some of the measures being promoted to arrest soil erosion and improve land productivity (GoM, 2017).

This study aims to explore policy options for promoting SLM and results are of regional relevance to the maize-mixed farming system of east and southern Africa (Figure 3.6). Of the 14 African farming systems, the maize-mixed system has one of the highest farming populations with smallholders accounting for more than 90% of the 91 million hectares under cultivation (Dixon et al., 2001; Garrity et al., 2012). Malawi has of late registered success in improvements in crop productivity due to fertilizer and improved seed input subsidies that targeted smallholder farmers (Figure 1.1). To supplement the reported successes, research and development programs on ecological intensification have been heightened to address the challenges of deteriorating ecosystem functions as a result of continuous cropping. Hence, Malawi is one of the candidates for testing scalability and sustainability of integrated farming systems.

Recently the study site, Nsipe, has been a pilot and primary out-scaling area for two research projects on SLM and SAI. The International Centre for Tropical Agriculture (CIAT) through the project called AGORA: Acting together now for pro-poor strategies against Soil and land degradation used transdisciplinary and participatory approaches to map social, economic and ecological drivers of adoption of SLM practices (Braslow & Cordingley, 2016; Mponela et al., 2018). The Michigan State University, through the Africa RISING project, uses farmer managed mother-baby trials to validate and enhance the role of nitrogen fixing legumes (ground nuts – *Arachis hypogaea* L., pigeon peas - *Cajanus cajan* (L.) Millsp. and soybean - *Glycine max* (L.) Merr.) for SAI of maize (*Zea mays* L.) dominated farming system (Snapp et al., 2018).

Since the nation-wide land resources evaluation between 1988 and 1992, soil fertility decline due to overuse started getting attention and has heightened due to increasing demand for productive land (Benson et al., 2016; Li et al., 2017). Apparently, most of the technologies being promoted under the banner of SLM are not entirely novel; they are often repackaged as systems innovations addressing multiple constraints (Adolwa et al., 2017). These include SLM, conservation agriculture (CA), agroforestry, climate smart agriculture (CSA), integrated soil fertility improvement (ISFM), cereal-legume integration, and inorganic fertilizers including subsidies –which apparently have similar technological sets (Pretty et al., 2011; Weltin et al., 2018).

For the past five decades, farmers' usage of the technologies promoted under the umbrella of these different "programs" has been selective and mixed depending on the farm and household resource endowments (Mponela et al., 2016; Ngwira, Aune, et al., 2014). In the following paragraphs, the usage of individual technologies as reported by various studies across the country, is presented in brief:

(1) Inorganic fertilizers. The use of nitrogen and phosphorus fertilizers and improved cereal and legume seeds have been widely promoted and adopted in Malawi (Conroy et al., 2006). This has been supported by fertilizer and improved seed subsidies and grants since 2005 (Figure 1.1). Although reports show that increased nutrient input through inorganic fertilizers have increased the yields and generated the required calories for the farming households (Dorward et al., 2011), there are concerns that continuous cropping has lowered SOC levels (Mpeketula, 2016) and net primary productivity (Messina et al., 2017), thereby affecting land productivity in the long-term. Substantial applications of inorganic fertilizers are needed but the application rates are lower than the recommended ones (Mutegi et al., 2015) and the response to inorganic fertilizer inputs for the main crop – maize – under smallholder farming is sub-optimal (Snapp et al., 2014).

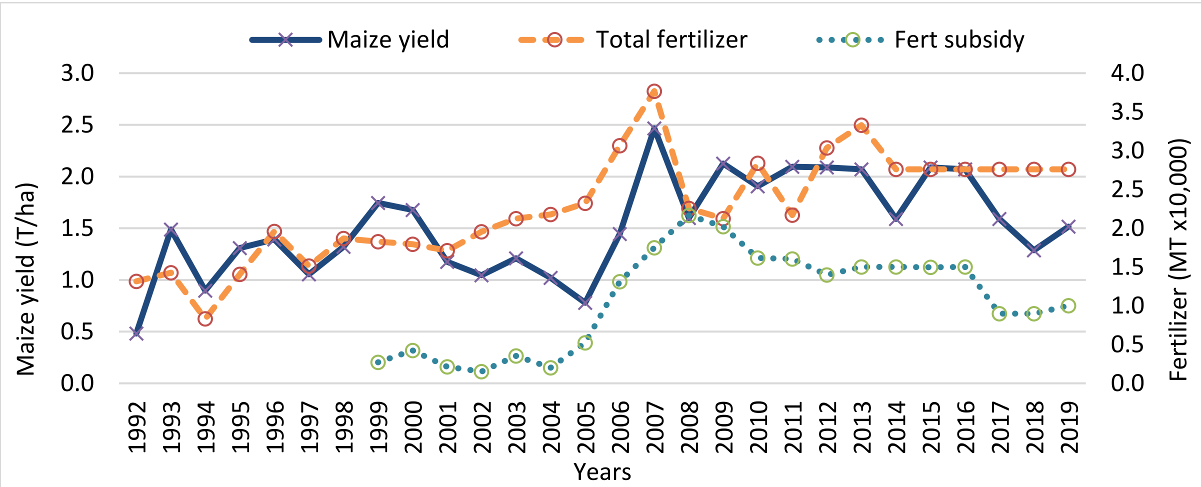


Figure 1.1 Trends of national average maize actual yields and fertilizer subsidy between 1999 and 2019 in Malawi.
Data sources: (Chirwa et al., 2011; Dorward & Chirwa, 2011; GoM, 2007; IFDC, 2013; Ricker-Gilbert et al., 2013)

(2) Crop rotation/grain legumes. Cereal-legume rotation or intercropping is a common practice in Malawi. Cereals are considered the staple crops whilst legumes (beans, soybeans, groundnuts, pigeon peas, cow peas) have been extensively grown as main protein source for the poor and are being promoted for both the local and export market (GOM, 2013). The legumes are an important source of biologically fixed nitrogen and the residues provide high quality organic matter (Mhango et al., 2017; Njira et al., 2017; Snapp et al., 2002).

(3) Residue incorporation. Situated in a sub-humid region, crop residues and natural biomass are abundant in Malawi. Some farmers incorporate them (Mponela et al., 2016; Snapp et al., 2002) and a few early adopters use them as mulch in conservation agriculture demonstration plots (Jumbe & Nyambose, 2016; Ngwira, Johnsen, et al., 2014). Whilst the residues of legumes could constitute soil fertility inputs, considerable number of farmers burn the aboveground biomass (Snapp et al., 2002), hence several institutions promote the practice of residue incorporation (Sosola et al., 2011). That being the case, most farmers still leave them in the fields and burn them during the long dry season to clean the fields (Snapp & Pound, 2008), are set ablaze by mice hunters (Ngwira, Johnsen, et al., 2014) or destroyed by stray fires from neighbouring grasslands (Campbell, 1996).

(4) Compost and farm yard manure application. After the harvest, demonstration of compost manure making is one prominent activity led by extension officers in Malawi. Use of farmyard manure is limited because of fewer livestock but composting of plant residues is on the rise (Chilimba et al., 2005).

(5) Trees on farm. With increasing pressure for land resources, woodlands have been deforested. However, trees (planted and exotic) are still the dominant feature of the farming landscapes in Malawi and continue to be an integral part of soil fertility replenishment (Kundhlande et al., 2017). Research and development programs introducing fertilizer trees on farms started in late 1980s (Beedy et al., 2013). Lately, the focus has shifted to propagate or tend indigenous trees (Beedy et al., 2010; Kundhlande et al., 2017). Efforts are made to scale-up systematic planting of trees as a recommended practice (Beedy et al., 2013), but adoption patterns are sporadic (Kakhobwe et al., 2016). This has resulted into a mosaic landscape of trees, either scattered on plots, lined along field boundaries or along the meandering streams, or as small remnant patches of woodlots protected by the government, communities or owned by land-rich farmers.

(6) Contour marker ridges/vetiver grass. Addressing soil erosion is increasingly becoming imperative as studies reveal hotspot areas (Vargas & Omuto, 2016). The eroded soils, in addition to causing sedimentation problems down the Shire River, drain the nutrients from farmlands. Farmers across the country, regardless of the soil type, topography and climate, construct ridges across the slope, which have been found to be inefficient to control erosion (Mohamoud & Canfield Evan, 1998). Among technologies developed to address erosion, marker ridges along contours, ridge realignment across slopes, and box ridges are most prominent (Jeffery & Vira, 2001). The marker ridges are reinforced with vetiver grass or agroforestry shrubs (Chigwiza & Kanazawa, 2008).

Research has shown that the aforementioned technologies and approaches have the potential to improve productivity. However, their on-farm performance is likely to vary as they are being adapted to biophysical limits bounded by socio-economic constraints (Mungai et al., 2016). As such, it is generally realised that there is no “single bullet” technology suitable for the spatially heterogeneous farms managed by the diversified rural society (Tittonell et al., 2010). At landscape and community level, assessing the continuum contribution of the technologies to ecological and human wellbeing, and not averages, has been challenging (Vanlauwe et al., 2016). The ES approach has been widely used as it provides an integrated framework for assessing performance of ecological systems (Simpson, 2011; Snapp et al., 2010; Tallis et al., 2008). In a smallholder farming system, ES that are directly influenced by SLM activities and are of relevance to the wellbeing of farmers have been considered for the assessment. The contribution to food and biomass fuel is considered under provisioning services, soil nutrient balance, soil erosion and primary production as supporting services and soil organic matter (SOM) as a regulating service (Alcamo et al., 2003).

1.3 Challenges and opportunities in modelling SAI for smallholder farms

Typically, the performance of technologies is evaluated and modelled under controlled environments and the dissemination pathways are based on the empirical relationships: out-scale to areas with similar conditions as the testing area (Komarek et al., 2018). This approach failed to replicate the results in smallholder farming systems of Africa. As observed by Giller et al. (2011), complexity in land conditions, land use and weak socio-political structures within smallholder rural landscapes constrain the use of universal models for technology dissemination. From these field observations, it is evident that the system is characterised by features that makes it challenging to choose the appropriate methods, approaches and techniques.

Weltin et al. (2018) highlights the need to understand the structural and functional forms of the farming system when conceptualising SAI assessments. In Malawi and the region, there exist sectoral policies and legal tools to frame implementation of sustainable land and natural resource management practices. Nonetheless, lack of consistent incentives coupled with insecurity of customary land tenure, weaknesses in the enforcement and social safeguards, and low public awareness has led to unregulated resource uses (Dalupan et al., 2015; Sambo et al., 2015). Smallholders have been operating without the need to comply with environmental policies and autonomously manage their fields in spontaneous response to the changes in soil quality, constrained by the environment and prevailing socioeconomic conditions (Sileshi & Akinnifesi, 2017; Smith et al., 2015; Stephens & Middleton, 2002). The resulting farming systems are diverse in both structure and functions. However, the actions by individual farmers on fragmented land parcels, when aggregated, emerge as mosaics of crops, farming practices, and soil conditions that tend to be complex to visualise and analyse at landscape level (Giller et al., 2011; Morton, 2007). Considering the complexity, SAI is being promoted for balanced production and ecological sustainability (Weltin et al., 2018), which have been challenging to land-users, agro-ecologists and policy makers (van Noordwijk & Brussaard, 2014).

SAI, as defined above, calls for an understanding of the linkages, compromises and synergies between alternative land use and practices at the management level of a smallholder farming household. As complex as the system is, there are commonalities in the way communities respond to natural processes such as soil erosion control and soil fertility enhancement. These social and ecological aggregates can be harnessed to test and promote packages of technologies using broader approaches such as SLM (Liniger et al., 2011). Due to the volume of data required, knowledge systems, and computing capabilities, a few typical farmers and farms from the centroids of the aggregates are often used to explore the performance of the social-ecological landscapes using bio-economic models (Giller et al., 2011, 2015; Mungai et al., 2016; Tittonell et al., 2010; Vayssières et al., 2011).

Building on bio-econometric models, systems for monitoring human action and resulting nutrient balance and land productivity trends have been established (Weltin et al., 2018). However, these are not enough for anticipating future changes in SAI as they lack feedback mechanisms. An improved understanding of feedback mechanisms and controlling factors is essential for pro-active decision-making and projection of future scenarios (Le, 2005). In line with declining productivity due to soil fertility loss, since 1990s there has been a growing need to understand the balance between nutrient input and output levels (Stoorvogel & Smaling, 1990). The latter authors grouped the flows into five inputs (IN_1: mineral fertilizers, IN_2: organic manure, IN_3: biological N fixation, IN_4: atmospheric deposition, and IN_5: sedimentation) and five outputs (OUT_1: harvested product, OUT_2: crop residues, OUT_3: erosion, OUT_4: leaching, OUT_5: gaseous losses) that are used as a frame (Equations 1.1 – 1.4 below). In addition to the underlying natural processes, the human actions are considered a significant force accelerating or decreasing the nutrient flows through the first three inputs and three outputs. Therefore, instead of being a static and stable processes, the nutrient balances are highly dynamic and differ over time and across space (Cobo et al., 2010).

$$Net(N) = IN1 + IN2 + IN3 + IN4 + IN5 - OUT1 - OUT2 - OUT3 - OUT4 - OUT5 \quad 1\text{-}1$$

$$Net(P) = IN1 + IN2 + IN4 + IN5 - OUT1 - OUT2 - OUT3 \quad 1\text{-}2$$

$$Net(K) = IN2 + IN4 + IN5 - OUT1 - OUT2 - OUT3 - OUT4 \quad 1\text{-}3$$

$$Net(C) = IN2 + IN2(fromOUT2) + IN5 - OUT3 - OUT5\ (decomposition) \quad 1\text{-}4$$

Due to the complexity of farms and owners, treating the nutrient balance as a static inflow and outflow is an oversimplification. It is crucial to explore how *Net nitrogen (N), phosphorus (P), potassium (K) and carbon (C)* influences farmers response to their target productivity output *OUT1* which in principle is linked to all others stocks and flows and their controlling factors. Hence, farmers through pursuit of their main objective of increasing productivity exert a dynamic force on soil nutrients and shift its equilibrium over temporal and time scales. Capturing these spatio-temporal changes continues to be a challenge.

In the past, nutrient balance studies have consistently reported negative trends. This stems from the notion that the tropical soils of Africa are either inherently less fertile (Tully et al., 2015) or degraded (Vlek et al., 2008), understating the variations across regions and within single farms (Cobo et al., 2010; Vanlauwe et al., 2006, 2016). A review by Cobo et al. (2010) showed that of the 57 studies included, 15% reported positive N, 24% positive K and 44% positive P balances. In mostly rural areas, the positive balances are distinct for wealthier households and for plots closer to houses (Vanlauwe et al., 2006). However, crop type, farm size and the accompanying soil management practices override these distinct social and physical boundaries, thereby creating a random distribution across landscapes. Even for those with negative balances, continued usage of the farmland (Braslow & Cordingley, 2016) is indicative of farmers' adaptive strategies that sustain these nutrient mining farming systems (Smaling et al., 1997).

Apparently, there are some positive deviants with positive balances. The study therefore has a particular focus on the balance of the major nutrients including nitrogen, phosphorus and potassium at household level as the main ecological objective for land managers. To-date, although the aims have been to come up with country budgets or in some cases for the whole of SSA, nutrient balances are still modelled at two micro spatial scales (Lord et al., 2002). The farm-gate scale accounts for major nutrient inputs and outputs for the household or community by aggregating plots. The soil-surface scale accounts for every land parcel, whether cultivated or not, and considers the internal nutrient cycling such as manure and crop residue management. Nutrient balance studies have shown that most farms tend to have negative partial budgets (Cobo et al., 2010; Kangalawe, 2014). The combined nutrient inputs, through inorganic and organic fertilizers, purchased feed, external grazing, crop residues, atmospheric deposition, sedimentation and biological nitrogen fixation tend to be lower than the total outputs through crop harvest, residue removal and potential soil erosion. At plot or farm level, it has been established that continued cropping without adequate inputs leads to nutrient depletion (Potter et al., 2010).

Considering the limited data on nutrient pools and flows in the region, there is growing interest to improve estimates. This is a prerequisite for development of appropriate interventions. Even more important is to have data that is spatially and temporary explicit. Generally, empirical measurements are made for material flows such as fertilizers, manure, and crop yield whilst the natural processes of volatilisation, deposition and denitrification are not included or simply estimated using transfer functions from the literature (Scoones & Toulmin, 1999).

Nutrient stocks-and-flows within farming systems is the major indicator of productivity and has been efficiently modelled using the NUTMON/MonQ (De Jager et al., 1998) and FarmDESIGN (Groot et al., 2012). These are bio-economic models and are limited to the use of static entities of agricultural enterprise and land units. Without the decision-making component, it is impossible to capture dynamic feedback, evolution of processes and system states and adaptive behaviour of smallholder farmers that are essential to understand social-ecological adaptation and the transitions (Boulanger & Bréchet, 2005). In addition, by averaging the nutrient transfers of a sample to a few archetypical farms, selected from centroids of the clusters and assumed to be representative, these models fail to allow for and capture autonomous actions by heterogeneous farmers that interact within landscapes and communities.

In an expert-based assessment of six common integrated system modelling approaches for supporting sustainable development, Boulanger & Bréchet (2005) found that a multi-agent system (MAS) approach is well-suited for understanding sustainable development involving social-ecological system (SESs). Many studies in the previous decade have demonstrated capabilities of MAS for comprehensive *ex ante* assessment of impacts of policy interventions on land use and cover (Berger, 2001; Le et al., 2008, 2010a; Miyasaka et al., 2017; J. Schindler et al., 2009; Villamor et al., 2014). Still, there are discussions on how to model dynamic reverse causality among human and ecological sub-components in smallholder farming systems (Giller et al., 2011, 2015; Mungai et al., 2016). Elsewhere, MAS has shown that although farmers opt for short-term options to improve yield, human induced nutrient losses from erosion and crop removal have long-term consequences (Quang et al., 2014; Schreinemachers & Berger, 2011).

Notwithstanding, the NUTMON framework and its survey tools have good value to capture nutrient flows for different farms. Therefore, this study builds on the NUTMON and FarmDESIGN frameworks and implement them within a MAS platform to capture variability and residual effects over time and farm types (Le, Scholz, et al., 2012). This facilitates the estimation of cumulative and emergent effects over temporal scales, as well as simulating the reverse causality between several components that static models such as NUTMON and FarmDESIGN fail to address.

Apart from the uncertainties in the estimations of nutrient pools and transfers, it is also important to reduce an attribution error when up-scaling from the basic measurement units of plot-household to the landscape-community levels. Instead of using one or a few average farmers selected through farm typologies and arbitrarily extrapolate over an entire area, the nutrient balances for the representative sample (above 30% of a village population) was collected through in-depth interviews. The interviews took place during farm visits, thereby allowing collection of detailed information for each plot. Since the aim of the study is to evaluate the determinant factors and empirically evaluate their effects, the sampled households and plots are first grouped into distinguishable types. The average and distribution of nutrient balances within the typologies, not the single observation from centroid farms, are then used to upscale to the population/landscape level.

1.4 Research questions and objectives

In view of the foregoing, this study aims to address the following questions:

1. What are the adoption pathways for land management technologies that improve crop productivity whilst sustaining the integrity of the natural resource base?
2. How can SLM policy options be best applied considering the heterogeneity of farming households and their farmlands?
3. Are there trade-offs or synergies in ecosystem services derived from land management practices in terms of sustainability of crop productivity and soil quality and the household's income?

To address these questions, the main objective of the study is to build a SES that spatially and temporally explicit capture, analyse, present soil nutrient balances and explore ex ante possible livelihood and ecological outcomes from alternative soil management practices to better inform smallholder farmers and other stakeholders whilst making their sustainability decisions. In achieving this goal, the study aims to realise the following specific objectives:

1. To develop a coupled human-ecological framework for modelling the co-evolution of soil fertility within smallholders' land parcels and food production dynamics framed by land management policies and other externalities in maize mixed farming systems of east and southern Africa.
2. To calibrate and verify factors that nudge agents of change (farmers) to decide on taking and intensifying soil fertility management (SFM) alternatives thereby creating shifts in the human sub-system. The decision models are based on the empirical data collected from households and their farms in a typical subsistence maize mixed farming system of Malawi.
3. To calibrate and verify factors that drive changes in land productivity within the farming landscape as an ecological sub-system. The nutrient input and output models are calibrated based on the nutrient transfer ecological models with data

collected from secondary sources and own surveys of existing farmlands in the study site.

4. By integrating parameterised models from 2. and 3., build a Multi-Agent System for Sustainable Agricultural Intensification (MASSAI) that simulate, *ex ante*, the possible economic and ecological outcomes of certain soil management and other related policies from dynamic interactions among farming households and their farmlands.

2 Review of Existing Literature

2.1 Household social-ecological livelihood diversity

Adopting the farming styles theory of van der Ploeg (1990) which postulates that, despite having structural and functional forces that increases complexity in agriculture, there is unity or cohesion in farming styles defined by three elements:

> *"... a set of strategic notions, values and insights shared by a particular group of farmers concerning the way farming ought to be organised*
>
> *... a specific structuring of the practice of farming that corresponds to the strategic notions or cultural repertoire used by these farmers*
>
> *... a specific set of interlinkages between the farm enterprise on the one hand and the surrounding markets, market agencies, government policy and technological developments on the other. These interrelations are structured in such a way that the specific farming practice can be reproduced over time"* (van der Ploeg & Long, 1990).

The livelihood strategies undertaken by farming household in the maize mixed farming system are construed to be dictated by production factors including human and ecological resources as well as institutional processes that influence how resources can be used to realize different household objectives. The large number of the pre-conditions despite providing a wide range of options, pose an analytical challenge in identification of the most limiting or enabling ones.

Several notions have been used to explain diversity in terms of differentiating factors among farming households. Some studies have used the factors of production that are considered limiting and therefore most important to the farmer decision making. Among them, land size has been the main discriminant used either as a sole attribute or as a determinant factor in regression equations. Prior to the 1990s, when large farms were established on prime agricultural lands, were sole beneficiaries of input credit and had unprecedented access to new varieties, the relationship between land size and productivity was positive (Dorward, 2002). From 1990s onwards, inverse relationships have been observed which has been attributed to the government supported farm input programs that enabled smallholders to access fertilizers and improved varieties (Matchaya, 2007). Labour which is linked to farm size - productivity relationship has also been found to discern the households into either subsistence or commercial (Douillet & Toulon, 2014).

Gender of the household decision maker has also been an important grouping variable for income levels (Brown et al., 1996; Peters, 2006). Some researchers used regional administrative boundaries for determination of agricultural opportunities (Simler, 1994) while others used anthropological panel analysis of household resources (Peters, 2006). Recent studies have considered a combination of agro-ecological and socioeconomic characteristics for determination of livelihood in terms of food access and coping strategies (Douillet & Toulon, 2014).

However, these approaches fail to capture the spatial and temporal dimensions of real situation of farms and households. In smallholder agricultural systems, diversified behavioural portfolios affect the pursuit of different livelihood strategies by households and dictates sustainability of the ecosystem (Shefrin & Statman, 2000). This study considers the Malawian maize-mixed farming community to be a typical SES and uses two frameworks: (1) The Sustainable Livelihood Framework (SLF) by Sconnes (1998) to identify structural factors that influence farmers' decisions and abilities to undertake practices for a particular livelihood strategy; and (2) The Human-Environmental System Framework (HES) by Scholz et al. (2011) to draw relations among factors in the social and ecological domains that yield functional farm household types.

2.2 Household typology

Development programs being implemented by various agencies have given rise to many farming sub-systems in the region. Intensification through use of inorganic fertilizers and modern improved varieties has been widely promoted to replicate the benefits observed during the Green Revolution in Asia. High poverty levels and the love for local varieties led to wide usage of local varieties or recycled improved seeds and locally available organic resources. The preference for local varieties is heightened due to their good aroma and taste ad well as storability. Researchers have also promoted the mixed farming of cereals and legumes to mimic the natural ecosystems and maximise land utilisation equivalence and nutritional returns from the small parcel holdings. In mixed plots, however, yield and quality of individual crops is usually compromised. It is also technically difficult to manage mixed farms. As a result, there is a growing trajectory of plots under sole crops.

SFM has also been widely diversified. Apart from the universal application of fertilizers, many farmers do not apply the recommended; hence, their input levels are different. Use of manures and other organic resources is also highly skewed. Of late, there has been strong debate and conflicting messages of incorporation of crop residues. Traditionally people used to burn the residues, which was supported by the concept of immobilisation where mostly the residues from maize and weeds in the region have low C:N ratios. The farmers benefited from the potash released during burning at the expense of organic matter depletion (Mpeketula, 2016) and consequential loss of organic nitrogen. Composting the residues before application has been the recommended practice but some farmers bury for reasons ranging from labour constraints to lack of consolidated circumstantial evidence to validate the biochemical processes. Management of residues has become more diversified with the outbreak of the fall army worm. One control measure has been to uproot and burn the residues. Although the government has put in place robust environmental legal frameworks, the ultimate decision is made by the farmer. Lack of incentives coupled with insecurity of customary land tenure, weaknesses in the enforcement and social safeguards and low public awareness have led to unregulated use (Dalupan et al., 2015). This caused emergence of mosaic land use patterns where individual farmers adjust the technologies to their own limitations (Coe et al., 2016; Sileshi & Akinnifesi, 2017). Consequently, often a complex mixture of crops, rotations and farming practices exist within adjacent fields.

Most research and development programs implemented on farmer's fields within the maize mixed smallholder farming systems are overwhelmed by the unmatched heterogeneity (Giller et al., 2006, 2011; Vanlauwe et al., 2016). Considering heterogeneity whilst making research outputs relevant to a wider society or environment has become one of the objectives for farming systems research. As such, characterizing farms into types has been proposed and used. Farming types facilitate targeting of technologies, dissemination of appropriate technologies to a larger scale, selection of prototype farms for detailed research and scaling-up to larger spatial and organisational scales (Garrity et al., 2012; Kamau et al., 2018; Le Bellec et al., 2011; Le et al., 2010c; McConnell J & Dillon L, 1997). In theory, the aspirations and actions of farming households are influenced by the underlying and exogenous factors including demographic, institutional as we well as availability of economic and biophysical resources.

Since the 1950s, voluminous research has been done in SSA to explore the levels of heterogeneity among the farmers and the farming practices undertaken. As indicated by Larson *et al.* (2012), heterogeneity in demographics, farm endowments and produce markets generate heterogeneity in applied technologies. Debate ensued as to whether the broad spectrum of heterogeneity should be considered when conducting research or a set of variables considered subtle should be chosen for detailed diagnosis. The latter has been widely preferred since in controlled experiments, the impact of a factor can be noted. However, focusing on the variables considered key tend to increase the likelihood of mis-representation with subsequent failure of technologies due to misfit with the real farm and farmer conditions. As a bounding natural resource, the inherent soils are quite variable across the landscapes (Hengl et al., 2017; Li et al., 2017; Njoloma et al., 2016; Snapp, 1998; Towett et al., 2015; Vågen & Winowiecki, 2013). The differences in soil fertility have also been recorded between fragmented plots belonging to the same farmer as well as within the plots (Vanlauwe et al., 2006). These background fertility gradients influence farmers' crop choices and SFM behaviours.

Advances in development and deployment of context specific technologies have focused more on the major crops, not taking into account the heterogeneity in landscapes and communities (Rware et al., 2014; Snapp et al., 2003). To enable better matching of crops to biophysical conditions, global and national stratifications of agricultural systems into agro-ecological zones (Fischer et al., 2002) and farming systems (Dixon et al., 2001) have been established. These are principally used in designing global products such as new cultivars of crops and for soil fertility, the inorganic fertilizers. Although these technology-based interventions have increased adoption and impacts on crop yield, there are still wide performance gaps among smallholder farmers (Vanlauwe et al., 2016). The overriding reason for this has been that when taking the 'best-bet' technologies to smallholders, the actual farm attributes, household resources, external pressures and their linkages have often been overlooked.

That said, the regional or sub-catchment zonation provides a guiding framework. Further zonation for exploring how the research outputs can be effectively implanted considering the differences among farmer and their farms. For example, farmers of different resource endowments and access to information have different decision-making abilities, so is their level of technology adoption. Consequently, the mechanisms by which they cope with risks and their ability to bounce back from stresses also differ. In attaining the livelihood goals, farmers generally react according to production rules, mimic neighbours but also optimise utility, giving rise to diversified and complex behaviour patterns and feedback loops (Le, Seidl, et al., 2012). Without detailed information related to the local context, planners and decision makers use one size fits all recommendations even when faced with farm/household heterogeneity. This often results in skewed outcomes. To reduce the gaps, there is need for explicit examination of the combined role of households' characteristics, farm and neighbourhood biophysical attributes and linked external institutional factors on farmers' behaviour. Lack of differentiation of the real situation of the households and their farms complicates development of decision support system for agricultural production planning (Riveiro-Valiño et al., 2009).

Therefore, the aim of this study is to understand the heterogeneity among farming households and classify them using household, plot and ecological variables into homogenous farm types. Ultimately, these typical household farm types are used as a basis to develop typology representative decision and production functions. The typology results further support farm type specific analysis of sustainability outcomes for targeting agricultural intensification interventions.

2.3 Soil fertility management usage in the context of household resource availability and land constraints.

Low adoption of SFM practices is the basic concern in smallholder farming systems of Africa. This is not expected as for the past six decades, farming systems research and development (R&D) has provided farmers with several technologies that potentially protect, maintain and build soil fertility (Vanlauwe et al., 2017). The last three decades have seen repackaging and disseminating sets of proven SFM technologies and promoting them as systems innovations. Among others, integrated SFM (ISFM) is touted to address multiple constraints. Farmers are expected to select a suite of technologies that fits their land, cropping system and socio-economic capacities.

In Malawi, farmers typically supplement inorganic fertilizers with locally produced organic resources such as animal manure, compost manure and crop residues (Palm, Gachengo, et al., 2001). Legume crops have been promoted for their triple benefits of nutrition, income from grain sales and soil improvement through biological nitrogen fixation (BNF) and used as a measure of functional crop diversification (Kankwamba et al., 2018). SFM usage by farmers could be considered a response to diminishing nutrient levels, which unfortunately, have to be addressed first before farmers can realise the benefits from other farming practices (Sanchez, 2002).

During the era of shifting cultivation and natural fallows, organic resources played a significant role in SFM through their short-term effects on nutrient supply and longer-term contribution to SOC (Chilimba et al., 2005; Palm, Gachengo, et al., 2001). At the time when farms transitioned into continuous cultivation in the 1960s, studies already established that without external inputs, SOC decreased by 41% in 25 years and maize grain yields by 57% in 10 years (Chilimba et al., 2005).

Clearly, inorganic fertilizer, organic manure and legumes are not novel technologies, farmers have used them for decades. Weak regulatory systems and fewer incentives make it hard for farmers to conform to minimum SFM standards (Dalupan et al., 2015). Beyond the experimental and dissemination phase, unregulated, farmers adjust the technologies to fit them to their farm conditions and household resource endowments (Coe et al., 2016). As such, the technologies and practices by individual farmers on fragmented land parcels lead to emergence of mosaic soil fertility management patterns that are complex to visualise, analyse and communicate at landscape level (Giller et al., 2011). Usually, a random mixture of crops, rotations and farming practices exist within adjacent fields.

During the last decade, research on SFM adoption and its impact on crop yield and household welfare has increased in volume. Organic manures and legumes are considered short-term drought adaptation strategies but usage is inconsistent after prolonged droughts (Katengeza et al., 2019). (Katengeza et al., 2019; Mungai et al., 2016; Sauer & Tchale, 2009; Silberg et al., 2017). However, previous studies use national sampling strategies and draw limited samples from individual villages, hence are not representative at village or even national level (Katengeza et al., 2019). These studies revealed trends relevant for national or regional planning such as for resilience to climate shocks (Katengeza et al., 2019) or shifts in labour markets (Sauer & Tchale, 2009) but are not representative enough to give insights into why farmers would not imitate their immediate neighbours.

In Malawi, the sloping rift valley escarpments are classified as having a medium agricultural potential (Li et al., 2017). They have been under continuous cultivation for over three decades and farmers experience various forms of land degradation (Braslow & Cordingley, 2016). These are areas with a lot of challenges and in urgent need for SFM activities, thus providing a prime opportunity to explore the village level scaling constraints. In the region, almost all farming activities are done by hand and haulage of heavy items by head (Amede et al., 2014). Hence, household demographics in terms of availability of labour and the number of dependants compared to workers in a household are important investment factors in rural communities.

Productivity of smallholder farmers remained low till 2005 when the Malawian government introduced farm input subsidies. Fertilizers and improved seeds supplied through the program have shifted and stabilized maize yields from 0.7 ton ha^{-1} in 2005 to around 2 ton ha^{-1} for 13 years (FAO, 2016a). Despite the observed stability, the current yield levels are not that different from the 1.7 ton ha^{-1} attained under normal rains in 1999 - 2002 and much lower than the potential yields for improved maize varieties in Malawi of 4-15 ton ha^{-1} (Tamene, Mponela, Ndengu, et al., 2016). As of 2015, some Malawian farmers were not applying inorganic fertilizers and the majority applied below the recommended rates (Mutegi et al., 2015).

Although manure have also been promoted since 2000 (Chilimba et al., 2005), the subsidy program has focused more on the inorganic fertilizers which probably has led to decline in SOC (Mpeketula, 2016) and land productivity (Messina et al., 2017). From as early as 1965, significant crop response could be observed when 5 ton ha^{-1} of farm yard manure were applied to maize (Chilimba et al., 2005). A study in Zimbabwe revealed that long-term manure application (>10 years) at the rate of 3-5 ton ha^{-1}

increased SOC to medium fertility whilst 10 ton ha^{-1} could replenish the SOC to pristine levels (Musinguzi et al., 2013). Considering the threshold levels set by these studies and SOC levels that are lower than the critical levels required for structural stability of 2.0% (Tamene et al., 2019), the observed organic inputs are insufficient to contribute to nutrient supply and SOC build-up (Chilimba et al., 2005). In these nitrogen limited soils and low input farming systems, the nitrogen fixed by legumes is a major source of the nutrient (Njira et al., 2017). The crop enumeration by the agricultural office in Ntcheu district revealed that the three common legumes: groundnuts, soybean and pigeon peas take 14% share of cultivated land (Ortega et al., 2016).

It is against this background that paper examine reasons behind non-adoption and drivers for upscaling of SFM technologies within a rural population in Malawi by addressing the following questions: (1) Why farmers do not adopt SFM technologies practiced by their neighbours? and (2) for those that adopted, what could be the household and plot level factors leading to varying levels of SFM usage?

2.3.1 Stocks of soil NPK, SOC and their thresholds in smallholder managed escarpments of Malawi.

Much as the declining soil fertility is the major biophysical factor threatening food production for rural smallholder farmers in Africa (Sanchez, 2002), the knowledge of the status and gaps at the management scale of a single plot is limited (Forkuor et al., 2017). This is because information needed to understand processes at detailed scale are overwhelming (Jenny, 1941). Soil fertility is a complex mix of biological, chemical and physical properties that centres around the stocks and cycling of nutrients, influenced by ecological and human factors. However, in most tropical extractive farming systems, it is increasingly concerning that if the nutrient and SOM stocks continue to decrease, the capacity of land to support agriculture would be compromised (Ayuke et al., 2019; Rattan Lal, 2015). Therefore, understanding the manageable physical properties and the associated soil nutrient cycling mechanisms is critical (Ayuke et al., 2019). In Malawi, recent studies have shown declining trends in primary productivity and SOM due to continued cultivation (Li et al., 2017; Messina et al., 2017; Mpeketula, 2016). These and other studies act as pointers as they were conducted either at national (Li et al., 2017; Messina et al., 2017) or point scales (Mpeketula, 2016), hence not representative of the soil conditions across and within farming landscapes (Forkuor et al., 2017).

Efforts have been made to take stock of the soil status in smallholder farms using samples collected at national scale (Njoloma et al., 2016; Snapp, 1998). These soil studies and the resulting management strategies in use are based on agroecological zones which are said to be delineated based on similarities in climate, topography and the major soil types. However, high within-site variations in soils with coefficient of variations of 52, 30, 67 and 69% for N, P, K and SOM, respectively, have been observed (Njoloma et al., 2016). In particular in Malawi, variations within the sample areas can be huge considering that the country is in the Rift Valley floor with varying degrees of terrain attributes that drives biogeochemical properties of soils. Consequently, a complex mixture of soil classes has been observed within 0.5 km distances across the landscapes (Garrity et al., 2012; Snapp et al., 1998). Soil mapping that captures such variability is therefore needed to support informed decision making.

To-date soil replenishment strategies still use the soil class polygons developed in the 1960s (Mutegi et al., 2015). Notably, newly acquired data are used to update the soil nutrient and biophysical attributes of the base soil polygons. At the same time, there have been improvements in spatial soil mapping associated with high spatial resolution proxy co-variates and analytical approaches. At continental scale, studies of soil forming factors (Towett et al., 2015) and digital mapping of soils (Hengl et al., 2014, 2015, 2017; Towett et al., 2015) utilised 28,500 legacy soil profiles and 160x60 locations from 60 sentinel sites across Africa to predict soil properties for 18.3 million km^2. The proxy co-variates used are the spectral signatures for the moderate resolution imaging spectroradiometer (MODIS) satellite imagery with spatial resolution of 62,500 m^2. These resolutions are finer for farms and places with homogenous soils over 6.25 ha land units and are quite coarse to spatially register and depict the differences between the 0.5 ha fragmented plots owned by most farmers in Malawi (Ichami et al., 2018). Although these new maps are widely used, the sampling intensity at national level is still low. In Malawi about 2,983 legacy points and two sentinel sites were included to represent 94,080 km^2 land surface. Moreover, outdated legacy data collected between 1964 and 1990 were used (Kempen, 2014), not reflecting the changed soil status (Mpeketula, 2016). As for the recent collections included in the continental mapping, the two sentinel sites were sampled by the Africa Soil Information Service (AfSIS) project from Nkhatabay in the north and Tchuchila in the south, which are not representative of soil conditions for the majority of areas in Malawi.

Notably, local scale variations in soils are more correlated with local factors such as topography, soil texture, land use and rainfall (Schillaci et al., 2017), covariates that have been overlooked by the said continental mapping attempt, shortcomings that render these recent maps still not very useful for planning spatial soil management for small-scale farming (Hengl et al., 2017). In principle, plot level soil information are pivotal for optimal decision making in terms of soil conservation and nutrient replenishment (Forkuor et al., 2017). Moreover, farmers having moderately fertile or infertile plots are less likely to adopt ISFM technologies (Mponela et al., 2016). In spite of its economic relevance, most nutrient balance studies do not take into account the existing nutrient stocks at the farm scale (Cobo et al., 2010). With small land parcels of 0.45 ha and below often managed differently, plot to plot differences can only be distinguished at higher spatial resolutions. Therefore, this study aimed to predict and map soil fertility attributes with higher resolution (100 m^2) along 10km x 10 km vegetation, topographical and geological gradients in maize mixed farming systems of Malawi. Methodologically, model specification tests are done to find the best predictor set for the soil attributes. After estimating the spatial distribution/dynamics of nutrients and SOC within smallholder farms, a conformity test was performed to check if the 6.25 ha resolution maps could comparably be sufficient for the attributes (see Appendix S1). Finally, the nutrient gaps are calculated as deviation of the estimated soil conditions from the thresholds for plant growth and ecosystem health.

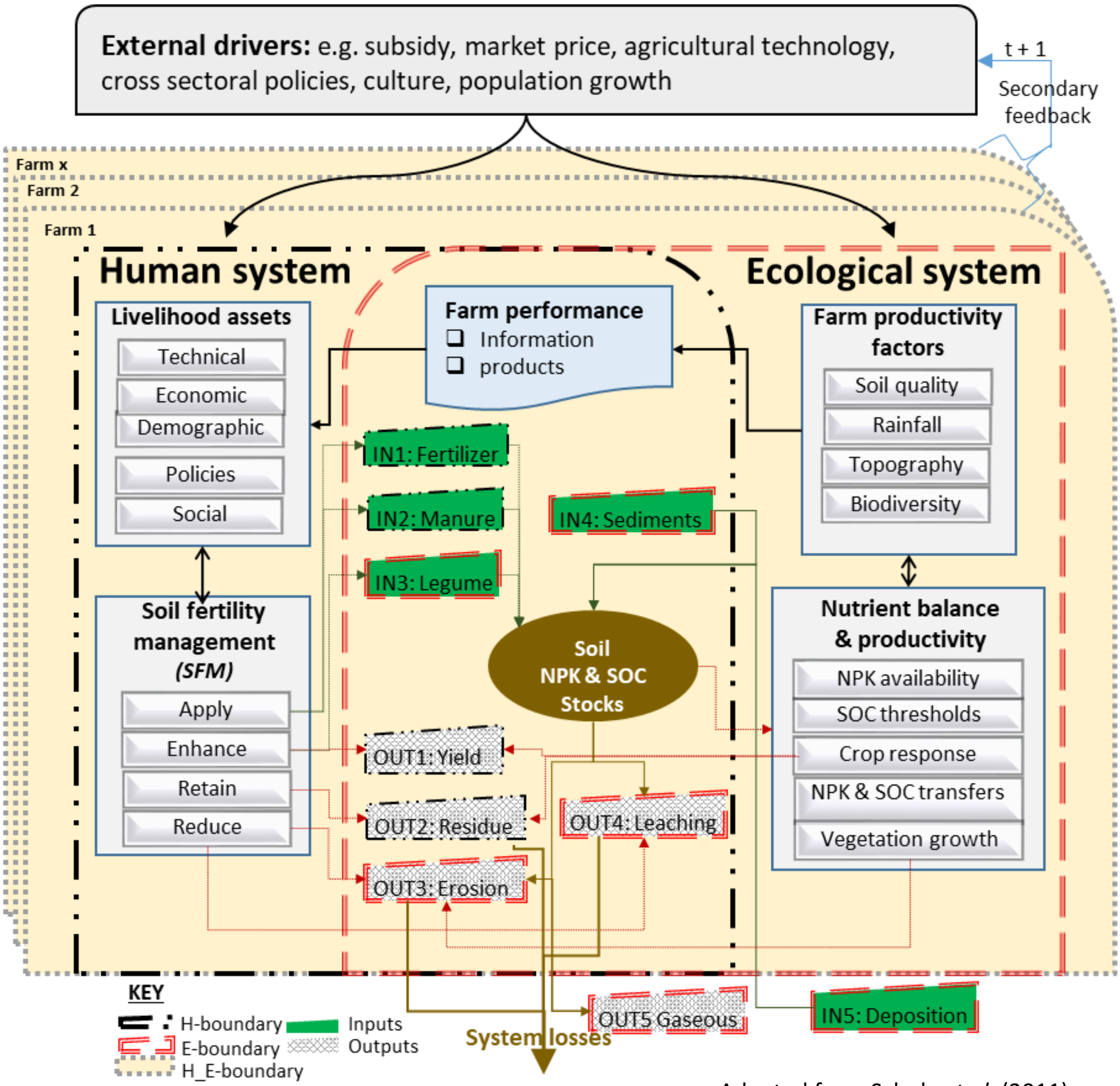
There is a growing body of research aiming at establishing linkages and feedbacks between government policies, development of rural livelihoods and sustainability of land productivity. Analytical frameworks such as that of Le et al (2012), outlines the empirical foundations for linkages and trade-offs among dynamic changes in human behaviour that alters nutrient input and extraction pathways. In their Multi-agent system model for exploring efficient smallholders' P use and management strategies (MAPU) framework, Le et al (2012) purports that overapplication of P is detrimental to the environment, leads to biodiversity loss and low profitability while under application leads to low productivity and soil degradation. Tracing the flow of nutrients and carbon in farming systems from the plot to the household and *vice versa*, external drivers such as input subsidies that aim at inducing human agentic behaviours, alter the rate and magnitude of flows which in turn influences land productivity and reshapes household livelihood portfolios and conditions the next code of action. In the absence of long-term representative panel data to model rate of change, detailed baseline survey data can used to conduct, *ex-ante*, comparative analyses of long-term impacts of multiple human-induced drivers on soil fertility, crop production and profitability of farming systems.

3 METHODOLOGY

3.1 Concepts, model framework and design

The maize mixed smallholder farming system of east and southern Africa is a Multi-Agent System (MAS) characterised by an environment, objects and agents (the agents being the only ones to act), relations between all the entities, a set of operations that can be performed by the entities, and the changes of the universe in time due to these actions (Ferber, 1999). The farming *environment* comprises of farming plots, uncultivated areas, built areas, socio-cultural and ecological features that provides ES. The *agents* are the farming households within the system that makes autonomous decisions regarding the ES they derive from the *environment*. Well explained by the theory of Farming Styles (van der Ploeg & Long, 1990), the farming households are biological organisms and social beings with transformational energy whose actions are interdependent with the environment they live in.

Schematically, the linkage between human and ecological components arises from the decisions by *agents* and the ecosystem benefits that shapes individual's reactions to the ecosystem conditions and processes (Figure 3.1). In real life, the current soil fertility and productivity informs the subsequent decision on types of activities and the perceptions about benefits that an individual has on the environment. To model spatially explicit evolutions of the maize-mixed farming system as a coupled human-environment system at a landscape level, the condition and interactions among farming households and the services and condition of the land patches are empirically calibrated. The work builds on the previous operationalisation of the HES frameworks to simulate social-ecological performance of farming and other managed natural systems that have been fully developed (Badmos et al., 2015; Le, 2005; Le et al., 2008, 2010b; Le, Seidl, et al., 2012; Quang et al., 2014; J. Schindler et al., 2009; Smajgl et al., 2011; Tsai et al., 2015; Villamor et al., 2014). This work builds operational MAS that simulate the sustainability outcomes from drivers of soil and land management interventions.



Adapted from Scholz *et al.* (2011).

Figure 3.1 The Human-Environmental Framework showing linkages and feedbacks between human actions and soil productivity.

Given the many notions and dimensions of SAI, the implementation framework ought to be based on the practical phenomenon (Weltin et al., 2018). In this paper, we develop a coupled human-ecological framework for simulating soil fertility and crop production dynamics framed by land management practices and other externalities in maize mixed smallholder farming systems of east and southern Africa. Applying the overview, design concepts, and details + decision (ODD + D) protocol (Railsback & Grimm, 2012), the study intent, characteristics of the study units and processes and the procedure (Table 3—1).

Table 3—1 The ODD+D documentation of the MAS for simulating SAI in rural Malawi

Guiding contents	ODD+D Model Description
Overview	
Purpose	
Question, problem or hypothesis (i.e. overall objective)	The purpose of the MASSAI is to understand how usage of chemical and biological soil management interventions impacts on short-term yields and their potential long-term nutrient balances and structural changes in farming systems and household food security.
Entities, state variables and scales	
Kind of entities, their attributes (include units), spatial and temporal resolutions	**Human agent**: farmers with heterogeneous demographic characteristics (age, number of members, labour), economic activities (income from farm and non-farm), land ownership and the parameters for soil fertility technology choice probabilities and usage intensities. **Ecological agent**: the top soil (0-30cm) actively explored by annual crops for nutrients and water. Its characteristics includes soil properties (NPK and SOC), which through transfer functions that capture nutrient dynamics responds to natural and human influence. The ecological agent set also include crops grown on a farm parcel and is characterised by parcel location, area, crop type, farming practices, yield and nutrient contents. It responds to dynamics in soil and human actions. These human and ecological agents and their attributes (see details in Table 3—2) make up the 'farm' with linked agronomic, ecological, social and economic performance indicators. They are georeferenced as points and raster with grid cells 10m x 10m and linked the heterogeneous environmental assets and drivers: natural (soil properties, topography, vegetation, and the co-efficient of variables for nutrient transfer dynamics sub-models), agricultural (land-use – cultivated or not use, area, agricultural yield, nutrient inputs, labour force, parameters for agricultural yield and nutrient transfer dynamics sub-models), and institutional (ownership, village)
Process overview and scheduling	
What entities does and in what order?	Soil nutrient input strategy choice and use intensity which is largely influenced by human behaviour to start or increase use of SFM practices. A brief introduction is given in Table 3—3 and Figure 3.2 and detailed in section 3.2.3. Nutrient stockpiles and transfer modules capture biophysical processes of nutrient supply, transformation, transport and export mediated by ecological factors and human activity. The overview is presented in Figure 3.1 and more detailed linkages between drivers and outcomes are presented in Figure 3.2, Figure 3.3 and Figure 3.4.
Design concepts	
Basic principles	
Concepts, theories and hypotheses underlie the model design	Theoretically, we adopt the farm styles theory which purports that the farm and landscape structural changes are spatio-temporal explicit dynamic processes emerging from individual farmers decisions and actions about farm inputs and their perceptions and control of farm outputs on each of the land parcels (van der Ploeg & Long, 1990). To explicitly capture and formalise the linkages and feedbacks between ecological processes and human actions, we adapt the Human-Environmental Framework (Scholz et al., 2011) as indicated in Figure 3.1. We use the indicators and the matrix of sustainable intensification for smallholder farmers (Smith et al., 2015), and the nutrient input and output transfer functions as formulated in farmDesign (Groot et al., 2012) and nutrient monitoring farm models (Smaling & Fresco, 1993). The farmer decision model is based on micro-economic theory, with the assumption that farmers are private entities and use the soil technologies to maximize utility from land units. To do so they either enhance productivity or reduce degradation hence the decisions are: (i) rational by maximizing yields; (ii) bounded rationality of input use according to average benefits; or (iii) ad-hoc rules such as applying due to availability and/or access to fertilizer. In some cases, there are abrupt system changes such as burning crop residues, which were previously incorporated, to control fall armyworms. The double hurdle model is chosen because the aim is to analyse the factors influencing household's probability and extent of soil fertility management. The SLF was used to take stock of factors that influence farmers' decisions and abilities to undertake practices for a particular livelihood strategy (Scoones, 1998). Data was captured through soil, crop yield, household surveys that were geo-linked using a common sample frame. The landscape dynamics and processes are formulated based on the principle of ecological equilibrium. As populated by (Stoorvogel & Smaling (1990), there are five inflows and outflows of inputs and outputs. Of the five ins and outs, 3 are archetypal ecological, and in

Guiding contents	ODD+D Model Description
	pristine ecosystem we envisage that nutrient flows and stocks are in ecological equilibrium. However, for managed agricultural systems, two of the inputs and two outputs are deliberate efforts by humans while others partially mediated by human action and the system's equilibrium shifts after a series of disturbances.
Expected variations in the model results when parameters change. Vis -a-vis results imposed by model rules.	Framed by constraints and opportunities, actions by individual smallholder farmers when aggregated over space and accumulated over time become an unformidable force that continually shapes the environment-community agricultural productivity. The recurrent low crop yield, hunger, poverty, low input use, low yield cycle is a typical phenomenon in most parts of Malawi, creating vicious cycles of poverty traps (Tittonell & Giller, 2013). Aggregation and shifts are expected in land use expressed as crop(s) planted, nutrient input and output, soil nutrient stocks, crop yields, farm incomes and number of farmers adopting nutrient input strategies.
How the individuals make decisions/ behavior to achieve objectives e.g. change cells	The agents change parameters or objective function defining the behaviour after acquisition of experience during the model lifespan (in this case 1 year). The parameters for the household's resource allocation for soil fertility improvement may change depending on farm performance. The farming household may transform (move to another farm type) when their objective function, therefore the resource allocation rules as well as the parameters, changes. For continuous cropping, the plant production function remains the same but parameters changes. In rotation, both parameters and production functions changes overtime.
Individuals success is a result of adaptive traits. Criteria used for ranking alternatives.	Agents exploit, control and consume entities and resources from their own plots and those of the surrounding environment to achieve organisational goal. Each farming household within the community is assumed to have specific state variables that enable them to make autonomous decisions regarding the improvement of soil fertility of their plots at a particular time. Proactively and opportunistically, take actions to achieve its goals given the dynamic and unpredictable environment. The households make decisions depending on the current and expected soil fertility of their farm bounded by the resources available with the aim of staying focused on achieving own objectives. Given many SLM options, the farmers allocate resources to either the one that maximises soil fertility improvement or the one that minimises risk of soil degradation.
Collective experience changing traits.	The individual farmer's actions influence others in the system although not explicitly modelled, but through social ties and shared landscapes, farmers tend to learn from and imitate those with the same typology.
Models for future conditions or consequences that individuals use for successful decision making	Perceptive: considered to have true scientific knowledge (or if information is limited, belief) of the environment. The household's perception (vision) of soil fertility though referenced by others in community, efforts to improve it are restricted by tenure to own plots and not the entire landscape. However, for landscape processes such as soil erosion on hill slopes, they also mainly recognise activities upslope and in rare cases downslope in case of an extending gully.
Sensing	
Internal, neighbours, and environmental state variables that individual's sense and consider in their behaviour. Local, networks or global levels. Information acquisition. Accuracy and uncertainty.	Households exploits the farms for ecosystem services that are dependent on soil status, radiation, rainfall and temperature. Humans controls the land through among many things, managing soil fertility and controlling soil erosion. Their search area is constrained by (im)mobility and, for the established settlements, by use rights. Therefore, the possibility of accessing and manipulating the existing environmental entities and resources is often bounded. Sensing is constrained within the agent class and also local neighbourhood (village boundary). The actions of the agents on real farms are non-deterministic with some degree of uncertainty. The main processes for soil fertility improvement and erosion control represent the environment. The choices of SLM technologies are depended on their knowledge about availability and performance of the technologies in the study area (Zambonelli et al., 2003).
Interaction	
Direct or indirect (e.g competition for mediating resource). Do they communicate?	Interactive: achieve goals by interacting with other agents in the environment the agent is situated. Individuals do not conform to set land management rules, but since some objectives of SLM are realised at larger community and landscape level, farmers co-operate with others through social ties or shared landscapes.

Guiding contents	ODD+D Model Description
	Given their socioeconomic capabilities and land potential farmers tend to mimic those with similar typology (Le, 2005). The interactions are bounded by village boundary and shared landscapes.
Heterogeneity	
Do the agents and landscapes differ	The study population is comprised of farmers with different demographic (sex, age), endowments (land, labour) and location (dwelling and plot ownership). The environmental attributes that are variable among pixels (plots) include topography, base soil fertility, and crop biomass/yield.
Stochasticity	
Processes assumed to be random. Used to reproduce variability for processes difficult to capture actual.	The remaining population and their plots are randomly allocated to non-sampled cultivated grids. Their attributes are drawn from the sampled households and plots to generate a population which mimics the distribution of statistics found in real farms. Some of the environmental attributes such as terrain are set to be static i.e. remain unchanged without the action of the agents. However, most of the entities and resources are active and dynamic with changes that are beyond the control/regulation of the individual agents. Geo-simulation of dynamic soil fertility is one required end but quite variable hence is randomly set in most cases using the estimated confidence limits.
Collectiveness	
Belonging or forming groups: defined by modeler or result of individual behaviours.	Households often apply soil fertility improvement technologies on individual plots as discrete entities. However, for technologies that aim at controlling soil erosion such as permanent vegetation cover, the environment needs to be viewed as a continuous entity with connected discrete entities. The households belonging to a typology have similar resource endowments that they use to pursue similar livelihood strategies (bounded by policies and institutions). The typologies are therefore used to initialise the agent population and households transition depending on resource accumulation or depletion by end of the activity calendar.
Observation	
Data collected from ABM	The simulated outputs, which is the projected development, are compared with the baseline and other past states in terms of changes in structure (e.g. distribution of farm types) and function (e.g. nutrient balance) of the farms.
Details	
Implementation	
How has the model been implemented	Using the multi-agent system platform Netlogo (Wilensky, 1999), the first step has been to adapt the Land Use Dynamics Simulator (LUDAS) modules (Le, 2005) for initialization, decision, agricultural production, and crop allocation. Empirical models and transfer functions have been used to frame local processes and estimate parameters for the study site.
Initialization	
Initial state of the model world at time 0 of a simulation run (exact or stochastically set).	Baseline soil nutrient levels, plot sizes, distribution and productivity, and farmer attributes, nutrient stocks, input and outputs (exact for the sample and randomly allocated for the rest). Ecological inputs and outputs estimated using transfer functions. For subsequent runs initial conditions are the same among simulations, the stochastic attributes are estimated using the random number generator with a certain confidence interval and random seed.
Input data	
External data files or models to represent processes	The external data and models are used to set initial state and for parameterization of processes. These include satellite imagery and parameters for the transfer functions. More details are provided in Table 3—2
Sub-models	
Detailed process overview and scheduling	Processes and sub-models are introduced in Figure 3.5, summarized in Table 3—3 and detailed in the respective methodology sections. The empirical results and estimations are systematically implemented in NetLogo using the main directory (subdirectory) structure (see Box S1 in the Appendix S3. The initialisation and static processes are one-step whilst the dynamic processes are set to run for a one-year cycle corresponding to the unimodal production season.

Table 3—2: Agent attributes, states and the processes affected during the simulation runs

Attribute	Attribute definition	Way the attribute will change in a simulation run (assumptions)	Linked sub-model(s)
Human agents			
$P_{SUBSIDY}$	Subsidized price of fertilizer ($-purchase/$-market)	Increase or decrease depending on alternative subsidy policy regimes	Fertilizer intensity Manure choice & intensity Legume choice & yield Tree-on-farm decision Maize yield
H_{AGEH}	Age of the household (HH) head (years)	Automatically increased for each time step	Household farmtype clustering
H_{EDULHM}	Education attainment of HH members (primary=1, secondary=2 or tertiary=3) computed as an index	constant	Fertilizer intensity Manure choice & intensity Legume choice & yield Tree-on-farm decision Maize yield
H_{LABOUR}	Total HH labour (man equivalent)	Constant	Fertilizer choice & intensity Manure choice & intensity Legume choice Tree-on-farm decision Household farmtype clustering
P_{LABOUR}	Labour invested on a plot (manhours ha^{-1})	constant	Maize yield Legume yield
H_{DEPR}	Dependency ratio = Number of workers / numbers of dependents that are below 16 and above 65	Constant	Fertilizer choice & intensity Manure choice & intensity Legume choice & yield Tree-on-farm decision Maize yield
H_{GENH}	Gender of household head (male=1, female=0)	Constant	Fertilizer choice
H_{WEAI}	Women empowerment in agriculture index	Constant	Fertilizer choice & intensity Manure choice & intensity Legume choice & yield Tree-on-farm decision Maize yield
H_{GMEM}	Group membership for HH members	Constant	Fertilizer choice Legume choice & yield Maize yield
H_{TLUN}	Tropical livestock units	Constant	Manure choice Tree-on-farm decision
H_{COMM}	Monetary value of phone and radio expressed as an index	Constant	Fertilizer, Manure and Legume choice Household farmtype clustering
H_{TRAN}	Monetary value of bicycles and wheelbarrows	Constant	Fertilizer choice & intensity
H_{HECT}	Land cultivated and managed (ha)	Constant	Fertilizer choice
H_{INCC}	Income from cash crops ($ ha^{-1})	Constant	Manure intensity
H_{INCL}	Income from livestock sales ($ ha^{-1})	Constant	Manure intensity Household farmtype clustering
H_{INMS}	Total annual income of the household ($ ha^{-1})	Constant	Fertilizer intensity
Ecological agents			
P_{HECT}	Size of the plot (ha)	Constant	Fertilizer choice & intensity Manure choice & intensity Legume choice and yield Tree-on-farm decision Maize yield
$P_{CULTYRS}$	Period the plots has been under cultivation (years)	Automatically increased for each time step	Fertilizer choice & intensity Legume choice & yield

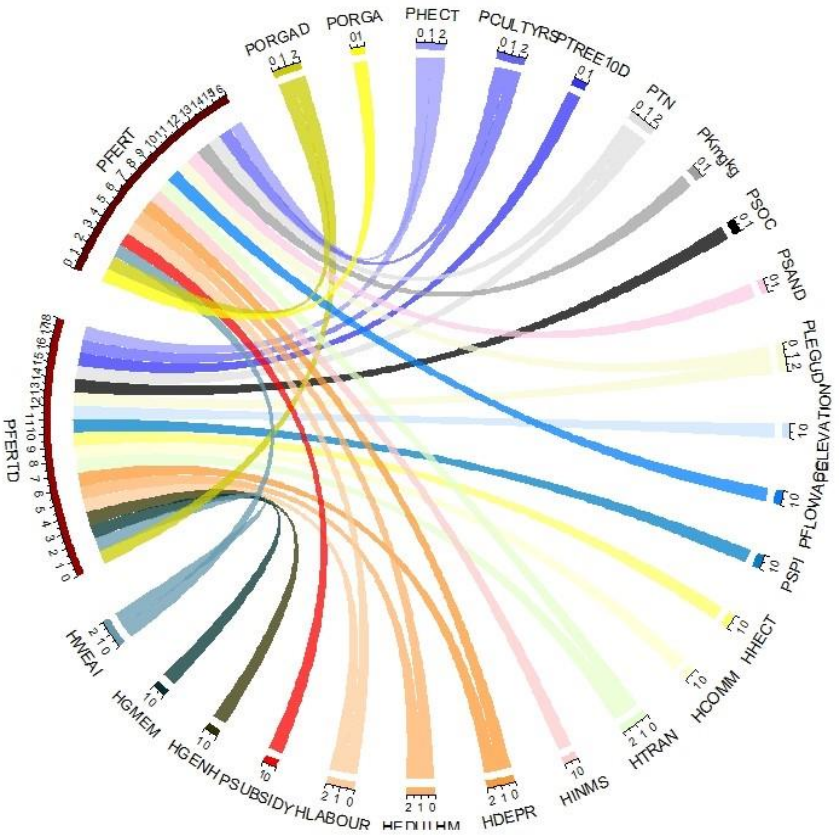
Attribute	Attribute definition	Way the attribute will change in a simulation run (assumptions)	Linked sub-model(s)
			Tree-on-farm decision Maize yield
P_{LEGUD}	Cropping Dummy (legume=1 & non-legume=0)	Probabilistic updated using logistic model in response to drivers of choice	Fertilizer choice & intensity Manure choice & intensity Tree-on-farm decision Maize yield Erosion & sedimentation
$H_{Leguarea}$	Plot area under legume integration (ha)	Based on cropping choice	Household farm-type clustering
P_{ORGAD}	Dummy for manuring (yes=1, no=0)	Probabilistic updated using logistic model in response to drivers of choice	Fertilizer choice & intensity Legume choice Tree-on-farm decision
P_{ORGA}	Amount of manure applied (kg/ha)	Deterministic updated using GLM	Maize yield Legume yield Leaching
H_{ORGA}	Total manure applied by the household (kg)	Computed from P_{ORGA}	Household farm-type clustering
P_{FERTD}	Dummy for fertilization (yes=1, no=0)	Probabilistic updated using logistic model in response to drivers of choice	Manure choice Legume choice Tree-on-farm decision
P_{FERT}	Amount of fertilizer applied	Deterministic updated using GLM in response to drivers of intensification decisions	Manure intensity Maize yield Legume yield Leaching
H_{FERT}	Total amount of fertilizer applied	Computed from P_{FERT}	Household farm-type clustering
$P_{TREE10D}$	Tree cover (>10%=1, <10%=0)	Probabilistic updated using logistic model in response to drivers of choice	Fertilizer choice Manure choice Legume choice Maize yield
H_{SWC}	Number of soil and water conservation technologies	Constant	Household farm-type clustering
P_{SLOPE}	Surface slope of the land pixel	Assumed to be static over time	Erosion & sedimentation Manure choice Legume choice
P_{MZkg} P_{LEGUkg}	Crop yield of the land pixel as nutrient output pathways	Deterministic updated using GLM in response to drivers of yield functions	Maize yield Legume yield Nutrient balance
$P_{SAND\%}$	Percentage of sand indicating land quality	Constant	Fertilizer intensity Manure choice & intensity Legume choice Maize yield Residue input Erosion & sedimentation Gaseous loss
$P_{COARSE\%}$	Percentage of course fragments indicating land quality	Constant	Maize yield
$P_{TN\%}$	Nitrogen content indicating indigenous nutrient stocks	Baseline plus or minus the balance for 10 cm top soil	Fertilizer choice & intensity
$P_{SOC\%}$	Soil organic carbon content indicating land quality	Baseline plus or minus the balance for 10 cm top soil	Manure intensity Legume choice Tree-on-farm decision Maize yield Erosion & sedimentation Gaseous loss
P_{Pmgkg}	Phosphorus content indicating indigenous nutrient stocks	Baseline plus or minus the balance for 10 cm top soil	Legume yield
P_{Kmgkg}	Potassium content indicating indigenous nutrient stocks	Baseline plus or minus the balance for 10 cm top soil	Fertilizer intensity Maize yield

Attribute	Attribute definition	Way the attribute will change in a simulation run (assumptions)	Linked sub-model(s)
$P_{ELEVATION}$	Elevation above sea level (m) for topographic position	Constant	Fertilizer choice Tree-on-farm decision Maize yield Legume yield
$P_{FLOWACC}$	Flow accumulation for topographic position	Constant	Fertilizer intensity Maize yield
P_{SPI}	Stream power index for topographic position	Constant	Fertilizer choice
P_{Ri}	Roughness index	Constant	Erosion & sedimentation
P_{RAIN}	Average annual rainfall	Constant	Residue input Atmospheric deposition Leaching Gaseous loss
P_{TEMP}	Average annual temperature	Constant	Residue input Gaseous loss
P_{SLOPEM}	Slope length	Constant	Erosion & sedimentation
$P_{UPSLOPE}$	Upslope contributing area	Constant	Erosion & sedimentation
$P_{CLAY\%}$	Soil clay content	Constant	Legume yield Erosion & sedimentation Leaching
P_{BD}	Soil bulk density	Constant	Erosion & sedimentation Gaseous loss
P_{CEC}	Soil cation exchange capacity	Constant	Leaching
Other variables			
P_{BOUND}	Boundary of the study area	Constant	Study boundary
$H_{VILLAGE}$	Village boundary	Constant	GIS proportional upscaled plot population
$P_{PROD\text{-}INDEX}$	Productivity index	Not updated despite some of its input variables such as nutrient stocks changing.	GIS proportional upscaled plot population
H_{CLUS}	Farmtype	The farm types remain unchanged despite the input levels changing. The type specific predictions and simulations are based on initial household profiles.	Input level specific upscaled plot population
P_{UNCULT}	Uncultivated patches	Land use remains the same but cover is assumed to change as reserved areas regain woody vegetation	Mask of uncultivated patches Erosion & sedimentation
P_{CULT}	Cultivated		

Table 3—3: Sub-models for nutrient input and output flows

Sub-model	Agent controlling Description of main task	Key functions	Input variables and linked sub-model(s)	Output variables and linked sub-model(s)
GENERATE REMAINING POPULATION	Developer upscaling based on agent attributes and states	Monte Carlo approximation & GIS-based proportional up-scaling	sample data. used to generate the population and landscape attributes and use patterns	Plot population of 2640 from the 451 sampled
DECISION1 (Choice of soil fertility management (SFM) including Fertilizer, Manure, Legume & Tree-on-farm)	Controlled by household agent Computing choice variable based on own livelihood profile, land productivity, complementary technologies and policies.	Bi-logit	Policies (input subsidy) Natural capital (farm size, soil nutrient stocks, topographic position, land quality, crop land, livestock) Livelihood profile (labour, gender, age, education, communication, transport, income) Social capital (group membership) Linkages among complementary SFM practices	Predicted probability to (or not to) apply fertilizer and manure, grow legumes or retain trees on farm
DECISION2 (Intensify SFM including Fertilizer & Manure)	Controlled by household agent Computing intensify variable on own livelihood profile, land productivity, complementary technologies and policies.	Generalized linear model (GLM)	Policies (input subsidy) Natural capital (farm size, soil nutrient stocks, topographic position, land quality, crop land, livestock) Livelihood profile (labour, gender, age, education, communication, transport, income) Linked to DECISION1 and to complementary SFM practices	Deterministic prediction of potential to increase or decrease application of fertilizer (kg ha^{-1}) and manure (kg ha^{-1})
CROP-YIELD (maize & legume)	Controlled by household agent, land parcel and their interactions Computing yield of the plot given the household livelihood profile, land productivity, SFM technologies and SFM policies	Generalized linear model (GLM)	Policies (input subsidy) Natural capital (farm size, soil nutrient stocks, topographic position, land quality, crop land, livestock) Livelihood profile (labour, gender, age, education, communication, transport, income) SFM practices Social capital (group membership)	Deterministic prediction of potential to increase or decrease maize yield (kg ha^{-1}) and/or legume yield (kg ha^{-1}).
EROSION	Controlled by land pixel (patch) Computing net soil loss in the land pixel based on main factors driving soil erosion	RUSLE adjusted by sedimentation delivering ratio. NSL = R*K*LS*C*P* (1 - SDR)	R-factor: estimated from empirical relationship of erosivity with annual rainfall K-factor: empirically estimated from erodibility defined by soil texture, structure and permeability LS-factor: empirically estimated from topographic contributing area and steepness. C-factor: empirically defined by land cover type P-factor: set as 0.65 which is commensurate with the existing soil conservation measures	Gross soil loss Net soil loss (NSL) Net soil deposited (kg ha^{-1} yr^{-1})

Sub-model	Agent controlling Description of main task	Key functions	Input variables and linked sub-model(s)	Output variables and linked sub-model(s)
			SDR: empirically derived from flow direction, flow length and surface roughness	
Nutrient balance (NPK and SOC)	Controlled by household agent, land parcel and their interactions Computing partial and full nutrient balance for the centroid pixel of the plot given the input and output flows	Transfer functions taken from NUTMON and FarmDESIGN and re-implemented	IN1: NP empirically predicted from DECISION1 & DECISION2 for fertilizer IN2: NPK & SOC empirically predicted from DECISION1 & DECISION2 for manure and the incorporated crop residues and roots IN3: N derived from atmosphere in incorporated legume residues and roots estimated after DECISION1 and from CROP-YIELD IN4: NPK and SOC estimated using sediment delivery ratio from EROSION model IN5: NPK computed from wet and dry atmospheric deposition in Malawi OUT1: NPK computed from CROP-YIELD model OUT2: NPK computed for removed residues derived from CROP-YIELD model OUT3: NPK and SOC computed in net soil loss OUT4: NK estimated from empirical relationship of leaching with rainfall, texture, CEC, nutrient inputs and decomposition of soil organic matter (SOM) OUT5: N empirically estimated from relationship with rainfall, inputs and SOC C empirically estimated from SOM degradation	NP from fertilizer NPK&SOC from organic inputs N fixed by legumes NPK and SOC in deposited sediments NPK from atmospheric deposition NPK in crop yield NPK in removed residues NPK and SOC in soil eroded NK leached N denitrification and volatisation SOC from SOM degradation



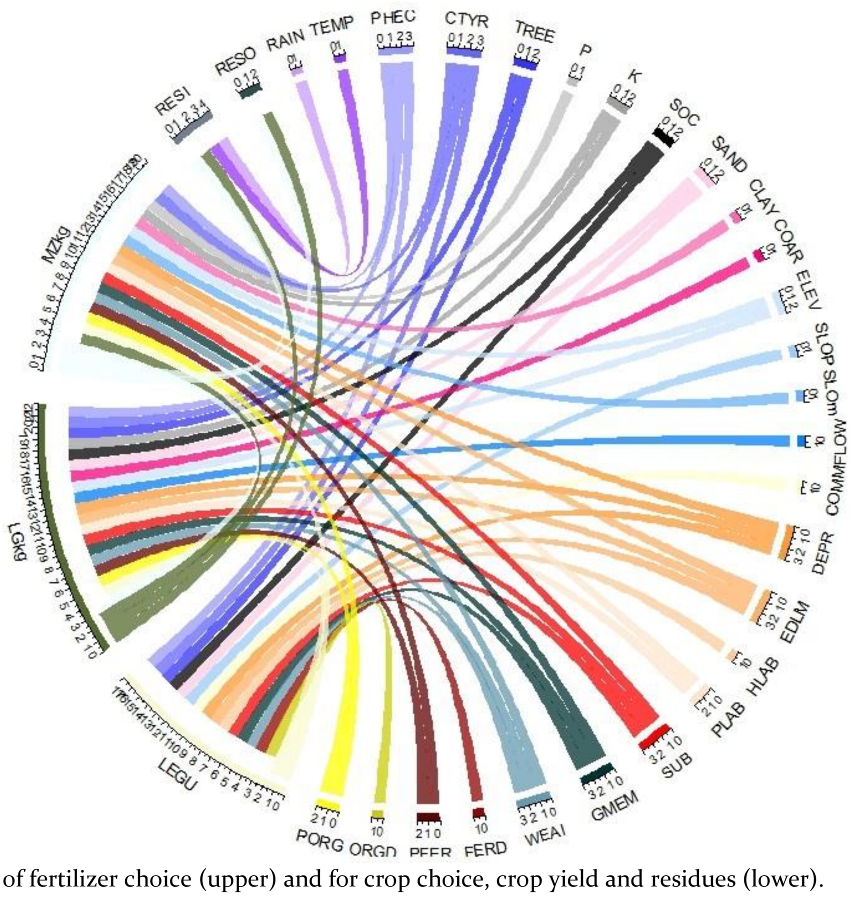


Figure 3.2 Drivers of fertilizer choice (upper) and for crop choice, crop yield and residues (lower).
NB: The flow has the same colour as the driver; and there's a white space between flow and the target variable. The definitions of the variables are presented in Table 3—2.

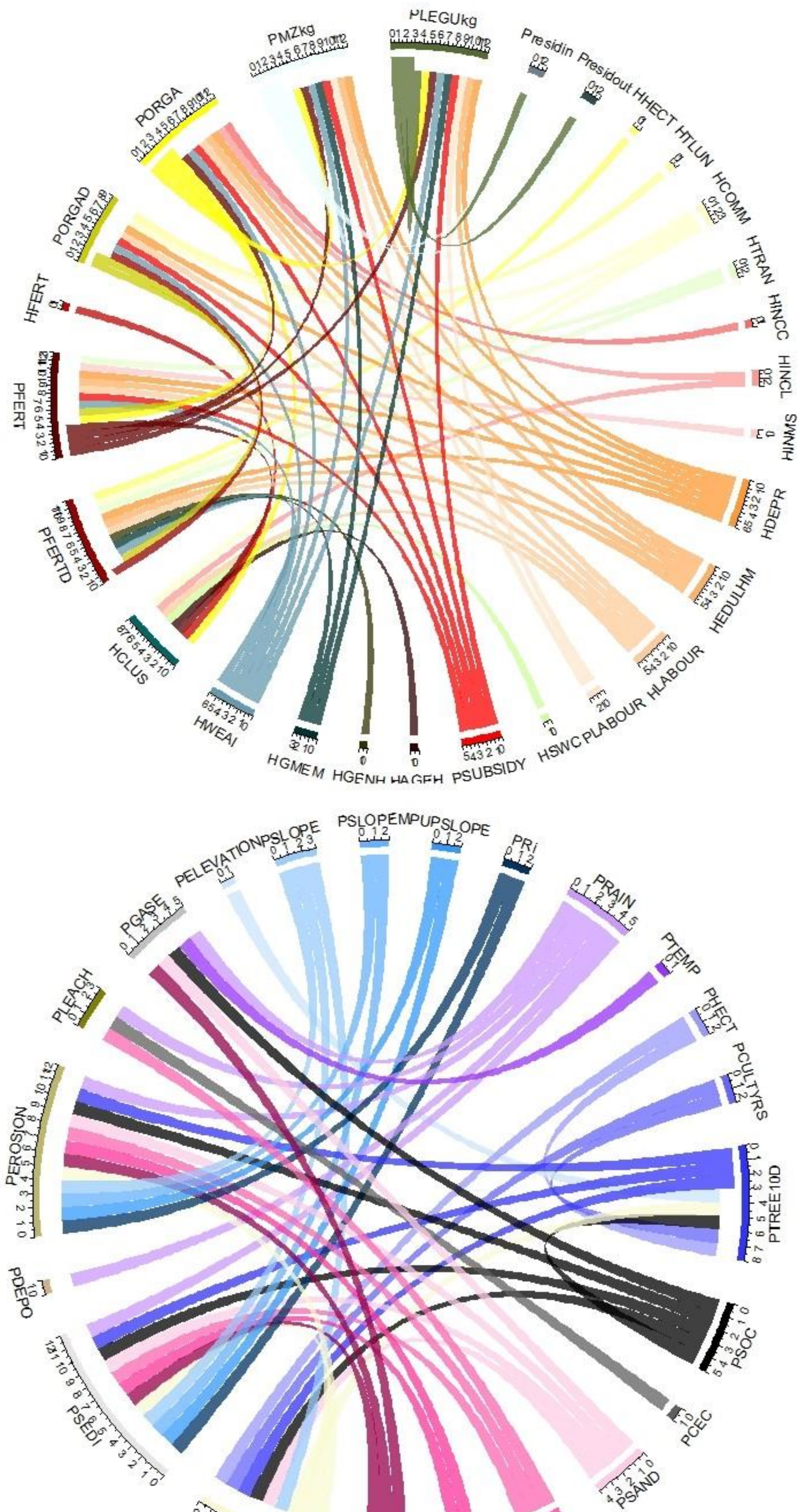


Figure 3.3 Human (upper) and ecological (lower) systems depicting drivers and outcomes

NB: The flow has the same colour as the driver; and there's a white space between flow and the target variable. The definitions of the variables are presented in Table 3—2.

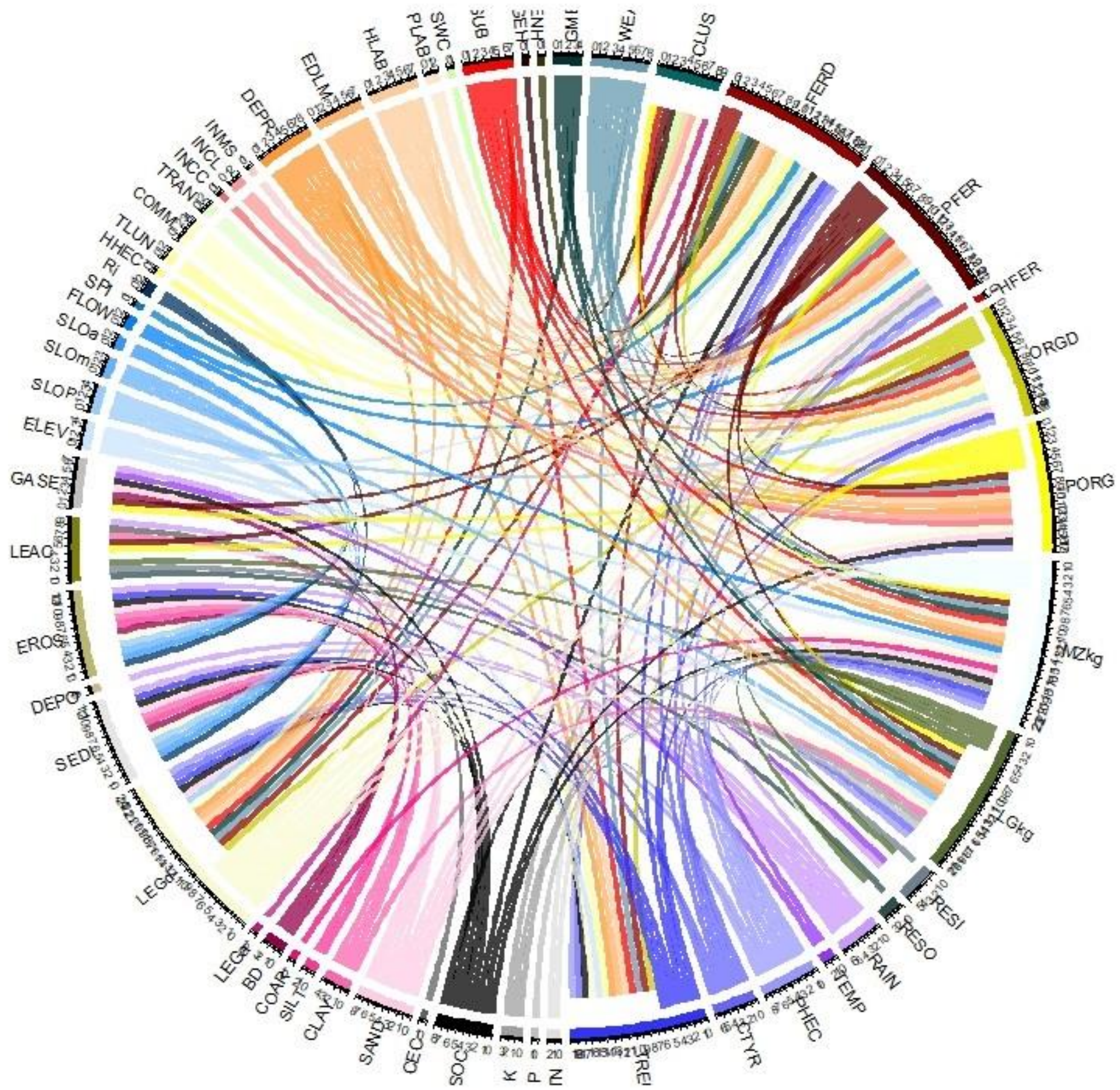


Figure 3.4:Integrated Human and ecological systems showing interlinkages among drivers and outcomes

NB: The flow has the same colour as the driver; and there's a white space between flow and the target variable. The definitions of the variables are presented in Table 3—2 (here without suffixes H_ for human and P_ for plot and abbreviated.

3.1.1 MASSAI structure and workflow

The study adopts the modules in the Land Use Dynamic Simulator (LUDAS) developed by Le (2005) for further customisation to the characterisation of the study site within the maize mixed farming system and the overall policy and governance setting of the community. Although the customisation inherits the concepts, module organization and principal codes of the LUDAS framework, further work was done to: (1) build new modules (e.g. nutrient balance) and (2) specify across all module's relevant variables, empirical parameterisation and verifications. The LUDAS falls into the class of MAS models (Bell et al., 2015), where both the human population and the landscape environment are all defined by interactive agents.

To initialise, simulate and validate the model, 8 steps are undertaken (Figure 3.5). The first two stages have been covered identifies the gaps and proposes a working solution. Stages 3 and 4 involves data and parameter generation for setting initial conditions upon which simulations of expected changes are built.

The components of land that regulates provision of ecosystem goods and services for human wellbeing including humans, topography, climate, soil, and the main land cover and management practices have been captured (stage 3 of Figure 3.5). These facets of the environment embody a wide range of processes that can be hierarchically modelled at different spatial and temporal scales (Le et al., 2008).Taking stock of existing indicators and metrics for sustainable agricultural intensification, Smith et al. (2015) gave an overview of agroecosystem functions at three spatial scales of plot, household and community. Although the basic processes such as plant growth and decay are largely regulated by environmental factors, the states of factors are altered by human actions resulting in trade-offs and synergies between ecological, economic and social benefits (Smith et al., 2017). In this study we attempt to integrate the multiple ecosystem services at the scale of a managed plot (Table 3—4).

Table 3—4: Sustainability domains, agro-ecological functions and their indicators

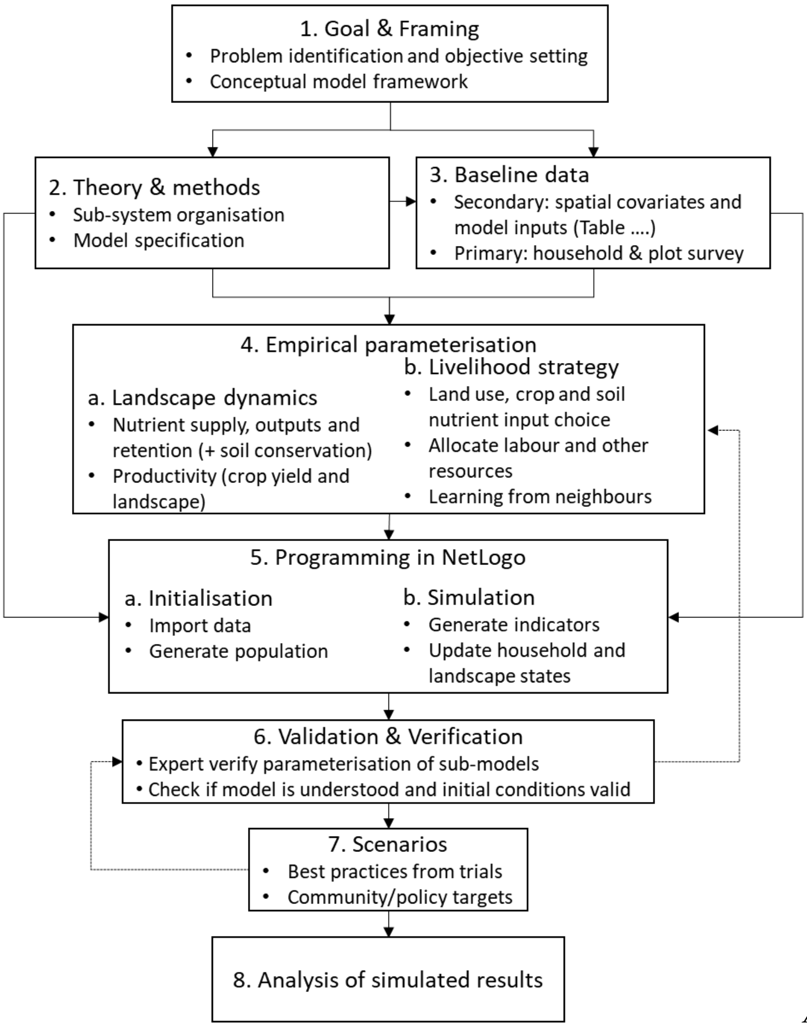
Sustainability domain	Agro-ecosystem function	Indicators	Metrics
Productivity	Food production	Biological inputs (manure and legume)	Farm generated inputs used (kg/ha)
		Input intensity (fertilizer)	Fertilizer (kg ha^{-1})
		Yield	Product (kg ha^{-1})
		Quality and macro nutrients of yield	NPK in product (kg ha^{-1})
	Raw material production	Residue quality (macro nutrients)	NKP in residues (kg ha^{-1})
		Residue biomass	Carbon produced (kg ha^{-1})
Environmental	Nutrient cycling	Nutrient cycling (full budget)	Total NPK import – total NPK export
		(partial budget)	NPK applied – NPK export in grain
	Erosion control	Erosion rate	Top soil lost (ton ha^{-1} year^{-1})
		Sediment delivery	Sediments deposited (ton ha^{-1} year^{-1})
		Aggregate stability	Structural stability index
	Greenhouse gas regulation	C sequestration	Soil organic carbon (ton C ha^{-1})
		Chemical input reduction	Inorganic fertilizer replaced (kg ha^{-1})
Social	Technology adoption	Adoption rate	% households adopting[c]
	Equity	Yield gap	Locally attainable yield – actual yield
		Income distribution	Lorenz curve and *Gini* coefficient
	Empowerment	Bargaining power	Women empowerment in agriculture index[h]
Economic	Agricultural income	Crop value	Product ($ ha^{-1})
		Labour intensity	Man-hours (time/ha) ~ ($ ha^{-1})
		Capital intensity	Expenses ($ ha^{-1})
		Profitability	Product ($ ha^{-1}) – expenses ($ ha^{-1})
		Productivity	Product ($ ha^{-1}) / expenses ($ ha^{-1})

[c]Community scale; [h]Household scale source: (Smith et al., 2017)

Achieving SAI is a broad goal aimed at ensuring efficient use of resources for improved productivity, with equal attention to equity and environmental services (The Montpellier Panel, 2013). Striking the balance is still considered as one of major research and development challenges since it requires an understanding of how ecosystems function under changing environment and socio-economic constraints (Matson et al., 1997). As Schreinemachers and Berger (2011) put it, sometimes a higher productivity in terms of yield is not the main goal for small-scale farmers. This is the case because smallholders are overly constrained by multiple factors including labour, capital, fertilizers, pesticides and herbicides and sometimes by insecurity over land that they have to overcome. Researching options for addressing these constraints is methodologically challenging with regards to (1) the complex human-environment interactions among factors, (2) uncertainties caused by that complexity, (3) the long-term perspective of sustainability research, and (4) externalities and trade-offs over space, time and social groups (Boulanger & Bréchet, 2005).

The study is built on the hypothesis that the agents (in this case smallholder farmers) despite being autonomous, interact as they share the environment and the observed phenomena is the aggregate emergent outcome. Therefore, farming can be considered to be an organised complex system and its analysis requires a hybrid of analytical hypothesis driven experiments used for organised simple systems and statistical data driven observational studies used for complex unorganised systems (Le, 2005). Although an empirical MAS for simulating complexity including the continuous flow of multiple environmental services, the platforms are built by integrating modules based on the principle of *ceteris paribus*.

The individual modules (Stage 4 of Figure 3.5) were parameterised using empirical models of technology choice and adoption, productivity, soil erosion control and nutrient retention using from household(field) data. Studies of these processes in the study area used limited sample sizes and different sampling frames (Chimtengo et al., 2014; Haile et al., 2017; Mponela et al., 2018; Mungai et al., 2016; Signorelli et al., 2016; Silberg et al., 2017; Timler et al., 2014). However, they provide useful guidelines for sampling and variable choices. Data and associated processes for the MAS need spatial signatures (Berger, 2001).



Adapted from Le (2005)

Figure 3.5 Overview of the MASSAI implementation process

3.1.2 Common sampling frame, methods and sources for environmental and household data

For integration of human and ecological sub-systems, the study landscapes and communities have been drawn using a common geo-referenced sampling frame (Berger & Schreinemachers, 2006). The farming units for ecological evaluations were spatially co-located and nested to the households that own or manages them using geographic information system (GIS). The framework for human sub-system includes the farming households within the five villages. The ecological framework was principally the land parcels within the villages but also surrounding landscapes that are managed or accessed by farmers from the study villages. Existing spatial data about soil types (Lowole, 1965), land cover and use (FAO, 2012) and topography was used for landscape stratification and sampling. The list of households for sampling of households and their farms was obtained from the extension planning area and village leaders. The study households, plots and sub-watersheds were drawn using multi-stage sampling. Since there is no record of ownership of plots or maps showing plot locations, first to be sampled were the households. Nested within the households, the plots belonging to those sampled were then enumerated for ecological studies.

Within the study area the households and their farmlands are clustered into traditional allotments called villages. Mostly, the households within each village are related by descent and are ruled by the village chief who come from the royal family lineage. The villages have clustered settlement pattern and generally separated by the physical features such as streams, hills and sometimes farmlands. Being a typical rural area, most of the land is owned and managed by the households that are resident in the study villages. During the time of the study, there were no records of plot ownership hence to link households with plots and other resources, households were sampled first then followed by enumeration of their plots. The boundaries of the plots for the surveyed households were mapped using a GPS. For the unsampled households, there could be potential conflicts and misunderstandings to arbitrary assign ownership (Deininger & Xia, 2017). Therefore, the plots for the remaining population were assigned to the remaining cultivated areas using proportional upscaling with the sample distribution and restricted within the respective village boundaries.

The basic unit for analysing ecological processes is the plot/pixel (considered as the *landscape agent*) that provides agro-ecosystem services and is directly managed by the farming household and the neighbouring community. Farm surveys of soil, biomass and crop yield, nutrient inputs and outputs were carried out to capture current farm performance, soil quality and their determinants. The household learning and technology usage surveys was nested within the ecological sampling units by using the common sampling frame. The study households were randomly sampled from a village list available at the village leader. This gives a hierarchical subject structure with households and their landholdings at lower level, constituting the village population and landscapes, and the entire study area is at the highest aggregate catchment level. The household choice was dynamic as no prior attributes were available and allowed for a plausible understanding of household's strategies (Tittonell et al., 2010).

Data were collected on households and their plots and the environmental attributes. Household surveys were conducted to generate socioeconomic data of the households' resource status and behaviour supporting seed production. Data was sourced on farmers wealth attributes, socioeconomic profile, how they carried out previous agronomic practices and marketing. Topographic data (slope, upstream contribution area etc) were extracted from the Shuttle Radar Topography Mission (STRM) 30m resolution digital elevation model (DEM) that was downloaded from United States Geological Survey Website (USGS, 2018).

3.2 The people, livelihood resources and soil fertility management strategies.

3.2.1 Case study area and sampling methods

This study was conducted among the households practicing smallholder farming in the villages of Malaswa, Phikani, Amosi and Hiwa in Nsipe Agricultural Extension Planning Area, Ntcheu District of Malawi (Figure 3.6). During the focus group discussions (Braslow & Cordingley, 2016), it was revealed that the people of these villages have related ancestors. The ancestors, about 6 generations, settled along the fertile basin sandwiched by two major streams Nsipe and Riviridzi. As population grew, some of their 3rd generation parents moved to the hills forming Malaswa village and recently, another group have moved further up forming Kaombe village which has fewer households and is still under the jurisdiction of village head Malaswa.

As of 2015, the average population density from village records was 137 persons/km² (Emerton et al., 2016). The villages on flat areas had higher density such as 397 persons/km² in Amosi while those situated in hills such as Malaswa had lower density of 83 persons/km². The average household size was 5.2 members. On average, the household cultivated 0.9 ha of land fragmented into 2 or more small plots (Mungai et al., 2016). Considering that farming is the main livelihood strategy, and low productivity levels the average land holding size per capita of 0.17 ha is insufficient. In order to improve land productivity, farmers in the study area use a variety of soil productivity enhancing technologies. Combined use of inorganic fertilizer and farmyard manure, residue incorporation, grain legumes, trees on farm are common for all whilst those with relatively larger land holdings rotate staple cereals with legumes and cash crops (Mponela et al., 2016). The main crops grown include maize, tobacco, groundnuts, soybean and sweet potatoes. Most households own small ruminants (goats and pigs) and chickens as a source of income. They supplement their food, housing and income needs by collecting natural resources and engaging in *ganyu*, a paid daily labour. Fuelwood obtained from woodlands, forests and croplands is the main source of energy for cooking. The production of charcoal has depleted fuelwood resources leading to negative fuelwood budgets for the households (Braslow & Cordingley, 2016).

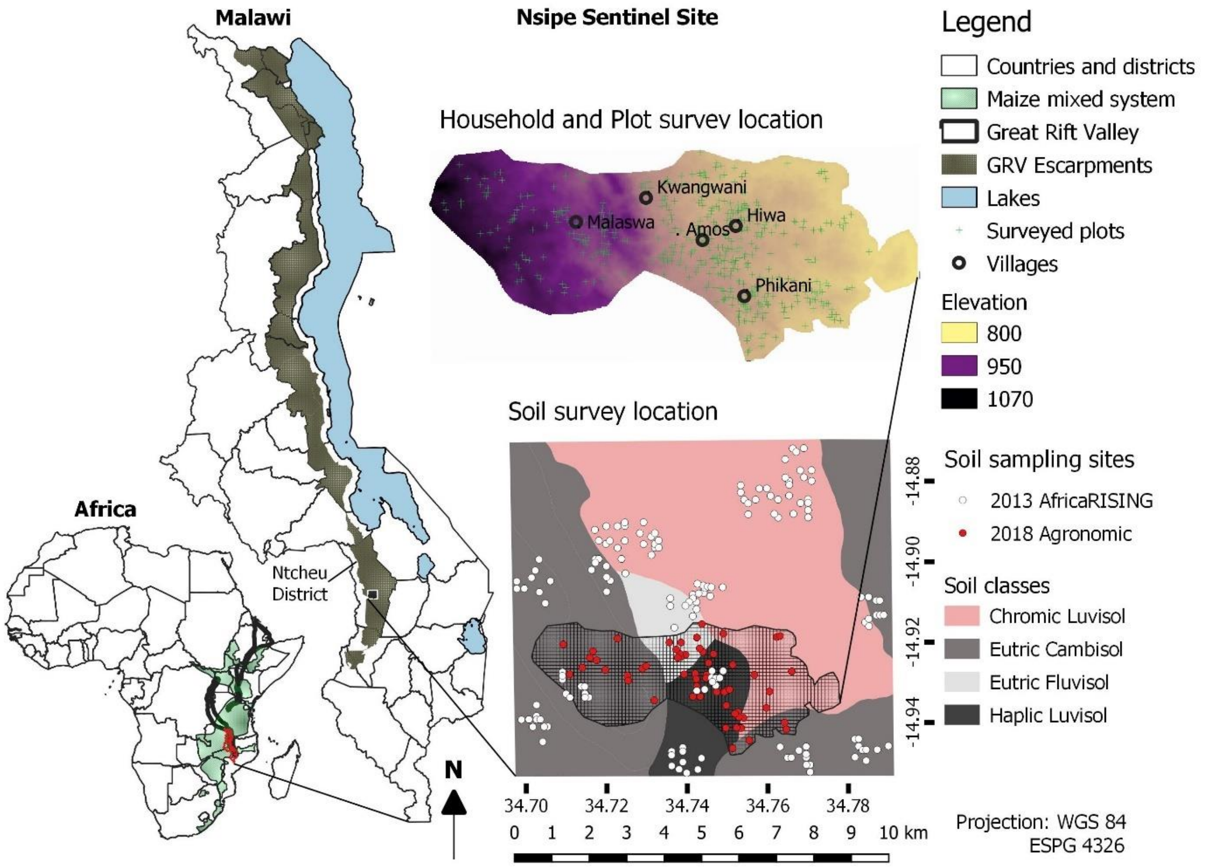


Figure 3.6 Location of study area within Ntcheu district of Malawi, southern Africa.

The sampling frame comprised of smallholder farming households in the five adjoining study villages (Malaswa, Amosi, Hiwa, Phikani, and Kwangwani). The main defining feature of smallholding is that farmers own and manage small plots of up to 5 ha and produce mainly for subsistence (Anseeuw et al., 2016). Study households were randomly sampled from a list made available by the respective village leaders. This gives a hierarchical subject structure with households and their landholdings at lower level, constituting the village population and landscapes, and the entire study area is at the aggregate catchment level. The random sampling of households was one as no prior attributes were available and therefore allowed for a plausible understanding of households' SFM strategies (Tittonell et al., 2010). To ensure that the estimates were representative of the study population, at least 30% of the households in each village were sampled.

3.2.2 Human livelihood analytical framework

As discussed above, earlier studies took either structural or functional approaches (Tittonell et al., 2010) but recently these have been integrated (Kamau et al., 2018). Understanding livelihood strategies indeed requires attention to the interactions among household and landscape state variables and the ways in which they may be clustered, sequenced or substituted to enable different livelihood production strategies (Scoones, 1998). As individual households pursue their goals, they are bounded by available assets and can be classified into distinct categories. One of the approaches commonly used to do so is through building typology (Le, Seidl, et al., 2012). Household types emerge as household tend to categorise itself into the most similar type, based on comparing and ranking dissimilarities in state variables of itself and its environment, with those of its neighbours.

The heterogeneity among farmers is a result of several factors. Therefore, it is imperative to identify a small set of variables that explain most of the variability. Principal components analysis (PCA) has been widely used to empirically identify the main factors differentiating the households (Douillet & Toulon, 2014; Kamau et al., 2018; Le, 2005). The PCs are optimal linear combinations of initial variables explaining the variance in descending order. As such, PCA enables reduction of a larger number of initial household and farm variables into a smaller set without losing important defining information (James et al., 2013; Le, 2005).

The PCA has been found to be influenced by the magnitude and scale of each variable hence variable scaling and centring at zero were done prior to conducting PCA (James et al., 2013). The observed and measured values for the selected variables were normalised by Z-scoring and centred so that they could be drawn to the same axis and used to characterise household types. The Z-scores were computed in excel as: $z_i = \frac{x_i - \overline{x}}{SD}$. Secondly, since data is for individual households, there is potential that variables have some degree of association. There is generally a problem of factor indeterminacy. To address this and achieve a simple structure, where important variables have high loading on single PC and lower loadings on all others, the Varimax orthogonal rotation and Kaiser Normalisation were used (Kaiser, 1958). The factor loadings/weights signify the importance of the variable for a particular component and those with highest loadings are most distinctive.

Since PCs are independent, variables with highest loading on each PCs are used in subsequent cluster analyses thereby addressing the problem of multi-collinearity (Naes & Mevik, 2001). With 238 households and 25 variables, the variable to sample size ratio was 9.52. The cut-off to select the number of PCs was determined by use the eigenvalues greater than 1.0. The meta-analysis of more than 800 substantive factor analysis studies found that mean percentage accounted for is around 54% for 21-30 variables, 55.6 for sample size of 176-300, and 58.2 for 6 or more principal components (Peterson, 2000).

Subsequently, determination of household types was done using the K-mean cluster analysis (K-CA) in R (Lesschen et al., 2005; R Core Team, 2018). K-mean is the most commonly used clustering algorithm which was developed by Mac Queen in 1967 and it is the most effective even for small data sets (Ghosh & Kumar, 2013; Le, 2005). Basically, K-Means clustering is a partitioning method that treats observations of the data as objects based on locations and distance between various input data points. Partitioning the objects into mutually exclusive clusters (K) is done in such a way that objects within each cluster remain as close as possible to each other but as far as possible from objects in other clusters. Each cluster is characterized by its centre point i.e. centroid and each of the records is assigned to the nearest cluster centre. A centroid is the point whose coordinates are obtained by means of computing the average of each of the co-ordinates of the points of samples assigned to the clusters. The K-means converges to one of many local minima because it minimises distance measure between each data and its nearest cluster centre thereby minimising the intra-cluster variances while maximising the intra-cluster distances. K-CA maximises the sum of the squared error (SSE) by measuring the total squared Euclidean distance of observations from the cluster centroids.

With highly heterogeneous dataset, number of centroids could be large resulting in several clusters. Optimal number of clusters k were determined using the knee method, with optimal k value on the inflexion point/bend on the curve of sum of distances of clusters from the centroid against the number of clusters (Salvador & Chan, 2004). To determine whether the clusters are conceptually and statistically distinguishable, the differences in state variables were tested using least significant differences (LSDs) after analysis of variance using unbalanced structure. The unbalanced structure run in Stata ensures complete fit of unbalanced sample sizes between clusters (StataCorp, 2017). After distinguishing the clusters, a meaningful name or label for each cluster was assigned using variables that adequately reflects the objects in the cluster.

3.2.2.1 Data sources, types and pre-processing

As pointed out in section 3.2.2 the household and plot attributed that are assumed to have an influence on the pursuit of livelihood strategies of the study population within its environment were surveyed. When choosing the variables for inclusion in the construction of farm types, it was assumed that they do not undergo rapid change/shift within the short to medium term. It is then further assumed that the household types arising from the interactions among these variables tend to be stable over the growing calendar year and whose changes may be inferred in the foreseeable future. This is crucial for types to be functionally different from one another and households/land parcels may be assigned to the closest cluster centre based on overall nature of the attributes rather than on a few unstable variables.

The sustainable agricultural practices transcend the plot boundaries and take long time to show positive results hence are information intensive. Education and the age of the household head as a decision maker has been used as a proxy for differences in knowledge of the technical aspects of the technologies and ability to engage in community discussions. In addition, the group membership has been used as indicator of community cohesion which is essential in pursuit of livelihood strategies that require co-ordinated efforts. Moreover, most of the sustainable agricultural practices are not only information intensive they are also labour demanding.

Gender has also been found to be a social capital with regard to decisions in farming. A women decision making index ($H_{decision}$) averaged over 5 main household production, resources and income decisions including food production, cash cropping, livestock, off-farm engagement, marketing of produce and reinvestment is used (Alkire et al., 2012). The respondents were asked about the role of women in idea generation, engaging in discussions and/or making ultimate decisions. For households with both male and female members the scale was 0 (no female involvement), 0.5 (jointly make decision with men in the household) or 1 (female members decide). Research has shown that joint decision making ensures higher involvement of both gender groups and has overall higher efficiency. However, the WEAI was used because for farming communities most short-term decisions regarding re-investments are made by men. In so doing, WEAI shows the bargaining power.

Physical assets include infrastructure, production equipment and technologies. Assets are relatively slow changing variables as they are usually accumulated over time and last longer. However, assigning values to assets is challenging as they lack comparability and mutual substitution. In most rural areas, the value of equipment, depreciation rates and their relative contribution to household livelihoods are largely unknown making it more complex. We used the monetary weighting approach where the number of assets was multiplied by their mean prevailing market prices (Moser & Felton, 2007).

Natural resource stocks and environmental services from which resource flows and services useful for livelihoods are derived considered include livestock units, natural resources and land holding production orientation. Within a largely agrarian society, access to land for agriculture and what people do with that land are the most important natural attributes. As is the case in most farming systems, household production orientations do not vary in the short to medium term. As a result, income derived from crops, livestock and wild collections have been used as proxies for production orientation and livelihood strategies. The livestock which comprise mainly of poultry and goats but fewer pigs and cattle were converted into standard livestock units (LU) using nutritional and feed requirement factors for sub-Sahara Africa (FAO, 2005).

Financial assets are the capital base in terms of cash, credit/debt, savings, and other which are essential for the pursuit of a livelihood strategy. Sell of selling livestock and livestock products is one of most coping strategies among households in times of risks and shocks. Therefore, households that sell livestock could be assumed to be resilient. Non-farm sources captured included formal employment and small trade, sell of *ganyu* labour and remittances from relatives.

3.2.3 Human decision analytical methods

SFM are promoted to address the very basic challenge of soil fertility decline which should be reversed if farmers in these fragile landscapes are to benefit from other technologies (Vanlauwe et al., 2015). Every year, at the time of planning to use the SFM technology or not, a household is assumed to derive utility from expected soil-productivity gains, which is conditioned by resource endowments. This paper adopts the sustainable livelihoods framework and assumes that in order for households to pursue their agricultural livelihoods strategies, in vulnerable rural contexts, they strive to add value to land as their base natural capital using the prevailing resource endowments.

We hypothesise that households with increasing labour, in which women are empowered to make farming decisions, and have more income and farm equipment would use SFM in increasing intensities. However, in rural areas with unskilled labour, poor soils and variable production, the elasticity of substitution among production assets is deemed to be high (Gavish & Kalay, 1983). For instance, much as larger land sizes would be associated with increasing input use and productivity (Dorward, 1999), inverse farm-size productivity relationships have been observed since early 2000s (Matchaya, 2007).

As is the case with most smallholding systems of Africa, a significant fraction of the community does not apply inorganic fertilizer and manure inputs a good number do not grow legumes. Therefore, the data contain zeros and is assumed to be continuously distributed over the positive values. Considering the low levels of nutrient inputs and increasing efforts by the government and non-governmental organisations to promote soil fertility enhancing technologies, we assume that the zero observations emanate largely from non-participation decisions. Therefore, for a household to be considered a participant, it has to cross two *hurdles* namely to (1) *choose* then (2) *intensify*.

Several empirical models are used to analyse the truncated *choice - intensify* phenomena. We adopt the disaggregated model by Cragg, the double-hurdle, which considers the fact that the observed zeros might also be linked to 'non-participation' decisions that could not be referred to as non-adoption (Cragg, 1971). Moreover, in some situations, the decision to invest in SFM and the amount of investment may not be so intimately related (Cragg, 1971). Considered that the SFM technologies have been practiced in the region for a long time, it could be possible that the zero responses could also arise from truncated sampling period - we have no information whether some of those that did not apply during the survey period dis-adopted.

The participation in SFM technology and the corresponding extent of usage can be expressed as underlying stochastic models where:

$$Y_{i1^*} = \alpha W_i + v_i \qquad \textit{Participation decision} \qquad 3\text{-}1$$

$$Y_{i2} = \beta X_i + \varepsilon_i \qquad \textit{Intensify decision}$$

$$Y_i = \beta X_i + \varepsilon_i \qquad \textit{if } y_{i1}^* = 1 \textit{ and } y_{i2} > 0$$

Where i is number of households or plots under observation, y_i is the dependent variable, W_i & X_i are vectors of independent variables, α & β are vectors of unknown coefficients, and v_i & ε_i are error terms.

3.2.3.1 Participation / adoption model

Given that some households decide to participate or adopt while others don't, given a set of conditions. The conditional probabilities are expressed as:

P(x|1) = probability that the i[th] household-plot apply/receive SFM practice

P(x|0) = 1 – P(X|1) = probability for not adopting SFM practice

The response probability function (log of the odds) is expressed as:

$$P(Y = 1|x) = ln(P(x|1)/P(x|0)) = ln((P(x|1)/(1- \left(P(x|1)\right) = z \qquad 3\text{-}2$$

Where X, denote full set of explanatory variables. The probabilities can be estimated from an underlying latent variable model (z), which assumes that the response function is linearly related to a set of parameters and is expressed as:

$$z = \beta_0 + x_i\beta_i + \varepsilon; \quad Y = 1\, if\; z > 0\; and\; Y = 0,\; if\; otherwise \qquad 3\text{-}3$$

The binary response function that is transformed algebraically from equation 3.2 and 3.3 gives the values of estimated probabilities between zero and one and used to predict conditional probabilities for SLM practices (Wooldridge, 2012). It is expressed as:

$$P(Y|x) = exp\,(z)/(1 - exp(z)) \qquad 3\text{-}4$$

where

$1 > P(Y|x) > 0;$ $\qquad +\infty > z > -\infty;$

$P(Y|x) \to 1\; as\; z \to +\infty;$ $\qquad P(Y|x) \to 0\; as\; z \to -\infty;$

$P(Y = 1|x)\; if\; P(Y|x) \geq 0.5;$ $\qquad P(Y = 0|x) if\; P(Y|x) < 0.5$

3.2.3.2 Double hurdle model for household SFM intensification

Using the aggregated plot data and the household attributes, the data tend be truncated at zero and have positive values. In this case it is assumed that there is an underlying stochastic index equal to $X_i\beta + \varepsilon_i$ which is observed only when it is positive. The expected value of y_i is:

$$Ey_i = X_i\beta * F(X_i\beta/\sigma) + \sigma f(z), \qquad 3\text{-}5$$

Where, $f(z)$ is unit normal density, and $F(X_i\beta/\sigma)$ is cumulative normal distribution function. Therefore, the expected value of y being above the limit, referred to as y^* is $X_i\beta$ plus the expected value of the truncated normal error term.

$$Ey_i^* = E(y_i|y_i>0) = X_i\beta + \sigma f(z)/F(z). \qquad 3\text{-}6$$

Therefore, the expected level of soil fertility management strategies computed as an index for all the sampled households, Ey_i can be expressed as a product of the expected value conditional upon having at least some practice, Ey^* and the probability of implementing more or intensifying SFM practices, F(z).

The elasticities for continuous variables are estimated by decomposing the effect of a change in an explanatory variable on dependent variable (McDonald & Moffitt, 1980). This implies that the total change expected in y is decomposed into two: (1) the change in y of those households above the limit, weighted by the probability of being above the limit; and (2) the change in probability of being above the limit, weighted by the expected value of y if above the limit which is expressed as:

$$\delta Ey/\delta i = F(z)(\delta Ey^*/\delta X_i) + Ey^*(\delta F(z)/\delta X_i) \quad 3\text{-}7$$

The elasticity of probability, $F(z)(\delta Ey^*/\delta X_i)$, indicates how a variable affects the probability of implementing the SFM technique. The elasticity of conditional level, $Ey^*(\delta F(z)/\delta X_i)$, indicates how a variable affects the intensification level given that farmers applied inputs or planted legumes. The unconditional elasticity, $\delta Ey/\delta i$, is the sum of the two which indicates the overall responsiveness of the household to a particular variable in the application of the SFM.

3.2.3.3 GLM for plot level intensification

Considering that a greater proportion farms that do not receive inputs while others receive in large amounts, data is truncated at zero with positive skewness. The data does not meet the assumptions of normal distribution as the variance is often larger, which is a biological and socioeconomic reality but a statistical problem called overdispersion. Therefore, plot level fertilizer and manure input levels were predicted using generalised linear models (GLM) following a similar approach as the one used for estimating crop yield output levels (see section 3.6.1.3).

The models were estimated in STATA 14 (StataCorp, 2017).

3.2.3.4 Data sources, the dependent variables and computation of indices

The *choice* was captured through a variable that asked whether a farmer used a SFM practice while the *intensity* as the amount of inputs (inorganic fertilizer, organic manure) used in the 2016/2017 growing season and the amount of land planted with legumes over a 5-year period (2012-2017). Farmer estimates were much easier for inorganic fertilizers as farmers access fertilizers in 50kg bags and for those that shared (as is mostly the case with subsidy), the stated proportions were used as divisor. To find the total fertilizer applied, a summation of the basal dressing - mostly a 23% N: 21% P: 0K+ 4% S (nitrogen, phosphorus and sulphur) fertilizer and top dressing - mainly 46% N UREA (nitrogen) was done. Organic manures comprised mostly of farmyard manure and household refuse. The quantities applied were estimated from either size of landfill or the transport used. Mostly, manures were transported on heads and shoulders using 50kg bags and 20 litter buckets and in a few cases, using oxcarts. Unlike fertilizer and manure, the area under legume was estimated as the size of the plots that had legumes averaged over a 5-year period. Until now, although there are attempts to use land equivalent ratios and farmers guesstimate, it is still difficult to estimate the land equivalence for each of the crops in an intercrop for smallholder farmers. Hence, we use the raw plot sizes and control for intercropping using the number of crops grown.

A desk review of literature on agricultural innovation adoption was done to identify factors that explain adoption and extent (Doss, 2006; Pattanayak et al., 2002). A questionnaire was then developed, and household surveys were conducted to generate socioeconomic data of the households' resource status, farm characteristics and farming practices for the cropping season 2016/2017. With a largely younger population, it is evident that increasing dependency ratio could significantly affect household labour allocation. Being in a matrilineal marriage system, role of women is assumed to influence technology adoption. The women empowerment in agriculture index (WEAI) was constructed using Alkire-foster method. It measures whether a woman is involved in four decision domains: production, resources, leadership and income (Alkire et al., 2012). Consequently, higher WEAI implies that woman in the household is empowered. Data on different livestock species were further converted into LUs using nutritional and feed requirement factors for Africa (Chilonda & Otte, 2006). The communication, transport, and farm items such as radios, bicycles and hoes were broadly standardised into monetary indices using the numbers owned and the prevailing market prices. Plot sizes were measured using the global positioning system (GPS). The model specifications are presented in Figure 3.2 and the resolutions for variable choices are presented in section 3.10.

3.3 The environment

3.3.1 Topography and climate

The study area is the middle course of the Riviridzi-Nsipe River, which is a major sub-watershed for the Shire River Basin. Within the 8 km distance, there are two distinct terrain patterns that bound the agro-ecological potential and livelihood strategies of the area. The eastern part is a plane along the banks of Riviridzi and Nsipe stream at elevation of 800-900 meters while the there is a 256m ascent on sloping hills westwards. As indicted on Table 3—12, about 44% of the study area has gentle slopes of <5%, 29% moderate slopes of 5-10% and 27% has strongly sloping to steep slopes with maximum slope of 55.51%. Other topographic characteristics presented in Table 3—5 shows that the study area has varied surface and hydrological characteristics. In terms of rainfall, the area is in the transition from semi-arid to sub-humid with annual rainfall ranging from as low as 700 mm to around 1100 mm.

Table 3—5 Topographic characteristics of the study area

Attribute	Min	Max	Mean	SD
Roughness index (Ri)	1	44	5.44	5.13
LS-erosion factor	0.03	9.98	1.15	1.30
Flow length	29	5912	1331	1330
Flow accumulation	1	2153	22	107
Upslope area	900	425256000	3429796	24246578

3.3.2 Soils

According to the soil mapping by Lowole (1965), there are four main soil classes (Figure 3.6). The dominant soil for the hilly areas is the Eutric Cambisols that are moderately deep (50-100cm depth limit) with gravely subsoil. These soils are shallow and prone to erosion hence are widely used for grazing or forestry (Driessen & Deckers, 2001). The densely populated flatlands around Gwauya and Phikani have Haplic Luvisols which are deep (>1500cm depth limit) and course to medium textured. Deep, reddish-brown and fine textured Chromic Luvisols dominate the stream valleys whilst the foot slopes have deep, brown and medium textured Eutric Fluvisols. The fluvisols receive sedimentary material at regular intervals from upslope. The eutric cambisols and fluvisols have an effective base saturation of more than 50% and ideal for cultivation. The luvisols have argic horizon with CEC of equal to or greater than 24 cmol(+) kg^{-1} clay starting from around 100cm to within 200 cm from the soil surface. These are easy to till with no impediments and base saturation of more than 50% but are affected by water erosion and loss in fertility. Nutrients are concentrated in top soil and have low levels of organic matter, typical of haplic luvisols (Driessen & Deckers, 2001).

The plane has been under continuous cultivation by more than three generations (Braslow & Cordingley, 2016). This is indicative of a relatively high agricultural potential. On the west, there are pockets of marginal lands stretching on hillslopes that are virtually not suitable for crop cultivation (Benson et al., 2016; Li et al., 2017). When demand for resources and land was minimal, these hilly areas were set aside as village woodlands. Lately, with increased demand for biomass energy and land for food production, marginal areas that were under forest cover in 1992 have been cut down for charcoal and fuelwood and open areas are being converted to farmland (FAO, 2012). The communities have been observing accelerated rate of degradation since 2005 when some community members started producing charcoal for the urban markets (Braslow & Cordingley, 2016). Subsequently, soil eroded from the bare catchments loads the sediments into the rivers which are negative externalities affecting stream flow and provision of environmental services downstream (Chimtengo et al., 2014). We constructed the period the plot has been by either asking for generational usage and if cleared by the current user, the year of opening to date. Hence for simulation, plots have different ages and progressively age, on the assumption that fallows and abandonment will not happen during the simulation period.

Soil classes, despite providing information necessary for regional and national planning, fall short of detail to guide plot level management. As effort has been made to set a baseline for the sentinel (learning) site of 10 km by 10 km that encompass the study community-landscape as detailed below.

3.4 Stocks of soil NPK, SOC and their thresholds

Soil is a multi-dimensional and complex entity. Its complexity arises from a range of natural and anthropogenic factors that influence the biogeochemical processes. In an ideal world, information for every cubic meter of land would be available, but sampling and measuring the range of soil attributes for every pedon is obviously an unsurmountable task. Promisingly, predictions of attributes at unvisited locations can be made from locations with known soil attributes.

Geo-statistics are used to extrapolate or predict values of land features (including soil classes, elemental composition and suitability for farming) from a set of points with known values of the target and auxiliary variables. The most popular approach is ordinary kriging (OK) where predictions for unvisited locations are made using weighted averages of the observations from known locations in proximity, given their map coordinates (Hengl et al., 2007). The OK is guided by the principle of spatial auto-correlation under the assumption that adjacent sites tend to have similar characteristics. Such maps are a result of extrapolation from limited samples, sometimes using fewer soil controlling factors that are relevant for planning at administrative block or agro-ecological zone level (Lowole, 1965; Mutegi et al., 2015). However, the soil information generated from this approach are of limited applicability to social-ecologically heterogeneous farms (Vågen et al., 2016).

With regression kriging, a relationship between the target and auxiliary variables at sample locations is used to predict for the unknown location, the target variable from the known auxiliary variables. Spectral signatures for land surfaces generated through satellite imagery have been widely used as proxy for geological and biological soil forming factors (Hengl et al., 2017; Schillaci et al., 2017). With the advancements in spatial data acquisition and prediction of surface features using spectral signatures, soil predictions have become more spatially explicit (Hengl et al., 2017; Vågen et al., 2016).

Empirical models for estimating soil parameters are chosen based on whether the properties being observed are randomly '*anisotropic*' or systematic '*isotropic*' distributed in relation to the soil forming factors (Jenny, 1941). Soil classification and mapping usually assume that the co-variates systematically vary along climo-sequencies (climate), chrono-sequences (time), bio-sequences (organisms including human activity) or topo-sequencies (terrain) and their combinations. Capturing the variability of soils across landscapes and over time is inherently challenging, especially so for subsistence small-holder farms where soil data are limited and the covariates are usually not readily available. Getting a representative soil sample and predictors are compromised by the large number of determinants of soil formation and subsequent transitions, including the parent material, climate, relief and living organisms that vary over space and time (Jenny, 1941). These may be as many as 75 at global level (Hengl et al., 2014). Hence, generating information on soil forming factors is in itself not feasible in most cases. As such, proxy variables are used.

This study uses the randomForest model which has been found to improve the predictions for data sparse regions (Forkuor et al., 2017; Grimm et al., 2008; Hengl et al., 2015, 2017; Polhill et al., 2008). The RF is a data mining method with the ability of modelling high-dimensional non-linear relationships and handles both categorical and continuous variables with resistance to overfitting and enhanced robustness (Breiman, 2001). The RF prediction was implemented in R using the Breiman's code (Liaw & Wiener, 2002).

3.4.1 Data sources and pre-processing

The 2013 soil data collected by the International Centre for Tropical Agriculture (CIAT – https://ciat.cgiar.org). The sampling site was purposively drawn with the mother trial areas for the Africa Rising project at the centre and represent the typical conditions of the Rift Valley escarpment agro-ecological zone (Mungai et al., 2016). It is also one of the sentinel sites for mapping and monitoring soil conditions in Africa (Tamene et al., 2019). Composite samples were taken from 160 plots of 0.1 ha each to the depth of 0-20 cm. All the samples were analysed by near infrared (NIR) and mid-infrared (MIR) diffuse reflectance spectroscopy. The NIR and MIR reflectance were calibrated using 10% of the samples analysed by standard laboratory procedures. P and K were analysed using Mehlich III extraction; pH using 1:2.5 soil–water suspensions; SOC and total N using thermal oxidation; bulk density using cumulative augering method; and texture using the laser diffraction method (Tamene et al., 2019).

The 2018 soils samples collected from the plots were sampled for yield estimation. The fields for soil surveying were sampled from the list of fields captured during the agronomic survey using the conditioned latin hypercube sampling (CLHS) (Minasny & McBratney, 2006). CLHS optimises the representativeness using auxiliary site conditions which included elevation, geographic coordinates, soil classes, and FAO land cover layers for 2010. The soils were analysed at the chemistry laboratory of the Agricultural Research and Extension Trust using Walkley & Black method for SOC, Bouyoucous for texture, Bray 1 for P and Mehlich III Extraction for K. The total nitrogen was derived from stoichiometric relation with SOC as high correlation of 90 was found from the 2013 data. The P estimates from Bray 1 were upscaled to the level of Mehlich III using the factor of 6% (Gutiérrez Boem et al., 2011).

The hybrid of the vegetation and soil spectral signatures from satellite imagery are used as proxies for vegetation, parent material and climatic soil forming factors Figure 3.7 and Table 3—6. In the study, we use the high-resolution satellite imagery, the topographic attributes associated with erosion and deposition from the SRTM-DEM, the soil geological classes and the surface reflection representing the parent material, and the grid latitudes and longitudes that are proxy for spatial associations.

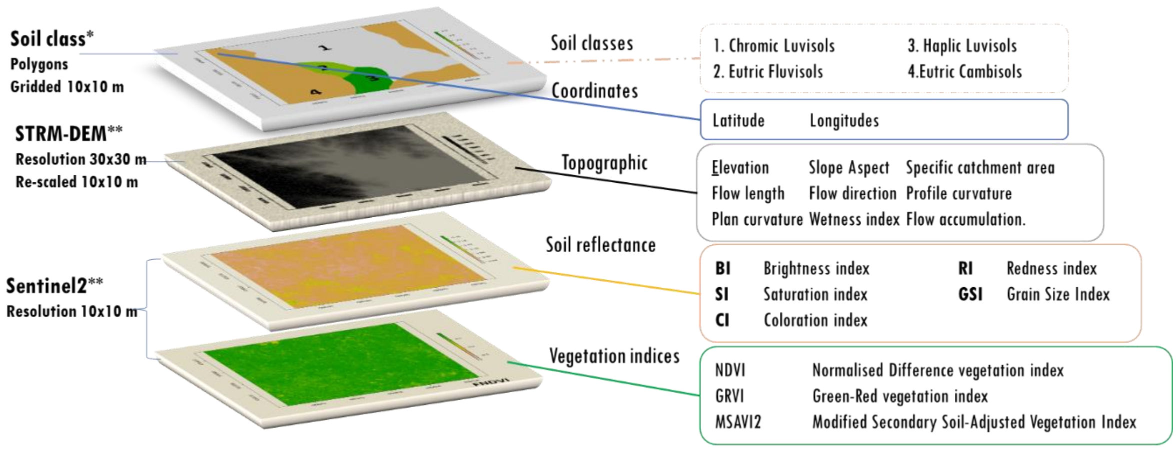


Figure 3.7: The spatial covariates comprising of soil classes, topography, and vegetation and soil reflectance.

Since its launch in 2002, the Sentinel2 imagery have been widely used to support spatial planning in agriculture and food security (ESA, 2015). After downloading from USGS Earth explorer, the tiles were spatially registered to WGS84-UTM-36S projection, merged and later clipped to form a complete mosaicking of the study area using QGIS version 2.18 (Team, 2016). To be able to detect differences in feature reflectance from the original bands, the soil and vegetation indices were calculated (Table 3—6).

Since reflectance is affected by a combination of land features, the spectral signatures from different months were used to correct for annual variations (Forkuor et al., 2017). The study region receives unimodal rainfall that falls between November/December and ends in March/April. The vegetation indices were derived from rain and dry season imagery, whereas the soil reflectance indices were derived from the dry season imagery. The dates of image acquisition for vegetation indices were synchronised with kernel fill for maize and pod filling for legumes, which ensured optimum reflectance from plants. The peak vegetative growth and ground cover for annual crops and grasses is between January and February. For the two seasons (2016-17 and 2017-18), it was only on 18 February 2017 when the sentinel imagery had no cloud cover. In 2018, almost all rainy season imagery had dense cloud cover plus cloud shadows covering entire scenes.

Table 3—6 Soil and vegetation indices derived from the sentinel 2 spectral bands

Index	Formula	Index property	Reference
Vegetation			
*Normalised Difference vegetation index (NDVI)	(NNIR-R)/(NNIR+R)	Health of vegetation	(Forkuor et al., 2017)
Green-Red vegetation index (GRVI)	(G-R)/(G+R)	Vegetation and soils	(Motohka et al., 2010)
*Modified Secondary Soil-Adjusted Vegetation Index (MSAVI2)	0.5* [(2NNIR+1) – sqrt[(2NNIR+1)2 – 8(NNIR-R)]]	Plant growth, yield, SOM, soil erosion	(Xue & Su, 2017)
Soil			
Brightness index (BI)	$((R^2 + G^2 + B^2)/3)^{0.5}$	Average reflectance	(Forkuor et al., 2017; Ray S. et al., 2004)
Saturation index (SI)	(R-B)/(R+B)	Spectral slope	
Coloration index (CI)	(R-G)/(R+G)	Soil colour	
Redness index (RI)	$R^2/(B^*G^3)$	Hematite content	
Grain Size Index (GSI)	(R-B)/ (R+B+G)	Texture (grain size composition)	(Xiao et al., 2006)

Spectral bands	Blue	Green	Red	NIR	NNIR
Designation	B02	B03	B04	B08	B8A
Median wave length (nm)	492.1	559.8	664.9	832.8	864.7
Spatial resolution (m)	10	10	10	10	20

*the NDVI and MSAV-2 were calibrated based on Landsat bands for which the NIR wavelength of 851-879 (nm) corresponds to the narrow NIR (NNIR) for the Sentinel2 imagery. 20m NNIR (B8A) was resampled to 10m x 10m pixel size.

During the dry season, crop fields are either bare or covered with dry weeds and/or crop residues. Moreover, the deciduous trees scattered on farmlands and in adjacent uncultivated fallows, woodlands and grasslands lose most of the leaves which are usually burnt leaving the ground bare. The dry season imagery captured on 2nd August, 1st September, 1st October and 10th November in 2016 were used for the bare ground reflectance and partial vegetation.

Topographic factors related to soil genesis, hydrological flow patterns and mass movements that lead to soil particle and nutrient redistributions were derived from the 30m SRTM DEM. The sink filled DEM tiles *s15_e034_3arc_v2* and *s16_e034_3arc_v2* downloaded from USGS Earth Explorer (USGS, 2018) were merged, then the watersheds delineated and eventually, the target watershed was clipped. From the target watershed, the topographic attributes computed included slope, aspect, specific catchment area, maximum flow path (flow length), flow direction, profile curvature, plan curvature, wetness index, and flow accumulation (stream power index (SPI)). The topographic layers of 30m x 30m were then down-resampled to 10m x 10m. Although highly debated, it has been found that the DEM resampling has little effect on hydrological flow estimations (Tan et al., 2015) and on the corresponding nutrient flow estimations (Lin et al., 2010).

Geological factors considered for the study included the soil classes which indicate the dominant underlying parent material and soil conditions (Lowole, 1965). The polygon soil layers obtained from the National Spatial Data Centre for Malawi were converted to raster at grid resolution of 10m x 10m. The hilly areas are dominated by Eutric Cambisols (Oxisols) that are moderately deep (50-100cm depth limit) with gravely subsoil. These soils are shallow and prone to erosion, hence are widely used for grazing or woodlots (Driessen & Deckers, 2001). The flatlands have Haplic Luvisols (Alfisols) which are deep (>1,500cm depth limit) and course to medium textured. The stream valleys are dominated by deep, reddish-brown and fine textured Chromic Luvisols whilst the foot slopes have deep, brown and medium textured Eutric Fluvisols (Entisols). The fluvisols receive sedimentary material at regular intervals from upslope and are ideal for cultivation. The luvisols have argic horizon, clayey at deeper depths of 100 to 200 cm, easy to till but greatly affected by water erosion and loss in fertility. With soil and water conservation measures, these soils would be most productive as the cation exchange capacity (CEC) is typically equal to or greater than 24 cmol(+) kg^{-1} and base saturation above 50%.

3.4.2 *Model specification and co-variance importance*

Considering that the covariates are proxies of soil forming factors, their associations and importance for explaining the soil attributes are likely to vary. We test the significance of the covariates using different model specifications and choose the one with the highest Out Of the Bag (OOB) R^2 for prediction (Table 3—7).

Table 3—7 Prediction accuracy using RF out of the bag error (AOB) and 3-fold cross validation

	OOB validation								3-fold cross validation							
	SOC		TN		P		K		SOC		TN		P		K	
	R2	RMSE	R2	RMSE	R2	RMSE	R2	RMSE	R2	RMSE	R2	RMSE	R2	RMSE	R2	RMSE
Full model	0.90	0.17	0.89	0.08	0.83	20	0.83	20	0.35	0.42	0.14	0.21	0.25	52	-0.25	56
Dry soil&veg	0.87	0.19	0.88	0.08	0.84	20	0.84	20	0.38	0.41	0.08	0.22	0.08	57	-0.21	55
Dry soil	0.87	0.20	0.87	0.09	0.74	25	0.74	25	0.36	0.42	0.00	0.23	-0.04	61	-0.28	57
Dry veg	0.86	0.20	0.85	0.09	0.83	20	0.83	20	0.36	0.42	0.08	0.22	-0.03	60	-0.48	61
Topography	0.77	0.25	0.80	0.11	0.81	22	0.81	22	0.00	0.52	0.05	0.22	0.21	53	-0.23	55
Dry&Feb veg	0.86	0.20	0.87	0.09	0.81	22	0.81	22	0.35	0.42	0.28	0.19	0.04	59	-0.37	59
Feb veg	0.76	0.26	0.79	0.11	0.72	26	0.72	26	0.04	0.51	0.07	0.22	-0.03	61	-0.27	56
Class&coords	0.53	0.37	0.54	0.16	0.59	32	0.59	32	0.23	0.46	0.03	0.23	0.15	55	-0.03	51

It was observed that, F_{all}, which is a combination of all the covariates was the best specification for predicting SOC and TN. The covariate set F_{dry}, which include soil and vegetation reflectance indices from dry season (Aug to Nov) were best at explaining the occurrence of Phosphorus and Potassium, and second best for SOC and TN. The three fold cross validation however, shows that using a reduced dataset (2/3rd) for training and the remaining one third for testing, the best specification for SOC is F_{dry}, for TN is F_{vg} (the dry season and optimal growth vegetation indices), for P is F_{all}, and none of the specifications fit the distribution of Potassium well.

3.5 Input flows of NPK and SOC

Considering the existing farmer management, the common nutrient replenishment approaches as introduced under section 1.2 include inorganic fertilizers, legumes and organic manures.

3.5.1 IN1: Inorganic fertilizers

The Nitrogen, Phosphorus elemental composition for fertilizers was proportionally estimated from the types and quantities for the two common inorganic fertilizers used. The 23:21:0+4S which is applied as a basal fertiliser had 23% N, 21% P_2O_5 and 0% KO_2 whilst the top-dressing Urea had 46% N. The subsequent nutrient inputs estimated as probabilities for a plot to receive fertilizer and the respective quantities have implications on N and P stocks and transfers.

3.5.2 IN2: Organic manure

The elemental (NPK) composition of manure was estimated using the concentrations established for compost made by Malawian smallholder farmers (Chilimba et al., 2005). The NPK contents were found to be in the range of 0.21 to 2.2%, 0.05 to 0.73%, 0.12 to 2.62%, respectively. The biomass conversion factors used to estimate residues from yield and the NPK stoichiometry were obtained from existing literature (see Table 3—8).

Carbon inputs include the retention of crop residues, crop roots and organic manure input. Crop residues are traditionary incorporated on the plot (Emerton et al., 2016) although the practice was halted during the survey period due to outbreaks of the exotic fall army worm *Spodoptera frugiperda* (J.E. Smith) (Lep.: Noctuidae), whose management involved burning of residues after harvest. Therefore, residue incorporation was conservatively set at 80% following earlier reports (Emerton et al., 2016). The decomposition is mediated by environmental factors such as soil texture (*ST*), moisture availability in terms of period of rainfall (*MP*), and annual average temperature (*T*) and whose combined effect (*f-env*) is empirically expressed following Groot and Oomen (2018) as:

$$f\text{-}env = 1 / (ST * MP / 365 * 2^{(T-9.5)/10}) \qquad 3\text{-}8$$

For the study region the *MP* and *T* are not variable but *ST* factor differs. Most studies set ST at value of 1 (corresponding to sand soil). Hence, we adjust ST between 1 to 0.5 since the soils in the study region have sand in the range of 16 to 71%.

The carbon input from crop residues (C_{CR}) is estimated from the organic matter in the residues incorporated plus roots (42% of the total residue input) adjusted for effective organic matter for stover of each target crop (EOMc). The greater part of the roots for groundnuts are uprooted. The EOMc is around 20 – 40% of residue organic matter (X. Wang et al., 2007). The effective organic matter is the organic matter remaining after one year.

$$C_{CR} = RESIDUE_{(KG/HA)} * 0.5 * EOMc / 1.724 * f\text{-}env \qquad 3\text{-}9$$

The amount of organic carbon in manure (C_{MN}) is derived from the amount of organic matter in the manure imported and the active organic matter content of manure. Organic matter of manure made by mixing livestock dropping, household and crop residues is around 34% in Malawi (Naohiro et al., 2016). The active organic matter content ($AOM_{MN\%}$) of 48 to 80% is degraded during the first year of application (X. Wang et al., 2007). This implies that after the first year, 20 to 52% of the organic matter from manure input remains on the plots. Due to the uncertainty in the actual decay rates, simple sensitivity was done by varying the annual decay rates in the simulation using either lower or higher rate.

$$C_{MN} = MANURE_{(KG/HA)} * 0.5 * ((100 - AOM_{MN\%}) / 100) / 1.724 * f\text{-}env \qquad 3\text{-}10$$

3.5.3 *IN3: Biological nitrogen fixation by leguminous plants*

The biological nitrogen fixation by leguminous crops was estimated from the yield and biomass estimates from household-plot surveys, the nitrogen derived from air for groundnuts grown in Malawi (Mhango et al., 2017), and the total nitrogen content and biomass conversion factors from other secondary sources (Table 3—8).

Table 3—8 Nutrient stoichiometry and conversion factors for maize and groundnuts for grain, shoot and roots

Nutrient	Attribute	Maize	Ground nuts
Nitrogen	Grain	0.0230 [a]	0.0410 [h]
		0.0088 [b]	0.0278 [i]
	Shoot	0.0213 [a]	0.0408 [a]
		0.0122 [c]	0.0278 [I]
			0.0201 [j]
Phosphorus	Grain	0.0038 [d]	0.0022 [a]
		0.0020 [b]	0.0009 [k]
	Shoot	0.0019 [a]	0.0032 [a]
		0.0012 [c]	0.0025 [j]
			0.0015 [k]
Potassium	Grain	0.0187 [a]	0.0214 [l]
		0.0043 [d]	0.0097 [j]
	Shoot	0.0205 [a]	0.0280 [a]
		0.0206 [c]	0.0101 [m]
Organic matter	Shoot	0.4200 [c]	
Moisture content $_{(1\text{-}MC)}$		0.8750	0.8200 [n]
Harvest index	Sole	0.3630 [e]	0.3520 [o]
	Intercrop	0.1100 [f]	0.2090 [o]
Shoot-root	Unfertilized	0.0800 [g]	0.2945 [p]
	Fertilized	0.4200 [g]	
Nitrogen (NdFA)			0.7550 [q]

[a](Van den Bosch et al., 1998),[b](Ganunga et al., 2005) [c](Partey et al., 2016), [d](Hgh-Jensen et al., 2013), [e](Hay & Gilbert, 2001), [f](Mas-ud et al., 2016), [g](Anderson, 1988), [h](Nautiyal, 2002), [i](Raverkar & Konde, 1988), [j](Sakonnakhon et al., 2005) [k](Madhuri et al., 2018) [l](Smartt, 1994), [m](Mupangwa & Tagwira, 2007) [n](African Institute of Corporate Citizenship, 2014), [o](Mas-ud et al., 2016), [p](Mohamad et al., 2018) [q](Mhango et al., 2017).

3.5.4 *IN4: Sedimentation*

Within the catchment, erosion and deposition (*IN4*) redistributes the nutrients from higher to lower elevation. With minimal soil erosion control measures and cultivation of slopes, there is potentially high erosion upslope. Not all eroded material from upper portions of the watershed are delivered to the valley bottom (Tamene et al., 2017a). Intermediate sediment deposition is one of the major nutrient input sources for the plots downhill. The deposition depends on the surface cover and terrain characteristics, and occurs at points where the momentum of the transporting water is insufficient to carry the eroded material downslope or along the channel. The potential for intermediate deposition also increases as the area of the watershed increases. To estimate the sediment deposited, we use the sediment delivery ratio (SDR). It is the proportion of eroded sediment that leaves a given parcel and relates to the sediment transported to a location in the channel system to the gross erosion from the drainage area above that point (Tamene et al., 2017a). For details on the estimation see section 3.6.2.7 as this is calculated during erosion estimation.

3.5.5 IN5: Atmospheric deposition

One major nutrient source that is often neglected and roughly estimated is the wet and dry atmospheric deposition (*IN5*). Using the existing transfer functions (Stoorvogel & Smaling, 1990), the inputs were largely underestimated. The studies in the region found that atmospheric deposition is a significant source of P and a major source of N (Bootsma et al., 1999). The dry deposition of potassium was estimated from the total global average of 4.1 $kg\,ha^{-1}yr^{-1}$ (Sardans & Peñuelas, 2015). Considering that the estimated atmospheric content in southern Africa due to Savannah biomass burning is around 0.4 $\mu g\,m^{-3}$ (Sinha et al., 2003), the dry potassium deposition could be potentially high for Malawi. NPK inputs were estimated using measured wet and dry nutrient deposition concentrations (Table 3—9).

Table 3—9 The nitrogen (N), phosphorus (P), and potassium (K) inputs from rain and dry deposition

Wet	$kg\,ha^{-1}$	=	$\mu mol\,l^{-1}$	x	μmol->μg	x	μg->kg	x	ml	x	ml->l ha^{-1}
	N_rain	=	13.13	x	14.007	x	1E-09	x	Rainfall	x	10000
	P_rain	=	0.78	x	30.974	x	1E-09	x	Rainfall	x	10000
	K_rain	=	1.53	x	39.063	x	1E-09	x	Rainfall	x	10000
Dry	$kg\,ha^{-1}$	=	; $\mu mol\,m^{-2}day^{-1}$	x	μmol->μg	x	μg->kg	x	m^2-> ha	x	days
	N_dry	=	343.00	x	14.007	x	1E-09	x	10000	x	365
	P_dry	=	10.50	x	30.974	x	1E-09	x	10000	x	365

Note: μmol were converted to μg using molar masses. Source: (Bootsma et al., 1999)

3.6 Output flows of NPK and SOC

3.6.1 OUT1 & OUT2: Crop yield and residues

First, a review of literature was conducted to tabulate the nutrient stoichiometry and biomass conversion factors for grain, shoot and roots in maize and groundnuts (Table 3—8). These are used to compute nutrient contents based on predicted plot yields. The NPK parameters have upper and lower values. The upper values are the widely used parameter values that were mostly established for Kenyan farms (Van den Bosch et al., 1998). The lower paramater values are from studies conducted mostly in Malawi and other tropical regions (Table 3—8). Sensitivity analysis was done using the sampled data before running simulations. It was oberved that the Kenyan parameters shifted the mean nitrogen and potassium losses downwards but shifted upwards the phosphorus buildup and carbon input (Table 4—14 vs Appendix S2).

3.6.1.1 Estimating crop yield

In spite of several efforts to increase productivity in the past six decades (Vanlauwe et al., 2017), Malawi still face food deficiency because the yields for the staple cereal crops have plateaued at 2 t ha^{-1} over the past decade (see Figure 1.1). Hence, identification of pathways for improving the productivity of smallholder farmers has become a major goal for reducing food and nutrition insecurity and eradicating rural poverty. However, the opportunities are eroded by the scarcity and unreliability of data including reliable yield estimates (Carletto, Jolliffe, et al., 2015). Reports and studies that use national datasets tend to aggregate results which masks the variations and opportunities within sub-regions, potential positive shifts in technological frontiers (Thirtle et al., 2003), and increased technical efficiency in practice by some farmers (van Ittersum et al., 2016). The yield estimates for the past two decades showed an initial positive shift as a result of farm input subsidy for the main crops (Chirwa et al., 2011). This success was short lived as recent studies have shown that the net primary productivity of farming landscapes during the same period has significantly declined thereby threatening the sustainability of the system (Messina et al., 2017) mainly because the majority of small-scale farms are on moderately and marginally suitable lands (Li et al., 2017).

It is envisaged that with the available technologies and management practices, the potential yield gains of 4.3 t ha^{-1} additional to the national average of 1.7 t ha^{-1} simulated for the climatic conditions of the year 2000 could be achieved (Mueller et al., 2012). The attainable yield analysis showed that yield gains of 4.38 t ha^{-1} above the average of 4.1 t ha^{-1} would be achieved within a productive site in central Malawi (Tamene, Mponela, Ndengu, et al., 2016). Studies have established that the observed variation in yield is a result of differences in inherent soil conditions, inefficient farmer practices and genetic-environmental misfit of the cultivars (Kihara et al., 2015; Tamene, Mponela, Ndengu, et al., 2016). However, limited attention is given to unravel village-landscape yield heterogeneity (Tamene, Mponela, Ndengu, et al., 2016).

3.6.1.2 Crop yield and residue data

Several methods exist for estimating crop yield ranging from farmer interviews, crop cuts to whole plots harvests (FAO, 1982). The validity of results depends on the level of precision in the estimation of the cultivated area as well as the quantity of produce obtained from that area which has been a challenge in most areas dominated by small-scale farms (M. Burke & Lobell, 2017; Deininger & Xia, 2017; Lobell, 2013; Sud et al., 2015). Often, the actual yields are estimated from a few randomly allocated farms and reported as an average of administrative units covering hundreds of thousands of fields. The yield potential, i.e. the difference between average yield and those obtained at research sites, is estimated by use of agronomic trials or crop simulation models. The main setback with the two approaches is that the sample sizes are a handful and research plot sizes are often quite small.

The standard approach is to harvest the whole field and measure the produce. For larger landscapes and populations, like in this study, the costs associated with whole plot harvests are prohibitive (Sud et al., 2015). As a result, small plot cuts are often used as pseudo replicates within the whole field. However, most studies still use farmer reported yields (FRY) obtained from household surveys. Apart from the self-reporting bias, farmers report whole numbers of the items they use for transportation or storage such as loads of oxcarts, tonnage of cars or number of 50 kg bags. Despite the inconsistencies in reporting the FRY, it has been established that when estimating yields over larger areas, the FRY are comparable to the crop harvests done by a team of trained personnel (M. Burke & Lobell, 2017; FAO, 1982; Fermont & Benson, 2011). The small plot cuts tend to have high level of measurement error due to within-field variations in small-scale farms (Tittonell et al., 2005). Considering that this study draws a large sample of randomly allocated households and plots with multiple crops, a combination of FRY and small plot cuts approaches were used.

Apart from estimation of produce, measurement of plot sizes has also been an underlying challenge. The farmer reported plot sizes, usually through guess work, have been used for convenience and cost-effectiveness. However, there is growing evidence that farmers with smaller plot areas of < 0.5 hectares tend to overestimate their fields by a factor of 5 (M. Burke & Lobell, 2017; Carletto, Gourlay, et al., 2015). Although the average plot size in the study region is 0.9 ha, with fragmentation the multiple plots owned by a household are ≤ 0.5 hectare. The differences between reported area and measured area are much bigger compared to the discrepancies in crop produce (M. Burke & Lobell, 2017). The research purpose of our visit was explained to the farmers; they were not expecting any follow up program and were less likely to inflate or underestimate the yields realised from the plots.

Therefore, for 2016/17 growing season the farmer reported yields were captured and the corresponding areas measured using a GPS. For 2017/18 growing season, a second round of survey was conducted with the same farmers. Addition to the survey, yield estimates were done using crop cuts.

For exploring the pathways for individuals or group of farmers towards sustainable agricultural intensification, spatially explicit data that links the estimated crop yield to management practices is a precondition. However, there are concerns regarding spatial and temporal representation of yields for farms not visited or years not surveyed. Lobell (2013) purported that it is challenging for yield gap analysts to generalize results from a small number of sites and years to a broader scale relevant for regional measures of agricultural performance. Use of productivity indices derived from remote sensing imagery has been widely used as covariates to predict the yields for unvisited sites. However in the savannah region, the grass and maize reflectance are indistinguishable (Forkuor, 2014) and different densities of trees are randomly distributed within cultivated fields and adjoining grasslands, leading to more spectral confusion.

Crop yields for the main crops in the area are used to determine agricultural production potential and constraining/enabling factors. The total yield per plot (including intercropping) was used to evaluate the effects of various yield influencing factors on overall benefits. As noted in earlier sections, farming in the region is subsistence oriented where several crops are generally mixed on the same plot. In addition, the crops grown in the region have different values in terms of weights, nutritive value and nature (which includes grains, roots, leafy, and cotton) and they could not be converted to equivalent dry matter content. In such a complex farming system, disaggregation and determination of yield per unit area of land for individual crops is impractical, so is the partitioning of inputs used.

Smallholder farms are characterised by variability in climate and soils as well as agronomic practices. Hence, yield estimation and extrapolation were then done using agronomic survey tools developed by Kihara et al. (2015) and Tamene et al. (2016). For each surveyed plot of the household, a record of agronomic practices including crop varieties, cropping systems, land management, and labour and nutrient inputs was made.

3.6.1.3 *Crop yield model*

Smallholder farming systems of Africa are largely considered nutrient extractive industries because even if nutrient uses are low, the input usage is inefficient. Being the principal product farmers derive from farmlands, the policy and management aim is not to reduce the output but to increase its production efficiency. Empirical estimates of aggregate production functions (the Cobb-Douglas and Frontier) are important theoretical basis for determining the elasticity of substitution between technical change, labour and capital (Knoblach & Stockl, 2019). These basic models have been extensively applied in smallholder agricultural systems that are characterised by heterogeneous farms and operational forms (Sau, 1971). The individual farms are considered as firms with considerably large cross-firm differences in productivity, low aggregate productivity and imperfect markets that lead to friction in resource allocation (Kaiji & Zheng, 2007).

The interventions aiming at increasing efficiency such as input subsidies and technical approaches can trigger cyclical changes not only in production but also in consumption and capital re-allocation. When allocating resources among the plots, farmers face little policy and market influence within a growing season. Hence, we explore the efficiency arising from application of largely subsidised fertilizer. Production on each plot depends on input of fertilizer, land (size and inherent fertility) and labour as well as un-expected shocks such as droughts, pest and diseases. At the time of choice of land and labour allocation, it is largely assumed that the farmer knows land productivity but only partially aware of the distribution of un-expected shocks. Since smallholder farming still rely to a greater extent on natural soil fertility, accounting productivity variations due to inherent land quality is essential. Heterogeneity in soil quality is observable but available datasets are at scales not too large to depict differences between smallholder farm sizes.

Since the subsidy coupons are received by a few households, community members usually share the subsidized fertilizers (50kg of NPS and 50kg Urea) and those with finances cover the fertilizer deficit with own purchases from the market. Given a household of two plots and average land size of 0.9 ha, the subsidised fertilizer is either applied to one of the plots or in most cases, mixed with own purchases. To evaluate efficiency arising from the subsidy as an input policy, we estimate the effect of increasing share of subsidised fertilizer on the yield derived from a plot. Consistent with policy, we assume that input subsidy boosts production. We are mindful however, that despite the removal of prize cap, the government continues with the deliberate policy of setting the prices of maize at minimum that is supposed to be affordable for the poor net maize buyers (Chirwa et al., 2008). This could potentially erode the gains from input subsidy.

In the absence of extension officers, there is little government support (Tamene, Mponela, Ndengu, et al., 2016). To-date, with no/minimal mechanisation and synthetic pesticides and herbicides use, production in smallholder maize mixed farming systems of East and southern Africa is a function of family labour, fertilizers, manure and seeds. Operating at such a scale, the majority of households tend to produce less than the potential yields due to various degrees of inefficiency (Tamene, Mponela, Ndengu, et al., 2016). In these systems, the existing resources in terms of soil quality, water availability and agronomic practices strongly influence variability in yield within the community-landscape. Adjustments have been made to factor production functions which enabled inclusion of nonconventional variables such as social capital and inherent land properties that have been found to shift the production function (Kihara et al., 2015; Tochombe, 2002). Including the local variables reduces the unexplained variation and improves parameter estimates for the inputs used in these highly diverse environments.

The model specification therefore follows the basic structure of the production function and includes the variables that take into account spatial variations (Gourlay et al., 2017). Since there is usually a proportion of farms that do not receive inputs while others receive them in large amounts, data are truncated at zero with positive skewness. The data from smallholder farms do not follow the assumptions of normal distribution as the variance is often larger, which is a biological and socioeconomic reality but a statistical problem called overdispersion. Data transformations, including log transformation, are widely used to successfully normalise. The main estimation challenge encountered using log transformations is when the dataset contains zero values. The practice here is to fudge the whole dataset by adding one, which distorts the estimates (O'Hara & Kotze, 2010).

To avert these problems, the Generalised Linear Model (GLM) was used and implemented using the procedure by Glick (2015) to select both the family and link functions. The link function directly characterises how linear combination of predictors is related to the prediction on original scale. The first run with gaussian family and identity link and the slope of residuals and predictions indicate the distribution with 0 being gaussian, 1 poisson, 2 gamma, and 3 inverse gaussian/wald. The link was chosen by comparing the specification tests including model convergence and the BIC and AIC of the different link specifications. Incidentally, all nutrient input and crop yield models had gamma distribution and the log link function with variance proportion to square-root of mean. This has been found to be robust at modelling data distributions with positive mass often clumped at zero (representing no impact) and a continuous density on the positive reals capturing the impact (Maindonald & Braun, 2010; McCullagh, 2007). In that way, the model estimates the joint impact of frequency and intensity of explanatory variables on the target variables.

3.6.1.4 Crop yield explanatory variables

Development of varieties that are suited to environmental constraints, coupled with increasing use of inorganic fertilizers, were the major break-through during the green revolution. However, it is now a common knowledge that resource limited farmers continue to produce sub-optimal yields despite using improved varieties. Consistent with theory, we assume that a farmer h produces crop yield (q_{hi}) on plot i from land (l_{hi}), family labour (h_{hi}) and nutrient inputs (k_{hi}).

Most if not all of the farms in the study area use family labour, which is typical in rural communities. Unlike the capitalistic farm, labour allocation at household level is guided by the utility of additional consumption made possible by its marginal productivity but also faces disutility among members after extended engagement in farming. This study corrects for this by not using total family size or available labour since larger households owning smaller farm sizes would have relatively excess labour which might not be utilised for production. The study captured the actual number of hours that the various labour categories worked. To correct for disutility, the contributions of different age classes and gender groups were converted into man-equivalents. The labour data was collected for number of hours worked by individual members disaggregated by sex and age class for each plot managed by the household and for each agricultural activity: land preparation, sowing, weeding, fertilizer application, manuring and herbicide/pesticide application, harvesting, transport and threshing for storage. The number of days and hours worked by hired labour was also collected for each sex and age class by plot and activity.

The survey captured mostly the two main fertilizers used: NKP (23% N: 21% P_2O_5: 0% KO_2) and Urea (46% N) as basal and top-dressing or mixed. For economic evaluation, the prevailing market price of fertilizer was obtained and used for estimation of potential cost of fertilizer and subtracted the reported purchase prizes to come up with the percentage that was subsidized. Although almost 30% of the households in the study villages receive coupons, community members contribute and share the fertilizer.

Weeds compete with crops for resources such as sunlight, moisture or nutrients. In this study the weeding frequency was used as a proxy where farmers do not weed (worst case), weed one or twice or three times. Although it could be possible that plots that were weeded three times are located in areas with more vigorous weed growth, we assume that inclusion of other plot variables reduces the effect and weeding twice or thrice are the best agronomic practices.

Since the study region's catena is characterised by sloping escarpments and flat plains, site productivity could vary along the topographic gradients. At landscape level, the major determinants of ecosystem's primary productivity that includes soil and water status have profound effect on production. Since spatial data on soil and water status covering the study site are non-existent, the drivers including climate and topography are used (Le, 2005). A topical study and intensive review of literature by Le (2005) provides a detailed overview of drivers of productivity at landscape level as summarised below. The soil properties that strongly vary across the catena and drives plant growth include: soil depth, soil moisture, soil carbon content, soil pH and total exchangeable bases. Several indices are used to depict the water flow and material transport capacity.

The mean rate of soil material transport capacity (SPI) at a given location is a product of water flow factor and the slope shape factor approximated using the upslope contributing area per unit of a contour length ($P_{upslope}$) and slope gradient (P_{slope}), respectively. These are estimated for grid cells of a 30m SRTM-DEM 1 Arc-Second Global (USGS, 2018). The water flow factor for a grid cell shows the accumulative potential of soil and water with a positive influence on soil productivity while slope gradient determines kinetic energy of water flow hence has a positive relationship with erosion as a degradation process. In other words, S is the measure of erosive power associated with flowing water. It is based on the assumption that discharge is proportional to the specific catchment area. It predicts net erosion in areas of profile convexity and tangential concavity (flow acceleration and convergence zones) as well as the net deposition in areas of profile concavity (zones of decreasing flow velocity). The balance between the deposition accumulation potential and erosion degradation risk defines the inherent site productivity and could be used to predict sustainability.

The inclusion of these long-term productivity drivers gives an indication of the long-run determinants of agricultural growth. Much of the variation in agricultural performance and in data collected for economic analyses is associated with short-term fluctuations that are sometimes considered as nuisance variations or potential biases and often averaged over temporal scales. Yet, such averaging distorts the estimates across sites and affects empirical results of productivity.

3.6.2 OUT3: Soil erosion

Soil erosion is one of the five major nutrient outputs from farming systems and could also be considered as an input to farms downstream through sedimentation (Stoorvogel & Smaling, 1990). In the sloping rift valley escarpments of Malawi, for decades cultivation has been extended to non-arable hill slopes without appropriate soil conservation measures (Farmer et al., 1977). The focus has been to estimate and curb soil loss (FAO, 2019; Vargas & Omuto, 2016). Downstream and intra catchment gains from sedimentation have been entirely neglected despite being one of the major nutrient sources (Tamene, 2005).

A six-year erosion trial conducted from 1984 on the escarpments with 44% slope measured as high as 80 t ha^{-1} yr^{-1} loss of topsoil a year after vegetation clearance with continued losses of 21.6 t ha^{-1} yr^{-1} in subsequent years (Banda et al., 1994). The erosion trial was managed as per the farmer practice of constructing earth ridges of 30 cm high across slopes. Despite soil erosion's rating of being the most serious form of land degradation impeding national development, the efforts to curb it are non-existential. Earlier estimates showed high soil loss rates of 10 - 43 t ha^{-1} yr^{-1} compared to the rate of soil formation of 12 t ha^{-1} yr^{-1} (Bishop, 1995; World Bank, 1992).

The loss of soils from farmlands entail loss of medium for crop growth, essential plant nutrients as well as disruption to downstream services. In Malawi, the total annual cost of land degradation due to vegetation loss and soil erosion is estimated at 320 million US$, which is equivalent to 7% of the country's Gross Domestic Product (Nkonya et al., 2016; World Bank, 2019). In addition to their direct contribution to the livelihoods of farmers upstream, the escarpments are essential sub-catchments for supply of irrigation and hydropower water in the Shire Basin and directly or indirectly influence the livelihoods of 22% of Malawi's population. With a population density of 230 persons km^{-2} in Malawi (GoM, 2018), which is quite high compared to other countries in southern Africa, there is enormous pressure on land resources. Consequently, land degradation is a major environmental problem in the country (Li et al., 2017).

The Ntcheu district in central Malawi is one of the areas in the country which is seriously affected by soil erosion, contributing significant sediment to the Shire and ultimately siltation of the major hydro-electric dams (Vargas & Omuto, 2016). Rills and gullies are relatively prominent features in the district (Davies et al., 2010). In 1992, erosion estimates showed that Ntcheu escarpments experienced soil loss of more than 30 t ha^{-1} yr^{-1} compared to the national average of 20 t ha^{-1} yr^{-1} (World Bank, 1992). The more refined estimates in 2014 also found that despite the earlier higher estimates that convinced the Malawian government to push for programs on land management in Ntcheu escarpments, the area continued to experience higher soil loss of around 10 t ha^{-1} yr^{-1}, which was double the estimated national average of 5.9 t ha^{-1} yr^{-1} (Vargas & Omuto, 2016). The soil erosion problem in the study sites are exacerbated by rapid increase in population, resulting in land scarcity, and putting even more pressure on the remaining land resources (GoM, 2018). Due to shortage of land, farmers resorted to cultivating marginal lands/unsuitable areas with no or inadequate conservation measures (Braslow & Cordingley, 2016; Nakhumwa, 2004).

Considering the severity of land degradation and its associated impacts, there is an urgent need to devise mechanisms that can minimize both the on- and off-site impacts of catchment erosion and sediment delivery. Integrated landscape management and restoration practices are essential to both improve existing system productivity and reduce land degradation, thereby sustaining gains both on- and off-site. Soil erosion and delivery processes do not occur uniformly across space, and it is currently not feasible to sustainably manage all areas affected. A targeted response is therefore employed where resources are directed to areas of high risk rather than spreading them equally across the landscape. In addition, limited financial resources as well as restrictions on land often exclude the application of conservation measures to all areas experiencing erosion (Tamene & Vlek, 2007). For appropriate targeting of relevant areas, there is hence a need to identify major sub-catchments that should be targeted due to not only their major sediment contribution downstream but also considering their potential in the provision of ecosystem services.

Distributed soil erosion models can be used to identify landscape positions with high rates of soil loss for both detailed investigation and applying suitable management options geared to tackle sediment yield. A wide range of models are used to estimate soil erosion at different scales including the Universal Soil Loss Equation (USLE) (Wischmeier & Smith, 1978); the Revised Universal Soil Loss Equation (RUSLE) (Renard et al., 1997); and the Soil Loss Estimation Model for Southern Africa (SLEMSA) (Elwel & Stocking, 1982). Despite the fact that soil erosion models are considered useful options to predict erosion/deposition processes for resource management applications (Nearing et al., 1989), there is, however, no clear agreement in the scientific community which kind of model is more appropriate for the simulation of natural processes. Selection of appropriate model(s) that can suit the areas under study considering the objective at hand, resources available and detail and scale of investigation is therefore crucial. In data sparse regions, the models/tools that are easily available and are easy to use are preferable. Provided that appropriate parameter specification is made before application, empirical models such as the USLE and its derivatives are best suited to such environments. However, it is also important to note that these models may not provide adequate soil loss information at landscape scale due to limitations such as inability to represent deposition and sedimentation processes, and account erosion from gullies and bank collapse (I. D. Moore et al., 1991). However, recent advances in the development of digital elevation models (DEMs), GIS and derivation of the slope-length component of the USLE has enabled their application in complex terrain landscapes and handle modelling hydrological processes of complex topography at larger geographical scales (Tamene, 2005).

Soil mapping for land use classification and land management needs to consider other parameters such as available land management options in terms of biophysical limits and economic feasibility (Farmer et al., 1977). Successful implementation of land management practices is viewed to depend on the aspiration of land users and the strategies by the local political systems by, for example, aggregating soil erosion at administrative units such as districts (Vargas & Omuto, 2016). The aggregates do not provide useful insights for targeting interventions, especially for most districts in Malawi that have high variability in land form and land uses. Micro-catchment-based options are thus needed. Without detailed soil erosion measurements, the study uses site specific parameters and aim at identifying priority areas that require prior management interventions through mapping hotspots areas of erosion and simulate the potential impacts of different land management options in reducing soil loss and sediment yield. The main task was to identify key sub-catchments to focus investments for integrated soil and water conservation in the Riviridzi Sub-watershed of Shire Basin employing geospatial techniques and soil erosion models. This was followed by analysis of the sediment yield reduction potential of SWC/SLM practices through scenario analysis. The escarpments of Ntcheu district of Malawi were selected considering the severity of soil erosion and its impact on downstream sedimentation of irrigation canals and HEP stations. The framework employed and results of this study contributes to the provision of quantitative information for targeting land and water management efforts, which in general are missing. It also fits well to the aspiration of the Malawian gorvernment to develop a baseline soil loss rate to help with the Agricultural Sector Wide Approach Program (ASWAp) indicator monitoring (Vargas & Omuto, 2016).

3.6.2.1 Model selection and derivation of parameters

Identification of "hotspot" areas of erosion for appropriate management interventions to tackle the major causative factors at their specific locations is imperative from both economic and ecological viewpoints. Although data are scarce, there is need for reliable information. A good model is one that can satisfy the requirements of reliability, universal applicability, ease of use with minimum data, comprehensiveness in terms of the factors and erosion processes included and the ability to take account of changes in land-use and management practices (Morgan et al., 1998). However, no single model can satisfy all these requirements or is the "best" for all applications and the choices of models generally depend upon the purpose for which they are needed, the accuracy and validity of the model, resources available and the scale and detail of application. In this study, the Revised Universal Soil Loss Equation (RUSLE)-based approach adjusted for sediment delivery ratio (SDR) was applied to identify areas that are at high risk of soil erosion and require priori management intervention. The RUSLE model was selected due to relative data availability. The basic RUSLE model (Renard et al., 1997) is expressed as :

$$RUSLE\ (t\ ha^{-1}y^{-1}) = RxKxLSxCxP \qquad 3\text{-}11$$

Where, R = rainfall erosivity (MJ mm ha^{-1} h^{-1} y^{-1}); K = soil erodibility (t ha^{-1} MJ^{-1} mm^{-1}); LS = slope length-steepness (-); C = land use/cover (-); and P = conservation/management (-) factors.

The data required by the model were collected from different sources and pre-processed to be used as model inputs. The parameters of the model and the methods of their extraction are discussed below.

3.6.2.2 The slope-length (LS) factor

Topography defines the effects of gravity on the movement of water in a watershed and therefore influences many aspects of the hydrological system. Terrain geometry and characteristics (slope, aspect, and curvatures) have significant impacts on the spatial distribution of erosion/deposition processes and are key inputs for hydrological models (I. D. Moore et al., 1991; Renard et al., 1997). Terrain attributes can easily be derived using GIS and other associated hydrological models provided that sufficiently detailed DEMs are available. The most common sources of DEMs are digitized contours from existing topographic maps. For the study site, the recently released 30-m DEM-STRM was used to estimate terrain attributes. Once the DEM was downloaded, appropriate pre-processing steps were undertaken to fill pits/sinks in order to "route" runoff to the catchment outlet without facing "unnecessary obstacles (Tamene et al., 2017a). After necessary pre-processing, the LS-factor at landscape scale were estimated based on the unit contributing area and slope steepness (I. D. Moore et al., 1991):

$$LS = (m+1)\pi \left[\frac{A_S}{22.13}\right]^m \left[\frac{sin\beta}{0.0896}\right]^n \qquad 3\text{-}12$$

Where, m (0.4 – 0.56) and n (1.2 – 1.3) are slope length and angle coefficients; A_s is the specific upslope contributing area per unit length of contour; β is the local slope gradient (degrees). For both A_s and β , it is assumed that within the 30m pixel distance, the water flow for land with ridges across the slope may not be significantly altered as ridge breakage has been observed (Mohamoud & Canfield Evan, 1998). The unit contributing area (A_s) is calculated by multiplying a flow accumulation grid with the cell size (I. D. Moore et al., 1991) as:

$$A_s = \frac{1}{b_i}\sum_i^n (a_i \mu_i) \qquad 3\text{-}13$$

Where, a_i is the area of the i^{th} grid cell; b is the contour width of the i^{th} cell (approximated by pixel resolution); μ_i is the weight depending upon the runoff generating mechanism and infiltration rates; N is the number of grid cells draining into the i^{th} grid cell. In this study we used μ=1 assuming that rainfall excess is generated uniformly over the landscape (I. D. Moore et al., 1991).

3.6.2.3 Rainfall erosivity (R) factor

Soil loss is closely related to rainfall through the detaching power of raindrops striking the soil surface and the transportation power of runoff (Wischmeier & Smith, 1978). The R-factor, defined as the product of kinetic energy and the maximum 30 minute intensity, is a very good representation of rainfall intensity and can be related to the erosivity of rainfall events (Wischmeier & Smith, 1978). In the study region, measurements of kinetic energy and raindrop size are not readily available, especially when large geographical areas are involved. Consequently, empirical relationships have been established between rainfall intensity and kinetic energy (Renard et al., 1997). In such cases, total annual or monthly rainfall can be used to estimate rainfall intensity provided that appropriate calibrations are made. Different researchers have tried to derive the R-factor based on monthly or annual rainfall data of representative stations. Some relationships developed for African conditions that are widely used include those for Botswana (van der Poel, 1980), eastern Africa (T. R. Moore, 1979) , and West Africa (Roose, 1977). To-date, there is no adequate literature available for Malawi that enables to derive the R-factor based on monthly/annual rainfall data. Hence, the equation by Moore was adopted and R-factor was estimated from annual rainfall as follows:

$$R = 0.029 * (11.36 * Rainfall - 701) - 26 \quad 3\text{-}14$$

where $(11.36 * Rainfall - 701) = RE$, based on the estimated correlation between rainfall energy RE and total annual rainfall for areas receiving less than 1,250 mm yr^{-1} like that of the study area which is as high as 0.962 (T. R. Moore, 1979). Rainfall records for nine years (2009 – 2018) were obtained from the weather station at the Nsipe Extension Planning Area, which is situated around 10 km north of the study site.

3.6.2.4 Soil erodibility (K) factor

The K-factor is defined as the rate of soil loss per unit plot reflecting the susceptibility of soil materials to the 'moving/shearing' forces of running water or rain droplet (Renard et al., 1997). Soil erodibility is an important factor that determines the relative easiness of the soil for detaching and transporting forces and is mainly a function of texture, organic matter (OM) content, structure and permeability (Wischmeier & Smith, 1978). The equations developed by Auerswald et al. (2014) were used to derive the K-factor values for the study area.

$$K = 2.77 * 10^{-5} \left(f_{si+vfsa} * (1000 - f_{cl})\right)^{1.14} * (12 - SOM) + 0.043 * (S - 2) + 0.033 * (P - 3) \quad 3\text{-}15$$

$f_{si+vfsa}$ is % silt + % very fine sand, f_{cl} is % clay, *SOM* is % soil organic matter, *S* is structure index from 1 - 4 increasing from very fine granular, blocky, platy, or massive, *P* is permeability index from 1-6 from rapid to very slow. The digital soil maps developed in section 3.4 were used for the texture and SOM distribution and for derivation of the permeability parameters. *P* was derived using the established relationship with bulk density as indicated in Table 3—10. S was determined by packing density (*PcD*) which is a function of bulk density (*BD*) and percent clay content (FAO, 2006a).

$$PcD = BD + \%clay * 0.009 \quad 3\text{-}16$$

If if $PcD < 1.4$, $S = 3$; if $PcD >= 1.4$ to 1.7, $S = 2$; and if $PcD > 1.7$, $S = 1$ (Jones et al., 2008). Here we take note that the *S* and *P* values are inverse indicating for instance that the higher the S index value, the lower the *PcD* resulting in higher K-factor and its larger contribution to erosion.

Table 3—10 Permeability Index estimated from Bulk Density (g cm3)

P	BD range	Ped shape (soil structure)	Field observation
6	<0.9	granular	Many pores, moist materials drop easily out of the auger. When dropped, sample disintegrates into numerous fragments
5	0.9-1.2	single grain, granular	Sample disintegrates at the instance of sampling, many pores visible on the pit wall
4	1.2-1.4	subangular blocky	When dropped, sample disintegrates into few fragments, further disintegration of sub-fragments after application of mild pressure
3	1.4-1.6	angular blocky	Knife can be pushed into the moist soil with weak pressure. Sample remains mostly intact when dropped, further disintegration possible after application of large pressure.
2	1.6-1.8	platy	Knife penetrates only 1–2 cm into the moist soil. Sample remains intact when dropped, no further disintegration after application of very large pressure.
1	>1.8	prismatic	Very large pressure necessary to force knife into the soil, no further disintegration of sample. Sample remains intact when dropped, no further disintegration after application of very large pressure.

Source (FAO, 2006b)

3.6.2.5 Cover-management (C) factor

Land-cover types play a significant role in the variability of infiltration capacity, runoff potential and erosion risk. When land is covered with vegetation, total roughness can be high, which can increase the runoff threshold and reduce erosion. When the land has poor surface cover, its roughness decreases, ultimately resulting in a lower runoff threshold and a quick response to rainfall. Mature forest parcels do not generate runoff and are therefore hydrologically isolated, while arable land areas can be considered as being hydrologically continuous (Desmet & Govers, 1996). Human interventions such as repeated cultivation and overgrazing can result in surface crusting and increase runoff potential. Accounting the spatial variability of surface cover is therefore one of the most decisive elements of soil erosion assessment. In the RUSLE model, the C-factor is used to handle the impacts of surface cover on soil erosion and redistribution processes and transform natural erosion potential due to rainfall, soils and terrain into to actual soil erosion risk involving human practices (Wischmeier & Smith, 1978). Typically, established C-factors within the regions are assigned to land use and cover classes. Since 1992, forested areas in the study area have been cleared creating a sparse tree mosaic with grass undergrowth (Braslow & Cordingley, 2016). Within smallholder farming systems of the escarpments, land use or cover classes are complex to distinguish.

The greening between cultivated and non-cultivated areas is usually indistinguishable over short distances (Forkuor et al. 2015). The natural vegetation is dominated by open woodlands with a layer of grass as an understorey interspersed with savannah grasslands (Campbell 1996). Moreover, trees are either left to grow or planted within the grassy (cereal) dominated farmlands, thereby creating an indistinguishable mosaic of trees and grass/crop as an understorey. Other sources of confusion include management (e.g. planting dates, plating patterns, weeds and fertilisation), background soil and other environmental conditions such as clouds. As a result, even within the same crop class, the spectral identities may change over time as the crops grow.

For the natural vegetation, the deciduous trees and grasses regain leaves during the start of the rainy season and reach peak vegetative growth at the same time as annual crops. The trees maintain their leaves until the onset of the dry season. Herbaceous plants including grasses and forbs have similar physiological growth pattern as annual crops. The differentiation between different fields may vary over years due to rotation and mixed cropping. There is temporal overlap of the signatures within and between classes. The national land cover or land use maps employ satellite imagery with coarser spatial resolution as co-variates (FAO, 2012). Hence, the maps may not be of practical use to capture cultivated land at the scale of smallholder farmers (Forkuor et al., 2017). We assume that non-cultivated lands within the landscapes are either inherently infertile or fallow lands where farmers have observed no productivity response.

Therefore, to avoid confusion when allocating unvisited farmlands to the remaining population, the mapping aimed at demarcating the boundaries for uncultivated fields and dwelling areas. Manual digitising was found to be problematic with overlapping or discontinuous boundaries even for medium-scale customary estates (Deininger & Xia, 2017). The field boundaries are fuzzy due to unregulated sharing, fragmentation and selective expansion to uncultivated areas.

Considering the complex configurations, sequential mapping was used (Braslow & Cordingley, 2016). The aim was to map and mask uncultivated areas and built-up residential areas. The first step involved visual identification and digitizing of parcel boundaries with the aid of the high resolution Google Earth imagery of between June 2016 to December 2017 (Google Earth, n.d.). At this stage, the typical confusions such as fields with moderate tree cover and an under-storey of either grass or cereal crops were somehow discerned. The mapped boundaries were printed and used for validation within each village with the residents, and where in doubt, verification was done through field visits. The land cover and land use therefore had three distinct categories and the C-factors established for African bushland (0.02-0.04), cereals (0.1-0.17), cereal intercropped with pulses (0.15) and built up residential (0.13) were used (Breetzke et al., 2013; Henao & Baanante, 1999).

3.6.2.6 Support practice (P) factor

The severity of erosion in an area is dictated by the degree of conservation practices in place or the magnitude of human influence exerted on it. The P-factor gives the ratio between the soil loss expected for a certain soil conservation practice to that with up- and down-slope ploughing (Wischmeier & Smith, 1978). The P-factor values are generally derived from spatially distributed land conservation or management data. Although reports indicate that some conservation measures exist within the study region and farmers continue to implement soil and water conservation measures (CIAT, 2016), there is no adequate information available about the spatial distribution of these management practices at the required resolution. Others report that no significant conservation practices have been implemented across the study watershed (Vargas & Omuto, 2016). However, during the study period, there were some efforts to introduce soil and water conservation measures (CIAT, 2016) but the scale was still too small to be captured during the survey. Since farming is done using a hand hoe and ridges of 30 cm height are aligned across the slope, we set the P-factor for cultivated fields at 0.9 (Tamene et al., 2017b). Since there is little disturbance to soils of uncultivated areas, a P-factor value of 0.65 was used, which is between that of forested area and strip cultivation (Tamene et al., 2017b).

3.6.2.7 Estimate sediment delivery ratio

The RUSLE is designed to predict annual soil loss without considering intermediate potential deposition (Tamene, 2005). However, not all of the soil eroded from the upper portions of a watershed are delivered to a point downstream (Stefano et al., 2005). Depending on surface cover and terrain characteristics, much of the material can be re-deposited at locations where the momentum of the transporting water is insufficient to carry the eroded material downslope or along the channel. Generally, the potential of intermediate deposition increases as the watershed area increases because there are more opportunities for eroded sediment to settle. To estimate the proportion of eroded sediment that leaves a given parcel, the sediment delivery ratio (SDR) which relates the sediment transported to a location in the channel system to the gross erosion from the drainage area above that point (Stefano et al., 2005; Tamene et al., 2017b) was calculated as:

$$SDR_i = exp\,(-\beta * \frac{L_i}{R_i S_i^{1/2}}) \qquad 3\text{-}17$$

Where, β is a routing coefficient; L_i is the length of i^{th} segment in the flow path and is equal to the length of the side or diagonal of a cell depending on the flow direction in the cell; R_i is coefficient based on surface roughness characteristics derived from the digital elevation model; S_i is the slope gradient.

The β coefficient represents 'watershed specific' parameter to characterize effects due to roughness and runoff along the hydrologic path and primarily depends on watershed morphological parameters (Ferro et al., 2003a). In this study, we used a β coefficient of 0.0014 which corresponds to the catchment size and LS-factors for the RUSLE estimated by Ferro et al. (2003b).

3.6.3 OUT4: Leaching

In addition to overland losses through erosion, nitrogen and potassium which are loosely held to soil particles leach to lower soil horizons. Data on nutrient leaching African soils are sparse. The amounts of nutrients leached (kg ha^{-1} yr^{-1}) vary with soil attributes (clay content and soil organic carbon), nutrient levels in inorganic and organic inputs as well as nutrient uptake by plants. The empirical relationship expressed by Lesschen et al. (2007b) as:

$$N\_leached = (0.0463 + 0.0037 * (P / (C * L))) * (F_N + D * S_N - U) \qquad 3\text{-}18$$

$$K\_leached = (-6.87 + 0.0117 * Rainfall + 0.173 * (K\_fert + K\_orga) + 0.265) * CEC \qquad 3\text{-}19$$

where

P = precipitation (mm year-1) , C = clay (%), L = layer thickness (auger depth) (m), F_N = mineral and organic fertilizer nitrogen (N_fert + N_orga) (kg ha^{-1} year^{-1}), D = decomposition rate of organic matter, S_N = amount of nitrogen in soil organic matter (SOM) which is assumed to be the total N (TN) in the soil as high correlations between SOM and TN have been observed (Tamene et al., 2019), U = Uptake by crop (N_Product + N_Residues) (kg N ha^{-1} year^{-1}) and CEC = cation exchange capacity (cmol kg^{-1}).

3.6.4 *OUT5: Gaseous losses*

Nitrogen gaseous losses through denitrification and volatilisation occur mostly under wet flooded conditions and in alkaline soils, respectively (Stoorvogel & Smaling, 1990). Therefore, in the Rift Valley escarpments, losses through these processes are minimal and yet to be measured. The *N_gase* (kg ha^{-1} yr^{-1}) is therefore estimated using the transfer function derived by Lesschen et al. (2007b). It is a function of annual rainfall (mm yr^{-1}), fertilizer and manure input and available soil organic carbon:

$$N\_gase = 0.025 + 0.000855 * Rain + 0.13 * (N\_fert + N\_orga) + 0.117 * P\_SOC \quad 3\text{-}20$$

Soil organic carbon (C_{SD}) is lost due to degradation of SOM by microbes under the environmental conditions (*f*) as computed for IN2. The C_{SD} is adjusted for soil active carbon (*ActiveC*) estimated for the study area of 0.034 ± 0.016% C (H. Wang et al., 2019). The soil organic matter degradation rate widely used of 2% is based on the initial studies by Van den Bosch et al. (1998) which could be lower. However, the study by Mpeketula (2016) found that between 1990 and 2013 soil organic carbon in croplands of Malawi declined from 10.3 to 9.5 g kg^{-1}, indicative of a loss rate of 0.0615 per annum. Hence the carbon loss due to degradation is estimated both via the annual decay rate established by Mpeketula (2016) and empirically using the model by Van den Bosch et al (1998), which is parameterised as:

$$C_{SD} = ACTIVEC(KG/HA) * BD * Depth * SOM_{DEGRATE} * f\text{-}env \quad 3\text{-}21$$

3.7 Full soil NPK and SOC balance

The nutrient balance has several implications. Of major concern is that low nutrient input leads to low food production and soil degradation. On the other hand, excessive application of major nutrients especially nitrogen and phosphorus can lead to nutrient drain into the environment causing pollution and low profitability. The nutrient (NPK) and SOC balance are computed as the difference between inputs and outputs. The total soil nutrient (X_{soil}) after a growing calendar year (*t+1*) is expressed as:

$$f(X_{soil}^{t+1}) = f\left(X_{soil}^{t} + \frac{1}{\rho CD}\left(\sum X_{in}^{t} - \sum X_{out}^{t}\right)\right) \quad 3\text{-}22$$

Where *X* includes the nutrients NPK and the SOC, whilst *ρ, C* and D are soil density, coarse fragments factor and profile depth, respectively. The depth is set at 10cm considering that the ridge spacing is 60, 75 and 90 cm and corresponding heights are 20, 25 and 30 cm (Mloza-Banda et al., 2014). Annually, farmers re-make the ridges by hand hoeing and minimise drudgery by scraping to the minimum depth possible. Moreover, fertilizer and manure are placed at 5 cm depth on the ridge as it has been established that such supplements largely contribute to nutrient stocks for the upper 0-10 soil layer (Ibrahim et al., 2015).

The plots and landscapes have different initial states at time *t*, as well as the *in* and *out* pathways. The NPK and SOC share pathways such as inputs of organic materials, crop residues and sedimentation and the output through soil erosion. The inorganic fertilizer has been the major source for N and P and only since recently, it has become a minor K source. Biological nitrogen fixation is a source of N when the farmer decides to grow legumes on a plot. Rainfall deposition adds to the NPK stocks whilst crop harvest and residues, if not returned, are major human pathway taking NPK out of the soil and farming system. Leaching is a significant output pathway for N and K and gaseous losses contributes to N output.

3.8 Economic costs, income, losses and benefits

As farmers invest their labour, land and other financial resources to produce food and attain food sufficiency, considerable amounts of resources and energy go to soil fertility management. Hence their decisions to undertake various forms and degrees of soil fertility management are strongly influenced by economic considerations. Over time, several programmes and policies have been made that have had a direct or indirect impact on pricing of both inputs and outputs. The notable example is the introduction of the Malawian farm input subsidy programme which subsidies inputs such as fertilizer and seeds with the aim of making them affordable and accessible to poor farmers. Using classical economics, it is assumed that farmers would decide to apply more inputs if benefits outweigh the costs. Such cost-benefit analyses have been widely used for measuring both short- and long-term economic impacts.

The paucity in data capturing initiatives and lack of standardised monitoring and evaluation frameworks make it difficult to precisely quantify the impacts on soil fertility depletion or replenishment (Vågen et al., 2016). Approaches such as potential loss of production and replacement costs of depleted nutrients has been used (De Jager et al., 1998), but still there are hidden and off-site costs such as irreversible eroded top soil that is hardly accounted for (Telles et al., 2013). It is even more challenging to project these into the distant future. Considering the spatial extent and availability of data, we estimate the onsite-costs of nutrient loss, yield loss and attempt to evaluate land values. We consider sedimentation as an externality that benefits downstream farms.

Land values are estimated from the average rental prices for the period 2002 – 2012 (Chamberlin & Ricker-Gilbert, 2016). These are adjusted using the productivity index (Table 3—12). Cost of erosion and benefit from sediment deposition are estimated as the value of crop yield drop due to erosion empirically established in Zimbabwe and west Africa averaging 8.2% per year (Panagos et al., 2018), linearly adjusted for net erosion or net deposition.

The economic benefit for deposition, leaching, nitrogen loss through gaseous and residues are estimated from the potential replacement cost for N and P. Price for N, P & K are derived proportional to the nutrient composition in Urea (46%N), 23:21:0+4s (N:P:S) and super-D (14N:28P:14K) fertilizers, respectively.

3.9 Integrated scenarios of structural and functional changes in the agro-ecosystem

At the doorstep of smallholder farmers, to maintain or improve soil fertility, they strive to either increase the nutrient input or reduce losses. The maize mixed farming system of East and Southern Africa is characterised as a nutrient mining system with overtly negative nutrient balances (Cobo et al., 2010). The problem is particularly grave for sloping escarpments, where despite the rugged landscapes and low productivity (Li et al., 2017), minimal measures are put in place to increase input or reduce nutrient output flow. In the past, land would be left under fallow to regain productivity. With increased pressure, even uncultivated lands are barely revegetated but rather continuously cultivated. Of late, farmers are increasingly becoming aware of the challenges, and there has been increased interest and efforts to restore landscape health and reduce the nutrient gap (Braslow & Cordingley, 2016).

Considering the sheer number of decisions that farmers have to make and ecological processes contributing to balance of nutrient at plot-household level, projecting farm sustainability requires an integrated approach. Much as integrated analysis is desired, it is challenged by complexity. Complexity is a result of several factors namely: *invisibility* in that soil nutrients are invisible and difficult to measure directly, *heterogeneity* in terms of number of casual factors and their outcomes, *non-linear* and *interdependent* processes, *nesting* of factors, processes and outcomes within social and ecological hierarchies, and *cumulative* effects leading to emergence of unforeseeable system states and equilibriums. These features pose both challenges and opportunities to modellers, policy makers and land managers as detailed in Le (2005) and Villamor (2012)and briefly indicated below.

In pursuit of their livelihood strategies, land managers aim at maximising products such as crop yield by judicious use of inputs. However, the chemical, biological and physical processes that transform inputs to outputs are largely *invisible* to them. In these smallholder farming systems, the paucity of basic data on farm inputs and outputs is indicative of little attention given to data generation.

From the account of drivers of farmers' decision to invest in soil fertility management, it is evident that farming communities significantly differ in characteristics and therefore take different actions (see section 2.1). For the three most commonly used inputs, the heterogeneity in resource endowments and social capital has risen into varying usage levels (section 1.2). The landscape too is also quite variable in terms of nutrient contents and topographic features that are directly or indirectly linked to both decisions: to re-invest and the anticipated output levels.

Since the input and output flows and stocks are directional, single reductionist models of cause–effect are assumed to be the most appropriate and widely used (Akinola et al., 2010; Lesschen et al., 2007b; Smaling & Fresco, 1993). However, there are interdependencies not only among the driving factors but also among the outcomes with the combinations and combinatorics which tend to be exceedingly large to disentangle (Mponela et al., 2018). In some instances, a single factor may even have opposing effects on the same or different input or output processes (Howard & Matheson, 2005; McDonald & Moffitt, 1980). In a single directional bio-economic model, the estimates for nutrient inputs are used as drivers for the nutrient output processes. Farmers are conditioned to change their behaviour over time depending on their experiences, knowledge and beliefs accumulated over time. Hence the output, from previous year, though endogenous, sets precedence for the next action. The interlinkages creates a casual web such that one casual variable drives one or several others and vice versa (Le, 2005).

Without considering the chemical reactions that happen at atomic and molecular level, even at farm level there are cobwebs of ecological entities and their process that aggregate, interact and emerge as dynamic landscape states and processes. Similarly, actions by single farmers aggregate at community level forming a plot/farmer –

landscape/community. The scale aggregation entails that nutrient flows initiated by farmers' actions and the resulting nutrient stocks are constrained by exogenous factors operating at higher level and controlled by endogenous factors operating at lower level. These nested hierarchies create a constraint envelope and instability, as a model that empirically establishes casual relationships at one level in space and time could be endogenous and multi-directional at a higher level or a later stage (Le, 2005). In a real farming system, all processes are interwoven and impossible to disentangle. For instance, it seems plausible to establish a partial budget and evaluate fertilizer input and its impact on yield. But not all fertilizer is used by the plant and not only fertilizer supply nutrients to the plant. This makes tractability a challenge especially when the aim is to forecast sustainability.

Forecasting in most cases depends on having prior knowledge of the system and based on histories, a possible future pathway is premised. Typically, scientific experimentation is needed to develop deterministic models. Through series of observations and experimentation, casual relationships have been developed which show input and output models but are usually simple and straight forward (Lesschen et al., 2007a; Smaling & Fresco, 1993; Van den Bosch et al., 1998). These mono-directional models were developed for small-scale single farms, homogenous landscapes (such as efficient large-scale farms) and for national aggregates that are useful to guide short-term planning. For heterogeneous landscapes with in-efficiencies in almost all production factors and high rates of factor substitution (Larson et al., 2012), the power of prediction is reduced. In addition to prediction and validation challenges, data for the associated drivers and even the outcomes are either limited or inconsistently captured for most smallholder farming systems, making the predictions even more uncertain.

Social-ecological transformations in farming systems emerge from complex and adaptive feedback loops between interventions by land users and the underlying ecological processes, which continuously regulate the resource base. The drivers of transformation can either be proximate or underlying (Geist & Lambin, 2002). The proximate causes include ecological features such as terrain, rainfall and soil texture and agricultural practices that directly influence the nutrient input and output flows and stocks. The underlying factors are those that drive the proximate factors and the resulting nutrient balances, and that can best be explained by multiple factors and drivers acting synergistically. The changes therefore do not just have a single or a partial factor-causation but full interplay of economic, institutional, technological, cultural, and demographic variables. Much of these factors do not directly influence nutrient balance but underlie the proximate causes such as land use patterns, resource utilisation and development as thematically indicated in Figure 3.8.

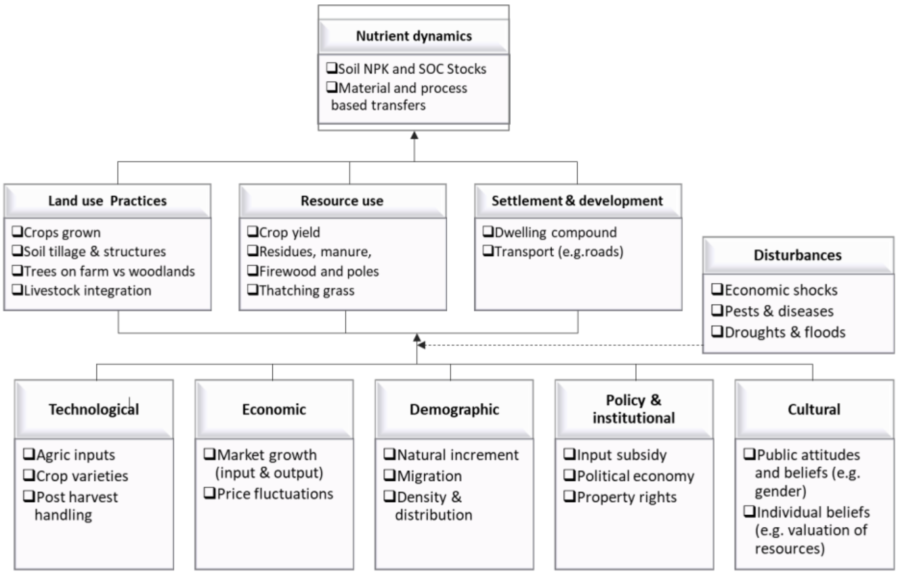


Figure 3.8: Underlying drivers (technological, economic, demographic, policies and culture) of human actions that induce dynamic changes of nutrient balances at farm level.

3.9.1 *Heuristic setting of conservation areas based on hotspot mapping*

Ensuring sustainable intensification at landscape level requires some deliberate policies to set aside degradation hotspot areas. Not all land is suitable for crop production. In Malawi due to biophysical limits and degradation levels, 28% of land is marginally suitable and 39% is unsuitable for crop production.

To map hotspot areas for conservation as well as productive areas for cultivation, we construct the land productivity index using the Liebig's linear scoring function (LLSF) (Liebig et al., 2001). LLSF is a multi-criteria evaluation procedure where land attributes are scored based on the range required for crop growth and environmental health on a scale of 0 – 1 (or a s a percentage). Values close to 1 are indicative of optimal conditions whilst close to 0 imply critical thresholds.

The scoring factors and weights are generated using empirical knowledge and fuzzy techniques such as the established critical and sufficiency levels of nutrient stocks, or topographic limits for allowable soil loss as set in national environmental policies. To be of practical relevance, the identification of hotspots for targeted management should be in tandem with community aspirations raised during focus group discussions and participatory resource mapping.

3.9.1.1 *Uncultivated vs cultivated areas*

Much as humans influence the entire landscape, for this analysis, the farmer's actions are concentrated on managed and cultivated areas. Hence, the need to crop out uncultivated areas and distinguish them from cultivated ones. The uncultivated areas include patches that are marginally poor, unsuitable for cultivation or fallows reserved for future cultivation. Although cultivated land is considered productive and has a higher factor scores there might be pockets that are degraded or too prone to degradation that need to be set aside.

3.9.1.2 *Rules for land conservation based on NPK and SOC critical values*

Although fallows have disappeared in the area due to land pressure, there are portions of land that are not cultivated either because they are inherently infertile or have reached a non-responsive state due to degradation (Kihara et al., 2016). For agricultural purposes, soil fertility has been evaluated in terms of critical and optimal plant growth requirements and ecological health of SOC, N, P, their stoichiometric rations and the functional soil quality indices (Table 3—11). Although the quantitative relationship between nutrient levels and land suitability for crop productivity vary among soil orders (e.g. texture), climatic regions and plant tolerance, the critical levels of soil conditions used in this study are generally regarded as limits below or above which substantial soil quality decline occurs (Tamene et al., 2019).

Apart from the direct mineral inputs, substantial amounts of nutrients in agricultural and other terrestrial ecosystems are stored by and released from organic matter by micro-organisms (Palm, Gachengo, et al., 2001). Microbial nutrient demand is determined by the elemental stoichiometry of microbial biomass in relation to environmental nutrient availability, which has been found to be consistently similar in terrestrial ecosystems (Cleveland & Liptzin, 2007; Frossard et al., 2016). As such, stoichiometric ratios are the basic bio-chemical thresholds at which microbial metabolic control in ecological systems switches from energy flow (C) to limiting nutrient flow (N, P), in the process either mineralising or immobilising N and P (Sinsabaugh et al., 2009). For instance, plants can be N-limited when foliar C:N > 25 whereas C:N < 20 implies C limitation (Stevenson & Cole, 1999). Most of the C:N:P research has focused on degrading plant material, and hence the thresholds are insightful especially for soil N which is mostly in organic form (Frossard et al., 2016; Tamene et al., 2019).

Table 3—11 Critical ranges of soil structural stability index (StI), soil organic carbon (SOC), total nitrogen (N), available phosphorus (P), and their stoichiometric ratios

Variable	Range	Condition	Study	Region and intent
SOC (%)	0.46 0.46-0.65 >0.65	No response Variable Responsive	(Musinguzi et al., 2013)	Fertilizer response in Arenosols (sandy soils) of Zimbabwe.
	0.7 1.0 1.6 2.1	Low fertility Moderate Most fertile Uncultivated	(Zingore et al., 2011)	Soil fertility zones based on gradients from no manure low fertilizer input, high manure input, to natural woodland in sandy clay loam soils of Zimbabwe
	1.0 1.5	Critical limit Threshold	(Rattan Lal, 2015)	Reducing soil degradation risks and reversing degradation trends in SSA.
	2.0	Threshold	(D. J.	Structural stability for English and Welsh
Total N (%)	<0.15 > 0.25	Low High	(Hazelton & Murphy, 2007)	Critical limits mostly from Australian soils.
	0.08 0.08-0.12	Very low Low	(Lowole, 1965)	Soil classification Malawi, fertilizer recommendation
P (mg/kg)	<11.0 >17.0	Low High	(Hazelton & Murphy, 2007)	
	7.42 10.6	Critical limit Threshold	(Aune & Lal, 1997)	
	15	Threshold	(Snapp, 1998)	
	4 13.5 15.6 25	Minimum Critical limit Threshold Maximum	(Bado et al., 2010)	The min-max range from original to fertilised Ultisols of West Africa. The critical limit maize yield dropped from 1 to 0.5 ton ha^{-1}. The threshold set at level below which P input is needed for yield >
K (mg/kg)	<125 >190	Low Sufficient	(van Biljon et al., 2008)	Threshold values and sufficiency levels for maize producing sandy soils of South
C:N ratio	<20 >25	N excess N limitation	(Mooshammer et al., 2014)	Terrestrial microbial decomposition of organic materials
	<20 >30	Net N gain Net N loss	(Stevenson & Cole, 1999)	Global values of net immobilisation or net mineralisation of crop residues
	<16 >23	N mineralisation N	(Enwezor, W., 1976)	Thresholds C:N and C:P rations for plant materials used as green manure as source
	16.1-23.4	Typical range	(Xu et al., 2013)	In shrubland soils of eastern Africa
C:P ratio	<200 >300	P excess C excess	(Stevenson & Cole, 1999)	
	338–797 286.5	Typical range Average	(Xu et al., 2013)	In shrubland soils of eastern Africa Boreal forests to deserts of the world
St (%)	<5 5-7 >7	Degraded High risk Low risk	(Pieri, 1995)	Sandy ferruginous tropical soils of semiarid francophone Africa

The structural stability index (StI) is indicative of the risk of soil structural degradation associated with SOM contents and texture (Pieri, 1992). The StI of ≤5% indicates that the soil is structurally degraded due to low levels of SOC and is highly susceptible to erosion, while StI >9% indicate sufficient SOC to maintain structural stability. The factor of 1.72 was used to convert the SOC to SOM and StI was computed as:

$$StI = 100 \times \frac{1.72 * SOC\ (\%)}{Clay\ (\%) + Silt\ (\%)} \qquad 3\text{-}23$$

3.9.1.3 Reducing the net nutrient loss through erosion

As one of the major nutrient loss pathways in Malawi, controlling erosion has received little attention (Sandram, 2018). The land use maps currently in use classify most of the patchy cultivated slopes as woodlands, hence land management planners do not devise active strategies for these fragile landscapes. To give an overview of potential nutrients savings that could be achieved if erosion was considered in landscape planning and management, the adjustments to erosion factors are envisaged. This also applies to cultivated plots where the crop types and vegetation cover are adjusted and the potential impacts in reducing soil and nutrient loss evaluated. In this study, we used the LAndscape Management and Planning Tool (LAMPT) developed using the RUSLE model adjusted for SDR to estimate net soil loss (Tamene et al., 2014). For the scenario analysis, we first identified the soil erosion factors that are influenced by changes in the crop cover and soil organic matter dynamics including C-, P-, and R-.

Much as farmers aspire to tackle soil loss through improving surface cover by planting trees and constructing contour bands (Braslow & Cordingley, 2016; Emerton et al., 2016), the progress is minimal. To the contrary, the sloping areas have experienced massive deforestation during the last two decades (Braslow & Cordingley, 2016; CIAT, 2016). Nonetheless, erosion experiments conducted 20 km north of the study area between 1986 – 1991 provided evidence that increasing vegetation cover with agroforestry shrubs (*Leucaena*) significantly reduced erosion by 19 folds and increased crop yields by 14 folds (Banda et al., 1994). During the 6 years of experimentation on a 44% slope, soil (and nutrient) loss was 78 - 21.6 t ha^{-1} yr^{-1} on sole maize plots whilst intercropping of maize with shrubs significantly reduced soil loss to around 3.8 – 1.5 t ha^{-1} yr^{-1}. The maize yield under agroforestry shrubs ranged between 1.5 - 4.1 t ha^{-1} yr^{-1} which was in sharp contrast to the rapid decline recorded under sole maize from 0.8 to a meagre 0.15 t ha^{-1} yr^{-1}.

On both cultivated and uncultivated lands, woody vegetation cover is sporadic due to selective retention and cutting, respectively. Despite being deciduous, trees regain leaves before the onset of the rains and provide partial surface cover. For soil conservation purposes a landscape with at least 10% woody vegetation cover is considered forested. This is in support with the agroforestry programme that supports "trees on farm", which is pioneered by the Government of Malawi and included in the 4.5 million ha pledge to the Bonn Challenge (GoM, 2017; Ministry of Natural Resouces Energy and Mining, 2017).

From the survey, the probability for a cultivated plot to be under tree cover of between >10% is estimated and the associated effects on both choice of input strategies and output flows are further examined. The cover's effect on reducing erosion on farmland was estimated as a binary tree cover of >10% vs <10% with corresponding cover factors (C-factors) assigned accordingly. For uncultivated areas, which include mostly areas that were either under fallow or low productive marginal lands on the steep slopes, the tree biomass and cover are estimated from the density and above ground biomass estimated for the study area (Tamene, Mponela, Sileshi, et al., 2016) and updated using natural growth rates established for the archetypal miombo vegetation (Chidumayo, 2019). The tree densities in the area were 201 (58-343) ha^{-1}, with height of 5.6 (5.0-6.1) m, diameter at 1.3 m height of 15.2 (12.9-17.5) cm and biomass of 183.5 (96.5-227.9) kg $tree^{-1}$. The above ground tree biomass stocks were 27.4 (11.5-35.5) ton ha^{-1}.

In the rift valley escarpments of Malawi, slope is a major constraint to productivity. Sleep slopes usually have shallow soil depths, fewer nutrients and are prone to erosion. According to the classification by FAO (2006a), land with slope of ≥ 30% is considered steep and unsuitable for cultivation. Land with slope gradient of 10 - 30% is considered sloping while 0 - 10% is considered level and suitable for cultivation. As the slope increases from 2 to 18%, a negative relationship between slope and crop yield has been observed (Al-kaisi, 2008). Therefore, areas with larger slopes are relatively less suitable for cultivation.

The focus group discussion (FGD) conducted with the aim of mapping resources, constraints and opportunities revealed that gully erosion is one of the key drivers of soil erosion in the study area (Braslow & Cordingley, 2016). Therefore, addressing gullies for watershed management should be a priority to contain mass nutrient loses. To achieve this, it is necessary to have information about the spatial distribution of major and ephemeral gullies. The potential location of gullies was delineated based on two conditions as (Tamene, 2005) as follows:

$$A_s tan\beta > 18 \qquad and \qquad ln\left(\frac{A_s}{tan\beta}\right) > 6.8 \qquad 3\text{-}24$$

Where, A_s is unit contributing area (m^2 m^{-1}); $tan\beta$ is tangent of local slope.

In addition, simulation can be run with terracing steep slope areas (P-factor = 0.6). The tool also provides an option to identify gullies and assign management options. For instance, one would explore the potential of protecting gullies through terracing and dense grass and the respective P- and C-factor values of 0.6 and 0.01 would be used as post-intervention defaults in the tool.

A fourth option targets erosion hotspots. Erosion studies revealed that the sloping escarpments experience high volume of soil loss which can be as high as 50 t^{-1} ha yr^{-1}. Cultivation of steep slopes in Ntcheu accelerated the rate of erosion to 80 t ha^{-1} yr^{-1}. These losses are exceedingly higher than the tolerable soil loss to maintain crop productivity which is close to the rate of soil formation of 12 t ha^{-1} yr^{-1} (Montgomery, 2007). This level of soil loss is assumed to be the tolerable soil loss limit considering surface lithology and soil thickness of the area

To construct a compound index, the factors and weights were applied as presented in Table 3—12. Among the management options suggested, setting aside areas that are less suitable for cultivation is primary. It is expected that if these areas are reserved during the simulation, they will regain tree cover and hence be at a reduced risk of erosion. However, it is unlikely that enclosures might turn into dense woodlands in the short-term owing to slow growth rates of indigenous miombo trees (Chidumayo, 2019).

Table 3—12 Soil productivity indicator rating and corresponding areas

Domain	Indicator	Description	Criteria	Scores	Area%
Land use	Cropping	Preferred	Cultivated	1	58
		Reserved	Uncultivated	0.5	42
Soil fertility	Total Nitrogen (%)	Very low	0.00- 0.08	0.5	58.08
		Low	0.08-0.12	0.6	42.03
		Moderate	0.12-0.15	0.7	0.19
		Optimal	0.15-0.25	1	0
		High	>0.25	1	0.19
	Phosphorus (g/kg)	Critical limit	0-7.4	0	0.12
		Low response	7.4-11.0	0.25	3.58
		Variable response	11.0-13.5	0.5	5.25
		Moderate response	13.5-17.0	0.75	9.78
		Optimal	>17.0	1	81.55
	Potassium (g/kg)	Deficient	0-125	0.5	20.48
		Moderate	125-190	0.7	69.85
		Sufficient	>190	1	9.95
	C:N ratio	Critically low	0-16	0.5	99.99
		Shrubland soil	16-20	0.7	0.01
		Optimum	20-25	1	0
		N-limit	25-30	0.5	0
		N-loss	>30	0	0
	C:P ratio	C-critical limit	0-100	0	5.21
		C-limit	100-200	0.5	20.02
		Optimum	200-300	1	14.96
		P-limit	300-500	0.7	26.42
		P-critically	500-700	0.5	20.43
		P-exceedingly	>700	0	12.97
Soil health	SOC (%)	No response	0-0.5	0	0
		Variable response	0.5-0.7	0.2	2.53
		Moderate response	0.7-1.0	0.4	45.22
		Optimal response	1.0-1.5	0.6	45.71
		Most fertile	1.5-2	0.8	6.71
		Sufficient	>2.0	1	0.49
	Structural stability Index (StI)	Degraded	0-5	0.5	97.48
		High-risk	5-7	0.8	2.52
		Low-risk	>7	1	0
Topography	Slope (%)	Level	0-1	1	16.14
		Very gentle slope	1-2	0.9	9.84
		Gentle slope	2-5	0.8	17.71
		Sloping	5-10	0.6	29.79
		Strongly sloping	10-15	0.2	13.52
		Moderately steep	15--30	0	11.64
		Steep	>30	0	1.69
Erosion	Ephemeral gullies		Yes	0.1	31.34
			No	1	68.66
	Soil eroded (t/yr/ha)	Slight	0-2	1	16.69
		Tolerable	2-10	0.9	19.46
		Manageable	10-20	0.7	22.45
		Moderately high	20-30	0.3	13.95
		High	30-50	0.1	12.66
		Severe	50-80	0	8.60
		Extreme	>80	0	6.18

*test crop for the responsiveness is maize for phosphorus and SOC

3.9.2 Simulation of social-ecological transformations of soil nutrient stocks

Several nutrient transfer models have been developed as decision support systems (DSS) that capture and integrate human actions and ecological processes (Bell et al., 2015). Methodologically, if data is readily available and there are high efficiencies in input-output flows, the projections can reliably be made using constructed histories. In which case, the projections are optimal solutions on the basis of proven inertial that drove flow of matter, energy and information in the past. However, in data sparse regions, such projections can seldom be constructed and there are methodological limitations. Therefore, ex-ante analysis of cross-sectional and spatially explicit data from real life phenomena and simulations of potential regime shifts due to policy are deemed optional and best alternative (Le, 2005).

Simulation have their origin in computer missile games where players could adjust the missiles curved path and speed ("Contributors," 1944). The same principle has been applied in science to design models of real systems, then conduct ex ante experiments to understand behaviour in terms of potential effective changes in the system or its mode of operation (J. Wang et al., 2008). The major strength is that simulations can be done for systems that already exist and those in the preliminary or planning stage of development and potentially capable of being brought into existence (Shannon, 1998). For smallholder farmers, the common mistake is to assume that non-existence of data implies that the technologies do not exist. As earlier pointed out, the soil improvement technologies under consideration have been applied for more than six decades, and by capturing the current usage and performance, we assume that the farms form a natural continuum from those that are yet to adopt the technologies to those that have used them for a long time. Among non-users, there could be some farmers that have never used them but for the majority, non-participants used them earlier but not during the study year. Hence, non-users comprise of non-adopters and dis-adopters.

This study mimics the natural environment by making use of parameters from decision, productivity and hydrological analyses as empirical evidence for initial and boundary conditions, as well as interdependences among agents and their environment. Framed by the unimodal rainfall pattern, farming has a one-year production cycle, hence a one-year time step in simulation is used. Depending on changes in state variables of both the environment and households, the social-ecological profiles of individual patches and households are expected to change as well. Due to the simulation process, for instance, the household types as empirically drawn, would have acquired/lost resources overtime and are expected to adopt the decision-making mechanism of the new group. In this way, the simulation does not use a virtual experimental approach with random datasets, it generates the population and landscape attributes and use patterns from sample data. The population attributes were generated using Monte Carlo approximation (Berger & Schreinemachers, 2006), whilst landscape level statistics were extrapolated using GIS-based proportional up-scaling from the measured sample data (Le, 2005). In order to identify pathways for sustainable land management, different scenarios are built based on existing evidence from research trials, government policies and community aspirations. The parameters from the empirical models were used to indicate probability and intensity of change in both human decision and environmental models.

Each farming calendar year, farmers make two decisions with regard to nutrient inputs: *First,* whether a household receiving input allocates them to a particular plot or not. For this first decision, the predicted probabilities are estimated as a sum of weighted average reflecting the factor distribution and allowing inference to the total population from which the data was drawn. This is unlike using the estimated *margins at mean values* whose inferences are only relevant to the observations closer to the mean and with dichotomous factors, the estimation is made for every observation. Using the sigmoid (z) function, the predicted probabilities $\widehat{P_{x_i}}$ lie between 0 (non-adopt) and 1 (adopt) and are estimated as an odds ratio given the following function:

$$\widehat{P_{x_i}} = \frac{\varepsilon^{\beta_0+\sum\beta_i X_i}}{1+\varepsilon^{\beta_0+\sum\beta_i X_i}} \qquad for\ 0<p<1 \qquad 3\text{-}25$$

Where β_o and β_i are estimated regression coefficients, X_i's are the observed or measured determinant factors and $\varepsilon^{\beta_0+\sum\beta_i X_i}$ is the estimated logit score for each household-plot.

Secondly: the subsequent input levels are based on the estimated probabilities for the whole population and usage levels for those implementing the technologies to increase nutrient input and reduce output, or otherwise. We used the GLM with binomial family and log link of the functional form:

$$logit\ E(Y) = \beta_0 + \beta_i x_i$$

As described earlier, the quantity of nutrient input and output from a farming unit is a result of cumulative and nested processes at ecological scale of *watershed (landscape (plot (pedon)))* and the social scale of *region (community (household (individual)))*. The drivers at any level can either be endogenous or exogenous depending on the hierarchies but their effects transcend the boundaries and scales. This study analyses nutrient balance at the plot-household level; hence, community and landscape level attributes are considered external whilst plot-pedon and household-individual level attributes are endogenous. In determining the dynamic changes in usage of nutrient input technologies and the nutrient output processes, we examine the changes in exogenous policy actions but also control for the main internal factors. Among the various policies that the government provide and communities appreciate as viable policy options to enable them to replenish lost nutrients and increase crop yields, the farm input subsidy program (FISP) is noticeably featured. The alternative subsidy scenarios constructed were based on trends and political agenda (Table 3—13).

At household level, differences in gender, family labour and education have a strong bearing on farmers' decision to increase nutrient inputs. Their current usage is estimated from real-farm household and agronomic survey data while the potential effects are empirically determined. There are other factors, such as the period that the plot has been under cultivation and age of the household head, that change progressively. Since in the dataset we captured the length of time since conversion from natural/ long fallow to cultivation, all plots have different cropping ages which is updated individually as the simulations run. For the age of household head, it is updated by taking into consideration the life expectancy and inheritance systems. A random sample of household heads who are above 63 years (the life expectancy for Malawi) will die and their households are inherited by offspring who enter into marriage.

Based on the projected potential effects described above, the following scenarios are constructed and implemented, as virtual experiments in MASSAI (Table 3—14). For the baseline, all policy and static variables are held at the current levels but progressively changing variables such as age of household head and the cultivation period for the plot are changed. To compare the current policy with alternative regimes, we used the Bonferroni multiple comparison test which tests significance based on individual *p*-values between pairs (Hochberg, 1988)

Table 3—13: Trends, aspirations and policy settings

Scenario	Description
Baseline (BAS)	Current trend, with changes resulting from internal factors that progressively change over time. These include household attributes such as age of household head, the natural nutrient input and output flows, and plot features such as time it has been under cropping. By bringing these factors to the baseline levels and project their progressive natural changes, the effect of policy interventions is therefore additional and conditioned on existing conditions.
SUBSIDY	In Africa, due to high levels of poverty and over-reliance on agricultural production, the governments support farmers with various forms of subsidies as social security policies. In Malawi, the government distributes coupons for subsidized fertilizer and seeds to almost a third of the population. Once received, in some villages, members share the cost of fertilizer and seeds. Although the subsidy program is a pro-poor strategy aimed at reducing the nutrient gap, the sharing of coupon price and inputs makes it difficult to discern recipients from non-recipients. From our survey, we captured the total cost farmers paid for the fertilized applied on a plot and divided with the market price to find the share of subsidized fertilizer. The share of subsidized fertilizer applied to maize for the 2016/2017 growing season averaged 28% (range 0-100). The shift in fertilizer prices due to changes in subsidy regimes is expected to induce correspondingly shifts in demand for inputs. The scenario spectrum for assuming the impacts of subsidy on nutrient inputs and output are *four*. These are drawn empirically and projected based on historical trends (Figure 1.1) and include SC (current subsidy), SR (incremental reduce by up to 28%), SZ (incremental reduce by 100%), and SU (incremental increase by 150%). These correspond to the political aspirations and alternative policies that aim at either increasing or reducing subsidy and those that call for removal and replacement of the current subsidy regime. Considering that farmers act autonomously with differing cognitive capabilities and resource endowments, the changes in input demand triggered by input price changes ought to be multi-directional and farm type dependent. Hence, diverse soil management behaviours ought to emerge, leading to changes in farm productivity and nutrient balances. *IN1 and OUT1*: Farmers receiving subsidy tend to find it cheaper to access fertilizer than non-recipients do. Hence, fertilizer subsidy is assumed to be associated with increase in the probability to apply more inorganic fertilizer. We ought to be aware that those purchasing from the market, hence having lower share of subsidized fertilizer, would be the well-off farmers and could even buy more than those that rely solely on subsidy. The input subsidy affects the input prices with potential effect on output levels and the price bargaining power depending on the level of subsidy received by the farmer. Farmers receiving subsidized "*free*" fertilizer might not be as committed to manage the crops as those who purchase the fertilizer, thereby having differentiated nutrient output flows and benefit. *IN2 and OUT1*: Moreover, those receiving subsidized fertilizers, by having increased likelihood to apply inorganic fertilizer, would be less likely to invest in alternative nutrient sources such as manuring. *IN3 and OUT1*: Since subsidy includes legume seeds, we hypothesize that farmers receiving subsidy would have higher probability to plant legumes or allocate more land to legumes than non-recipients would.
Women empowerment	Input use choices are sensitive to changes in the role of women in decision-making. The scenario spectrum for assessing the impacts of women empowerment are *two*. The first spectrum of scenarios Weai_S0 (current scenario) and Weai_S1 (increased women involvement by 212% from the current average of 0.16 to 0.5).
Labour and dependency	The second most constraining factor at household level is labour availability. It affects the households' decisions to adopt the SFM technologies and the level of care for crops, hence affecting both nutrient input and output efficiencies. Households labour allocation decisions are affected by their responsibility to care for the members. For households with more dependants,

	they have lesser opportunities to invest in inputs for SF enhancement, and are also constrained to invest in crop management. The labour, therefore, is represented by the number of hours the family members allocate their effort to a plot and the dependency ratio. The changes in number of hours are derived from the utility function based on the expected benefit in crop yield relative to labour input and output and are thus empirically estimated from decision and crop yield models. From the survey farmers reported to use on average 96 manhours / ha. The recommended labour input is 64 manhours per hectare. Hence, we explore the effect of increasing labour productivity by reducing the total hours worked by 33% over the simulation period. The dependency ratio, though empirically linked to decision and yield models, is dependent on generational changes in demographic structure. The population of Malawi grows at an average rate of 2.9% per annum (GoM, 2018) which increases pressure on land and alters the dependency ratio. We should have used population growth model and age dependent death rates to estimate probability of loss of household members among workers and non-workers. The birth rates could contribute to the younger age class while the adjustments in other age groups would be estimated from the crude population growth rates.

Table 3—14 Policy settings for developing integrated SFM scenarios

	Subsidy	Gender	Family labour		Progressive	Held constant
Scenario	Share of fertilizer subsidised $(\beta\text{-}P_{SUBSIDY})$	Women empowerment $(\beta\text{-}H_{WEAI})$	Labour input $(\beta\text{-}P_{LABOUR})$	Dependency ratio HH $(\beta\text{-}H_{DEPR})$	Age HH head, cultivation period	Elevation, sand, slope, income, education etc.
units	% USD	WEAI	Man-hours	Workers/ dependants		
Value	$\bar{x} \pm \bar{x} * r$	$\bar{x} + \bar{x} * r$	$\bar{x} - \bar{x} * r$	$\bar{x} - \bar{x} * r$	$\bar{x} + \bar{x} * r$	$\bar{x}$
Share of subsidised fertilizer per plot:						
Sub_S0 (current)	28.6	0.16	96.26	1.78	$\bar{x} + \bar{x} * r$	$\bar{x}$
Sub_S1 (decrease)	20	0.16	96.26	1.78	$\bar{x} + \bar{x} * r$	$\bar{x}$
Sub_S2 (…to zero)	0	0.16	96.26	1.78	$\bar{x} + \bar{x} * r$	$\bar{x}$
Sub_S3 (universal)	70	0.16	96.26	1.78	$\bar{x} + \bar{x} * r$	$\bar{x}$
Gender (role of women in household decision making):, labour and dependency						
S0 (current)	28.6	0.16	96.26	1.78	$\bar{x} + \bar{x} * r$	$\bar{x}$
Weai_S1 (increase)	28.6	0.5	96.26	1.78	$\bar{x} + \bar{x} * r$	$\bar{x}$
Labour_S1 (decrease)	28.6	0.16	64	1.78	$\bar{x} + \bar{x} * r$	$\bar{x}$
Dep_S1 (decrease)	28.6	0.16	96.26	1	$\bar{x} + \bar{x} * r$	$\bar{x}$

$\bar{x}$ is the initial state for each attribute (2016/17 baseline), *r* is the annual rate of change to the alternative scenario.

3.10 Model validity, sources of variability, stochasticity and sensitivity analysis

The sample dataset used for empirical estimations and predictions was obtained from 238 households with 451 plots, hence for the whole sample, the number of subjects per variable (SPV) is adequate, even if the rule of thumb of 10 PSV is followed. For instance, observation (451) to variable (21) ratio for maize yield estimation and prediction was 21:1 which potentially gives reliable estimates. With yield model's SPV of 4, 8 and 6 for farmtypes 1, 2 and 3 respectively, the type-specific results may not reliably predict outside the dataset. For Monte Carlo simulations used in the simulations for this study, the recent studies have shown that the minimum SPV of two is adequate for estimation of regression coefficients, standard errors and confidence intervals (Austin & Steyerberg, 2015). Hence, for the type-specific analyses, the inclusion of a number of variables ensured that most of the causes of variations were considered and the simulations with 10 replications thereof on a population of 2640 plots, are assumed to give a more nuanced representation of yield responses for the study population and the landscape. However, caution should be exercised if they are to be used outside the study region.

In addition to observation to variable ratios, the inclusion of variables in models from a whole set of 64 variables was done by differing specifications. The software we used for estimations, Stata, has inbuilt feature that drops one of the explanatory variables if two factors are highly correlated i.e. there is multicollinearity. Potential multi-collinearities among continuous variables were screened and only variables that were less correlated were included in the alternative model specifications (see Appendix S4). The Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were used to select the model forms and functions that best modelled farmer SFM decisions and crop productivity (Vrieze, 2012). Several other tests were done to ensure validity of the models while capturing variability and at the same time addressing stochasticity. Validity and fitness of the empirical models were determined using measures of goodness of fit. For soil nutrient and SOC spatial distribution, the out of bag error was used (Breiman, 2001).

Both the human and ecological attributes on which the empirical models are developed are inherently variable. The households differ in demographics, resource endowments and nutrient input use, and are categorized into low input (*farmtype1*), medium input (*farmtype2*) and high organic input (*farmtype3*) farms. The plot and landscape features also vary, which coupled with household attributes frame the differentiated decisions, soil nutrient distributions, yields, nutrient balances and economic benefits among human and ecological agents.

The variability leads to yet another analytical challenge, stochasticity in estimated parameters. Stochasticity existed in the models for estimating nutrient distributions, and for predicting probabilities and intensity of input use, and crop production for the agent-based model (ABM). To ensure that the models are a representation of the real-world phenomena, several combinations of variables and parameters (replications) are used through machine learning algorithms. Uncertainty or stochasticity in soil estimation using the machine learning randomForest model was essentially managed by running tens to hundreds of replications for random subsets based on knee bend method that indicates stability and validity (Liaw & Wiener, 2018). The simulation results were also checked for model drift based on differences between the predicted and sampled attribute values. The predictions were adjusted to the baseline real farm conditions using data winzorisation and the model drift coefficients (Donkin et al., 2017).

The stochasticity in the predicted input and output flows for the ABM were expressed using random bounded functions with parameters allowed to vary randomly within the estimated confidence intervals than would be the case if the coefficient and margins-at-mean were used (Le, 2005; Villamor, 2012). For each scenario and replication, the stability was achieved by setting random seeds thereby ensuring same initialization and same set of parameters. We run 10 replications which is the minimum number of recommended for complex ABMs with non-linear relationships between the input parameters and simulation output (Thiele et al., 2015). The simulated outputs were averaged for each scenario and the uncertainty was evaluated using confidence interval which is also indicative of the statistical differences between scenarios.

To explore effects of policy actions on adoption and intensification of SFM usage and the resulting impacts on productivity, nutrient balance and profitability, a causal relationship ought to be established. However, in cross-sectional studies of real-life phenomena tend to have many confounders (S. Greenland, Robins, et al., 1999). These are factors that correlate with both the dependent and independent variables making it difficult to discern between association and effect (Wooldridge, 2012). Several design and analytical approaches are used, that either identify the instrumental variables that are associated with the dependent variable but not the output variable or by stratifying based on the confounder and having counterfactual subjects who are not exposed to the treatment included in the analysis (Wooldridge, 2012). Analytical approaches include propensity score matching methods, which address the problem of selection bias that is inherent in these itemized or informational distribution by government and local agents.

However, for MAS models with the aim of establishing causal relationships, feedback loops, synergies and trade-offs among several factors, confounders are inherent and a major threat to validity of inferences (Villamor, 2012). There tend to be a triadic reciprocal causation, where resource endowments, behaviours and environmental states and processes all operate as interacting determinants that influence each other bidirectionally (Bandura, 1986). To avoid type I errors (false positive) of indicating a casual effect, several methodological approaches are used. In MAS, casual diagrams, which are theoretical frameworks or impact pathways that provide a visual or theoretical model for distinguishing causation from association are widely used (S. Greenland, Pearl, et al., 1999). Building on these, process-based models are specified with input and potential outcome model structure (Villamor, 2012) as schematically represented in Figure 3.2, 3.3 and 3.4.

4 RESULTS AND DISCUSSION

4.1 Household social-ecological livelihood types

4.1.1 Exploring the differentiating factors among the farming households

The results in Table 4—1 shows key attributes for the sampled households. Most of the households settled in the lower altitude and own productive land within the fluvial plains. Those situated on the hills, about 24% of the sample, also have majority of their plots on the hill slopes. Topographic position could be assumed as the main physical differentiating factor in terms of productivity. With regard to variables related to resource endowments there is high variability. Distribution density plots showed that data for most continuous variables was censored at zero with positive skewness and a few outliers. The common practice is to transform the data and drop or winsorize the outliers. With greater number of zeros present in the data, transforming has been found to be inefficient and sometimes a source of interpretation and hypothesis testing errors (O'Hara & Kotze, 2010). After revising the raw data, it was observed that the data points falling outside the rage are genuine and were retained for the analysis on the assumption that they are positive deviants and might form archetypical farm types.

The average farmland holding is slightly less than a hectare fragmented into 2 plots. Several sustainable agricultural practices are being employed. About 70% of the households rotate legume with cereals and households use 1 or more soil and water conservation measures. In a year each household plant more than 2 crops. Farmers also apply both organic and inorganic fertilizers. On average, the households use around 100kg of inorganic fertilizer and 180kg of organic manures per year.

Prior to conducting PCA using SPSS v21, a few minimum requirements tests were done. Tests for adequacy of sampling showed that the variables entered had high sampling adequacy with anti-image correlation and commonalities of more than 0.5, the determinant of 0.002 and the combined test using Kaiser-Meyer-Olkin Measure of 0.52 (Table 4—2). This validates the use of the variables in cluster analysis and that their combination gives a robust result with significant Bartlett's Test of Sphericity ($P<0.001$). Variable independence test using pairwise correlation revealed low levels of correlations among most variables $r < 0.3$, warranting the use of Varimax rotation on the assumption that variables are independent.

Table 4—1: Description of household survey data

	Code	Descriptives Mean	SD	Anti Image Correlation	Communalities
Landscape position (1=upper; 0=lower)	H_{LSPN}	.24	.43	0.456	.701
Age of the household head (years)	H_{AGEH}	47.55	16.48	0.533	.644
Education of HH head (years)	H_{EDUH}	5.76	3.58	0.625	.608
… HH members (0=0;1=1-8; 2=9-12; 3>12)	H_{ELHM}	1.45	.78	0.650	.564
Gender of head (1=male; 0=female)	H_{GENH}	.52	.50	0.716	.579
Total labour (manequivalent)	H_{LAB}	2.94	1.49	0.727	.733
Dependency ratio (18-65)/(<18+>65)	H_{DEPR}	1.75	1.57	0.610	.599
Tropical Livestock units	H_{TLUN}	.54	1.47	0.676	.529
Income from crop sales ($/yr)	H_{INCC}	64.08	186.92	0.668	.444
… livestock ($/yr)	H_{INCL}	7.04	21.80	0.594	.688
… other sources ($/yr)	H_{INCO}	149.95	244.98	0.593	.602
… natural resources ($/yr)	H_{INCR}	2.23	19.94	0.489	.281
Communication index	H_{COMM}	.47	.53	0.823	.684
Transport index	H_{TRAN}	.12	.18	0.827	.620
Farm implements index	H_{IMPL}	2.06	1.08	0.765	.681
Women Empowerment Index	H_{WEAI}	.15	.18	0.652	.262
Group membership (1=yes; 0=no)	H_{GMEM}	.36	.48	0.465	.581
Number of plots cultivated	P_{PLOT}	2.10	1.02	0.515	.620
Land cultivated (ha)	P_{HECT}	.94	.74	0.552	.851
… allocated to legumes (ha)	P_{LEGU}	.36	.70	0.482	.882
Number of crops grown	P_{CROP}	2.26	1.12	0.578	.559
Rotation last 5 yrs (1=yes;0=no)	P_{ROTA}	.70	.46	0.648	.676
Soil water conservation	P_{SWCM}	1.03	.77	0.561	.625
Organic inputs (kg)	P_{ORGA}	174.72	329.20	0.724	.667
Fertilizer applied (kg)	P_{FERT}	97.03	89.22	0.647	.703
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.				.627	
Bartlett's Test of Sphericity		Approx. Chi-Square		1423.108	
		df		300	
		Sig.		.000	
Determinant				0.002	

From Table 4—2, 9 PCS explain about 62% of the variation of the original independent variables which is adequate. After varimax rotation, the 9 variables with highest loading on each PC was used to distinguish household types. In Table 4—2 the principal component 1 (PC1) is highly correlated to physical assets including communication facilities (*H_{COMM}*, loading *b*= *0.746*) and transport *(H_{TRAN}*, *b=0.685*). Worth noting is that this factor has also high loading for household head (*b=0.611*). This component accounts for 16% of the total variance and has been named physical assets factor. Since SAPs are information intensive, the variable communication is therefore the best representative and most important physical asset for distinguishing household types.

Table 4—2. Loadings for the first 9 components after Varimax rotation with Kaiser Normalisation

Variable	Principal Component PC1 Physical assets factor	PC2 Legume integration factor	PC3 Fertlizers factor	PC4 Demo graphic factor	PC5 Labour factor	PC6 Organics factor	PC7 Topography factor	PC8 Livestock income factor	PC9 SWC factor
H_{LSPN}	-.022	-.147	.165	.066	.028	.106	.642	.163	.443
H_{AGEH}	-.159	.132	.097	.749	.027	.138	-.058	.022	-.086
H_{EDUH}	.611	.010	-.086	-.400	.033	.060	.062	-.015	.241
H_{ELHM}	.169	.097	.045	-.194	.646	-.083	.102	.036	.225
H_{GENH}	.542	.124	-.052	-.171	.031	.069	.222	-.393	-.167
H_{LAB}	.039	.157	.105	-.042	.813	.145	-.013	-.015	-.108
H_{DEPR}	-.083	.131	-.047	.676	-.250	-.192	-.071	-.013	.111
H_{TLUN}	.135	-.026	.022	.236	.119	.467	.006	.460	-.103
H_{INCC}	.303	-.016	.308	-.048	.082	.010	-.447	-.154	.157
H_{INCL}	.175	.042	.102	-.106	-.046	-.049	.131	.782	-.042
H_{INCO}	.047	.055	-.024	-.037	.320	.099	-.406	.256	.502
H_{INCR}	.036	.169	.016	-.093	.073	-.058	.483	-.001	-.006
H_{COMM}	.746	.126	.192	-.031	.034	.087	-.014	.193	.162
H_{TRAN}	.685	.149	.131	.043	.133	.104	-.100	.265	-.026
H_{IMPL}	.520	.187	.188	.045	.432	.267	-.227	-.023	-.170
H_{WEIA}	-.013	.151	.280	-.139	-.285	.047	-.051	.221	.079
H_{GMEM}	-.121	.259	.137	-.524	-.016	-.104	-.424	.111	.037
P_{PLOT}	.046	.118	.744	.056	.082	.167	-.032	.088	.053
P_{HECT}	.204	.854	.164	.117	.195	.010	.027	-.032	.002
P_{LEGU}	.138	.919	-.016	.052	.079	.000	.079	.037	.047
P_{CROP}	.017	.386	.534	.014	-.128	.266	.006	.185	.050
P_{ROTA}	-.038	.023	.379	-.243	-.093	.639	.190	-.111	.073
P_{SWCM}	.089	.052	.046	-.025	-.061	-.020	.079	-.119	.766
P_{ORGA}	.282	.019	-.064	.075	.124	.736	-.140	.011	.028
P_{FERT}	.218	-.128	.747	.003	.175	-.213	.039	-.028	-.053
Initial Eigenvalues									
Eigen value	4.06	2.04	1.73	1.54	1.42	1.25	1.19	1.09	1.06
%variance	16.23	8.14	6.94	6.14	5.68	5.01	4.77	4.38	4.24
Cumulative	16.23	24.37	31.31	37.45	43.13	48.14	52.91	57.29	61.53
Varimax rotation with Kaiser normalisation									
Eigen value	2.39	2.06	1.93	1.71	1.69	1.53	1.42	1.34	1.32
%variance	9.54	8.24	7.72	6.85	6.75	6.14	5.66	5.35	5.28
Cumulative	9.54	17.78	25.49	32.35	39.10	45.24	50.90	56.24	61.53

The PC2 accounts for 8% of the variance and is highly correlated with total land cultivated by the household in 2017 (H_{HACT}, b=0.854) and the average land devoted to legume cropping for the past 5 years (H_{LEGU}, b=0.919). Since the farming systems are dominated by cereals, the high loading of land allocated to legumes is important indicator for adoption of legumes as one of the SAPs. The PC3 which accounts for 7% of the total variance explained largely by plot variables of total inorganic fertilizer applied (P_{FERT}, b=0.747) and total number of fragmented plots (P_{PLOT}, b=0.744). Use of inorganic fertilizers is still below the recommended with wide variations among households. The factor is therefore names fertilizer and used to distinguish the household types.

The PC4 which has high loading of the two demographic variables accounts for 6% of the total variance. It is highly correlated with age of the household head (H_{AGEH}, $b=0.749$) and the dependency ratio (H_{DEPR}, $b=0.676$). In a predominantly youthful society where youth have been found to either not having access to productive resources or shun farming for other activities, it is interesting to see if age has the discriminating influence among the farm types. PC5 is also demographic as it is highly related to household available labour (H_{LABO}, $b=0.813$) and highest education level attained by any member of the household (H_{EDHM}, $b=0.646$).

The remaining PCs (6 to 9) are highly correlated by individual variables. PC 6 is correlated with organic manure inputs (P_{ORGA}, $b=0.736$), PC7 by whether the household is located on the flatter plains or in the hill slopes (H_{LSPN}, $b=0.642$) PC8 by income from livestock sales (H_{INCL}, $b=0.782$) whilst the variance for PC9 of 4% is mainly accounted for by number of soil and water conservation measures being implemented by the household (P_{SWCM}, $b=0.766$). These PCs have therefore been named after these variables which were then used in differentiating the farm types.

4.1.2 Identification of household types using key variables

The k-mean was performed in Stata Version 15 (StataCorp, 2017) using Euclidean distance dissimilarity measure among the standardized scores for the nine principal components from PCA results. To ensure that that the clusters are reproducible, random seed of *123* was used and the initial cluster centres were determined from *K* unique random observations among the K^{th} *(where k=1, …, 238* standardized scores for the nine principal components) (StataCorp, 2017). The number of clusters were determined using the optimisation of within cluster sums of squares, popularly known as the elbow method.

Table 4—3: Comparison among the 3 farm types using descriptive statistics of the original variables

	C	N	mean	sd	se	cv	min	max	95% CI		P	Remark
H_{AGEH}	1	64	50.05	18.92	2.36	0.38	22	84	45.32	54.77	b	older
	2	90	52.63	13.32	1.40	0.25	23	85	49.84	55.42	b	older
	3	84	40.20	15.02	1.64	0.37	20	84	36.94	43.46	a	younger
H_{LAB}	1	64	2.20	1.01	0.13	0.46	0.56	5.28	1.95	2.46	x	fewer
	2	90	3.79	1.62	0.17	0.43	0.56	11.21	3.45	4.13	y	more
	3	84	2.58	1.20	0.13	0.46	0.17	8.19	2.32	2.84	x	fewer
H_{INCL}	1	64	13.20	35.33	4.42	2.68	0	177.24	4.38	22.03	a	
	2	90	5.53	14.71	1.55	2.66	0	74.48	2.45	8.61	a	
	3	84	3.95	11.66	1.27	2.95	0	68.97	1.42	6.48	a	
H_{COMM}	1	64	0.48	0.53	0.07	1.10	0	2.85	0.35	0.62	x	
	2	90	0.39	0.50	0.05	1.29	0	1.98	0.28	0.49	x	
	3	84	0.56	0.56	0.06	1.01	0	2.37	0.43	0.68	x	
P_{WILE}	1	64	0.48	1.16	0.15	2.43	0	6.90	0.19	0.77	a	
	2	90	0.39	0.49	0.05	1.26	0	2.72	0.29	0.49	a	
	3	84	0.24	0.26	0.03	1.08	0	1.20	0.19	0.30	a	
P_{SWCM}	1	64	0.91	0.56	0.07	0.61	0	2	0.77	1.04	x	fewer
	2	90	0.78	0.65	0.07	0.84	0	3	0.64	0.91	x	fewer
	3	84	1.39	0.88	0.10	0.63	0	5	1.20	1.58	y	more
P_{ORGA}	1	64	72.42	122.79	15.35	1.70	0	600	41.75	103.09	a	less
	2	90	135.44	219.78	23.17	1.62	0	1500	89.41	181.48	a	less
	3	84	294.73	471.06	51.40	1.60	0	2500	192.51	396.96	b	more
P_{FERT}	1	64	60.86	67.31	8.41	1.11	0	400	44.05	77.67	x	less
	2	90	139.17	98.79	10.41	0.71	0	425	118.48	159.86	y	more
	3	84	79.44	74.70	8.15	0.94	0	420	63.23	95.65	x	less

NB: C=farmtype; N=number of households; sd=standard deviation, se=standard error of the mean; cv=coefficient of variation; min,max=minimum and maximum values, P=within variable means with different letters are significantly different.

The emerging farm types significantly differ in terms of the observable household and plot variables (Table 4—3 and Figure 4.1). Of the eight livelihood indicator variables, only 5 significantly distinguish the farm types. These are age of household head, labour, soil and water conservation measures and amounts of organic and inorganic fertilizers. Using these five, and absolute but not statistical differences of the other three 3 variables, the following comparisons of the farm types can be established.

The *farmtype1* is comprised of middle-aged household heads with slim workforce, moderate communication facilities, income from livestock sales and land allocated to legumes and use fewer soil and water conservation structures and apply lesser inorganic fertilizer and lesser organic inputs. It is therefore called a low input farm group.

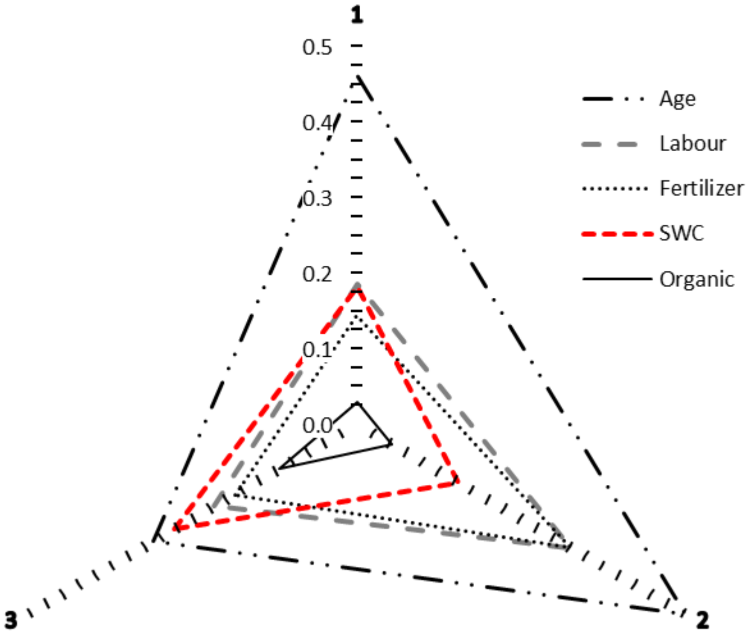


Figure 4.1: Farmtype differentiation by age, labour, nutrient inputs and soil and water conservation measures

The households in *farmtype2* are headed by relatively older people in their late 40s, with more available labour and applies relatively more inorganic fertilizers. They tend to have fewer communication facilities and use few soil and water conservation measures. They receive moderate income from livestock allocate moderate amount of land to legumes and apply moderate amounts of organic manures. This is a group of households with medium input farms.

The third farmtype is comprised mainly of younger households (up to early 40s) with moderate labour, communication facilities and levels of inorganic fertilizer application. This group get lower incomes from sell of livestock and allocate less land to legumes. However, on average, they use more soil and water conservation measures and larger amounts of organic inputs.

4.2 Stocks of soil NPK, SOC and their thresholds

4.2.1 Status of soil organic carbon and major nutrient concentrations

At farm level, it is generally anticipated that nutrients ought to be managed in a holistic way. However, the crop nutrition principle based on Liebig's law of the minimum purports that growth is limited by the most deficient nutrient or the most limiting soil condition. Based on this, the tendency has been to focus on the most deficient (nitrogen); unwittingly leading to unregulated extraction and degradation of SOC, P, K and other nutrients in the long run (Mutegi et al., 2015). The results from the sentinel site for the major three nutrients and SOC shows that the soils are deficient total nitrogen (TN) but have low to adequate SOC, P and K; with pronounced spatial variations (Figure 4.3 and Figure 4.4).

The internal validation yielded high prediction accuracy of 0.88 to 0.94 (Table 3—7). Validation using an independent sample (3-fold), shows that the randomForest model predicted SOC, TN, and P with moderate accuracy but had low prediction accuracy for K. It is evident that the distribution and concentration of soil nutrients and conditions are driven different geographical, biological, geologic and topographic sequences (see supplementary Figure S1.2:a-f).

For the greater part of the mapped area, SOC concentrations are within the moderate range for maize productivity. Studies in Zimbabwe and West Africa showed that soils with SOC > 6.5 and 8.0 g kg^{-1}, respectively, had steady response to nutrient inputs without accompanying SOM amendments (Musinguzi et al., 2013; Pieri, 1995). However, in terms of structural stability, the SOC levels (i.e. organic matter) is generally insufficient to critically low (range 0.45-2.0), rendering the soils more susceptible to degradation risks (Musinguzi et al., 2013). With sub-optimal organic matter input, these soils could be on a downward degradation spiral (Rattan Lal, 2015). Optimal levels of SOC, as shown in Figure 4.3, are concentrated along streams and valley floors. This could be the pull factor for cultivation of stream banks and valley floors, a phenomenon that exacerbates river bank erosion, resulting in river bed siltation and drying of the streams (Chimtengo et al., 2014; Sandram, 2018).

Distribution of nitrogen follows that of SOC, more concentrated along streams and valley floors. The maximum predicted N stock of 0.14 is lower than the lower limit, implying that soils of the area are N deficient. Nitrogen therefore continues to be the most limiting of the major nutrients. There is a wide gap to reach sufficiency levels and nitrogen demand for crops will rely on significant yearly inputs.

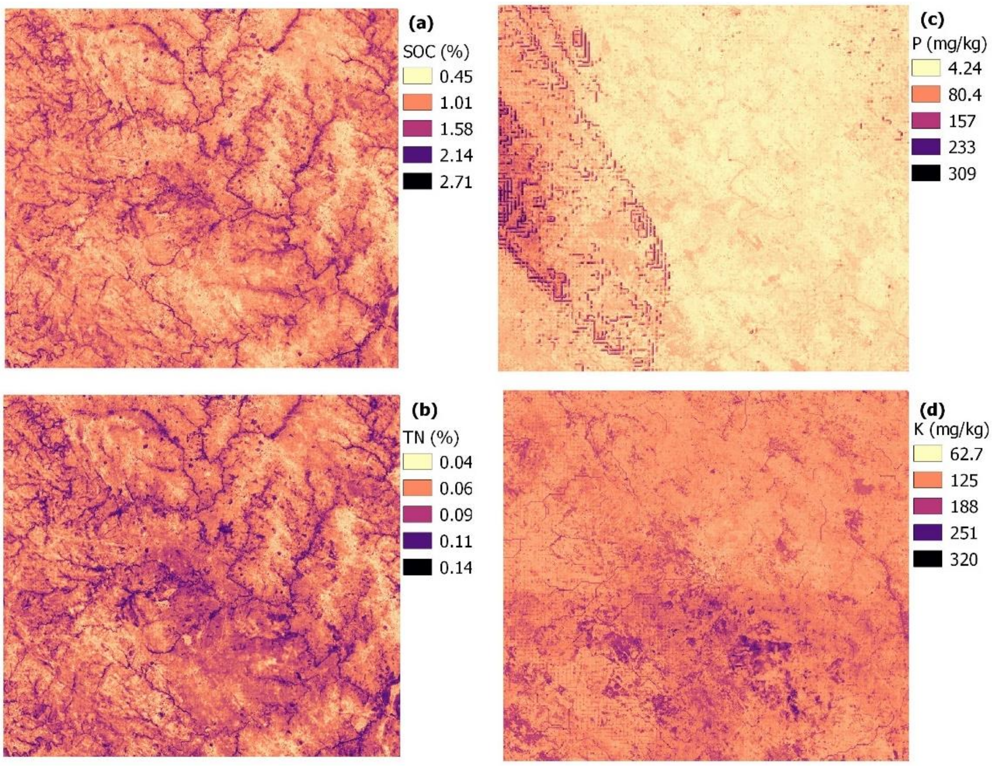


Figure 4.2. Spatial distribution of (a) Soil organic carbon (SOC), (b) Total nitrogen (TN), (c) Phosphorus (P) and (d) potassium (K) in the Nsipe sentinel site.

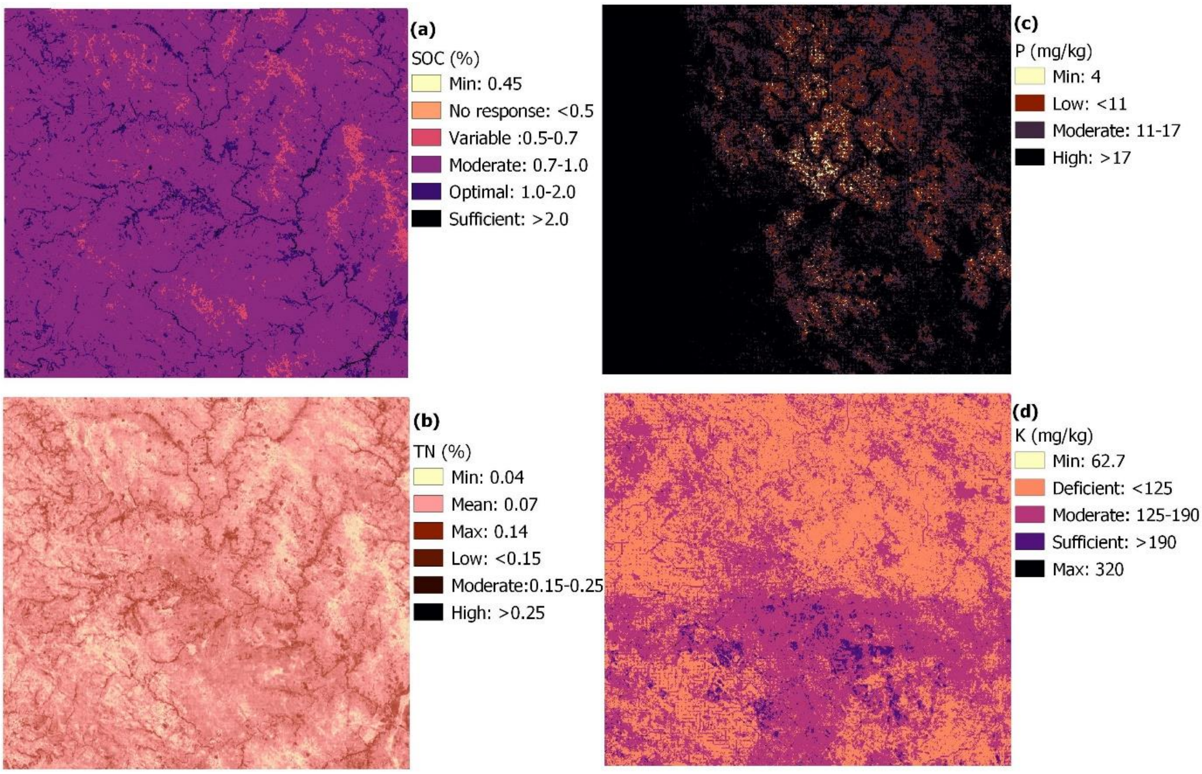


Figure 4.3 Soil health and plant growth limits for SOC, TN, P, K.

Distribution of phosphorus (P) is highly skewed with moderate to higher concentrations clustered on the western hills (Figure 4.2c). The eastern flat plains that have been under cultivation for decades have low to medium levels of P (Figure 4.3c). These spatial differences entail that indigenous P uptake by crops would be different affecting input decisions and yields across the landscape.

The medium low to moderate levels of potassium (K) observed indicate the need for potassium fertilizer. Significant maize response to application of K fertilizer was observed when the soil exchangeable K reached below 190 mg kg^{-1} in sandy soils of South Africa (van Biljon et al., 2008). For more than three decades, Malawian soils were considered to have sufficient amounts of K. Previously, the government through the 1970 (edited in 1996) Fertilizers, Farm Feeds and Remedies Act (Cap. 67:04) endorsed a few fertilizer mixtures of which 23N:21P:0K +4s and Urea (40N) have been widely used for maize production due to being nitrogen rich but have zero K (IFDC, 2013). The omission of K had no significant effect on crop yields but with continuous extractive cropping, there has been an increased risk of K depletion (IFDC, 2013). Elsewhere, studies have shown that with zero or insufficient K input, continued cropping of rice in Asia (Dobermann et al., 2003) and of seven major crops in Australia (Brennan & Bell, 2013) in soils that were initially considered not deficient, led to K depletion. The most plausible approach is to actively address the K depletion before critical levels are reached (F. V Schindler et al., 2005) and based on such revelations since the 2018/2019 growing season, the Malawian government has introduced K containing fertilizers.

4.2.2 Soil structural stability and stochiometric ratios

Apart from being an important source of nutrients, organic matter has been found to play a significant role in structural stability in most African soils (Ayuke et al., 2019). The low levels of SOC prevalent in the sentinel site of this study renders these soils structurally unstable (Figure 4.4a and Figure 4.5a). Greater portion of the land under study have soils that are structurally degraded and, can easily be washed down. Already, substantial erosion and sedimentation has been observed in the catchment, leading to nutrient losses through runoff and siltation of streams (Chimtengo et al., 2014; Sandram, 2018). Therefore, there is an urgent need to improve structural stability by, among many options, planting biomass cover crops, recycling the organic matter and substantial application of organic nutrient inputs. The gap is wide and studies have found that with low input usage, as practiced in the study area, no significantly improvements in soil stability can be achieved (Ayuke et al., 2019). For the majority of soils with StI <5, it is imperative to improve organic inputs/ recycling so as to increase StI to 7-9 in the medium term with the ultimate goal to increase SOC levels for the entire site and achieve StI of >9.

The dependence ratios among C:N:P shows that although N is considered the most limiting nutrient, it is in fact C that is more limiting. The normal C:N range is 20-25, but the results show that C:N is in the lower range of 6-15, implying that relative to C, N is in excess (Figure 4.5b). This could be due to continued government efforts to promote use of N fertilizers while not giving due attention to management of soil organic carbon. Recent studies show that the maize response to fertilization has levelled off and is becoming increasingly dependent on management of soil organic carbon (Kopper et al., 2020). This calls for N and C co-management, as efforts to improve N may in the long run be curtailed by erosion of less structurally stable soils and leaching.

Microbial solubilizing and immobilisation of P is increasingly getting attention and shows that C have significant influence on bioavailability and retention of P (Alori et al., 2017; Tamene et al., 2019; Zhang et al., 2018). C to P ratio (Figure 4.4c and supplementary Figure 4.5c) is indicative that most of the cultivated soils on the flood plains have P limitations, where as those on the western end have C limitations. For areas with P-limitation, P mobilising microbes such as biofertilizers could be promoted, which work by mobilising available P and in the process making it available for plant root uptake (Alori et al., 2017; Njira, 2013). On the other hand, for areas with net P-mineralisation, management options should include increase in organic matter input such as crop residues, organic manure, domestic wastes and cover crops; to accelerate microbial P immobilisation and ultimately reduce losses though leaching. The N:P ratio has been widely studied in plant biomass with limited focus on establishing the critical soil levels. Studies though point to its influence on soil microbial activity (Tamene et al., 2019).

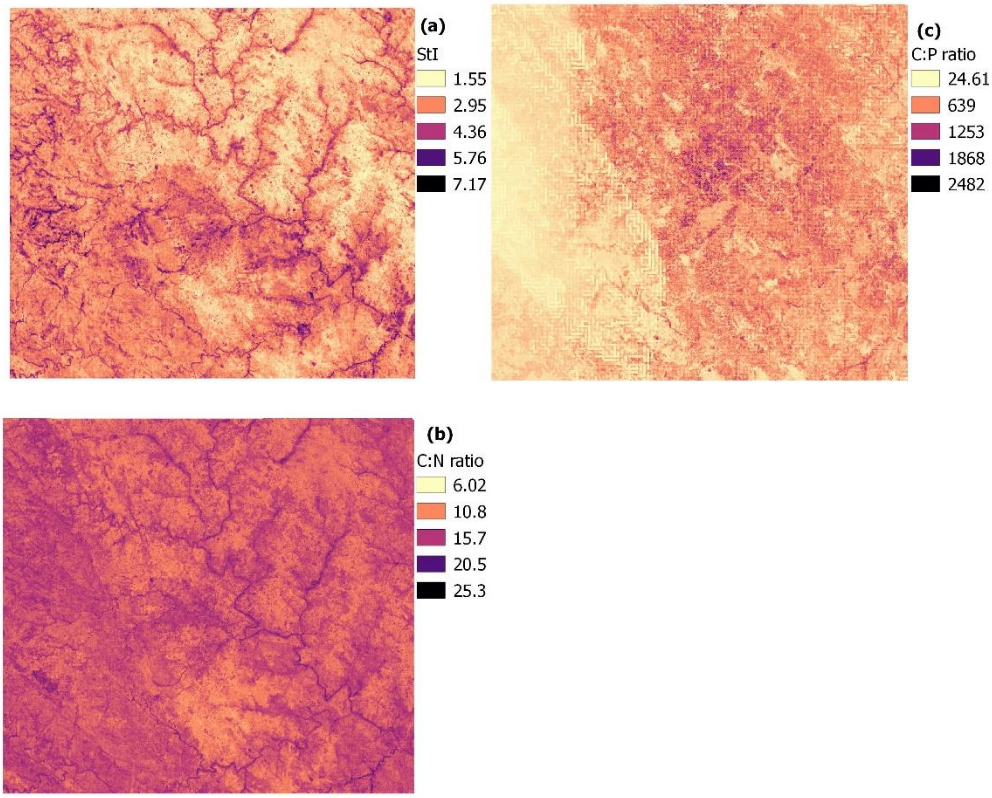


Figure 4.4 Spatial associations between soil organic carbon and (a) soil texture indicating structural stability (StI); and stoichiometric ratios with (b) total nitrogen (C:N), and (c) phosphorus (C:P) indicating nutrient dependence

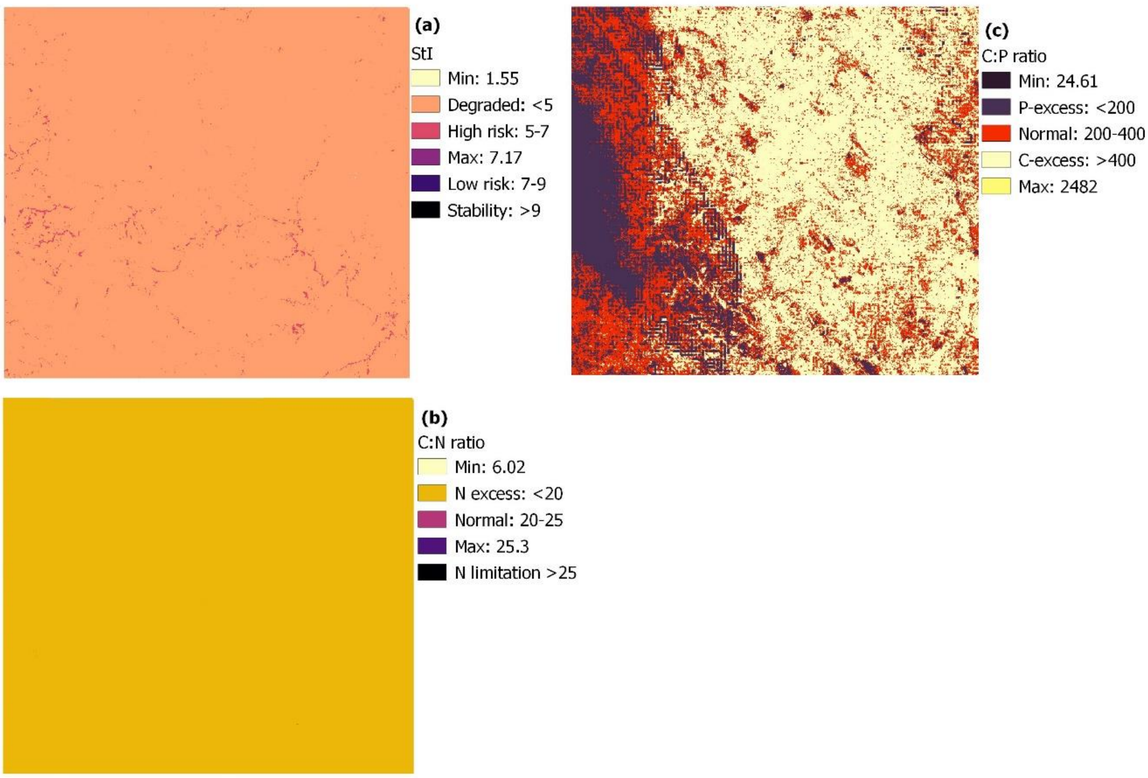


Figure 4.5 Extent of structural degradation and stoichiometric co-limitations (C:N & C:P).

The differences in stoichiometric associations among soil nutrients have implications on decisions for fertilisation and its effects on bio-availability of plant nutrients. There is emerging evidence on the effects of N-containing fertilizers on C:N:P ratio that in turn significantly influences the microbial community composition and enzyme activities (Shen et al., 2019). Since the pioneering studies of nutrient balance on Kenyan farms in 1990s, it is evident that, with low level of crop outputs, N is somehow adequately replenished. However, for these managed and tilled landscapes, large amounts of carbon are estimated to be lost through organic matter degradation than it is replenished which results in a wider negative balance. We also observe that the concentration of nutrients follows topographic gradient. Although both SOC and TN are concentrated in depressions, relative to TN, SOC is more concentrated on lower elevation and footslopes. Higher CN along streams could be due to combination of high vegetation cover along streams and sediment deposition from upslopes. Although it has been established that there are areas with normal C:P ratios, studies have shown that critical level of individual soil attributes dictate the responsive range. For instance, even when the N:P ratio was optimal, the critical N content determined the optimal range for the corresponding P for litter decomposition (Güsewell & Verhoeven, 2006).

Managing soil nutrients solely based on the Liebig's law of the minimum is considered a palpable theory for a steady system and where one or a limited identifiable number of limiting nutrients exist. Research trials in controlled environments use the tenets of the theory to supply other conditions in optimal ranges and omit or vary only the target variable. In real farm environments with variable conditions and multiple limiting nutrients, the actual limiting nutrient can abruptly change over spatial and temporal scales (Sperfeld et al., 2012). In such cases, a multiple limitation theory has been established considering that the nutrients interact, affecting availability and uptake of other nutrients (Gleeson & Tilman, 1992).

4.3 Household SFM choice

Inorganic fertilizer usage in the study region is widespread, yet amounts vary greatly. Among the surveyed farming households, 90% applied on average (± standard deviation) 106 (± 42) kg inorganic fertilizer during the 2016/17 growing season (Table 4—4). Almost one third of the households applied <50kg, 28% applied 50-100 kg, 20% 100-200kg, 6% 200-300kg whilst only 3% applied more than 300 kg of fertilizer. More than half of the households (55%) applied organic manures, which included mostly farmyard and household waste. Out of the 55, 16 applied <100 kg, 32 applied 100-500 kg whilst only seven applied > 0.5 tonnes. The proportion of farmers that planted legumes was 72 percent. The average amount of land under legumes for only those that planted was 0.45 (± 0.64) ha.

Descriptive statistics results show that adopters have an edge over non-adopters in most of the factors considered (Table 4—4). As illustrated by the respective confidence intervals, adopters of inorganic fertilizers compared to non-adopters have significantly more plots ([2.0-2.2] vs [1.0-1.9]), larger land sizes ([0.9-1.1 ha] vs [0.5-0.8 ha]) and the majority practice crop rotation ([70-80%] vs [30-70%]). Adopters of organic manures and legumes also grow more crops and have more agricultural implements. In addition, adopters of legumes have more communication facilities and apply more organic manure ([158-266 kg] vs [26-124 kg]).

Table 4—5 and Figure 4.6 presents the results for the double hurdle model. For each SFM practice, the first table column and first figure bars show the estimated coefficient and the elasticity of probability that the household used a SFM practice. Similarly, the third column and the second bar show the estimated coefficient and elasticity for those above zero. The third bar shows the summed up unconditional elasticity.

In general, the results reveal household and plot that significantly differentiate the discreet and intensification choices among households in the five study villages (Table 4—5). In terms of level of responsiveness, inorganic fertilizer and organic manure usage are quite inelastic (elasticity < 1) to respective unit percentage changes in the household and plot attributes. The larger response is expected from differences in usage of legumes. Varying household land holding size (H_{HECT}) has a significantly and larger influence on the area under legumes. Increasing the number of plots significantly enhances the probability of cropping legumes.

It has also been found that households having a relatively larger household labour increases household's capability to apply large quantities of fertilizer. However, increased number of dependant individuals compared to working members in a household have a decreasing effect on fertilizer intensification. A percentage point increases in available labour is associated with increase in in quantity applied by current users by 0.2 percentage points but reduces the probability by 0.13 percent points. Increasing dependency ratios also decreases the probability to apply organic manure. It has also been revealed that for households with higher levels of women involvement in decision-making, the level of organic manure application tend to decrease while the area under legume cropping increased.

Increasing farm sizes by 1.0 percentage point increases the probability to grow legumes by 0.5 percentage points and the use of inorganic fertilizers by 0.1 but has a negative effect on the extent of inorganic fertilizer usage by current users of 0.5 percentage points. A percentage point increase in number of plots (fragmentation) is associated with 0.4% increase in the probability to grow legumes by 0.4%. On the other hand, further plot fragmentation is associated with reduction in area under legume by 0.56% (Figure 4.6c).

The amount of fertilizer applied is enhanced by increasing income from sale of cash crops, and the income from livestock and livestock products enhances the probability to apply inorganic fertilizer, whereas the income from sale of natural resources reduces the probability to apply inorganic fertilizer. We found that a 1% increase in level of education would lead to a 0.22% increase in quantity of manure applied, *ceteris paribus*.

Table 4—4 Means and proportions (95% CI) for the household attributes, resource endowments and farming practices among adopters and non-adopters

	Inorganic fertilizer (kg)						Organic manures (kg)						Legume cropping (land%)					
	non-adopt (n=24)			Adopter (n=214)			non-adopt (n=104)			Adopter (n=134)			non-adopt (n=65)			Adopter (n=173)		
	mean	[95% CI]		mean	[95% CI]		mean	[95% CI]		mean	[95% CI]		mean	[95% CI]		mean	[95% CI]	
Demographic																		
H_{AGEH}							46	[43	50]	48	[46	51]						
H_{EDUH}	4.8	[3.4	6.2]	5.9	[5.4	6.4]	5.4	[4.7	6.1]	6.1	[5.5	6.7]	4.9	[4.1	5.8]	6.1	[5.5	6.6]
H_{GENH}	0.4	[0.2	0.6]	0.5	[0.5	0.6]	0.5	[0.4	0.6]	0.6	[0.5	0.7]	0.5	[0.4	0.6]	0.5	[0.5	0.6]
H_{DEPR}							1.9	[1.6	2.3]	1.6	[1.3	1.9]	2.7	[2.3	3.0]	3.0	[2.8	3.3]
Resource endowment and income																		
H_{TLU}							0.3	[0.1	0.5]	0.7	[0.4	1.0]						
H_{INCC}	30	[0.0	62]	68	[42	94]	57	[25	90]	69	[35	104]	61	[12	110]	65	[38	92]
H_{INCL}	1.4	[0.0	4.3]	7.7	[4.6	10.7]	4.0	[1.3	6.8]	9.4	[4.9	13.8]						
H_{INCR}	0.0	[0.0	0.0]	2.5	[0.0	5.3]	5.0	[0.0	10.8]	0.1	[0.0	0.2]						
H_{COMM}													0.3	[0.2	0.4]	0.5	[0.5	0.6]
H_{TRAN}	0.1	[0.0	0.2]	0.1	[0.1	0.1]	0.1	[0.0	0.1]	0.2	[0.1	0.2]	0.1	[0.1	0.1]	0.1	[0.1	0.2]
H_{IMPL}	2.0	[1.6	2.4]	2.1	[1.9	2.2]	1.8	[1.6	2.0]	2.3	[2.1	2.4]	1.7	[1.5	1.9]	2.2	[2.0	2.4]
H_{WEIA}							0.2	[0.1	0.2]	0.1	[0.1	0.2]	0.1	[0.1	0.1]	0.2	[0.1	0.2]
H_{GMEM}													0.3	[0.2	0.5]	0.4	[0.3	0.4]
Farm configuration and practices																		
H_{PLOT}	1.6	[1.3	1.9]	2.2	[2.0	2.3]	1.9	[1.7	2.1]	2.2	[2.1	2.4]	1.6	[1.4	1.7]	2.3	[2.1	2.5]
H_{HECT}	0.6	[0.5	0.8]	1.0	[0.9	1.1]	0.8	[0.7	0.9]	1.0	[0.9	1.2]	0.7	[0.6	0.8]	1.0	[0.9	1.2]
H_{WILE}	0.2	[0.1	0.4]	0.4	[0.3	0.5]										0.5	[0.4	0.6]
H_{CROP}	1.9	[1.6	2.3]	2.3	[2.1	2.4]	1.9	[1.7	2.1]	2.5	[2.3	2.7]	1.4	[1.2	1.5]	2.6	[2.4	2.8]
H_{CROT}	0.5	[0.3	0.7]	0.7	[0.7	0.8]	0.6	[0.5	0.7]	0.8	[0.7	0.9]	0.3	[0.2	0.5]	0.8	[0.8	0.9]
H_{ORGA}	196	[0.0	401]	172	[131	213]				310	[244	376]	75	[26	124]	212	[158	266]
H_{FERT}				107	[96	119]	100	[81	119]	95	[81	109]	82	[60	105]	103	[89	116]

H_{AGEH} = age of household head (years); H_{EDUH} = formal education of the HH (years); H_{GENH} = Gender of the household head (1=male, 0=female); H_{DEPR} = Dependency ratio (age 18-65/(18<age>65)); H_{TLU} = Tropical livestock units; H_{INCC} = income from cash crops (US$/year); H_{INCL} = income from livestock (US$/year); H_{INCR} = income from natural resources (US$/year); H_{COMM} = communication index; H_{TRAN} = transport index; H_{IMPL} = farm implements index; H_{WEIA} = women empowerment in agriculture index; H_{GMEM} = group membership (1=yes, 0=no); H_{PLOT} = number of plot fragments; H_{HACT} = farm size (hectares); H_{WILE} = hectares under legume (%); H_{CROP} = number of crops; H_{CROT} = crop rotation (1=yes, 0=no); H_{ORGA} = organic manure applied (kg); H_{FERT} = inorganic fertilizer applied (kg).

Table 4—5 Double-hurdle estimates of the probability to apply inputs or plant legumes and intensification

	Inorganic fertilizer (kg)				Organic manure (kg)				Legume land share (%)			
	Apply		Intensity		Apply		Intensity		Plant		Intensity	
	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.
H_{AGEH}	0.01	0.01	1.03	0.99	0.01	0.01	1.52	1.13				
H_{EDUH}	0.07 *	0.04	7.42 *	4.11	0.02	0.03	12.17 **	5.40	0.083 **	0.04	-0.013	0.01
H_{GENH}	0.53 *	0.29	-31.22	32.93	0.16	0.20	-21.56	33.71	0.067	0.25	-0.006	0.08
H_{LABA}	-0.01	0.09	19.30 **	8.10					-0.087	0.11	0.011	0.02
H_{DEPR}	0.01	0.08	-20.10 **	9.81	-0.13 **	0.06	-17.39	11.57				
H_{TLUN}					0.01	0.07	10.90 **	5.40				
H_{INCC}	5E-4	9E-4	0.12 **	0.05	-4E-4	5E-4	0.03	0.06	-0.002 **	7E-4	14E-5	1E-4
H_{INCL}	0.02 **	0.01	-0.27	0.41	0.01	0.01	-0.06	0.53				
H_{INCO}	8E-4	6E-4	-0.14 **	0.06								
H_{INCR}	1.75	1.08	-0.16	0.29	-0.13 *	0.07	18.46	27.65				
H_{COMM}									-0.096	0.38	0.17 *	0.10
H_{TRAN}	-2.55 ***	0.99	88.12	91.51	0.74	0.69	21.46	78.09	-1.902 **	0.91	0.048	0.21
H_{IMPL}					0.13	0.11	35.56 **	16.75	0.435 ***	0.17	-0.03	0.04
H_{WEIA}	0.16	0.69	43.88	71.32	-0.64	0.50	-179.57 **	89.48	2.305 ***	0.90	0.154	0.20
H_{GMEM}									-0.203	0.25	0.042	0.09
P_{PLOT}	0.50 ***	0.19	62.64 ***	14.37	0.11	0.11	18.72	15.89	0.316 **	0.16	-0.29 ***	0.07
P_{HACT}	0.82 ***	0.32	-12.18	16.74	0.05	0.14	-39.45 **	19.16	0.648 **	0.32	0.977 ***	0.03
P_{CROP}					0.28 ***	0.09	-10.93	14.09	1.052 ***	0.18	0.183 ***	0.04
P_{ROTA}					0.42 **	0.21	100.21 **	44.28	1.349 ***	0.24	-0.515 ***	0.09
P_{FERT}					25E-4 **	12E-4	-0.11	0.20	2E-5	16E-4	17E-5 *	4E-4
P_{ORGA}	-5E-4	3E-4	-0.05	0.05					-0.001	7E-4	-24E-5	1E-4
_cons	-1.27 *	0.76	-196.1 ***	75.43	-1.35	0.49	-30.96	92.15	-4.187 ***	0.62	-0.256 *	0.13
	Log pseudolikelihood			-952.8				-1246.7				-16.18494
	Wald chi2			44.7				30.9				109.8
	Prob > chi2 =			0.0002				0.006				0.000

Significant at: p<0.1*, p<0.05**, p<0.01***

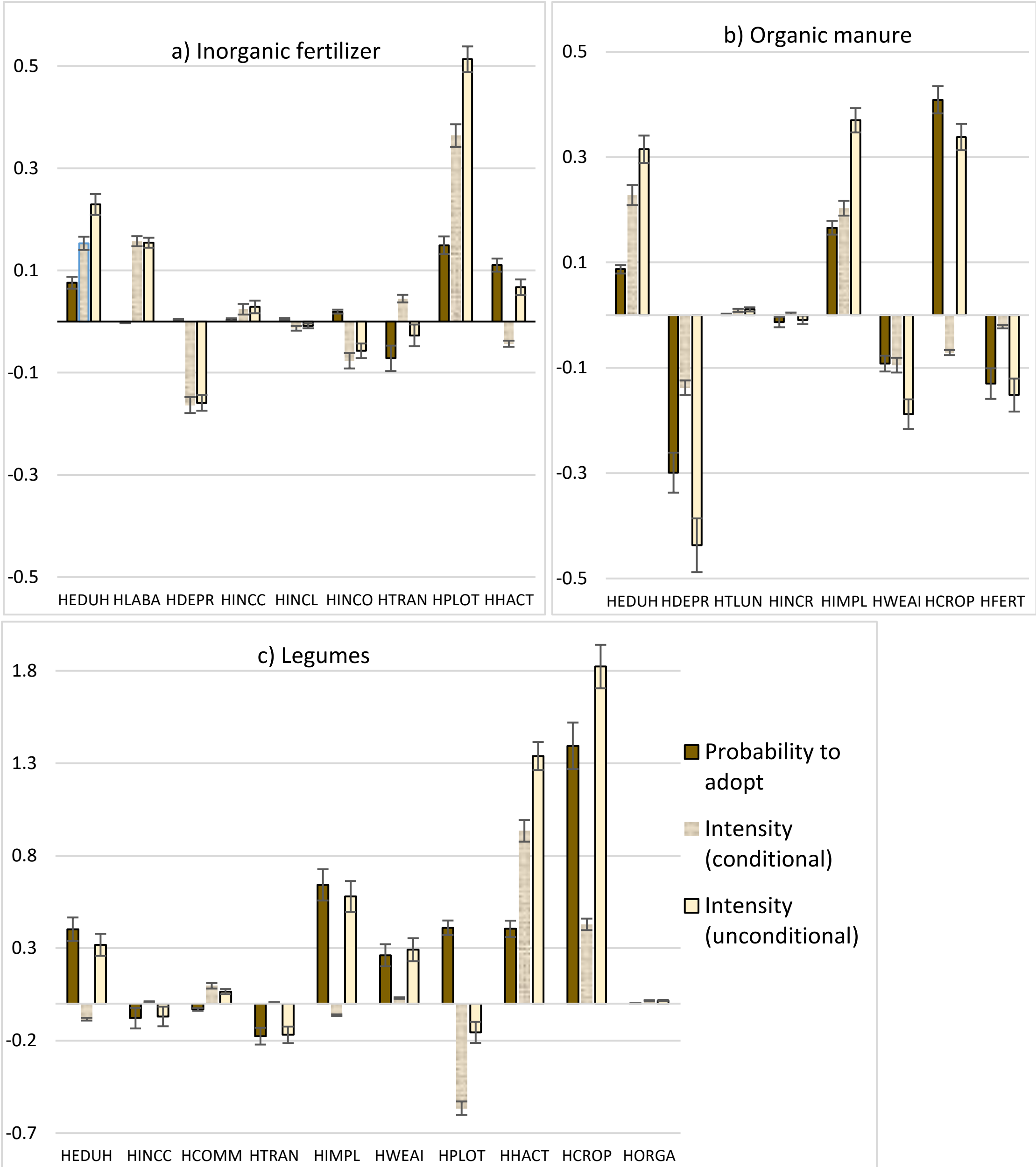


Figure 4.6 Elasticities for the probabilities to adopt and intensify (a) inorganic fertilizers, (b) organic manure, and (c) allocate land to legumes.

4.3.1 Bounding demographic and social conditions for household SFM usage

With the very limited transport facilities and predominance of using heads as primary method for transporting heavy products (Amede et al., 2014), manure application is bound to continue constraining labour allocation decisions. Increasing number of dependants means that the household need to spend more on food and other household necessities at the expense of the re-investing in fertilizer purchases. To meet the needs of household dependants, workers prioritise allocating labour to immediate needs and those with limited labour would hardly allocate to manure management.

Our finding on effects of gender on manure usage are in agreement with the notion that women find the preparation and transportation to be labour demanding

(Mustafa-Msukwa et al., 2011). The main legume grown in Nsipe is groundnut (*Arachis hypogaea L.*) for which some studies in the region found that households headed by women allocated less land than their male counterparts (Waldman et al., 2016). However, the effects of gender as a social factor are contextual and Waldman's research did not account for the influence of women in male-headed households. The marriage system in Nsipe is matrilineal and consequently, the role of women in household decision making is profound. Women-headed households tend to have more information about legumes than male-headed ones (Snapp et al., 2002). These results are also consistent with the findings by Pircher (Pircher et al., 2013) who found that even in participatory research, women farmers have higher preferences to integrate legumes in their maize farms compared to men farmers.

Since the technologies are generally promoted as an integrated basket, gender would have disproportionate effects on usage of organic manures and legumes. It has been purported that social aspects such as gender may not be important driver for a particular soil fertility technology, but could have indirect effects through its influence on complimentary or alternative technologies (Doss & Morris, 2000). A gender segregated study in Ghana showed that the propensity to apply fertilizer by women-headed households was positively associated with farming experiences, whereas for male-headed counterparts, income from other sources reduced investment in fertilizer (Mensah et al., 2018). Gender is invisible and its effects underpin other underlying factors. For instance, women's adoption of practices that require physical assets such as organic manure, is premised on them having access to resources first.

One major challenge in scaling technologies that require understanding of chemical, biological and physical processes happening below one's feet in rural farming systems is the prevailing high illiteracy levels. In Nsipe, the average education level attained by household heads is 6th grade which has a strong bearing on their ability to processing agricultural information. The education level has consistent positive effects on both probability and extent of inorganic fertilizer usage. Considering that average education level attained by households is 6th grade (incomplete primary), shifting the level of education by one grade to 7th grade represents 17% which would be associated with a 3.91% increase in overall fertilizer application. It is further envisaged that improvements in formal education from the current average of grade 6 by 33% to grade 8 (complete primary) or by 100% to grade 12 (complete elementary) would be associated, respectively, with increases in manure application from the current 244 kg to 261 kg and 325 kg. These results highlighting that an emphasis on education could lead to intensified inorganic fertilizer and organic manure application.

4.3.2 Resource endowments shaping SFM uptake and intensification

Considering that land is major resource, this study has found that increasing land sizes has larger positive elasticity on the decision start applying fertilizer than the negative elasticity the amounts applied by current fertilizer users. The overall effect on inorganic fertilizer is therefore positive. The positive effects of farm size were expected since the general tendency is that large farm sizes are associated with increased availability of financial capital, which can make investment in ISFM more feasible (Akinola et al., 2010). In addition, farmers with small plots might not benefit from 'economies of scale' when using more inputs as the returns tend to be too small especially in rainfed farming system (Harris & Orr, 2014). Unlike the established positive farm size - input use - productivity relationship (Dorward, 1999; Muyanga & Jayne, 2019), in Malawian smallholder farming systems, inverse farm size – productivity relationship has been observed since 1990s (Matchaya, 2007). This has been attributed to governments policy that has enabled farmers with small plots to have access to inorganic fertilizers and explains the inverse farm size - inorganic fertilizer relationship found in this study. This result is further confirmed by our earlier study within the same area. It was found that farmers with small landholdings use more inputs in an effort to maintain or improve productivity that declines due to continuous cropping (Mponela et al., 2016).

Plot fragmentation also increases the probability to start growing legumes but deter the current legume growers to increase area. The increase in area under legumes of be new users with more plots is overshadowed proportionally higher reduction in area under legume by current legume growers. The overall fragmentation effect on legume cropping is negative. As plot fragmentation increases farmers would more likely grow sole crops. Since maize is the staple crop and takes greater share, farmers are more likely to allocate smaller portions to sole legumes.

Farmers who derive higher incomes from crop sales tend to re-invest in farming through purchase of inorganic fertilizers. Since farming is a seasonal enterprise, farmers usually have multiple enterprises and they re-invest the revenues from one into the other (Chalmers & Agar, 2015). They also invest more in enterprises that provide immediate benefits and often fail to forecast into distant future, possibly because of their low formal education. In Nsipe, the dependence on income from other sources and natural resource extraction reduces the reliance on legumes and time for collecting and producing manures, respectively. The benefits of using legumes and farmyard manure for soil fertility are less compared to inorganic fertilizer and may take time to build.

As they earn a living from subsistence farming which is seasonal and low yielding, farmers like most small entrepreneurs manage uncontrollable downside risks by moving resources from enterprises affected by such a risk and invest in activities that sustain their livelihoods (Chalmers & Agar, 2015). Conversely, there is a significantly lower probability that farmers in in Nsipe would purchase more inorganic fertilizer if they get substantial income from other sources. Our results also show that income from cash crops like tobacco reduces the chances to allocate land to legumes. Therefore, farmers who engage in non-farm activities are less likely to invest their income, labour and land in these soils enhancing technologies. This is against the backdrop that farming is considered a major livelihood strategy for rural communities in Malawi.

Natural resource extraction in many instances contributes significantly to the livelihood needs in rural farming communities. Since organic manures are not purchased, the plausible reason for the observed negative influence could be due to labour allocation. Manure collection, storage, processing and application demand more labour (Mustafa-Msukwa et al., 2011) and families who rely on natural resources for income devote much of their time to natural resource collection and sale.

Interestingly, income from cash crops exert a negative influence on the probability to start legume cropping. In the Malawian rift valley escarpments, legumes are mostly grown for home consumption and surplus is sold. It is therefore expected that farmers generating more cash from legumes would grow more. However, a good number of farmers on the lower elevation grow tobacco for sale. Most farmers also sell surplus maize which fetches better prices as the region supplies to the food insecure neighbouring semi-arid and arid regions (Amede et al., 2014).

Situated in the maize mixed farming systems of the rift valley escarpments, livestock is a secondary farming activity to crop production. In Nsipe, the density of livestock is not as high as in adjoining arid areas (Amede et al., 2014), but many people own goats, chickens and pigs. Since manure is not usually traded, farmers who own livestock tend to have a higher propensity to apply more manure. this is a wakeup call for integration of livestock in the dominantly crop-based farming system as this could more likely increase usage of farm yard manure.

Although not usually considered, ownership of farming tools, including hoes has a significant impact on some farming operations. Lack of simple farming tools have been found to be a major limitation to proper farming in Malawi to the extent of endangering people's lives as they use legs and hands for some farming operations (Murray et al., 2016). Although mechanisation is being promoted as best alternative for agricultural productivity in Africa, the hand-held tools especially the hoe are still used for almost all farming operations. Improving efficiency of hand-held tools by diversifying would facilitate uptake of SFM technologies as shovels, spades and wheel-barrows could ease manure collection, preparation and transportation whilst smaller hoes could be efficient in weeding and planting legumes.

Existing cropping practices also considerably shape the decisions of farmers. We found empirical evidence that increasing fertilizer application tend to decrease the probability that farmers would start organic manure application. Given the ongoing efforts by the Malawian government to promote inorganic fertilizers which can give instant yield benefits, there is less likelihood that farmers would take on manures which require preparation and even if well prepared, they have lower and slower benefits in terms of productivity. Yet Nsipe farmers who apply more organic manures are more likely to allocate more land to legumes. Hence promoting organic manures could indirectly enhance legume cropping.

4.4 Plot SFM allocation: drivers of fertilization, organic manuring and legume cropping

In addition to the household's decision to acquire and use a particular SFM technique (presented and discussed in previous chapter), individuals make further decisions as to which plots should receive the SFM in question. These results are further used to parametrise the models for simulating, at plot level, the agent behaviours and impacts of policy alternatives on nutrient balance, productivity and farm incomes.

4.4.1 Probability to apply inorganic fertilizer

Empirical results for drivers of choice for the most common nutrient input, i.e. inorganic fertilizer, are presented in Table 4—6.

Table 4—6 Logistic results for propensity to use inorganic fertilisers on land units owned by households with different attributes

Variables	aggregate	*farmtype1*	*farmtype2*	*farmtype3*
H_{LABOUR}	-0.059	0.163	-0.150	-0.210
	(0.103)	(0.431)	(0.142)	(0.342)
H_{DEPR}	-0.022	0.247	-0.143	-0.034
	(0.077)	(0.224)	(0.134)	(0.225)
H_{WEAI}	-0.041	1.317	-0.142	-0.251
	(0.353)	(0.923)	(0.570)	(0.771)
H_{GENH}	0.445*	1.286*	0.846**	0.365
	(0.235)	(0.720)	(0.430)	(0.498)
H_{EDULHM}	0.093	0.579	0.039	0.220
	(0.142)	(0.505)	(0.206)	(0.337)
H_{GMEM}	0.518**	1.068	0.821*	0.462
	(0.252)	(0.767)	(0.427)	(0.553)
H_{COMM}	0.323	-0.302	0.132	0.774
	(0.229)	(0.580)	(0.467)	(0.622)
H_{TRAN}	-1.210***	-3.200	-1.384	-1.060*
	(0.425)	(2.142)	(1.281)	(0.597)
H_{HECT}	-0.322*	-0.488*	-0.326	-1.031
	(0.187)	(0.265)	(0.339)	(0.668)
P_{HECT}	1.460***	1.769	1.428**	2.847**
	(0.472)	(1.385)	(0.636)	(1.184)
$P_{CULTYRS}$	0.005	0.013	0.012**	0.001
	(0.003)	(0.010)	(0.005)	(0.007)
P_{LEGUD}	0.655***	0.959	0.709*	0.913**
	(0.240)	(0.628)	(0.403)	(0.416)
P_{ORGAD}	0.298	0.162	0.543	0.242
	(0.238)	(0.679)	(0.405)	(0.447)
$P_{TREE10D}$	-0.134	-0.832	0.221	-0.112
	(0.258)	(0.802)	(0.434)	(0.530)
$P_{TN\%}$	-7.727	-52.636	-16.380	-5.283
	(17.601)	(41.144)	(26.669)	(30.604)
$P_{SOC\%}$	0.714	6.397***	-0.193	1.761
	(0.786)	(2.253)	(1.038)	(1.627)
P_{SPI}	0.023	-0.255***	0.152**	-0.035
	(0.033)	(0.095)	(0.060)	(0.053)
$P_{ELEVATION}$	-0.000	0.049***	-0.005	0.001
	(0.003)	(0.017)	(0.004)	(0.005)
Constant	-0.315	-45.839***	5.451	-2.244
	(2.939)	(16.123)	(4.668)	(5.201)
Observations	468	93	217	158

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

The propensity to acquire and apply inorganic fertilizer increases in favour of male headed over women headed households. Significant gender impacts are noticeable among the low input and the medium input farms of farmtypes1&2, respectively. The results also show that households with members that belong to agricultural groups are more likely to apply inorganic fertilizers than those that do not join groups. Group membership is more discerning among medium input farms. The probability to apply fertilizer is higher for larger plot sizes tend and we observe that the effect of plot sizes significantly influence differentiation in fertilizer usage among medium and high input farms. There results show that for medium input farms, the propensity to apply fertilizer significantly increases if the plots have been under cultivated for longer period.

Among the low input farms, the propensity to apply inorganic fertilizer is higher for plots with higher soil organic carbon contents and for the ones at higher elevation. It is worth noting that legume incorporation tends to be associated with application of inorganic fertilizers. This could be because legume seeds are given together with fertilizer and improved maize seed in the subsidy input package. The legume effect is significant among medium and high input farms. Much as larger individual plots have higher probability to be fertilized, when accumulated at household level, larger total land holdings deter fertilization. The land holding's negative relationship is particularly significant for low input farms, implying that among this resource poor group, those with smaller plots have higher propensity to apply fertilizer. We also observe that an increase in transport facilities is associated with a significant reduction in propensity to apply fertilizer especially for high input farms. Interestingly, the stream power index has a positive influence among medium input farms (*farmtype2*) but reduces the propensity for the low input farms (*farmtype1*).

The second hurdle that farmers have to cross with regard to fertilizer application is the amount of inputs used. Table 4—7 shows the results from the generalised linear model. Increasing fertilisation is positively associated with amount of manure applied, increase in women empowerment (WEAI) especially among medium input households, sand content of plots for the low input farms and flow accumulation especially for plots of medium input farms. The positive association with manure across farm types is indicative of complementary effects that farmers observe if they supplement fertilizer and manure in larger quantities. Another interesting finding is that women empowerment induces increased usage of inorganic fertilizer signifying the role of increased women's bargaining power in farm investment. It is expected that with the plots down the slope that have a higher flow accumulation receive increasing sedimentation and would require less supplementation. However, the plains of the study area not the valley bottoms, most of the sediments are transported out of the sub catchment (Chimtengo et al., 2014; Sandram, 2018). Since these plains are prime agricultural lands that have been under cultivation for generations (CIAT, 2016), farmers proactively act to boost productivity by increasing fertilizer.

Increasing the level of fertilizer subsidy, plot sizes, whether a plot was supplied with manure or not and indigenous soil nitrogen are associated with decrease in amount of fertilizer applied. The increased level of subsidy has significant negative influence among low input and medium input farms. This implies that those that increasingly rely on subsidised fertilizer apply less fertilizer than that those that supplement or purchase from markets. Much as the larger plot sizes are positively associated with higher probability for those that are not using fertilizer to start using, farmers with large plots apply less fertilizer per capita than those with small plots. these results are consistent with the observed inverse relationship between input use intensity and plot sizes (Sheahan & Barrett, 2017). It is also interesting to note that despite increasing manuring having a positive association with the quantity of fertilizer applied, among medium input farms, those that do not apply manure use comparatively more inorganic fertilizer than those that apply manure. The results also show that farmers who own plots with relatively higher soil nitrogen content apply significantly less inorganic fertilizer, although the effect in not discernible for farm types.

Table 4—7 Generalised linear model results for factors influencing amount of fertilizer used

Variables	aggregate	*farmtype1*	*farmtype2*	*farmtype3*
H_{WEAI}	0.2734*	0.3660	0.5138***	0.0459
	(0.1422)	(0.2628)	(0.1722)	(0.2543)
H_{LABOUR}	0.0083	0.0322	-0.0168	-0.0057
	(0.0354)	(0.1557)	(0.0444)	(0.0909)
H_{DEPR}	-0.0062	-0.0469	-0.0029	-0.0523
	(0.0335)	(0.0578)	(0.0418)	(0.0552)
P_{SUBIDY}	-0.5635***	-0.7981**	-0.8353***	-0.2161
	(0.1800)	(0.3134)	(0.2601)	(0.3273)
H_{EDULHM}	-0.0596	-0.1360	-0.0008	0.0016
	(0.0557)	(0.1555)	(0.0587)	(0.0881)
H_{INMS}	-0.0001	0.0002	-0.0001	0.0000
	(0.0001)	(0.0003)	(0.0001)	(0.0003)
H_{TRAN}	-0.0254	0.8127	-0.0439	-0.6784
	(0.2164)	(0.5850)	(0.2300)	(0.5419)
P_{HECT}	-0.7796***	-0.5294***	-1.3756***	-2.2310***
	(0.0840)	(0.0728)	(0.1426)	(0.3129)
$P_{CULTYRS}$	0.0022	-0.0023	-0.0005	0.0022
	(0.0015)	(0.0025)	(0.0019)	(0.0021)
P_{LEGUD}	-0.1114	-0.0912	0.0418	0.0223
	(0.0989)	(0.2042)	(0.1305)	(0.1555)
P_{ORGAD}	-0.2567**	-0.1484	-0.3397**	-0.1528
	(0.1005)	(0.2514)	(0.1433)	(0.1860)
P_{ORGA}	0.0002***	0.0005**	0.0003**	0.0001**
	(0.0000)	(0.0002)	(0.0001)	(0.0001)
$P_{SAND\%}$	0.0029	0.0220*	-0.0019	0.0060
	(0.0054)	(0.0114)	(0.0061)	(0.0098)
$P_{TN\%}$	-7.2138*	12.8142	-8.4527	-10.9217
	(4.1555)	(8.6822)	(5.8842)	(6.7437)
P_{Kmgkg}	-0.0022	-0.0041	-0.0009	-0.0007
	(0.0015)	(0.0028)	(0.0018)	(0.0026)
$P_{FLOWACC}$	0.0005***	-0.0008	0.0007***	-0.0036
	(0.0002)	(0.0006)	(0.0002)	(0.0049)
Constant	6.4556***	3.9304***	6.9545***	6.8576***
	(0.5283)	(1.1546)	(0.6943)	(0.9140)
Observations	348	70	166	112

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

4.4.2 Probability to apply and intensify organic manures

As presented in Table 4—8, the drivers with significant positive influence on use of organic manures include higher fertiliser subsidy among high input farms (*farmtype3*), having more livestock units – with plot specific effects discernible only among medium input farms (*farmtype2*), larger plot size whose farm type specific effect is discernible only among low input farms (*farmtype1*) and whether a plot receives complementary fertilizer or not for medium input farms (*farmtype2)*. The positive effect of subsidy is indicative that farmers who increasingly rely on subsidy have high probability to apply manure especially if they belong the relatively high fertilizer and high manure input farm type. The effect of livestock is as expected since manure is not traded. Therefore, those with livestock have an added advantage to use high value manure made from animal droppings (Emerton, 2016).

On the other hand, significant reductions in the propensity for households to apply organic manures have been estimated to be associated with increase in number of dependants compared to workers for the aggregate and among the low input farms and with legume integration for the aggregate and among the medium input farms. The results also show that tree cover tend to have farm type specific effects: higher tree cover is associated with increase in probability to apply manure for the low input farms but there is a negative association among moderate input farms.

Table 4—8 Logistic results for propensity to use organic manures on land units owned by households with different attributes

Variables	aggregate	*farmtype1*	*farmtype2*	*farmtype3*
H_{LABOUR}	0.034	-0.227	0.153	0.000
	(0.081)	(0.338)	(0.117)	(0.234)
H_{DEPR}	-0.145**	-0.349*	-0.084	-0.187
	(0.066)	(0.195)	(0.103)	(0.147)
H_{WEAI}	-0.482	-0.042	-0.831	-0.161
	(0.311)	(0.888)	(0.524)	(0.618)
$P_{SUBSIDY}$	0.258	-0.032	-1.067	2.673***
	(0.451)	(1.195)	(0.778)	(0.882)
H_{TLUN}	0.110*	0.342	0.187**	0.211
	(0.065)	(0.316)	(0.086)	(0.153)
H_{EDULHM}	-0.088	0.498	-0.269	0.196
	(0.131)	(0.417)	(0.195)	(0.257)
H_{COMM}	-0.212	-0.057	-0.511	-0.415
	(0.174)	(0.376)	(0.338)	(0.366)
P_{HECT}	0.366*	0.444**	0.502	1.083
	(0.204)	(0.210)	(0.431)	(0.780)
P_{LEGUD}	-0.353*	0.923	-0.698**	0.026
	(0.208)	(0.614)	(0.327)	(0.405)
P_{FERTD}	0.264	0.361	0.748*	-0.374
	(0.241)	(0.621)	(0.390)	(0.443)
$P_{TREE10D}$	0.018	1.085*	-0.845**	0.560
	(0.224)	(0.608)	(0.405)	(0.398)
$P_{SAND\%}$	-0.010	0.012	-0.006	-0.025
	(0.012)	(0.037)	(0.018)	(0.025)
P_{SLOPE}	0.011	-0.032	-0.020	0.034
	(0.022)	(0.059)	(0.040)	(0.038)
Constant	0.218	-1.885	-0.021	0.388
	(0.650)	(2.079)	(0.928)	(1.350)
Observations	468	93	217	158

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

The expectation is that households with increased dependent household members will be constrained to allocate labour to manure making, but these results points to them being more likely to switch to manure application. Such tendency needs further exploratory study to unravel the conditions that lead these labour constrained households to be more likely to resort to labour demanding manuring. We also observe that for the moderate fertilizer and high manure farms (aka medium input farms), legume integration lowers the propensity for farmers to apply manure on plots where they are currently not applying.

For the plots that received manure (46% of those sampled), the application levels varied. About 23% received < 0.5 ton ha^{-1}, 16% between 0.5 and 1.0 ton ha^{-1} whilst 6.4% received ≥ 1.5 ton ha^{-1} during the 2016 / 2017 growing season. Empirical results from generalised linear model show that increasing plot sizes has a significant negative influence on the amount of manure applied across farm types (Table 4—9). The opposing effects of plot size on the decision to start which is positive and the second intensification hurdle which is negative mirror that of fertilisation and supports the inverse plot size vs input intensity relationship (Sheahan & Barrett, 2017). These results are indicative that farmers with smaller plots sizes use comparatively larger quantities of manure.

Table 4—9 Generalised linear model results for factors determining intensity of manure applied to a plot.

Variables	aggregate	*farmtype1*	*farmtype2*	*farmtype3*
H_{WEAI}	0.1775	-0.552*	-0.3689	0.3218
	(0.3594)	(0.323)	(0.3122)	(0.5668)
H_{LABOUR}	0.0417	0.465**	0.2835***	-0.0700
	(0.0520)	(0.193)	(0.0503)	(0.1552)
H_{DEPR}	-0.0637	-0.052	0.0223	-0.0442
	(0.0644)	(0.071)	(0.0575)	(0.0960)
$P_{SUBSIDY}$	-0.1235	-0.461	-0.6382*	-0.0444
	(0.3393)	(0.773)	(0.3323)	(0.4395)
P_{FERT}	0.0011***	0.003***	0.0002	0.0012**
	(0.0004)	(0.001)	(0.0002)	(0.0005)
H_{EDULHM}	0.0170	-0.767***	-0.2545***	0.1273
	(0.0723)	(0.201)	(0.0921)	(0.1389)
H_{INCC}	-0.0005	-0.005	-0.0005	0.0002
	(0.0004)	(0.004)	(0.0007)	(0.0009)
H_{INCL}	-0.0007	0.006**	0.0087**	-0.0144**
	(0.0024)	(0.003)	(0.0037)	(0.0073)
P_{HECT}	-0.8210***	-0.603***	-1.5192***	-0.9253**
	(0.0702)	(0.057)	(0.2267)	(0.4378)
P_{LEGUD}	0.1026	0.846***	0.0152	0.1205
	(0.2093)	(0.267)	(0.1770)	(0.3328)
$P_{SOC\%}$	0.0014	0.995	0.3339	-0.5575
	(0.3869)	(0.615)	(0.4631)	(0.6692)
$P_{SAND\%}$	-0.0025	-0.019	-0.0114	0.0192
	(0.0120)	(0.016)	(0.0096)	(0.0224)
Constant	6.6970***	5.272***	6.2591***	6.7504***
	(0.7513)	(1.004)	(0.5931)	(1.5710)
Observations	185	32	78	75

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Differentiated factor effects are observed among the different farm types. Increasing subsidy would induce significant reduction in levels of manuring for the moderate manure and high fertilizer farms (*farmtype2*). Increases in the total fertilizer used per plot is associated with increase in manuring from the aggregated analysis, for the low input farms (*farmtype1*) as well as for the high input farms (*farmtype3*). As the women bargaining power increases, the likelihood to increase manure application tend to be significantly lower for moderate input farms. Increase in family labour and livestock units tend to be associated with increases in manure input for low and medium input farms but reduces the likelihood to intensify manuring for high input farms. Conversely, increase in education level among household members reduces the probability to increase manuring for *farmtypes1&2* but increases the probability for high input farms. Moreover, for low input farms, increasing income from crops has a negative influence whilst legume cropping and indigenous soil carbon tend to have a positive and significant influence on intensity of manuring.

4.4.3 Probability for switching from sole cereal to legume integration

Table 4—10 4—11 Logistic results for propensity to plant legumes on a plot given household and plot attributes.

Variables	aggregate	*farmtype1*	*farmtype2*	*farmtype3*
P_{LABOUR}	-0.08	-0.72**	-0.11	-0.27
	(0.10)	(0.36)	(0.13)	(0.30)
H_{DEPR}	-0.03	-0.37**	-0.04	0.12
	(0.07)	(0.18)	(0.12)	(0.14)
H_{WEAI}	0.56*	0.73	0.78	0.76
	(0.33)	(0.95)	(0.52)	(0.67)
P_{SUBIDY}	0.27	-2.43*	0.02	2.08**
	(0.51)	(1.38)	(0.86)	(0.90)
H_{EDULHM}	0.31**	1.07**	0.33	0.49
	(0.15)	(0.52)	(0.21)	(0.32)
H_{COMM}	0.03	0.53	0.03	-0.26
	(0.20)	(0.41)	(0.34)	(0.41)
H_{GMEM}	-0.42*	-0.42	-0.40	-0.48
	(0.22)	(0.65)	(0.35)	(0.41)
P_{HECT}	1.32***	0.73***	1.28***	3.30***
	(0.34)	(0.28)	(0.50)	(0.81)
$P_{CULTYRS}$	0.01***	0.03***	0.00	0.01***
	(0.00)	(0.01)	(0.00)	(0.00)
P_{FERTD}	-0.66***	-0.16	-0.72*	-1.24**
	(0.25)	(0.67)	(0.40)	(0.49)
P_{ORGAD}	0.33	-1.27**	0.73**	-0.23
	(0.21)	(0.64)	(0.33)	(0.42)
$P_{TREE10D}$	-0.25	0.43	-0.82**	0.22
	(0.23)	(0.69)	(0.39)	(0.42)
$P_{SAND\%}$	0.01	0.06	0.02	-0.03
	(0.01)	(0.04)	(0.02)	(0.03)
$P_{SOC\%}$	-0.22	-0.24	-0.18	0.60
	(0.47)	(1.01)	(0.74)	(0.92)
P_{SLOPE}	-0.00	-0.05	0.03	-0.01
	(0.02)	(0.06)	(0.04)	(0.04)
Constant	-1.72**	-3.16	-2.00	-1.05
	(0.84)	(2.34)	(1.36)	(1.61)
Observations	468	93	217	158

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Legume integration in cereal dominated farming systems is considered to be a gateway to sustainable agricultural intensification (Gilbert, 2004). Despite their multiple benefits that include nutrition and income, their extent of cultivation is still low. From our empirical analysis, the results in Table 4—10 shows that the propensity for households to switch to legume integration is negatively associated with increase in available household labour for low input farms (*farmtype1*) and increase in number of dependants over workers among low input farms, group membership for the entire range of farms (aggregate), whether a plot receives inorganic fertilizer or not (among the moderate input farms of *farmtype2,* high input farms of *farmtype3* and the aggregate) and if there are trees on farm among moderate input farms.

On the other hand, the likelihood for legume integration increases with women empowerment in decision-making for aggregate, higher educational attainment by household members, and having larger plot sizes and if land has been cropped for longer periods. Type specific effects were detected for fertilizer subsidy and manuring. Increasing subsidy decreases the likelihood to plant legumes among the low input farms but increases the likelihood among the high input farms. The odds to plant legumes increase with application of organic manure among low input farms but increase for medium input farms.

4.4.4 Probability to retain trees on farm

Availability of trees on farms is one key feature of smallholder farmlands that influence nutrient stocks directly through processes such as erosion control (Banda et al., 1994) and indirectly through its association with the decisions by farmers to plant certain crops or apply inputs (Kuyah et al., 2019). The trees are either deliberately planted or retained on farms (GoM, 2017). Table 4—12 shows that the higher the number of dependants in relation to workers (among medium input and high input farms and the aggregate), higher education level of household members (among medium input farms and the aggregate), more livestock (among medium input farms and the aggregate), and larger plot areas (among high input farms) increase the propensity to retain higher cover of trees.

However, higher availability of family labour, women empowerment and legume intensification lowers the likelihood to retain higher tree cover on farm. The negative effect of women empowerment on retention of trees is not as expected. In the region, marriage system is matrilineal and women own land hence, as long-term natural resources, trees are often used to claim and safeguard ownership (German et al., 2009). Moreover, as woodlands are cut and fuelwood become scarce, trees on farm are a viable source of energy for cooking. The empowered women land owners are therefore expected to be more likely retain trees on farm. The negative influence may be due to other factors such as information on the benefits of trees and other purposes that the trees on farm are intended for. Generally, the large crown trees are either the planted fruit trees such as *Mangifera indica* or the indigenous poles and timber species, which are polarded. Women get the pollards, but it is men who decide on the density and species to retain.

The effects of elevation and manuring are farmtype specific. Manuring has a positive association with the probability to retain trees for high input farms (*farmtype3*) but a negative one for medium input farms (*farmtype2*). The probability to retain trees increases for plots on higher elevation among the medium and high input farms as well as for the entire sample. Conversely, the probability to retain trees increases in favour of plots situated on lower elevation among the low input farms (*farmtype1*). Much as higher elevation landscapes are prone to erosion and controlling degradation could spur farmers to reserve trees, the rationale for the low input farms on lower elevation to retain trees could be their potential contribution to soil fertility improvement and as source of fuelwood, poles and timber. The low level of manure and fertilizer application indicates that these are resource poor who, in pursuit of options for improving soil fertility and livelihoods, would be retaining more natural resources such as trees.

Table 4—12 Logistic results for propensity to retain trees above 10% threshold on a plot given household and plot attributes.

Variables	aggregate	*farmtype1*	*farmtype2*	*farmtype3*
H_{LABOUR}	-0.270**	0.411	-0.305*	-0.712**
	(0.106)	(0.394)	(0.156)	(0.313)
H_{DEPR}	0.152**	0.124	0.182	0.313**
	(0.073)	(0.170)	(0.130)	(0.147)
H_{WEAI}	-0.581*	-1.275*	-0.780	0.280
	(0.345)	(0.754)	(0.628)	(0.640)
P_{SUBIDY}	-0.096	-1.859	-0.462	0.120
	(0.501)	(1.429)	(0.853)	(0.906)
H_{EDULHM}	0.295**	-0.751	0.504**	0.194
	(0.142)	(0.554)	(0.198)	(0.366)
H_{TLUN}	0.203***	0.504	0.274***	0.098
	(0.059)	(0.734)	(0.080)	(0.214)
P_{HECT}	0.317	0.015	0.644	1.621*
	(0.203)	(0.251)	(0.509)	(0.908)
$P_{CULTYRS}$	-0.002	0.001	-0.008	-0.003
	(0.003)	(0.007)	(0.005)	(0.005)
P_{LEGUD}	-0.364	0.553	-1.168**	0.003
	(0.248)	(0.758)	(0.457)	(0.429)
P_{FERTD}	-0.019	-0.006	0.423	-0.146
	(0.269)	(0.620)	(0.451)	(0.534)
P_{ORGAD}	0.082	0.903	-0.878**	0.789*
	(0.224)	(0.683)	(0.392)	(0.412)
$P_{SOC\%}$	0.692	0.991	0.773	-0.190
	(0.453)	(1.281)	(0.799)	(0.938)
$P_{ELEVATION}$	0.008***	-0.037**	0.010**	0.010***
	(0.002)	(0.019)	(0.004)	(0.004)
Constant	-8.137***	29.123*	-10.190***	-9.627***
	(2.190)	(16.027)	(3.632)	(3.702)
Observations	468	93	217	158

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

4.5 Maize yield productivity under heterogenous household and farm conditions

Closing the yield gaps in African farming systems has been one of the major agricultural policy goals for the past decade. Using Malawi as an example, despite the increases in yields following introduction of subsidised mineral fertilizers and improved seeds, yields have stagnated for the past 15 years and gaps are still wide. The attainable yield for maize varieties in Malawi range from 4-15 t ha^{-1} (Tamene, Mponela, Ndengu, et al., 2016) but in our study the yields achieved by farmers in the Rift Valley escarpments for the 2016-2017 growing season were around 1.4 t ha^{-1}.

Table 4—13 shows results from the generalised linear model indicating that yield variations are associated with baseline soil fertility factors, nutrient and labour inputs and management practices. From the full sample analysis (aggregate), manure input, plot size, cropping system, amount of course fragments and elevation significantly affect maize yields. Moreover, labour and fertilizer inputs constrain the productivity for some of the moderate input farms (*farmtype2*).

According to the yield trends for the past 30 years, it is evident that rainfall variability and other events such as pest outbreaks, largely mediates interannual yield variations (Figure 1.1). For a region with rugged topography, the position of plots across the landscape and inherent soil conditions and nutrient levels have significant impact on land productivity. Controlling for fertilizer input and management, yield decreases follow the topographic and nutrient gradients. Yield tend to increase with increase in elevation indicating that during the study time, plots on higher elevation had comparatively higher productivity. This is expected since farmers recently cleared-up virgin woodland soils on slopes located at higher elevation (Braslow & Cordingley, 2016). Similarly, increase in flow accumulation is also associated with decreasing yield, significantly so for *farmtype3* (Table 4—13).

Our study shows that, despite the high variability in fertilizer input, its effects are indeterminant signalising that maize in the study region has become non-responsive to fertilization. Quite low and insignificant effects of nitrogen fertilization has also been found by Burke et al (n.d.) who analysed a 4-year panel data. Although we controlled for inherent fertility, agronomic practices and farmer ability, there is need for further exploration on input responses as various studies have found inconsistent effects. A pioneering study by Tamene et al (2016) in a high productive site within the region found that maize yield was responsive to basal dressing with phosphorus and sulphur containing fertilizers, which is in sharp contrast to the negative effects observed by Burke et al (n.d.) for soils with low SOC of <0.94%. A recent study by Kopper et al (2020) in the same region found that maize was responsive to N fertilization, and was more profitable for low fertility farms with SOC <1.3% than productive ones. Burke et al (n.d.) also found that maize is subtly responsive when fertilization is coupled with weeding, and is slightly higher for farms with low SOC of < 0.94% than for those with higher SOC. Whilst Han Wang et al. (2019), in their study within the same study sites, found that N fertilization was more effective for high productive sites. Although Kopper et al (2020) and Burke et al (n.d.) purport that effects are disaggregated by productivity, they could well be site-specific as data was composited from different agro-ecological zones that inherently differ in productivity (Li et al., 2017) and SOC and soil nutrient contents (Tamene et al., 2019).

The results (Table 4—13) also show that maize significantly responded to manure input, especially for medium input farms of *farmtype2*. In tandem, the results also show that maize was more yielding under soils with higher SOC, especially for low input farms. These results highlight the significance of enhanced SOC and organic inputs especially for marginal sites and farms (H. Wang et al., 2019). From our results and others in the region, we observe that, to be productive, fertilization strategies need to take into account site productivity, variations in SOC levels, manure input, weeding and other complementary factors. Although labour is generally regarded as a constraint, in smallholder farming systems, farmers manage considerably small plots and tend to have unlimited labour supplies. Their behaviour are well articulated in the works of Lewis (1954), who purports that for subsistence farmers, their labour allocation behaviour cannot be fully explained by capitalistic theories that assume competitiveness. This observation still holds for Africa where there is some level of unawareness on the margins from different labour investment portfolios and households tend to establish higher standards for themselves by focusing more on building technical knowledge than enhancing capital. Correspondingly, in our study, the effects of labour on yield improvement are noticeably significant only for low input farms, which is negative. The study by Kopper et al, (2020) showed that maize yields were not responsive to weeding labour although weeding labour had higher profitability for high productive farms with 1.3 to 4.3% SOC who apparently applied relatively lower fertilizers than the low fertile farms.

Table 4—13 Real farm maize yield productivity under heterogenous households and farm conditions

Variables	aggregate	*farmtype1*	*farmtype2*	*farmtype3*
P_{FERT}	-0.0005	-0.0009	0.0002	-0.0006
	(0.0004)	(0.0011)	(0.0005)	(0.0006)
P_{FERT}^2	3.79e-07	1.73e-06*	-3.93e-07	3.21e-07
	(2.94e-07)	(9.43e-07)	(4.62e-07)	(3.64e-07)
P_{ORGA}	0.0003**	0.0004	0.0014***	-0.0001
	(0.0001)	(0.0006)	(0.0003)	(0.0001)
P_{ORGA}^2	-2.58e-08	-2.41e-07	-3.73e-07***	2.35e-08
	(1.96e-08)	(2.00e-07)	(1.06e-07)	(1.68e-08)
P_{LABOUR}	-0.0007	-0.0032*	0.0001	-0.0002
	(0.0010)	(0.0022)	(0.0015)	(0.0018)
P_{LABOUR}^2	2.85e-06	7.81e-06***	1.66e-06	1.60e-07
	(2.03e-06)	(2.47e-06)	(3.19e-06)	(1.86e-06)
P_{HECT}	-0.5599***	-0.6852***	-0.5568***	-0.6870**
	(0.1214)	(0.1133)	(0.1783)	(0.3067)
$P_{CULTYRS}$	-0.0014*	-0.0021	-0.0007	-0.0017*
	(0.0012)	(0.0026)	(0.0015)	(0.0016)
P_{SUBIDY}	-0.0467	0.2743	-0.3716	0.0073
	(0.2614)	(0.5344)	(0.2564)	(0.2650)
P_{LEGUD}	-0.2473***	-0.0564	-0.2393*	-0.3293**
	(0.0902)	(0.1867)	(0.1403)	(0.1505)
$P_{TREE10D}$	0.0745	-0.7676***	0.3814**	0.1325
	(0.1108)	(0.1930)	(0.1488)	(0.1415)
P_{Kmgkg}	0.0005	0.0014	0.0002	0.0028
	(0.0012)	(0.0026)	(0.0014)	(0.0020)
$P_{SOC\%}$	0.1833	1.0207**	0.0728	0.0365
	(0.2158)	(0.5053)	(0.2153)	(0.2887)
$P_{SAND\%}$	0.0081	0.0199*	0.0027	0.0174
	(0.0073)	(0.0114)	(0.0071)	(0.0127)
$P_{COARSE\%}$	-0.0449***	-0.0464*	-0.0341*	-0.0246
	(0.0162)	(0.0260)	(0.0191)	(0.0299)
$P_{ELEVATION}$	0.0017*	0.0052	0.0018	0.0010
	(0.0010)	(0.0031)	(0.0013)	(0.0015)
$P_{FLOWACC}$	-0.0003	-0.0003	-0.0002	-0.0016**
	(0.0003)	(0.0009)	(0.0003)	(0.0008)
H_{DEPR}	0.0196	0.0199	-0.0272	0.0551
	(0.0359)	(0.0513)	(0.0365)	(0.0740)
H_{EDULHM}	0.0379	0.2935*	-0.0354	0.0858
	(0.0676)	(0.1603)	(0.0746)	(0.1138)
H_{GMEM}	0.0301	0.3585*	-0.1780	0.0366
	(0.0944)	(0.1923)	(0.1236)	(0.1364)
H_{WEAI}	0.2335	0.0849	0.2326	0.3534
	(0.1505)	(0.2712)	(0.1796)	(0.2583)
Constant	5.8926***	1.2931	5.8727***	5.7731***
	(1.0605)	(3.2600)	(1.3047)	(1.6643)
Observations	390	88	169	133

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

4.6 Nutrient balance

As indicated in Table 4—13, from the full balance analysis, the annual C : N : P : K inputs (kg ha^{-1} yr^{-1}) averaged 306 : 92 : 6 : 8 against outputs of 1124 : 70 : 4 : 86, giving average balances of -818 : 22 : 3 : -77, respectively. Considering the human inflows and outflows, the corresponding partial budgets are -28 : 8 : -59 kg ha^{-1} yr^{-1}. The N : P : K balances correspond with results obtained in Kenyan studies of -76 : -5 : -10 (Van den Bosch et al., 1998), -112 : -3 : -70 (Smaling & Fresco, 1993), and -73 : 3 : -16 (De Jager et al., 1998). These results indicate that the output pathway, crop products, take considerably higher N and K than what is added through fertilization and manuring. The physical outputs, over which current human actions have less control, such as erosion, leaching and gaseous losses lead to significant losses of N, K and C.

The major input for nitrogen is inorganic fertilizer followed by biological nitrogen fixation, and atmospheric deposition while the outputs include gaseous, erosion, crop produce and residues. The N balance, however, is much lower because of the higher fertilization, nitrogen derived from air by legumes and lower nutrient stoichiometric values used for estimation of N in crop outputs. Notably, even with full nutrient budget, 35% of plots have positive nitrogen. Among the farm types, the low fertilizer - low manure (*farmtype1*) and high fertilizer - low manure (*farmtype2*) have negative average balances with relatively fewer farms with positive balances compared to moderate fertilizer - high manure farms (*farmtype3*). Looking at the range of balances for the surveyed plots, it is evident that there exist farms with extremely high nitrogen losses of up to -401 kg ha^{-1} yr^{-1} but also some farms with net positive nitrogen balances of up to 589 kg ha^{-1} yr^{-1}.

For phosphorus, the main input source is also inorganic fertilizer while outputs include crop produce and erosion. Again, some farms are estimated to lose up to -120 kg ha^{-1} yr^{-1} while others could be experiencing phosphorus build-up of up to 81.6 kg ha^{-1} yr^{-1}. More than 50% of the farms are estimated to have a net positive phosphorus balance. Potassium is obtained from organic manure and atmospheric deposition. With almost all surveyed farms estimated to have net potassium losses, these inputs are quite low to counter the high output levels through leaching and residues.

Table 4—14 Baseline carbon and NPK flows and balances (kg ha^{-1} yr^{-1})

	Nitrogen					Phosphorus					Potassium					Carbon				
	$\bar{x}$	sd	min	max	0>%	$\bar{x}$	sd	min	max	0>%	$\bar{x}$	sd	min	max	0>%	$\bar{x}$	sd	min	max	0>%
Inputs (population)																				
IN1$_{fertilizer*}$	48	52	0	222		6.4	11.2	0.0	85											
IN2$_{organic*}$	3	6	0	32		1.0	2.1	0.0	11		4	8	0	39		11	24	0	133	
IN2$_{residues*}$																237	277	0	2496	
IN3$_{legume\text{-}bnf*}$	14	36	0	238																
IN4$_{sediment}$	3	8	0	84		0.3	1.9	0.0	61		1	3	0	49		58	196	0	2914	
IN5$_{atm\text{-}depo}$	14	0	14	15		1.4	0.0	1.4	1		4	0	4	4						
Total-in	92	95	14	686		6.4	7.9	0	41		8	8	4	60		306	344	0	3461	
Outputs (population)																				
OUT1$_{product*}$	15	17	0	143		3.3	0	27	27		7	8	0	66						
OUT2$_{residues*}$	15	18	0	138		1.3	1.5	0	13		19	22	0	191						
OUT3$_{erosion}$	26	31	0	257		1.7	4.2	0	40		6	9	0	89		406	612	1	6279	
OUT4$_{leaching}$	4	1	2	7							54	12	34	74						
OUT5$_{gaseous}$	19	0	19	19												718‡	139	444	1219	
Total-out	70	47	13	303		3.9	4.6	0	40		86	33	35	335		1124	623	508	7260	
Full balance																				
Population	22	94	-256	595	52	3	12	-36	81	55	-77	32	-302	3	0.0	-818	695	-6853	1779	4
Farmtype1	5	65	-138	381	45	1	9	-30	52	45	-77	29	-283	-22	0.0	-729	564	-5718	1580	3
Farmtype2	34	96	-172	569	61	6	13	-31	78	68	-75	30	-302	-19	0.0	-813	594	-4852	1544	4
Farmtype3	16	105	-256	595	46	2	11	-36	81	46	-81	35	-282	3	0.1	-875	858	-6853	1779	5
*Partial balance**																				
Population	35	61	-203.24	334	71	2.4	9	-36	39	50	-22	28	-230	35	5.7	248®	282	0	2587	94
Farmtype1	20	52	-164.41	298	68	0.3	8	-35	39	37	-24	26	-219	18	2.4	241®	242	5	2184	100
Farmtype2	44	62	-165.1	217	76	4.2	9	-32	31	59	-20	26	-229	13	4.8	216®	278	0	2587	92
Farmtype3	32	63	-203.24	334	68	1.4	9	-36	27	45	-24	32	-230	35	8.7	293®	303	0	2477	95

$\bar{x}$ = mean; sd = standard deviation; 0>% = percentage of plots with positive balances; *partial balance estimated from flows influenced by human action; ‡SOM degradation, ®inputs only.

However, the mechanism by which the large stocks of unavailable potassium in rock solids and minerals, and the slowly available reservoirs trapped in clay is mineralised is largely unknown. Hence this natural enrichment is not included in these estimations leading to under estimation of inputs to the rooting zone. Nonetheless, with continuous cropping potassium deficiencies have been identified (see section 4.2.1) and pot experiments with Alfalfa grass registered crop responses to potassium fertilisation (Lakudzala, 2013). These responses have, however, not been significant and consistent for field studies of maize. A multilocational study conducted in 9 sites in East, West and southern Africa - 3 of which were in Malawi - showed that omission of K does not significantly reduce crop yields (Kihara et al., 2016). Considering the deficiency levels, unless the natural enrichment is estimated, the current interventions and usage levels of organic manures are insufficient to replenish even the losses from crop harvest products and residues.

The main carbon inputs are crop residue retention, sediment deposition and organic manure which fall short to counter the large losses through erosion and decomposition of SOM. The full balance analysis shows that only 4% of the plots have positive net balance, with some plots having a net loss of up to -6,853 kg ha^{-1} yr^{-1}. Previous field studies found no response to carbon enrichment (Kihara et al., 2016) mainly because with the SOC is critically low for maize fertilizer response (Figure 4.3). Soils with sub-optimal SOC levels (Figure 4.3), and the low structural stability indicating that soils are prone to degradation (Figure 4.5). These negative balances, if not addressed, could subject the soils to structural instability and accelerate loss of soil and the nutrients (Mpeketula, 2016).

The extreme negative and positive inputs, outputs and balances are astounding. For instance, the maximum N input from fertilization is 222 kg N ha^{-1} which deviates too much from the mean of 48 kg N ha^{-1} and could emanate from several sources. Possibly, farmers with small plots apply exceedingly too much fertilizer as some scoop with hands, or the reported fertilizers could have been bloated. Since the data for fertilization was collected at the end of the growing season, it is possible that some farmers may have forgotten the actual amounts applied to a plot or may have given incorrect estimates.

The current efforts aimed at increasing fertilization, manuring and legume cropping have the potential to improve the N balance but falls short in offsetting losses of K and C through ecological processes. These results highlight the implications of focusing only on human induced flows and calls for consideration of ecologically driven nutrient flow management. Although sedimentation redistributes nutrients from upper to lower landscapes, without land management measures to control erosion, the nutrient losses through erosion continue to be enormous. Therefore, we expect that the plot-to-plot differences to be high, possibly resulting from inter-related policy, household and plot factors which would warrant an integrated agent-based spatial-temporal analysis.

4.7 Impact of fertilizer subsidy on sustainable agricultural intensification.

4.7.1 *Subsidy and productivity: fertilizer & organic manure input and crop yield*

The baseline scenario has been constructed to mimic real farm conditions and shows predicted fertilizer and manure input and maize yield output (Figure 4.7). These predictions are based on the current subsidy regime (averaging 28% of fertilizer purchase price) and allowing other dynamic idiosyncratic factors such as period the farm has been under cultivation to progressively change. The results for the baseline case demonstrate increasing, stabilising and declining trends across the three productivity indicators as well as among the three farm types. The projected fertilizer usage (kg ha^{-1}) increases for the whole population and high input farms (*farmtype3*) but slightly decreases for low and medium input farms (*farmtypes 1& 2*). The manure input is projected to have a stabilising trend while maize yields are more likely to decrease.

Factor elimination revealed that the decreasing trends in maize yields are strongly associated with cultivation period, among other factors. The baseline dataset comprised of farms with varying cultivation periods and the empirical analysis showed a negative association with maize yield (Table 4—13). The maize yield gradient is expected as plots that were opened up recently in hillslopes are considered more fertile (Braslow & Cordingley, 2016). But without soil conservation measures (CIAT, 2016), productivity is likely to decline if cultivation on these sensitive slopes is continued (Banda et al., 1994). Although we have not included climatic variables (assumed to be constant for our training site of 4 km x 8 km), the national input subsidy program coincided with better rainfall, such that the reported doubling of yields could not be entirely attributed to increased fertilisation induced by subsidy (Denning et al., 2009). Moreover, in recent years, although total fertilizer usage remained high, the maize yields fell to the pre-subsidy levels due to either recurrent droughts or the 2016/2017 fall armyworm outbreak (see Figure 1.1).

Given the autonomy that farmers have to decide on nutrient replenishments on their farms (Sambo et al., 2015), they tend to emulate those with similar statuses and aspirations and plan future actions based on their shared experiences. Hence, the farmtype specific effects observed in this study could be explained by behavioural economics focusing on the way farmers learn from one another and from their experiences. Generally, one group of farmers in the community ascribe performance by members of another group rather to their advantaged or disadvantaged positions than to the external forces such as fertilizer subsidies (Grant & Ashford, 2008). Since the agentic behaviours to anticipate, plan and act are drawn from experiences, their consequences are self-evident. From these results, we observe that although farmers attribute the behaviours to idiosyncratic abilities, their proactive behaviours are mediated by situational causes such as declining soil fertility where policies such as subsidies may have varying effects.

Juxtaposing the trend graphs in Figure 4.7, and from the empirical analyses in Table 4—13, maize is much more responsive to manuring than fertilization. Maize yields for the farms that received both higher fertilizer and higher manure (*farmmtype3*) yielded more maize than those that only received higher fertilizer and moderate manuring (*farmtype2*). This has long-term implications on productivity, nutrient balances and profitability.

Compared to the current regime that subsidizes fertilizer by an average of around 28%, there are stark deviations associated with increasing subsidy to reach all farmers i.e. *universal subsidy* as well as a regime that would aim at enabling farmers to graduate from subsidy i.e. *zero subsidy* (Figure 4.7). The Bonferroni multiple comparison test shows that increasing subsidy to an average of 70% (SU-SC in Table 4—15) is associated with a decreasing shift in total amount of fertilizer used by an average of -8, -17 and -18 kg ha^{-1} yr^{-1} for low input farms, medium input farms and for the whole population, respectively. This translates to a range of between -165 and -360 kg ha^{-1} fertilizer under-application over the 20-year simulated period.

Table 4—15 Comparative analysis of average fertilizer input under the four fertilizer subsidy scenarios

			Subsidy scenarios (SC = current (28%), SR= reduced to 15% as per trend, SZ= reduced to zero, SU = universal increase to 70%)					
Indicator	Group	statistic	SR - SC	SZ - SR	SZ - SC	SU - SC	SU - SR	SU - SZ
Fertilizer	Farmtype1	contrast	1.71	2.30	4.01	-8.25	-9.96	-12.26
(kg ha^{-1} yr^{-1})		sig.	1.000	0.632	0.033	0.000	0.000	0.000
	Farmtype2	contrast	2.38	5.67	8.05	-16.52	-18.90	-24.57
		sig.	1.000	0.102	0.005	0.000	0.000	0.000
	Whole	contrast	3.46	5.84	9.30	-18.01	-21.47	-27.32
		sig.	1.000	1.000	0.672	0.016	0.002	0.000
Manure	Farmtype2	contrast	0.31	-1.89	-1.58	15.69	15.39	17.28
(kg ha^{-1} yr^{-1})		sig.	1.000	1.000	1.000	0.000	0.000	0.000
	Farmtype3	contrast	12.34	14.07	26.41	-31.09	-43.44	-57.50
		sig.	1.000	1.000	0.083	0.024	0.001	0.000

The alternative policies established following the declining subsidy trends in previous years to an average of 20% (SR-SC in Table 4—15) does not induce significant shifts in fertilizer input levels while further reduction to zero% i.e. full market price (SZ-SC in Table 4—15) significantly raises the fertilizer input by 4, 8 and 9 kg ha^{-1} yr^{-1} for low input farms, medium input farms and for the whole population, respectively. This translates to a range of between 80 to 180 kg ha^{-1} additional fertilizer applied. Considering the two polar policies of either increasing or decreasing fertilizer subsidy (SU-SZ in Table 4—15), the respective fertilizer differences over the 20-year simulation period would be around 245, 480, and 546 kg ha^{-1} for low input farms, medium input farms and for the whole population

Increasing fertilizer subsidy to 70% has a significant positive effect on manuring for medium manure and high fertilizer farms (*farmtype2*) but a negative effect for high fertilizer and high manure farms (*farmtype3*). *Farmtype2* is characterised by higher fertilizer input and less manuring, and tend to increase manuring with increase in subsidy. On the other hand, farmers of *farmtype3* apply comparatively moderate fertilizer plus more manure, and it has been observed that increasing subsidy reduces the likelihood of manuring. Yet, manure input is still lower than the levels needed to supply nutrients (Chilimba et al., 2005) and far below the requirements to rejuvenate soils (Zingore et al., 2011).

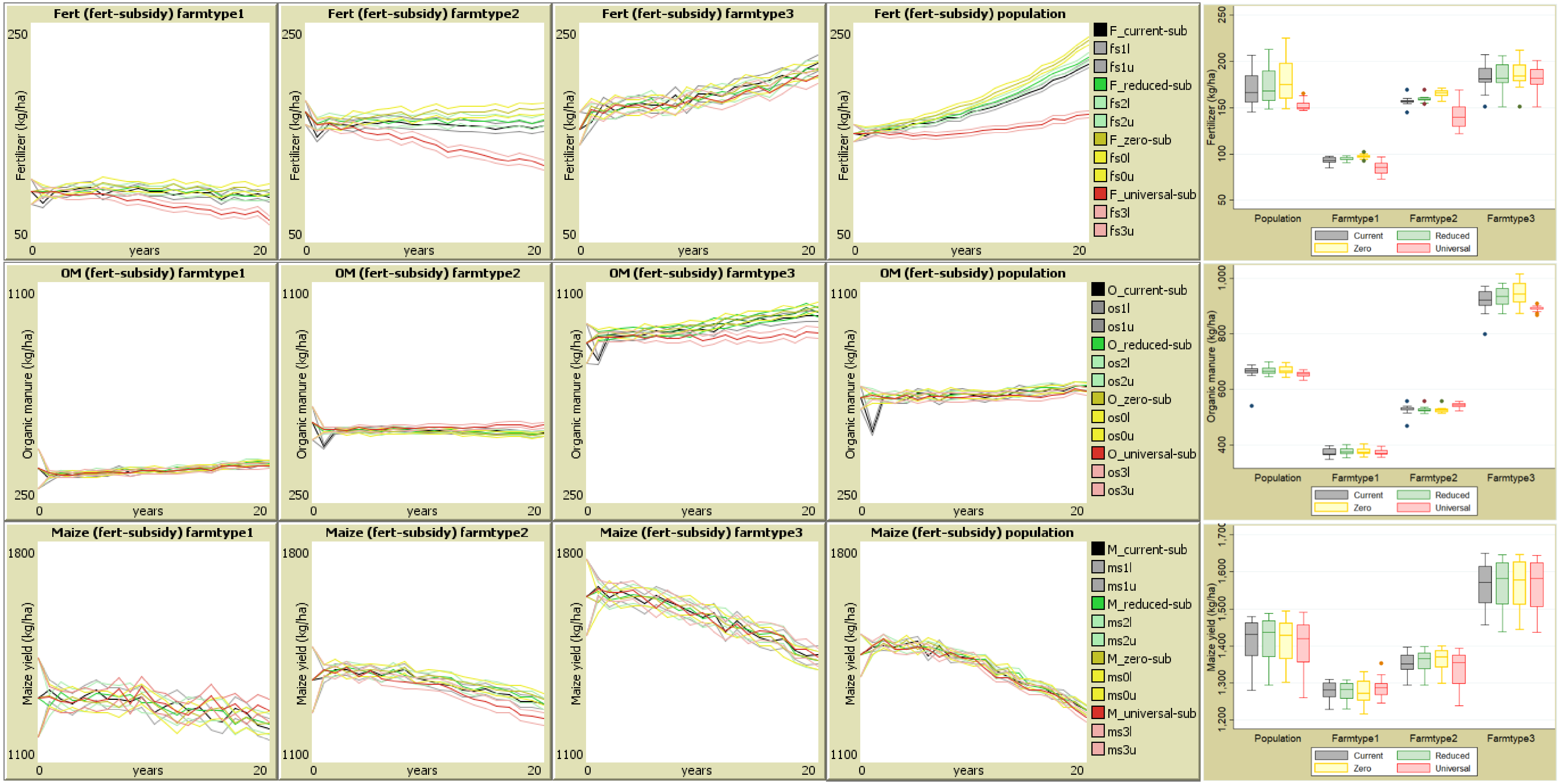


Figure 4.7: Yearly effects of subsidy regimes on nutrient flows.

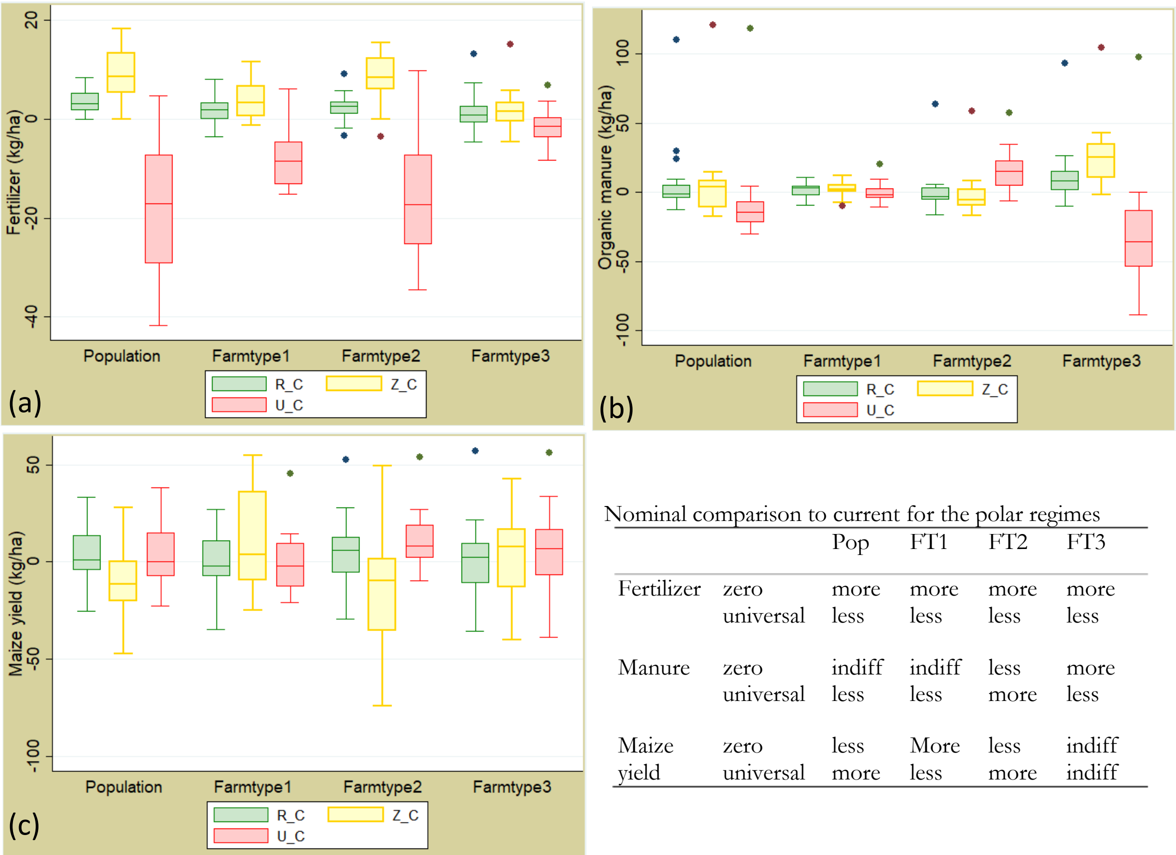
NB: input 1 (inorganic fertilizer- upper panel), input 2 (organic manure, OM – middle panel) and output 1 (maize yield – lower panel) for the entire population (second panel from right), and the farm types 1, 2 & 3. The darker middle lines are mean estimates for subsidy scenarios (_current-sub = current (28%), _reduced-sub= reduced to 15% as per trend, _zero-sub= reduced to zero, _universal-sub = universal increase to 70%). The lighter lines are 5% (_l) and 95% (_u) confidence intervals. The appended box plots on the right shows the median and range of average predictions over the 20-year period.

These results are in sharp contrast to the policy expectations. They also depart from existing knowledge established by earlier studies that subsidy induced increases in fertilizer input and economic benefits (Arndt et al., 2016; Holden & Lunduka, 2013; Komarek et al., 2017). The recent study by Komarek et al. (2017) used the prevailing prices and total nitrogen fertilizer in two scenarios: *full subsidy*, in which case all farmers are assumed to pay zero price, and *zero subsidy* where farmers pay market price. Unlike Komarek's study, the scenarios herein are set based on the assumption that the increases or reductions will be gradual and variable into the foreseeable future, and also contingent on other progressively changing variables being updated as well. Hence, the baseline is set as a trajectory given the prevailing conditions. Worth noting is that during the simulated period, there is a possibility for sudden subsidy regime shifts, which smoothens out over time as has been the case with the previous inducements (Figure 1.1). This corresponds to the findings by Ricker-Gilber and Jayne (2017) who used yearly panel data for 2003 to 2010 to estimate temporal effects of subsidy. They found that during these eight years, farmers were conditioned to purchase more fertilizer after three consecutive subsidies. The subsidy initially aimed at inducing farmers to use improved input sources such as inorganic fertilizer. We constructed a subsidy continuum from zero to universal (set at 70%), and estimate the empirical relationship between level of subsidy and fertilizer use. This study, done 15 years after the onset of the current subsidy regime, points to the likelihood that after prolonged exposure to subsidy, some farmers are increasingly attached and become reliant on the program for their fertilizer demand.

Unfortunately, the distribution and acquisition of subsidised fertilizer has been ambiguous. Instead of benefiting the targeted group of the ultra-poor, subsidy has largely benefitted the traders leaving farmers in despair and uncertainty (Holden & Lunduka, 2013). The ambiguity induced by subsidy has been an antecedent for proactive behaviours within the communities. Farmers, in their pursuit of farming as a major livelihood enterprise, develop agentic capabilities and strive to reduce ambiguity (Grant & Ashford, 2008). While some farmers exclusively rely on subsidy, some lie in between the spectrum with varying levels of subsidy and own purchase, while others proactively acquire marketed fertilizer.

The results are incriminatory that those without financial resources and reliance on subsidy, acquire comparatively less fertilizer. Hence, they proactively supplement with locally available nutrient sources and increasingly engage in manuring. The proactive behaviours enable farmers along the spectrum to predict, understand and determine how to influence their farming environment in advance. Both, those purchasing fertiliser on the market and those who increasingly use manure, accumulate the respective finances and materials for making manure before the growing season. Since fertilizer subsidy distribution and supply ambiguity still persists, it is assumed that farmers are likely to engage in further proactivity in their continued efforts to reduce uncertainty.

The effects of subsidy regimes on manure input are not as stark. Reducing subsidy to zero leads to mean average differences of around 26 kg ha^{-1} yr^{-1} of manure for the high input farms of *farmtype3*. Increasing subsidy to 70% has contrasting effects: positive for medium input farms of *farmtype2* but negative for the relatively higher manure input *farmtype3*. Increasing subsidy would see medium input farms increasing their manure input to an average of 314 kg ha⁻¹ but this would move the average downwards by 620 kg ha⁻¹ for high input farms. The difference between the polar subsidy regimes are projected to be 328 kg ha⁻¹ in favour of universal subsidy for medium input farms while the same would result in average -1,150 kg ha⁻¹ less manure compared to reducing subsidy to zero for high input farms.



Nominal comparison to current for the polar regimes

		Pop	FT1	FT2	FT3
Fertilizer	zero universal	more less	more less	more less	more less
Manure	zero universal	indiff less	indiff less	less more	more less
Maize yield	zero universal	less more	More less	less more	indiff indiff

Figure 4.8: Twenty-year moving average marginal differences between the baseline subsidy (current) and alternatives (reduce, universal and zero) for (a) fertilizer, (b) organic manure and (c) maize yield. **NB**: The estimates are made for each *farmtype (1,2&3)* and for the whole population. The inner line represents median, the box is the 25th and 75th quantiles and whiskers show the boundary for outliers. The inset table describes in brief the deviation from current subsidy as either more, less or indifferent.

Although the Bonferroni multiple comparison test showed that there are statistically insignificant deviations between the current subsidy regime and policy alternatives in terms maize yield (Figure 4.8.c), noticeable differences exist. As shown in Figure 4.8, despite the median differences being close to zero, some scenarios shift heavily on the positive side with the simulated results being higher for the alternatives than the current regime. Also, there exist scenarios such as universal vs current for maize yield among medium input farms of *farmtype2*, the median is below zero indicating that, though not significant, the increase in subsidy slightly shift the maize yield downwards. On the contrary, the maize yields associated with increase in subsidy tend to be positive among low input farms of *farmytpe1* but quite indifferent among high input farms of *farmtype3*.

4.7.2 Subsidy impacts on balance of soil organic carbon and major nutrients (NPK).

The effects of increasing or decreasing subsidy regimes on nutrient input and output flows have varying implications on nutrient and SOC balance. Given the current subsidy regime and interannual changes in dynamic variables, there is a general increasing trend for NPK balances but a decreasing trend for SOC balances (Figure 4.9). The Bonferroni multiple comparison test showed that there are significant differences in the resulting partial N and P balances. The negative impacts on N and P balances from increasing subsidy emanate from the negative effects on fertilization, while the positive shift on K balances among medium input farms is associated with the positive effect on manuring. This underscores the significance of fertilizer as main source of N and P while organic manure is the main source of K (Chilimba et al., 2005). The subsidy policies therefore have potential to impact N and P stocks through fertilization and K stocks through manuring.

The average annual N losses that could be associated with increasing subsidy would be around 7.5 kg ha^{-1} yr^{-1} for medium input farms, which translates to 150 kg ha^{-1} over the simulation period. Increasing the subsidy to universal shifts the average N balance slightly positive with differences from the baseline of 35 and 79 kg ha^{-1} over the simulation period for low input and medium input farms, respectively. The N balances projected for the two polar subsidy regimes favour subsidy removal over universal subsidy with differences of 121, 230 and 259 kg ha^{-1} for low input farms, moderate input farms and the entire population, respectively. Similarly, the P balances are significantly and positively associated with subsidy reduction. On the other hand, universal subsidy leads to a downward shift in P balances for low and medium input farms and for the whole population.

Table 4—16 Bonferroni comparison of N and P balances among alternative fertilizer subsidy policies

			Subsidy scenarios (SC = current (28%), SR= reduced to 15% as per trend, SZ= reduced to zero, SU = universal increased to 70%)					
Indicator	Group	statistic	SR vs. SC	SZ vs. SR	SZ vs. SC	SU vs. SC	SU vs. SR	SU vs. SZ
Nitrogen Balance (kg ha^{-1} yr^{-1})	Farmtype1	contrast	0.58	-4.91	-4.32	1.75	1.16	6.07
		sig.	1.000	1.000	1.000	0.045	0.429	0.000
	Farmtype2	contrast	1.27	-8.79	-7.52	3.95	2.69	11.48
		sig.	1.000	0.000	0.000	0.006	0.13	0.000
	Population	contrast	1.63	10.22	-8.60	4.38	2.75	12.98
		sig.	1.000	0.030	0.104	1.000	0.000	0.003
Phosphorus Balance (kg ha^{-1} yr^{-1})	Farmtype1	contrast	0.06	0.15	0.20	-0.42	-0.47	-0.62
		sig.	1.000	0.456	0.080	0.000	0.000	0.000
	Farmtype2	contrast	0.10	0.28	0.39	-0.77	-0.87	-1.16
		sig.	1.000	0.021	0.001	0.000	0.000	0.000
	Population	contrast	0.18	0.32	0.50	-0.94	-1.12	-1.44
		sig.	1.000	1.000	1.000	0.078	0.020	0.001

As for the partial K balance, which is calculated as the difference between organic manure input minus the K in crop produce and the removed residues, statistical differences could not be detected for the shifts from current subsidy regime but the absolute values are quite variable (Figure 4.10). The whole population analysis shows that decreasing subsidy is associated with a negative absolute mean K balance. However, the difference between universal subsidy and reduced subsidy regimes tend to be negative. For low input farms of *farmtype1*, the average K balance is indeterminant - negative for both alternative subsidy regimes. Whereas an inverse association is observed among medium input farms and high input farms, increasing subsidy is associated with an absolute increase among medium input farms but a decrease among high input farms. Similarly, decreasing subsidy is associated with an absolute decrease in K balance for *farmtype2* but an increase for *farmtype3*. These trends follow the type-specific effects of subsidy on manuring.

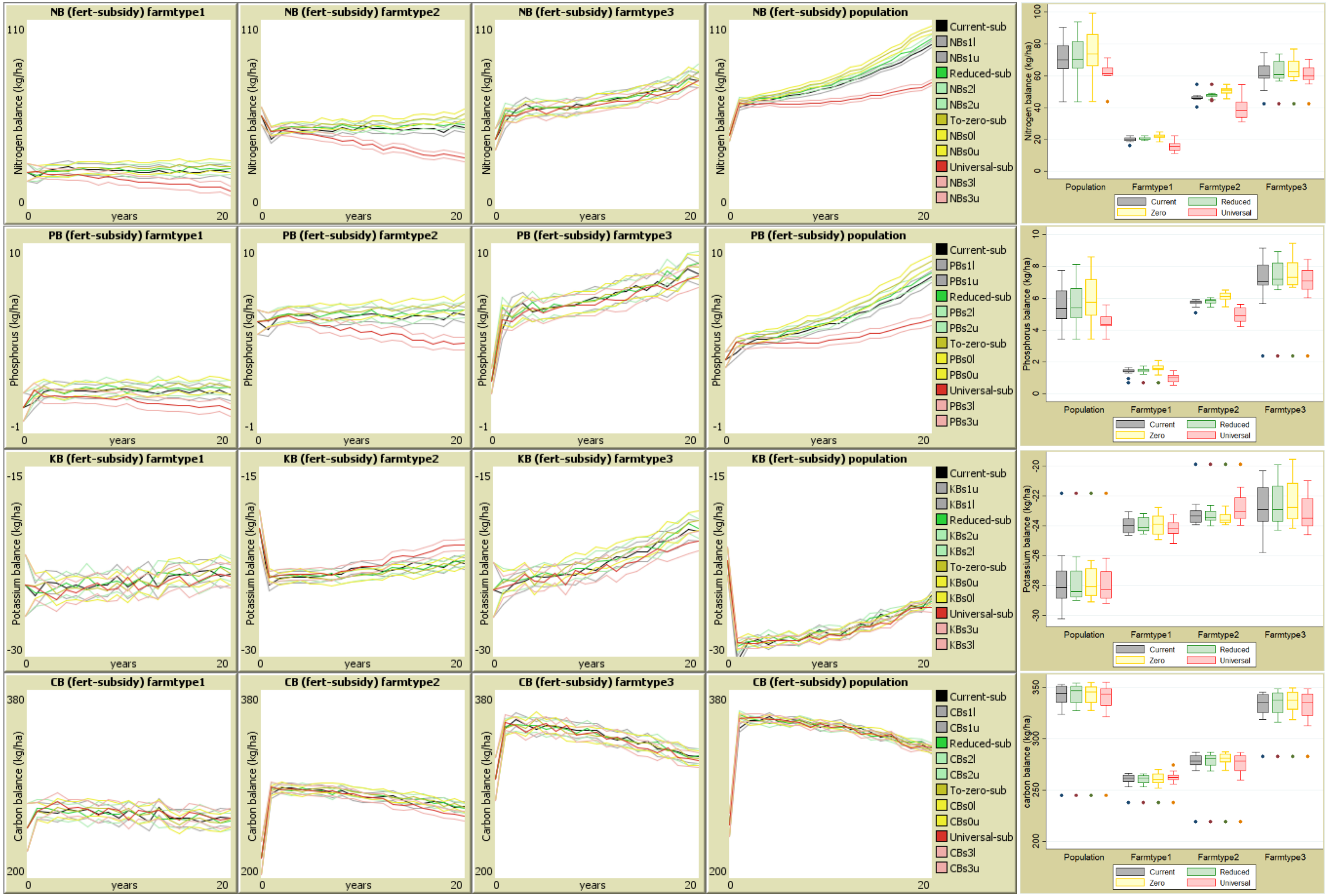


Figure 4.9 Interannual effects of subsidy regimes on partial nutrient balances.

NB: Nitrogen- upper panel, phosphorus – second, Potassium – third and carbon – lower panel: for farm types 1, 2 & 3 and the entire population. The darker middle lines are mean estimates and the lighter lines are confidence intervals. The appended box plots on the right shows the median and range of average predictions over the 20-year period.

Although the multiple comparison test could not detect the impacts of subsidy regimes on C balance the absolute values show an association with residue biomass yield. Increasing subsidy tend to reduce, in absolute terms, the C balance for *farmtype2*, *farmtype3* and the whole population but with higher tendency to be positive among low input farms of *farmtype1* (Figure 4.10a).

The effects of subsidy on the input and output flows given the current and potential subsidy regimes tend to be inelastic and where significant, the associated changes are subtly small. This has substantial implications for nutrient build up or depletion over medium to long term as the country continue to rely on fertilizer subsidy. From our results, it is evident that the current N and P inputs offsets the output flows for more than half of the farms. Therefore, harnessing the positive effects of subsidy on inducing farmers to take on fertilization or manuring has the potential to ensure N and P build up that is essential for the sustainability of the farms.

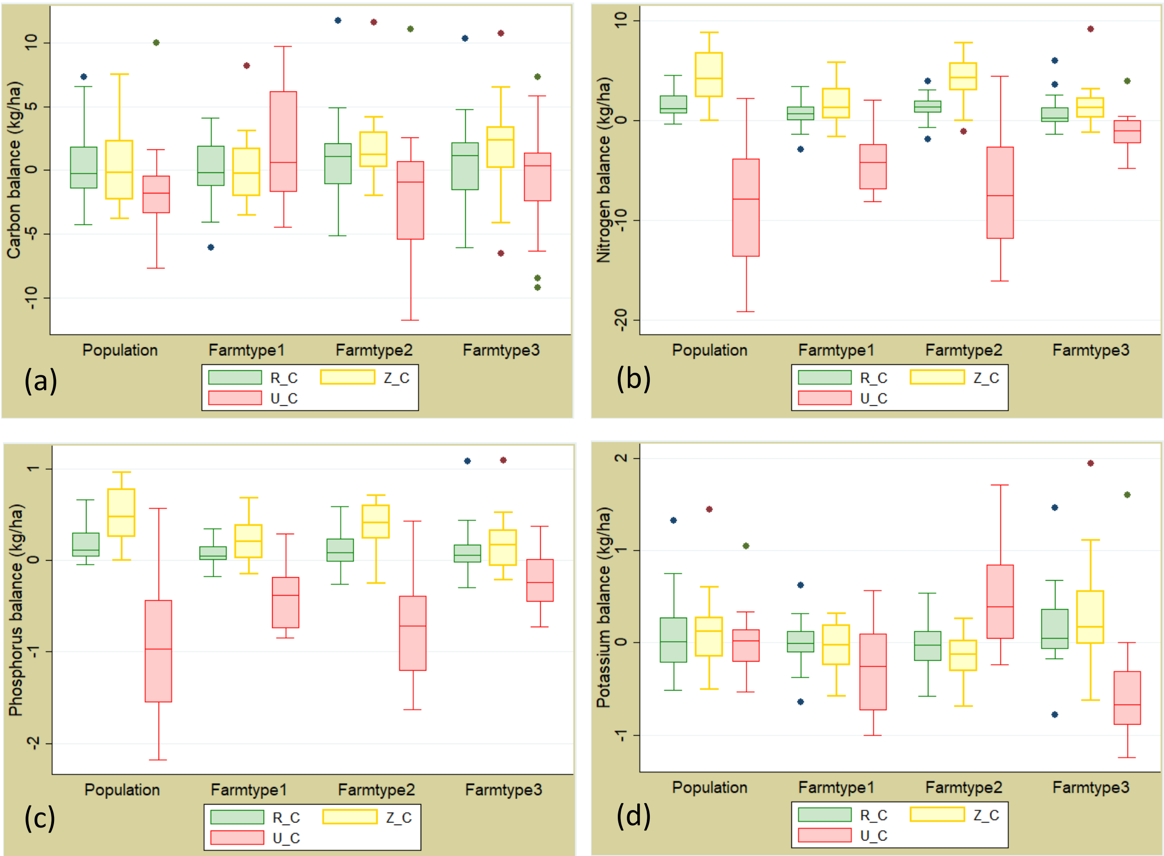


Figure 4.10 Yearly moving average marginal differences between the baseline subsidy (current) and alternatives (reduce, universal and zero) for balances of (a) carbon, (b) nitrogen and (c) phosphorus and (d) potassium.
NB: The estimates are made for each *farmtype (1,2&3)* and for the whole population. The inner line represents median, the box is the 25th and 75th quantiles and whiskers show the boundary for outliers. The table describes in brief the deviation current subsidy.

Although fertilizer recommendations were established based on national wide response trials that were set up in 1975 (Mutegi et al., 2015), farmers continued to apply varying levels due to past experience, economic instances, soil productivity and nonenforcement of policies. Five decades ago, the soils were naturally productive with 86% of the farms realising economical yields without fertilizer input because soils had sufficient P and K (Mutegi et al., 2015; Snapp, 1998). But as farming was intensified and transitioned from subsistence to market orientation in 1980s, nutrient mining was observed. The government using early results of the trial then, recommended the blanket application of 93 kg N and 40 kg P_2O_5 (87 kg ha^{-1} DAP and 175 kg ha^{-1} UREA) (Snapp, 1998). In the 1990s it was noted that at full blanketly recommended rate, 40% of the farms could not recover the cost of fertilizer (Benson, 1996). Hence, using trial data generated over a decade, production orientation (subsistence or market) and agroecological zone specific recommendations were established (Benson, 1999). Reduced rates of 69 N 21 P2O5 or 35 N and 10 P2O5 were set up for the study region that required application of 100 or 50 kg NPS fertilizer (23:21:0+4s) and 100 or 50 kg UREA, amounting to 200 or 100 kg ha^{-1} yr^{-1} of NP fertilizer. In the 2000s, with increasing recognition of leguminous nitrogen source, the lower rate of 35 N and 10 P_2O_5 was adopted and blanketly promoted. However, the recent revelation that soil nitrogen is critically low (< 0.15%), area specific recommendation for the study region has been set at 92 Kg N, 10 kg ha^{-1} P_2O_5 (4.3 kg P) for soils with moderate extractable P and 10 kg ha^{-1} K_2O (0.83 K) for soils with moderate extractable K (Mutegi et al., 2015). These recommendations are set without considering the supply of NPK from organic manure despite being the major source of K and P in Malawi and the greater part of tropical sub-Sahara Africa (Chilimba et al., 2005; Palm, Giller, et al., 2001).

Nonetheless, the usage of organic manure that is rich in K and SOC is below the levels required to counter large nutrient output through natural ways. Hence, there is need for increased manuring which from our empirical evidence could not be achieved with the subsidy policy alone. This is worrisome as studies have found a decadal decreasing trend in SOC (Mpeketula, 2016) and general decline of land productivity (Messina et al., 2017) within Malawian farmlands. With the less than optimal SOC levels currently available in the study region (section 4.2) and largely negative balances recorded under full balance analysis (especially if ecological flows are accounted for as indicated in Table 4—14), their sustainability might be abridged. The current manure and residue inputs of <1 ton ha^{-1} and <5 ton ha^{-1} are below the 8 to 10 ton ha^{-1} annual organic inputs required to bring the soils to levels essential for fertilization response and structural stability (Musinguzi et al., 2013, 2016). Manure usage ought to be increased by 8-folds to 5 and 12-folds to 8 ton ha^{-1} yr^{-1} to supply the sufficient levels of K and P (Chilimba et al., 2005) and of SOC (Zingore et al., 2011) required for optimum maize productivity, respectively. These estimations are based on average values, but there are few farms that supply 5-ton ha^{-1} of manure and could be self-reliant in terms of P and K, and potentially supply the SOC needed for structural stability and long-term productivity of the farms. The individual farm states and transitions over the simulation period is presented in appendix S5. From these distribution maps, farms that are projected to experience significant positive shifts and the hotspots with extremely low input, productivity and nutrient balances can be identified. A detailed analysis through typical farm analysis (Feuz & Skold, 1992) is needed to understand success accelerators for positive deviants and constraining factors for laggards.

5 CONCLUSION

We have found that within the fragile landscapes of Malawi, farmers strive to utilise the commonly available soil fertility management options as evidenced by wide but low usage. Nine in every ten households used inorganic fertilizers, a third planted legumes and almost half applied manures of various forms. From the empirical and simulated results, it is indicative that the maize mixed smallholder farming system in Malawi has become inelastic to changes in input policies. Despite the changes in subsidy, the total amount of fertilizer applied has levelled off (Figure 1.1).

The evaluation of proximate and underlying drivers of choice and intensification of inorganic fertilizers, organic manures and legumes has shed light on the diffusion pathways and limitations. There exist some drivers with similar effects on both decisions, but no single driver is consistently associated with both the choice and intensification decisions of these three input flows. The trade-offs of factor effects on the two decisions for each technology and among the three technologies can guide formulation of targeted research and development programs for both non-adopters and adopters.

Apparently, formal education continues to be a main factor influencing farmers to grow legumes and increase usage of organic manures. Additionally, increases in number of basic hand-held farm implements reduce drudgery and increases the probability to diversify cropping (e.g. by planting legumes) and making it easier to prepare and apply manure. As a major source of manure, increasing efforts to integrate livestock in these cereal dominated systems could lead to greater manure application. The study has revealed that as much as improvements in contribution of women in decision-making widens the scope for legume cropping, it could negatively affect manuring. With the notable opposing effects from women empowerment, addressing challenges that women face in manuring could offer greater opportunities for ISFM.

Though soil management in smallholder farming systems aims at addressing the most critical nutrient(s), the results from this study show that the soils in some areas are deficient in all three major nutrients (N, P, K) and SOC. Bringing these nutrient deficient soils to productivity, therefore, requires to first raise their levels to normal range. Bearing in mind the inter-dependence as expressed by stoichiometry, continuous monitoring and adjustments need to be made to create optimal conditions for plant uptake. The estimation was done at 10x10m pixel size, and enabled to predict within farm differences. For smallholder farmers with limited nutrient sources and in little amounts, a more judicious option would be to have knowledge of soil nutrient gradients within farms and target the amendments to hot spots. Considering the limited capacity on both demand for knowledge of soils by smallholder farmers and the supply of information by agricultural extension services, there is need to raise awareness, capitalise on digital tools to disseminate pixel-based soil information and input these into site-specific nutrient balance and crop yield models.

Subsidy as a major soil nutrient management strategy was viewed by many as a panacea for productivity improvement, a replica of Asia's Green Revolution (Denning et al., 2009). The initial impact of subsidy seem to be waning: the fertilizer input levels are still below optimal and the mean yields are far below the optimal or locally attainable yields (Tamene, Mponela, Ndengu, et al., 2016). Unfortunately, the government is complacent at ending hunger and continue to use national maize requirement and not the potential and attainable yields to set national policy agenda. Consequently, farmers' actions have been reactive rather than proactive. Previous studies constructed histories and found, *ex post*, positive effects of subsidy on maize yields during the initial period (Ricker-Gilbert & Jayne, 2017). This study highlights the current trends that maize yield has become inelastic to changes among the prevailing interventions.

We use behavioural economics and explore, *ex ante*, the agentic behaviours of farmers when faced with ambiguity in fertilizer acquisition and maize yield. The study determined that those relying heavily on subsidy are less likely to be caught in a nutrient depletion trap, they rather resort to measures to address their deficiency. Moving out of the nutrient depletion traps depend on harnessing complementarity and substitutability of the various nutrient input sources. Previous studies and anecdotal observations did not seem to focus on the very processes that transform subsidy to welfare and ecological health. To address this, we used the proportion of fertilizer subsidised to explore the complementarity and substitutability between fertilization and manuring. Our results suggest that subsidy does not induce farmers to substantially increase fertilization. The plausible explanation is that in the 15 years subsidy has been implemented, farmers internalised it in their fertilizer expenditure plan, some exclusively relying on subsidy while others sourcing increasing amounts from the market and are becoming self-reliant. Those that rely on limited fertilizer acquired through subsidy also engage in proactive behaviours to reduce the nutrient gap by increasingly investing in manuring.

5.1 Limitations and areas for further research

The aim of the study is to build a SES that spatially and temporally explicitly capture, analyse, and present soil nutrient balances and explore, *ex ante,* possible livelihood and ecological outcomes from alternative soil management practices to better inform smallholder farmers and other stakeholders whilst making their sustainability decisions. This goal sets out the scope of work and this study has managed to explore, analyse and present soil nutrient balances and explore alternative soil management regimes, but the feedback mechanisms to inform policy making are yet to be fully established as indicated below.

1. The parameterisation and the empirical models required for establishing the nutrient stocks and input and output flows are voluminous. Efforts were made to extensively review the literature for parameters, and a detailed survey data from representative farms have been used to customize the estimates for the study site which could be scaled up for the entire maize mixed farming system and beyond. Much as we have tried to improve the estimations by using parameters and transfer functions from Malawi as well as from neighbouring countries in east and southern Africa, there are still cases where parameters established in other regions

and generalised transfer functions have been used. As research in the region progresses, there is need to update the parameters and functions.

2. To evaluate the impact of subsidy, most of the household attributes, including labour, women empowerment, household resource endowments as well as plot attributes such as sizes, were not altered or updated. The village population is also assumed to remain constant which, if updated would affect labour, dependency ratios and education attainments. In this model, a sample of senior village citizens die, and their household and land are inherited by offspring. Although their inclusion ensured that most of the causes of variations were captured and that the simulations represented the effects of subsidy, their constancy does not reflect the real farm situation over a 20-year simulation period. In the current form, stochasticity in these variables is captured by random assignment of their coefficient in the SFM choice and production models. Further research should therefore build rules for updating these and test their effects by having different scenarios. In this thesis, we have provided the rules for updating labour and gender which are being incorporated in the model.
3. The modelling environment used, Netlogo, although widely used by ecological and social scientists to model human induced emergent phenomena, its computing capabilities for these small-sized heterogenous farms is limited. There is need to improve the model speed by implementing it in R or python, acquire computers with higher computing capabilities or use cloud computing to extend the model for application to larger landscapes and for other integrated socioecological processes.
4. Instead of running simulations spatially explicit with overlay of household and policy attributes, the centroids of plots were used to estimate the policy implications on human behaviour. This was linked to the respective patch landscape processes and upscaled to other patches/pixels regardless of their varied ecological features. Hence, the variations within plots were smoothened-out.
5. Related to model requirements are data. To fully build a functional real-life representative model, data requirements are huge. Considering inherent errors associated with data capturing in these data scarce regions, triangulation of multiple data sources is required in the short term and efforts to improve data acquisition through approaches such as Land-Potential Knowledge System (LandPKS) (Herrick et al., 2013) need to be intensified and integrated.
6. At the time of the compilation of this thesis, the effects of policy option such as subsidy is limited to productivity and its impacts on nutrient balance and profitability as ecological and economic sustainability indicators for a micro sentinel landscape in Malawi, and for mainly two crops: maize and ground nuts. Several other crops are integrated in farming landscapes and also the different non-cultivated patches have varied parameters. Efforts are ongoing to extend the model to unravel the effect of changing subsidy regimes on social sustainability indicators such as benefit distribution and vulnerability. Impacts of social policies such as gender empowerment are also being explored.
7. In its current form, the model is yet to present the changes in household livelihood portfolios due to their choices of technologies and resulting nutrient balances. The household profiles (farm types) are differentiated by input levels in the sample data. Although the plot level nutrient inputs are updated (using choice and

intensify decisions) they have not been used to update farm types. The trajectories are based on initial household profiles. Hence there is need to update the farm types. The feedback is limited to updates in fertilizer and manure inputs, legume cropping (maize plots are not rotated nor fallowed), and soil nutrient and SOC stocks. Hence, there is need to relax assumptions on household attributes by updating some idiosyncratic factors such as incomes, other crop types, fallowing (in some marginal areas as productive areas are continuously cropped), household sizes, land sizes and household profiles.

8. The study also calculated a compound productivity index that has been used to locate the remaining unsampled plots to patches with similar productivity. With this calculation, it would be interesting to develop ecological profiles and compare sustainability among the low, mid and high productive areas. A further combined index with household clusters is possible and would enable development of comprehensive social-ecological profiles that could be used to unravel the differences among low-input low-productive; high-input low-productive; low-input high-productive, and high-input high-productive integrated farm profiles. With a larger dataset, such analyses would be more robust.
9. Based on the findings and limitations of this research, a further project idea is to extend the model by using up-to-date datasets that cover a wider region such as that of Africa RISING for four sentinel sites representing different agroecological zones or the georeferenced Living Standards Measurement Study – integrated Household Survey (LSMS-IHS). With these and other datasets, a robust and comprehensive model with multiple approaches can be developed to assess SOC, nutrient dynamics and land productivity in Malawi and other smallholder farming systems in Africa.

REFERENCES

Adolwa, I. S., Schwarze, S., Waswa, B., & Buerkert, A. (2017). Understanding system innovation adoption: A comparative analysis of integrated soil fertility management uptake in Tamale (Ghana) and Kakamega (Kenya). *Renewable Agriculture and Food Systems*, 1–13. https://doi.org/10.1017/S1742170517000485

African Institute of Corporate Citizenship. (2014). *Harmonized Groundnut Production Manual for Malawi* (Issue April).

Akinola, A. A., Alene D., A., Adeyemo, R., Sanogo, D., Olanrewaju, A. S., Nwoke, C., & Nziguheba, G. (2010). Determinants of adoption and intensity of use of balance nutrient management systems technologies in the northern Guinea savanna of Nigeria. *Quarterly Journal of International Agriculture*, *49*(1). https://www.agrar.hu-berlin.de/de/institut/departments/daoe/publ/qjia/contents/2010/1-10/Akinola

Al-kaisi, M. (2008). *Soil Erosion , Crop Productivity and Cultural Practices*. *September*, 1–4. https://store.extension.iastate.edu/Product/Soil-Erosion-Crop-Productivity-and-Cultural-Practices-PDF

Alcamo, J., Bennett, E. M., & (Program), M. E. A. (2003). *Ecosystems and Human Well-being: A Framework for Assessment*. Island Press. https://books.google.de/books?id=oJGLgfndzoYC

Alkire, S., Mainzem-Dick S., R., Peterman, A., Quisumbing R., A., Seymour, G., & Vaz, A. (2012). *The Women's Empowerment in Agriculture Index.* International Food Policy Research Institute. http://ebrary.ifpri.org/cdm/ref/collection/p15738coll2/id/127557

Alori, E. T., Glick, B. R., & Babalola, O. O. (2017). Microbial Phosphorus Solubilization and Its Potential for Use in Sustainable Agriculture. *Frontiers in Microbiology*, *8*. https://doi.org/10.3389/fmicb.2017.00971

Amede, T., Tamene, L. D., Harris, D., Kizito, F., & Xueliang, C. (2014). *The Chinyanja Triangle in the Zambezi River Basin, southern Africa: status of, and prospects for, agriculture, natural resources management and rural development*. International Water Management Institute (IWMI). http://www.iwmi.cgiar.org/Publications/wle/r4d/wle_research_for_development-learning_series-1.pdf

Anderson, E. L. (1988). Tillage and N fertilization effects on maize root growth and root:shoot ratio. *Plant and Soil*, *108*(2), 245–251. https://doi.org/10.1007/BF02375655

Anseeuw, W., Jayne, T., Kachule, R., & Kotsopoulos, J. (2016). The Quiet Rise of Medium-Scale Farms in Malawi. *Land*, *5*(3), 19. https://doi.org/10.3390/land5030019

Arndt, C., Pauw, K., & Thurlow, J. (2016). The Economy-wide Impacts and Risks of Malawi's Farm Input Subsidy Program. *American Journal of Agricultural Economics*, *98*(3), 962–980. https://doi.org/10.1093/ajae/aav048

Arneth, A., Barbosa, H., Benton, T., Calvin, K., Calvo, E., Connors, S., Cowie, A., Davin, E., Denton, F., van Diemen, R., Driouech, F., Elbehri, A., Evans, J., Ferrat, M., Harold, J., Haughey, E., Herrero, M., House, J., Howden, M., … Zommers, Z.

(2019). Climate Change and Land: Summary for Policymakers. *An IPCC Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse Gas Fluxes in Terrestrial Ecosystems*, 43. https://doi.org/10.4337/9781784710644

Auerswald, K., Fiener, P., Martin, W., & Elhaus, D. (2014). Use and misuse of the K factor equation in soil erosion modeling: An alternative equation for determining USLE nomograph soil erodibility values. *CATENA*, *118*, 220–225. https://doi.org/10.1016/j.catena.2014.01.008

Aune, J. B., & Lal, R. (1997). Agricultural productivity in the tropics and critical limits of properties of Oxisols, Ultisols, and Alfisols. *Tropical Agriculture*, *2*(74), 96–103. https://journals.sta.uwi.edu/ta/index.asp?action=viewPastAbstract&articleId=2488&issueId=251

Austin, P. C., & Steyerberg, E. W. (2015). The number of subjects per variable required in linear regression analyses. *Journal of Clinical Epidemiology*, *68*(6), 627–636. https://doi.org/10.1016/j.jclinepi.2014.12.014

Ayuke, F. O., Zida, Z., & Lelei, D. (2019). Effects of Soil Management on Aggregation and Organic Matter Dynamics in sub-Saharan Africa. *African Journal of Food, Agriculture, Nutrition and Development*, *19*(01), 13992–14009. https://doi.org/10.18697/ajfand.84.BLFB1002

Badmos, B., Agodzo, S., Villamor, G., & Odai, S. (2015). An Approach for Simulating Soil Loss from an Agro-Ecosystem Using Multi-Agent Simulation: A Case Study for Semi-Arid Ghana. *Land*, *4*(3), 607–626. https://doi.org/10.3390/land4030607

Bado, B. V., Lompo, F., Sedogo, M. P., & Cescas, M. P. (2010). Establishment of the Critical Limit of Soil-Available Phosphorous for Maize Production in Low Acidic Ultisols of West Africa. *Communications in Soil Science and Plant Analysis*, *41*(8), 968–976. https://doi.org/10.1080/00103621003646055

Banda, A. Z., Maghembe, J. A., Ngugi, D. N., Chome, V. A., Banda Z., A., Maghembe A., J., Ngugi N., D., & Chome A., V. (1994). Effect of intercropping maize and closely spaced Leucaena hedgerows on soil conservation and maize yield on a steep slope at Ntcheu, Malawi. *Agroforestry Systems*, *27*(1), 17–22. https://doi.org/10.1007/BF00704831

Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory*. Prentice Hall.

Barrett, C. B., & Bevis, L. E. M. (2015). The self-reinforcing feedback between low soil fertility and chronic poverty. *Nature Geoscience*, *8*(12), 907–912. https://doi.org/10.1038/ngeo2591

Beedy, T. L., Ajayi, O. C., Sileshi, G. W., Kundhlande, G., Chiundu, G., & Simons, A. J. (2013). Scaling up Agroforestry to Achieve Food Security and Environmental Protection among Smallholder Farmers in Malawi. *Field Actions Science Reports. The Journal of Field Actions*, *Special Issue 7*. https://factsreports.revues.org/2082

Beedy, T. L., Snapp, S. S., Akinnifesi, F. K., & Sileshi, G. W. (2010). Impact of Gliricidia sepium intercropping on soil organic matter fractions in a maize-based cropping system. *Agriculture, Ecosystems & Environment*, *138*(3–4), 139–146. https://doi.org/10.1016/j.agee.2010.04.008

Bell, A. R., Robinson, D. T., Malik, A., & Dewal, S. (2015). Modular ABM development for improved dissemination and training. *Environmental Modelling & Software*, *73*, 189–200. https://doi.org/10.1016/j.envsoft.2015.07.016

Benson, T. D. (1999). Area-specific fertilizer recommendations for hybrid maize grown by Malawian smallholders. In *Ministry of Agriculture and Irrigation*.

Benson, T. D. (1996). A geographical and economic analysis of past nitrogen response trial results for maize in Malawi. In J. K. Ransom, A. F. E. Palmer, B. T. Zambezi, Z. O. Mduruma, S. R. Waddington, K. V. Pixley, & D. C. Jewell (Eds.), *Maize productivity gains through research and technology dissemination* (pp. 191–196). CIMMYT. https://repository.cimmyt.org/xmlui/bitstream/handle/10883/746/64589.pdf?sequence=1

Benson, T. D., Mabiso, A., & Nankhuni, F. (2016). Detailed crop suitability maps and an agricultural zonation scheme for Malawi: Spatial information for agricultural planning purposes. In *FOOD SECURITY POLICY INNOVATION LAB*. International Food Policy Research Institute. http://ebrary.ifpri.org/cdm/ref/collection/p15738coll2/id/130494

Berger, T. (2001). Agent-based spatial models applied to agriculture: a simulation tool for technology diffusion, resource use changes and policy analysis. *Agricultural Economics*, *25*(2–3), 245–260. https://doi.org/10.1016/S0169-5150(01)00082-2

Berger, T., & Schreinemachers, P. (2006). Creating Agents and Landscapes for Multiagent Systems from Random Samples. *Ecology and Society*, *11*(2). https://doi.org/10.5751/ES-01736-110219

Bishop, J. (1995). *The Economics of Soil Degradation: An Illustration of Change in Productivity Approach to Valuation in Mali and Malawi.* (No. 95–02).

Bootsma, H. A., Mwita, J., Mwichande, B., Hecky, R. E., Kihedu, J., & Mwambungu, J. (1999). *The atmospheric deposition of nutrients on Lake Malawi/Nyasa* (H. A. Bootsma & R. E. Hecky (Eds.)).

Boulanger, P.-M., & Bréchet, T. (2005). Models for policy-making in sustainable development: The state of the art and perspectives for research. *Ecological Economics*, *55*(3), 337–350. https://doi.org/10.1016/j.ecolecon.2005.07.033

Bouma, J., Montanarella, L., & Evanylo, G. (2019). The challenge for the soil science community to contribute to the implementation of the <scp>UN</scp> Sustainable Development Goals. *Soil Use and Management*, sum.12518. https://doi.org/10.1111/sum.12518

Braslow, J., & Cordingley, J. (2016). *Participatory mapping in Ntcheu district, Malawi. A case study.* International Center for Tropical Agriculture (CIAT). http://ciat-library.ciat.cgiar.org/Articulos_Ciat/biblioteca/Participatory_mapping_in_Ntcheu_district_Malawi_A_case_Study.pdf

Breetzke, D. G., Koomen, E., & Critchley, W. R. S. (2013). GIS-Assisted Modelling of Soil Erosion in a South African Catchment: Evaluating the USLE and SLEMSA Approach. In R. A. Wurbs (Ed.), *Water Resources Planning, Development and Management* (pp. 53–71). InTech. https://doi.org/10.5772/52314

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brennan, R. F., & Bell, M. J. (2013). Soil potassium—crop response calibration relationships and criteria for field crops grown in Australia. *Crop and Pasture Science*, *64*(5), 514. https://doi.org/10.1071/CP13006

Brown, D. G., Reutlinger, S., & Thomson, A. M. (1996). Malawi: Food Security in a Market Oriented Economy. In *Agricultural policy Analysis Project, Phase III (APAP III)*. University of Florida. http://ufdc.ufl.edu/UF00072240/00001

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*, *114*(9), 2189–2194. https://doi.org/10.1073/pnas.1616919114

Burke, W. J., Snapp, S. S., & Jayne, T. S. (n.d.). An in-depth examination of maize fertilizer response in Central Malawi reveals low profits and too many weeds. *Agricultural Economics*.

Campbell, B. (1996). *The Miombo in transition: woodlands and welfare in Africa*. Center for International Forestry Research (CIFOR). http://www.cifor.org/publications/pdf_files/books/miombo.pdf

Carletto, C., Gourlay, S., & Winters, P. (2015). From Guesstimates to GPStimates: Land Area Measurement and Implications for Agricultural Analysis. *Journal of African Economies*, *24*(5), 593–628. https://doi.org/10.1093/jae/ejv011

Carletto, C., Jolliffe, D., & Banerjee, R. (2015). From Tragedy to Renaissance: Improving Agricultural Data for Better Policies. *The Journal of Development Studies*, *51*(2), 133–148. https://doi.org/10.1080/00220388.2014.968140

Chalmers, G., & Agar, J. (2015). *Portfolio ownership among agribusiness SMES in Malawi*. USAID. https://www.microlinks.org/sites/default/files/resource/files/Malawi_SME_Assessement_Brief_1.13.15_508_Compliant.pdf

Chamberlin, J., & Ricker-Gilbert, J. (2016). Participation in rural land rental markets in Sub-Saharan Africa: Who benefits and by how much? Evidence from Malawi and Zambia. *American Journal of Agricultural Economics*, *98*(5), 1507–1528. https://doi.org/10.1093/ajae/aaw021

Chidumayo, E. N. (2019). Management implications of tree growth patterns in miombo woodlands of Zambia. In *Forest Ecology and Management* (Vol. 436, pp. 105–116). https://doi.org/10.1016/j.foreco.2019.01.018

Chigwiza, J., & Kanazawa, H. (2008). Survey result on Soil Erosion Control Technologies in Malawi. In *Community Vitalization and Afforestation in Middle Shire*. Japan International Cooperation Agency. https://www.jica.go.jp/project/english/malawi/001/materials/pdf/working_paper_02.pdf

Chilimba, A. D. K., Shano, B., Chigowo, M. T., & Komwa, M. K. (2005). *Quality assessment of compost manure produced by smallholder farmers in Malawi*. Department of Agricultural Research Services, Republic of Malawi. http://www.ndr.mw:8080/xmlui/handle/123456789/350

Chilonda, P., & Otte, J. (2006). Indicators to monitor trends in livestock production at national, regional and international levels. *Livestock Research for Rural Development*, *18*(8). http://www.lrrd.cipav.org.co/lrrd18/8/chil18117.htm

Chimtengo, M., Ngongondo, C., Tumbare, M., & Monjerezi, M. (2014). Analysing changes in water availability to assess environmental water requirements in the

Rivirivi River basin, Southern Malawi. *Physics and Chemistry of the Earth, Parts A/B/C*, *67–69*, 202–213. https://doi.org/10.1016/j.pce.2013.10.007

Chirwa, E. W., Dorward, A., & Matita, M. (2011). *Initial Conditions and Changes in Commercial Fertilizers under the Farm Input Subsidy Programme in Malawi: Implications for Graduation*. Future Agricultures. https://www.researchgate.net/publication/236945516_Initial_Conditions_and_Changes_in_Commercial_Fertilizers_under_the_Farm_Input_Subsidy_Programme_in_Malawi_Implications_for_Graduation

Chirwa, E. W., Kumwenda, I., Jumbe, C., Chilonda, P., & Minde, I. (2008). Agricultural Growth and Poverty Reduction in Malawi : Past Performance and Recent Trends. *Development*, *8*, 31. http://www.resakss.org/index.php?pdf=39397

CIAT. (2016, April 18). *Let's Conserve the Environment by finding solutions to end poverty* [Online video]. Youtube; International Center for Tropical Agriculture (CIAT). https://www.youtube.com/watch?v=0EZD5lv_xAQ

Cleveland, C. C., & Liptzin, D. (2007). C:N:P stoichiometry in soil: is there a "Redfield ratio" for the microbial biomass? *Biogeochemistry*, *85*(3), 235–252. https://doi.org/10.1007/s10533-007-9132-0

Cobo, J. G., Dercon, G., & Cadisch, G. (2010). Nutrient balances in African land use systems across different spatial scales: A review of approaches, challenges and progress. *Agriculture, Ecosystems & Environment*, *136*(1–2), 1–15. https://doi.org/10.1016/j.agee.2009.11.006

Coe, R., Njoloma, J., & Sinclair, F. (2016). Loading the dice in favour of the farmer: reducing the risk of adopting agronomic innovations. *Experimental Agriculture*, 1–17. https://doi.org/10.1017/S0014479716000181

Conroy, A., Blackie, M., Whiteside, A., Malewezi, J., & Sachs, J. (2006). *Poverty, AIDS and Hunger: Breaking the Poverty Trap in Malawi*. Springer. https://books.google.de/books?id=BgBaCwAAQBAJ

Contributors. (1944). *Proceedings of the IRE*, *32*(4), 247–249. https://doi.org/10.1109/JRPROC.1944.230664

Cragg, J. G. (1971). Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods. *Econometrica*, *39*(5), 829–844. https://www.jstor.org/stable/1909582

Dalupan, M. C. G., Haywood, C., Wardell, D. A., Cordonnier-Segger, M. C., & Kibugi, R. (2015). *Building enabling legal frameworks for sustainable land-use investments in Zambia, Tanzania and Mozambique: A synthesis*. Center for International Forestry Research (CIFOR). http://www.cifor.org/library/5753/building-enabling-legal-frameworks-for-sustainable-land-use-investments-in-zambia-tanzania-and-mozambique-a-synthesis/

Davies, G. M., Pollard, L., & Mwenda, M. D. (2010). Perceptions of land-degradation, forest restoration and fire management: A case study from Malawi. *Land Degradation and Development*, *21*(6), 546–556. https://doi.org/10.1002/ldr.995

De Jager, A., Nandwa, S. M., Okoth, P. F., Kariuku, I., Matiri, F. M., Odendo, M., Wanyama, J. M., Van den Bosch, H., Gitari, J. ., Ogaro, V. ., Maobe, S., Vlaming, J., De Jager, A., & Vlaming, J. (1998). Monitoring nutrient flows and economic performance in African farming systems (NUTMON). *Agriculture, Ecosystems &*

Environment, *71*(1–3), 63–80. https://doi.org/10.1016/S0167-8809(98)00130-3
Deininger, K., & Xia, F. (2017). *Assessing effects of large scale land transfers: Challenges and opportunities in Malawi's estate sector*. World Bank Group. http://documents.worldbank.org/curated/en/794751505916933601/pdf/WPS8200.pdf
Denning, G., Kabambe, P., Sanchez, P., Malik, A., Flor, R., Harawa, R., Nkhoma, P., Zamba, C., Banda, C., Magombo, C., Keating, M., Wangila, J., & Sachs, J. (2009). Input subsidies to improve smallholder maize productivity in Malawi: toward an African Green Revolution. *PLoS Biology*, *7*(1), e1000023. https://doi.org/10.1371/journal.pbio.1000023
Desmet, P. J. J., & Govers, G. (1996). A GIS procedure for automatically calculating the USLE LS factor on topographically complex landscape units. *Journal of Soil and Water Conservation*, *51*(5), 427–433.
Dixon, J., Gulliver, A., & Gibbon, D. (2001). *Farming systems and poverty: improving farmers' livelihoods in a changing world* (M. Hall (Ed.)). FAO & World Bank. http://www.fao.org/3/a-ac349e.pdf
Dobermann, A., Witt, C., Abdulrachman, S., Gines, H. C., Nagarajan, R., 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., Simbahan, G. C., Adviento, M. A. A., & Bartolome, V. (2003). Estimating Indigenous Nutrient Supplies for Site-Specific Nutrient Management in Irrigated Rice. *Agronomy Journal*, *95*(4), 924. https://doi.org/10.2134/agronj2003.0924
Donkin, E., Dennis, P., Ustalakov, A., Warren, J., & Clare, A. (2017). Replicating complex agent based models, a formidable task. *Environmental Modelling & Software*, *92*, 142–151. https://doi.org/10.1016/j.envsoft.2017.01.020
Dorward, A. (1999). Farm size and productivity in Malawian smallholder agriculture. *Journal of Development Studies*, *35*(5), 141–161. https://doi.org/10.1080/00220389908422595
Dorward, A. (2002). A Typology of Malawian Rural Households. In *Institutions and Economic Policies for Pro-Poor Agricultural Growth*. DFID. http://eprints.soas.ac.uk/id/eprint/8876
Dorward, A., & Chirwa, E. (2011). The Malawi Agricultural Input Subsidy Programme: 2005-6 to 2008-9. *International Journal of Agricultural Sustainability (IJAS)*, *9*(1), 232–247. http://eprints.soas.ac.uk/9598/
Dorward, A., Chirwa, E. W., & Jayne, T. S. (2011). Malawi's Agricultural Input Subsidy Program Experience over 2005–09. In P. Chuhan-Pole & M. Angwafo (Eds.), *Yes Africa can : Success stories from a dynamic continent* (p. 289). World Bank. http://siteresources.worldbank.org/AFRICAEXT/Resources/258643-1271798012256/YAC_chpt_17.pdf
Doss, C. R. (2006). Analyzing technology adoption using microstudies: limitations, challenges, and opportunities for improvement. *Agricultural Economics*, *34*(3), 207–219. https://doi.org/10.1111/j.1574-0864.2006.00119.x
Doss, C. R., & Morris, M. L. (2000). How does gender affect the adoption of agricultural innovations?: The case of improved maize technology in Ghana. *Agricultural Economics*, *25*(1), 27–39. https://doi.org/10.1111/j.1574-

0862.2001.tb00233.x
Douillet, M., & Toulon, A. (2014). *Developing a typology of agricultural holdings for improved policy design: a preliminary case study of Malawi*. Fondation pour l'Agriculture et la Ruralité dans le Monde (FARM). http://www.fondation-farm.org/zoe/doc/farm_201409_doctrav6_wawmalawi.pdf
Driessen, P., & Deckers, J. (2001). *LECTURE NOTES ON THE MAJOR SOILS OF THE WORLD*. FAO. http://www.fao.org/docrep/003/Y1899E/y1899e00.htm#toc
Elwel, H. A., & Stocking, M. A. (1982). Developing a simple yet practical method of soil loss estimation. *Tropical Agriculture*, *59*(1), 43–48. https://ueaeprints.uea.ac.uk/id/eprint/45746
Emerton, L. (2016). *Assessing the economic costs, benefits and drivers of sustainable land management for farmers*. International Center for Tropical Agriculture (CIAT).
Emerton, L., Snyder, K., Banda, R. G., Braslow, J., Chisinga, M., Cordingly, J., Kapalamula, M., Laizer, L., Mlenga, G., Mponela, P., Muhogora, F., & Ndengu, G. (2016). *Assessing the economic costs, benefits and drivers of sustainable land management for farmers in Ntcheu District, Malawi & Lushoto District, Tanzania.* CGIAR CGSpace. https://cgspace.cgiar.org/bitstream/handle/10568/77786/AGORA_WP1_2018.pdf?sequence=4&isAllowed=y
Enwezor, W., O. (1976). The mineralization of nitrogen and phosphorus in organic materials of varying C:N and C:P ratios. *Plant and Soil*, *44*(1), 237–240. https://www.jstor.org/stable/42932546
ESA. (2015). *Sentinel-2 User Handbook*. European Space Agency. https://earth.esa.int/documents/247904/685211/Sentinel-2_User_Handbook
FAO. (1982). Estimation of crop areas and yields in agricultural statistics. In *FAO economic and social development paper*. FAO. http://www.fao.org/fileadmin/templates/ess/ess_test_folder/World_Census_Agriculture/Publications/FAO_ESDP/ESDP_22_Engv2.pdf
FAO. (2005). *Livestock sector brief - Malawi*. FAO. http://www.fao.org/ag/againfo/resources/en/publications/sector_briefs/lsb_MWI.pdf
FAO. (2006a). *Guidelines for soil description* (4th ed.). FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS. http://www.fao.org/3/a-a0541e.pdf
FAO. (2006b). *Guidelines for soil description* (4th ed.). FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS.
FAO. (2012). *Atlas of Malawi, land cover and land cover change, 1990–2010.* FAO.
FAO. (2015). *Regional overview of food insecurity: African food security prospects brighter than ever.* FAO. http://www.fao.org/3/a-i4635e.pdf
FAO. (2016a). *FAOSTAT*. Food and Agriculture Organization of the United Nations. http://www.fao.org/faostat/en/#data
FAO. (2016b). *State of the world's forests 2016. Forests and agriculture: land-use challenges and opportunities*. FAO. http://www.fao.org/3/a-i5588e.pdf
FAO. (2016c). *The state of food and agriculture. Climate change, agriculture and food*

security. Food and Agriculture Organization of the United Nations. http://www.fao.org/3/a-i6030e.pdf

FAO. (2019). *Soil erosion: the greatest Challenge for sustainable soil management* (D. Pennock, C. Lefèvre, & R. Vargas (Eds.)). FAO.

Farmer, H. B., Brinkman, R., Moss, R. P., Wright, J., Symth, J. A., Greenland, J. D., Pullan, A. R., Curtis, F. L., Young, A., & Goldsmith, P. F. (1977). Soil Survey and Land Evaluation in Developing Countries a Case Study in Malaŵi: Discussion. *The Geographical Journal*, *143*(3), 431–438.

Ferber, J. (1999). *Multi-Agent Systems: An Introduction to Distributed Artificial Intelligence* (1st ed.). Addison Wesley Longman.

Fermont, A., & Benson, T. D. (2011). *Estimating Yield of Food Crops Grown by Smallholder Farmers: A Review in the Uganda Context*. International Food Policy Research Institute. http://www.ifpri.org/publication/estimating-yield-food-crops-grown-smallholder-farmers

Ferro, V., Di Stefano, C., Minacapilli, M., & Santoro, M. (2003a). Calibrating the SEDD model for Sicilian ungauged basins. *IAHS-AISH Publication*, *279*, 151–161. http://hydrologie.org/redbooks/a279/iahs_279_0151.pdf

Ferro, V., Di Stefano, C., Minacapilli, M., & Santoro, M. (2003b). Calibrating the SEDD model for Sicilian ungauged basins. *IAHS-AISH Publication*, *279*, 151–161.

Feuz, D. M., & Skold, M. D. (1992). Typical Farm Theory in Agricultural Research. *Journal of Sustainable Agriculture*, *2*(2), 43–58. https://doi.org/10.1300/J064v02n02_05

Fischer, G., van Velthuizen, H., Shah, M., & Nachtergaele, F. (2002). *Global Agro-ecological Assessment for Agriculture in the 21st Century: Methodology and Results*. International Institute for Applied Systems Analysis. http://pure.iiasa.ac.at/id/eprint/6667/1/RR-02-002.pdf

Fleskens, L., & Chilima, C. (2013). *Development of a green water credit scheme in the shire river basin*. UNDP-Malawi. http://www.nccpmw.org/index.php/documentation/doc_download/37-development-of-a-green-water-credit-scheme-in-the-shire-river-basin.

Forkuor, G. (2014). Agricultural Land Use Mapping in West Africa Using Multi-sensor Satellite Imagery. *Thesis*, 191.

Forkuor, G., Hounkpatin, O. K. L., Welp, G., & Thiel, M. (2017). High Resolution Mapping of Soil Properties Using Remote Sensing Variables in South-Western Burkina Faso: A Comparison of Machine Learning and Multiple Linear Regression Models. *PLOS ONE*, *12*(1), e0170478. https://doi.org/10.1371/journal.pone.0170478

Frossard, E., Buchmann, N., Bünemann, E. K., Kiba, D. I., Lompo, F., Oberson, A., Tamburini, F., & Traoré, O. Y. A. (2016). Soil properties and not inputs control carbon : nitrogen : phosphorus ratios in cropped soils in the long term. *SOIL*, *2*(1), 83–99. https://doi.org/10.5194/soil-2-83-2016

Ganunga, R. P., Yerokun, O. A., & Kumwenda, J. D. T. (2005). Contribution of tithonia diversifolia to yield and nutrient uptake of maize in malawian small-scale agriculture. *South African Journal of Plant and Soil*, *22*(4), 240–245. https://doi.org/10.1080/02571862.2005.10634714

Garrity, D., Dixon, J., & Boffa, J.-M. (2012). Understanding African Farming Systems:

Science and Policy Implications. *Food Security in Africa: Bridging Research and Practice*. http://aciar.gov.au/aifsc/sites/default/files/images/understanding_african_farming_systems_report_for_aifsc_conference.pdf

Gavish, B., & Kalay, A. (1983). On the Asset Substitution Problem. *The Journal of Financial and Quantitative Analysis*, *18*(1), 21. https://doi.org/10.2307/2330802

Geist, H. J., & Lambin, E. F. (2002). Proximate Causes and Underlying Driving Forces of Tropical Deforestation. *BioScience*, *52*(2), 143. https://doi.org/10.1641/0006-3568(2002)052[0143:pcaudf]2.0.co;2

German, G., Akinnifesi, F. K., Edriss, A. K., Sileshi, G., Masangano, C., & Ajayi, O. C. (2009). Influence of property rights on farmers' willingness to plant indigenous fruit trees in Malawi and Zambia. *African Journal of Agricultural Research*, *4*(5), 427–437.

Ghosh, S., & Kumar, S. (2013). Comparative Analysis of K-Means and Fuzzy C-Means Algorithms. *International Journal of Advanced Computer Science and Applications*, *4*(4). https://doi.org/10.14569/IJACSA.2013.040406

Gilbert, R. A. (2004). *Best-Bet Legumes for Smallholder Maize-Based Cropping Systems of Malawi BT - Green Manure/Cover Crop Systems of Smallholder Farmers: Experiences from Tropical and Subtropical Regions* (M. Eilittä, J. Mureithi, & R. Derpsch (Eds.); pp. 153–174). Springer Netherlands. https://doi.org/10.1007/1-4020-2051-1_6

Giller, K. E., Andersson, J. A., Corbeels, M., Kirkegaard, J., Mortensen, D., Erenstein, O., & Vanlauwe, B. (2015). Beyond conservation agriculture. *Frontiers in Plant Science*, *6*. https://doi.org/10.3389/fpls.2015.00870

Giller, K. E., Rowe, E. C., de Ridder, N., & van Keulen, H. (2006). Resource use dynamics and interactions in the tropics: Scaling up in space and time. *Agricultural Systems*, *88*(1), 8–27. https://doi.org/10.1016/j.agsy.2005.06.016

Giller, K. E., Tittonell, P., Rufino, M. C., van Wijk, M. T., Zingore, S., Mapfumo, P., Adjei-Nsiah, S., Herrero, M., Chikowo, R., Corbeels, M., Rowe, E. C., Baijukya, F., Mwijage, A., Smith, J., Yeboah, E., van der Burg, W. J., Sanogo, O. M., Misiko, M., de Ridder, N., ... Vanlauwe, B. (2011). Communicating complexity: Integrated assessment of trade-offs concerning soil fertility management within African farming systems to support innovation and development. *Agricultural Systems*, *104*(2), 191–203. https://doi.org/10.1016/j.agsy.2010.07.002

Gleeson, S. K., & Tilman, D. (1992). Plant Allocation and the Multiple Limitation Hypothesis. *The American Naturalist*, *139*(6), 1322–1343. https://doi.org/10.1086/285389

Glick, H. (2015). *Identifying an Appropriate Link and Family for Generalized Linear Models*. http://www.uphs.upenn.edu/dgimhsr/documents/ispor15.glmworkshop.glick.2.pdf

GoM. (2007). *National Fertiliser Strategy*. Ministray of Agriculture. http://fsg.afre.msu.edu/mgt/caadp/format_for_national_fertilizer_strategy9.pdf

GoM. (2013). *Shire river basin management program (phase I) project final environmental and social assessment report*. Ministry of Water Development and

Irrigation.
GoM. (2017). *National forest landscape restoration strategy*. The Ministry of Natural Resources, Energy and Mining. https://www.iucn.org/sites/dev/files/content/documents/2017/malawi_-_national_forest_landscape_restoration_assessment.pdf

GoM. (2018). *Malawi population and housing census report* (Vol. 8). http://www.nsomalawi.mw/images/stories/data_on_line/demography/census_2018/2018 Malawi Population and Housing Census Main Report.pdf

GOM. (2013). *Malawi National Export Strategy*. Ministry of Industry and Trade. http://www.moit.gov.mw/downloads/quick/National Export Strategy.zip

Google Earth. (n.d.). *[Google Earth imagery of June 2016 to December 2017 for nsipe area, Ntcheu Malawi]*. Retrieved December 27, 2017, from earth.google.com/web/

Gourlay, S., Kilic, T., & Lobell, D. (2017). Could the Debate Be Over? Errors in Farmer-Reported Production and Their Implications for the Inverse Scale-Productivity Relationship in Uganda. In *Development Economics*. World Bank Group. https://openknowledge.worldbank.org/handle/10986/28369

Grant, A. M., & Ashford, S. J. (2008). The dynamics of proactivity at work. *Research in Organizational Behavior*, *28*, 3–34. https://doi.org/10.1016/j.riob.2008.04.002

Greenland, D. J., Rimmer, D., & Payne, D. (1975). DETERMINATION OF THE STRUCTURAL STABILITY CLASS OF ENGLISH AND WELSH SOILS, USING A WATER COHERENCE TEST. *Journal of Soil Science*, *26*(3), 294–303. https://doi.org/10.1111/j.1365-2389.1975.tb01953.x

Greenland, S., Pearl, J., & Robins, J. M. (1999). Causal Diagrams for Epodemioloc Research. *Epidemiology*, *1*(10), 37–48. http://ftp.cs.ucla.edu/pub/stat_ser/r251.pdf

Greenland, S., Robins, J. M., & Pearl, J. (1999). Confounding and collapsibility in causal inference. *Statistical Science*, *14*(1), 29–46. https://doi.org/10.1214/ss/1009211805

Grimm, R., Behrens, T., Märker, M., & Elsenbeer, H. (2008). Soil organic carbon concentrations and stocks on Barro Colorado Island — Digital soil mapping using Random Forests analysis. *Geoderma*, *146*(1–2), 102–113. https://doi.org/10.1016/j.geoderma.2008.05.008

Groot, J., & Oomen, G. (2018). *Farm DESIGN Manual* (Issue April).

Groot, J., Oomen, G. J. M., & Rossing, W. A. H. (2012). Multi-objective optimization and design of farming systems. *Agricultural Systems*, *110*, 63–77. https://doi.org/10.1016/j.agsy.2012.03.012

Güsewell, S., & Verhoeven, J. T. A. (2006). Litter N:P ratios indicate whether N or P limits the decomposability of graminoid leaf litter. *Plant and Soil*, *287*(1–2), 131–143. https://doi.org/10.1007/s11104-006-9050-2

Gutiérrez Boem, F. H., Rubio, G., & Barbero, D. (2011). Soil Phosphorus Extracted by Bray 1 and Mehlich 3 Soil Tests as Affected by the Soil/Solution Ratio in Mollisols. *Communications in Soil Science and Plant Analysis*, *42*(2), 220–230. https://doi.org/10.1080/00103624.2011.535072

Haile, B., Azzarri, C., Roberts, C., & Spielman, D. J. (2017). Targeting, bias, and expected impact of complex innovations on developing-country agriculture: evidence from Malawi. *Agricultural Economics*, *48*(3), 317–326. https://doi.org/10.1111/agec.12336

Harris, D., & Orr, A. (2014). Is rainfed agriculture really a pathway from poverty? *Agricultural Systems*, *123*, 84–96. https://doi.org/10.1016/j.agsy.2013.09.005

Hay, R., & Gilbert, R. (2001). Variation in the harvest index of tropical maize: evaluation of recent evidence from Mexico and Malawi. *Annals of Applied Biology*, *138*(1), 103–109. https://doi.org/10.1111/j.1744-7348.2001.tb00090.x

Hazelton, P., & Murphy, B. (2007). *INTERPRETING SOIL TEST RESULTS WHAT DO ALL THE NUMBERS MEAN?* (2nd ed.). CSIRO PUBLISHING. https://www.publish.csiro.au/book/7386/

Henao, J., & Baanante, C. (1999). *Estimating Rates of Nutrient Depletion in Soils of Agricultural Lands of Africa* (No. 48). IFDC. http://pdf.usaid.gov/pdf_docs/Pnadw837.pdf

Hengl, T., de Jesus, J. M., MacMillan, R. A., Batjes, N. H., Heuvelink, G. B. M., Ribeiro, E., Samuel-Rosa, A., Kempen, B., Leenaars, J. G. B., Walsh, M. G., & Gonzalez, M. R. (2014). SoilGrids1km — Global Soil Information Based on Automated Mapping. *PLoS ONE*, *9*(8), e105992. https://doi.org/10.1371/journal.pone.0105992

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. D., & Tondoh, J. E. (2015). Mapping Soil Properties of Africa at 250 m Resolution: Random Forests Significantly Improve Current Predictions. *PLOS ONE*, *10*(6), e0125814. https://doi.org/10.1371/journal.pone.0125814

Hengl, T., Heuvelink, G. B. M., & Rossiter, D. G. (2007). About regression-kriging: From equations to case studies. *Computers & Geosciences*, *33*(10), 1301–1315. https://doi.org/10.1016/j.cageo.2007.05.001

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*. https://doi.org/10.1007/s10705-017-9870-x

Herrick, J. E., Urama, K. C., Karl, J. W., Boos, J., Johnson, M. V. V., Shepherd, K. D., Hempel, J., Bestelmeyer, B. T., Davies, J., Guerra, J. L., Kosnik, C., Kimiti, D. W., Ekai, A. L., Muller, K., Norfleet, L., Ozor, N., Reinsch, T., Sarukhan, J., & West, L. T. (2013). The global Land-Potential Knowledge System (LandPKS): Supporting evidence-based, site-specific land use and management through cloud computing, mobile applications, and crowdsourcing. *Journal of Soil and Water Conservation*, *68*(1). https://doi.org/10.2489/jswc.68.1.5A

Hgh-Jensen, H., M., F., Kamalongo, D., & Ngwir, A. (2013). The Bean – Naturally Bridging Agriculture and Human Wellbeing. In *Food Industry*. InTech. https://doi.org/10.5772/53164

Hochberg, Y. (1988). A sharper bonferroni procedure for multiple tests of significance. In *Biometrika* (Vol. 75, Issue 4, pp. 800–802). https://doi.org/10.1093/biomet/75.4.800

Holden, S. T., & Lunduka, R. W. (2013). Who benefit from Malawi's targeted farm input subsidy program? *Forum for Development Studies*, *40*(1), 1–25. https://doi.org/10.1080/08039410.2012.688858

Holt-Giménez, E., Miguel, A., & Rosset, P. (2006). Ten Reasons Why the Rockefeller and the Bill and Melinda Gates Foundations' Alliance for Another Green Revolution Will Not Solve the Problems of Poverty and Hunger in Sub-Saharan Africa. In *Food First*. Food First/Institute for Food and Development Policy. https://foodfirst.org/wp-content/uploads/2013/12/PB12-Ten-Reasons-Why-AGRA-Will-not-Solve-Poverty-and-Hunger-in-Africa.pdf

Howard, R. A., & Matheson, J. E. (2005). Influence Diagrams. *Decision Analysis*, *2*(3), 127–143. https://doi.org/10.1287/deca.1050.0020

Ibrahim, A., Pasternak, D., & Fatondji, D. (2015). Impact of depth of placement of mineral fertilizer micro-dosing on growth, yield and partial nutrient balance in pearl millet cropping system in the Sahel. *Journal of Agricultural Science*, *153*(8), 1412–1421. https://doi.org/10.1017/S0021859614001075

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*. https://doi.org/10.1007/s10705-018-9958-y

IFDC. (2013). *Malawi Fertilizer Assessment*. https://ifdcorg.files.wordpress.com/2015/04/malawi-fertilizer-assessment.pdf

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning. In *Springer Texts in Statistics* (Vol. 103). Springer New York. http://link.springer.com/10.1007/978-1-4614-7138-7

Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science*, *353*(6301), 790–794. https://doi.org/10.1126/science.aaf7894

Jeffery, R., & Vira, B. (2001). *Conflict and Cooperation in Participating Natural Resource Management*. Springer. https://books.google.de/books?id=tC5-DAAAQBAJ

Jenny, H. (1941). *Factors of soil formation: A system of quantitative pedology* (1st ed.). McGrawHill Book Company, Inc. http://netedu.xauat.edu.cn/sykc/hjx/content/ckzl/6/2.pdf

Jones, R. J. A., Verheijen, F. G. A., Reuter, H. I., Jones, A. R., & (eds). (2008). Environmental Assessment of Soil for Monitoring Volume V: Procedures & Protocols. In *EUR 23490 EN/5, Office for the Official Publications of the European Communities: Vol. V*. Joint Research Centre. https://doi.org/10.2788/94366

Jumbe, C. B. L., & Nyambose, W. H. (2016). Does Conservation Agriculture Enhance Household Food Security? Evidence from Smallholder Farmers in Nkhotakota in Malawi. *Sustainable Agriculture Research*, *5*(1), 118. https://doi.org/10.5539/sar.v5n1p118

Kaiji, C., & Zheng, S. (2007). *Financial Friction , Capital Reallocation and Expectation-Driven Business Cycles Financial Friction , Capital Reallocation and News-Driven Business Cycles* (Munich Personal RePEc Archive, Issue 3889). https://mpra.ub.uni-muenchen.de/3889/1/MPRA_paper_3889.pdf

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, *23*(2), 187–200.

Kakhobwe, C. M., Kamoto, J. F., Njoloma, J. P., & Ozor, N. (2016). Scaling Up Agroforestry Farming System in Malawi: A Case of Malawi Agroforestry Extension Project. *Journal of Agricultural Extension*, *20*(1).

http://aesonnigeria.org/ajm/index.php/jae/article/view/873
Kamau, J. W., Stellmacher, T., Biber-Freudenberger, L., & Borgemeister, C. (2018). Organic and conventional agriculture in Kenya: A typology of smallholder farms in Kajiado and Murang'a counties. *Journal of Rural Studies*, *57*, 171–185. https://doi.org/10.1016/j.jrurstud.2017.12.014
Kangalawe, Y. M. R. (2014). Nutrient budget analysis under smallholder farming systems and implications on agricultural sustainability in degraded environments of semiarid central Tanzania. *Journal of Soil Science and Environmental Management*, *5*(3), 44–60. https://doi.org/10.5897/JSSEM13.0390
Kankwamba, H., Kadzamira, M., & Pauw, K. (2018). How diversified is cropping in Malawi? Patterns, determinants and policy implications. *Food Security*, *10*(2), 323–338. https://doi.org/10.1007/s12571-018-0771-x
Katengeza, S. P., Holden, S. T., & Fisher, M. (2019). Use of Integrated Soil Fertility Management Technologies in Malawi: Impact of Dry Spells Exposure. *Ecological Economics*, *156*, 134–152. https://doi.org/10.1016/j.ecolecon.2018.09.018
Kempen, B. (2014). *Soil and Terrain Database for Malawi (ver. 1.0) (SOTER_Malawi)*. http://edepot.wur.nl/395242
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. https://doi.org/10.1016/j.agee.2016.05.012
Kihara, J., Tamene, L. D., Massawe, P., & Bekunda, M. (2015). Agronomic survey to assess crop yield, controlling factors and management implications: a case-study of Babati in northern Tanzania. *Nutrient Cycling in Agroecosystems*, *102*(1), 5–16. https://doi.org/10.1007/s10705-014-9648-3
Kirui, O. K. (2016). Economics of Land Degradation and Improvement in Tanzania and Malawi. In E. Nkonya, A. Mirzabaev, & J. von Braun (Eds.), *Economics of Land Degradation and Improvement – A Global Assessment for Sustainable Development* (pp. 609–649). Springer International Publishing. http://link.springer.com/10.1007/978-3-319-19168-3_20
Knoblach, M., & Stockl, F. (2019). What determines the elasticity of substitution between capital and labor? A literature review. In *Economic Outlook* (No. 1784; Discussion Papers, Vol. 4, Issue 9). https://www.diw.de/documents/publikationen/73/diw_01.c.612331.de/dp1784.pdf
Komarek, A. M., Drogue, S., Chenoune, R., Hawkins, J., Msangi, S., Belhouchette, H., & Flichman, G. (2017). Agricultural household effects of fertilizer price changes for smallholder farmers in central Malawi. *Agricultural Systems*, *154*(September 2016), 168–178. https://doi.org/10.1016/j.agsy.2017.03.016
Komarek, A. M., Koo, J., Haile, B., Msangi, S., & Azzarri, C. (2018). Trade-offs and synergies between yield, labor, profit, and risk in Malawian maize-based cropping systems. *Agronomy for Sustainable Development*, *38*(3). https://doi.org/10.1007/s13593-018-0506-6
Kopper, S. A., Jayne, T. S., & Snapp, S. S. (2020). Sifting through the weeds: Understanding heterogeneity in fertilizer and labor response in Central Malawi. *Ecological Economics*, *169*(August 2019), 106561.

https://doi.org/10.1016/j.ecolecon.2019.106561
Kundhlande, G., Winterbottom, R., Nyoka I., B., Reytar, K., Ha, K., & Behr Chandrasekharan, D. (2017). *Taking to Scale Tree-Based Systems that Enhance Food Security, Improve Resilience to Climate Change, and Sequester Carbon in Malawi*. PROFOR. https://www.profor.info/sites/profor.info/files/TakingToScaleMalawi_low_0.pdf
Kuyah, S., Whitney, C. W., Jonsson, M., Sileshi, G. W., Öborn, I., Muthuri, C. W., & Luedeling, E. (2019). Agroforestry delivers a win-win solution for ecosystem services in sub-Saharan Africa. A meta-analysis. *Agronomy for Sustainable Development*, *39*(5), 47. https://doi.org/10.1007/s13593-019-0589-8
Lakudzala, D. D. (2013). Potassium Response in some Malawi Soils. *International Letters of Chemistry, Physics and Astronomy*, *13*(2), 175–181. https://doi.org/10.18052/www.scipress.com/ilcpa.13.175
Lal, Rattan. (2015). Restoring Soil Quality to Mitigate Soil Degradation. *Sustainability*, *7*(5), 5875–5895. https://doi.org/10.3390/su7055875
Larson, D. F., Otsuka, K., Matsumoto, T., & Kilic, T. (2012). *Should African rural development strategies depend on smallholder farms? An exploration of the inverse productivity hypothesis*. World Bank Group. http://documents.worldbank.org/curated/en/684821468010489088/pdf/wps6190.pdf
Le Bellec, F., Cattan, P., Bonin, M., & Rajaud, A. (2011). Building a typology of cropping practices from comparison with a technical reference: first step for a relevant cropping system redesigning process – results for tropical citrus production. *Fruits*, *66*(3), 143–159. https://doi.org/10.1051/fruits/2011026
Le, Q. B. (2005). Multi-agent system for simulation of land-use and land cover change: a theoretical framework and its first implementation for an upland watershed in the central coast of Vietnam. In *Ecology and development series* (1. Aufl, Issue 29). Cuvillier.
Le, Q. B., Nkonya, E., & Mirzabaev, A. (2016). Biomass Productivity-Based Mapping of Global Land Degradation Hotspots. In E. Nkonya, A. Mirzabaev, & J. von Braun (Eds.), *Economics of Land Degradation and Improvement – A Global Assessment for Sustainable Development* (pp. 55–84). Springer International Publishing. http://link.springer.com/10.1007/978-3-319-19168-3_4
Le, Q. B., Park, S. J., & Vlek, P. L. G. (2010a). Land Use Dynamic Simulator (LUDAS): A multi-agent system model for simulating spatio-temporal dynamics of coupled human-landscape system. 2. Scenario-based application for impact assessment of land-use policies. *Ecological Informatics*, *5*(3), 203–221. https://doi.org/10.1016/j.ecoinf.2010.02.001
Le, Q. B., Park, S. J., & Vlek, P. L. G. (2010b). Land Use Dynamic Simulator (LUDAS): A multi-agent system model for simulating spatio-temporal dynamics of coupled human-landscape system. 2. Scenario-based application for impact assessment of land-use policies. *Ecological Informatics*, *5*(3), 203–221. https://doi.org/10.1016/j.ecoinf.2010.02.001
Le, Q. B., Park, S. J., & Vlek, P. L. G. (2010c). Household typologies and socio-ecological determinants of their land-use choices in Vietnam forest margins: implications for

empirical multi-agent modeling. *Sustainable Land Use And Rural Development In Mountainous Regions Of Southeast Asia*, 43. https://www.researchgate.net/publication/280742513_Household_typologies_and_socio-ecological_determinants_of_their_land-use_choices_in_Vietnam_forest_margins_Implications_for_empirical_multi-agent_modeling

Le, Q. B., Park, S. J., Vlek, P. L. G. G., & Cremers, A. B. (2008). Land-Use Dynamic Simulator (LUDAS): A multi-agent system model for simulating spatio-temporal dynamics of coupled human–landscape system. I. Structure and theoretical specification. *Ecological Informatics*, *3*(2), 135–153. https://doi.org/10.1016/j.ecoinf.2008.04.003

Le, Q. B., Scholz, R. W., Seidl, R., Vu, Q. M., & Tran, T. M. (2012). Phosphorus Use and Management Strategies: Exploring Scenarios of Smallholder's Soil Fertility, Food Production and Livelihood with a Multi-Agent System Model. *International Congress on Environmental Modelling and Software*, *6*.

Le, Q. B., Seidl, R., & Scholz, R. W. (2012). Feedback loops and types of adaptation in the modelling of land-use decisions in an agent-based simulation. *Environmental Modelling & Software*, *27–28*, 83–96. https://doi.org/10.1016/j.envsoft.2011.09.002

Lesschen, J. P., Stoorvogel, J. J., Smaling, E. M. A., Heuvelink, G. B. M., & Veldkamp, A. (2007a). A spatially explicit methodology to quantify soil nutrient balances and their uncertainties at the national level. *Nutrient Cycling in Agroecosystems*, *78*(2), 111–131. https://doi.org/10.1007/s10705-006-9078-y

Lesschen, J. P., Stoorvogel, J. J., Smaling, E. M. A., Heuvelink, G. B. M., & Veldkamp, A. (2007b). A spatially explicit methodology to quantify soil nutrient balances and their uncertainties at the national level. *Nutrient Cycling in Agroecosystems*, *78*(2), 111–131. https://doi.org/10.1007/s10705-006-9078-y

Lesschen, J. P., Verburg, P. H., & Staal, S. J. (2005). Statistical methods for analyzing the spatial dimension of changes in land use and farming systems. In *LUCC Report*. The International Livestock Research Institute & Wageningen University.

LEWIS, W. A. (1954). Economic Development with Unlimited Supplies of Labour. *The Manchester School*, *22*(2), 139–191. https://doi.org/10.1111/j.1467-9957.1954.tb00021.x

Li, G., Messina, J. P., Peter, B. G., & Snapp, S. S. (2017). Mapping Land Suitability for Agriculture in Malawi. *Land Degradation & Development*. https://doi.org/10.1002/ldr.2723

Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. In *R News* (2(3)). Berkerly, University of Califormia. https://www.stat.berkeley.edu/~breiman/RandomForests/

Liaw, A., & Wiener, M. (2018). *Breiman and Cutler's Random Forests for Classification and Regression* (4.6-14). CRAN. https://doi.org/10.1023/A:1010933404324

Liebig, M. A., Varvel, G., & Doran, J. (2001). A Simple Performance-Based Index for Assessing Multiple Agroecosystem Functions. *Agronomy Journal*, *93*(2), 313–318. https://doi.org/10.2134/agronj2001.932313x

Lin, S., Jing, C., Chaplot, V., Yu, X., Zhang, Z., Moore, N., & Wu, J. (2010). Effect of DEM resolution on SWAT outputs of runoff, sediment and nutrients. *Hydrology and Earth System Sciences Discussions*, *7*(4), 4411–4435.

https://doi.org/10.5194/hessd-7-4411-2010
Liniger, H. P., Mekdaschi Studer, R., Hauert, C., & Gurtner, M. (2011). Sustainable Land Management in Practice – Guidelines and Best Practices for Sub-Saharan Africa. In *TerrAfrica*. World Overview of Conservation Approaches and Technologies (WOCAT) and Food and Agriculture Organization of the United Nations (FAO).
Lobell, D. B. (2013). The use of satellite data for crop yield gap analysis. *Field Crops Research*, *143*, 56–64. https://doi.org/10.1016/j.fcr.2012.08.008
Lord, E. I., Anthony, S. G., & Goodlass, G. (2002). Agricultural nitrogen balance and water quality in the UK. *Soil Use and Management*, *18*(4), 363–369. https://doi.org/10.1079/SUM2002145
Lowole, W. M. (1965). Soil Map of Malawi. In *National Soil Maps*. Department of Agricullutral Research. http://esdac.jrc.ec.europa.eu/content/soil-map-malawi-0
Madhuri, K. V. N., Latha, P., Vasanthi, R. P., John, K., Reddy, P. V. R. M., Murali, G., Krishna, T. G., Naidu, T. C. ., & Naidu, N. . (2018). Evaluation of groundnut genotypes for phosphorus efficiency through leaf acid phosphatase activity. *LEGUME RESEARCH - AN INTERNATIONAL JOURNAL*, *00*. https://doi.org/10.18805/LR-3927
Maindonald, J., & Braun, J. W. (2010). *Data Analysis and Graphics Using R – an Example-Based Approach* (3rd ed.). Cambridge University Press.
Mas-ud, M., Seutra, J., Ofori, K., & Salifu, G. (2016). Relative Planting Dates Effect on the Agronomic Performance of Maize (Zea Mays L .) and Groundnut (Arachis Hypogea L) in an Intercrop System. *American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS)*, *16*(1), 262–276.
Matchaya, G. C. (2007). Does size of operated area matter? Evidence from Malawi's agricultural production. *International Journal of Agriculture and Rural Development*, *10*(2), 114–125. https://mpra.ub.uni-muenchen.de/11948/
Matson, P. A., Matson A., P., Parton J., W., Power G., A., & Swift J., M. (1997). Agricultural Intensification and Ecosystem Properties. *Science*, *277*(5325), 504–509. https://doi.org/10.1126/science.277.5325.504
McConnell J, D., & Dillon L, J. (1997). Farm Management for Asia: a Systems Approach. In *FAO Farm Systems Management Series*. FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS. http://www.fao.org/3/a-w7365e.htm
McCullagh, P. (2007). *Generalised Linear Models I. Advanced School and Conference on Statistics and Applied Probability in Life Sciences.* Department of Statistics, University of Chicago.
McDonald, J. F., & Moffitt, R. A. (1980). The Uses of Tobit Analysis. *The Review of Economics and Statistics*, *62*(2), 318. https://doi.org/10.2307/1924766
Mensah, M., Villamor, G., & Vlek, P. L. G. (2018). Gender specific determinants of inorganic fertilizer adoption in the semi-arid region of Ghana. *West African Journal of Applied Ecology*, *26*(Special Issue), 179–192.
Messina, J. P., Peter, B. G., & Snapp, S. S. (2017). Re-evaluating the Malawian Farm Input Subsidy Programme. *Nature Plants*, *3*(4), 17044. https://doi.org/10.1038/nplants.2017.13
Mhango, W., Snapp, S. S., & Kanyama-Phiri, G. Y. (2017). Biological nitrogen fixation

and yield of pigeonpea and groundnut: Quantifying response on smallholder farms in northern Malawi. *African Journal of Agricultural Research*, *12*(16), 1385–1394. https://doi.org/10.5897/AJAR2017.12232

Minasny, B., & McBratney, A. B. (2006). A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Computers & Geosciences*, *32*(9), 1378–1388. https://doi.org/10.1016/j.cageo.2005.12.009

Ministry of Natural Resouces Energy and Mining. (2017). *Forest Landscape Restoration Opportunities Assessment for Malawi* (Issue June). https://portals.iucn.org/library/sites/library/files/documents/2017-029.pdf

Miyasaka, T., Le, Q. B., Okuro, T., Zhao, X., & Takeuchi, K. (2017). Agent-based modeling of complex social–ecological feedback loops to assess multi-dimensional trade-offs in dryland ecosystem services. *Landscape Ecology*, *32*(4), 707–727. https://doi.org/10.1007/s10980-017-0495-x

Mloza-Banda, M. L., Cornelis, W. M., Mloza-Banda, H. R., Makwiza, C. N., & Verbist, K. (2014). Soil properties after change to conservation agriculture from ridge tillage in sandy clay loams of mid-altitude Central Malawi. *Soil Use and Management*, *30*(4), 569–578. https://doi.org/10.1111/sum.12139

Mohamad, M. A. A., Assar, A. H. A., & Idress, S. E. (2018). Genetic Diversity and Root Characters of Yield Components of Some Selected Groundnut (Rachis Hypogeal L .) Genotypes under. *IOSR Journal of Agriculture and Veterinary Science (IOSR-JAVS)*, *11*(7), 37–42. https://doi.org/10.9790/2380-1107023742

Mohamoud, Y., & Canfield Evan, H. (1998). *Development of a Rapid Assessment and Monitoring Technique for Soil Erosion in Malawi Based on Analysis of the Occurrence of Broken Ridges. I. Development of a Method for Qualitative Assessment of Soil erosion from Smallholder Farmers' Fields in Malawi.* Malawi Environmental Monitoring Programme. https://cals.arizona.edu/oals/malawi/Reports/Full_Report3.html

Montgomery, D. R. (2007). Soil erosion and agricultural sustainability. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(33), 13268–13272. https://doi.org/10.1073/pnas.0611508104

Moore, I. D., Grayson, R. B., & Ladson, A. R. (1991). Digital terrain modelling: A review of hydrological, geomorphological, and biological applications. *Hydrological Processes*, *5*(1), 3–30. https://doi.org/10.1002/hyp.3360050103

Moore, T. R. (1979). Rainfall Erosivity in East Africa. *Geografiska Annaler. Series A, Physical Geography*, *61*(3/4), 147. https://doi.org/10.2307/520909

Mooshammer, M., Wanek, W., Zechmeister-Boltenstern, S., & Richter, A. (2014). Stoichiometric imbalances between terrestrial decomposer communities and their resources: mechanisms and implications of microbial adaptations to their resources. *Frontiers in Microbiology*, *5*. https://doi.org/10.3389/fmicb.2014.00022

Morgan, R. P. C., Quinton, J. N., Smith, R. E., Govers, G., Poesen, J. W. A., Auerswald, K., Chisci, G., Torri, D., & Styczen, M. E. (1998). The European Soil Erosion Model (EUROSEM): a dynamic approach for predicting sediment transport from fields and small catchments. *Earth Surface Processes and Landforms*, *23*(6), 527–544. https://doi.org/10.1002/(SICI)1096-9837(199806)23:6<527::AID-ESP868>3.0.CO;2-5

Morton, J. F. (2007). The impact of climate change on smallholder and subsistence agriculture. *Proceedings of the National Academy of Sciences*, *104*(50), 19680–19685. https://doi.org/10.1073/pnas.0701855104

Moser, C., & Felton, A. (2007). The Construction of an Asset Index Measuring Asset Accumulation in Ecuador. In *Chronic Poverty Research Centre*. The Brookings Institution. http://www.chronicpoverty.org/uploads/publication_files/CP_2006_Moser_Felton.pdf

Motohka, T., Nasahara, K. N., Oguma, H., & Tsuchida, S. (2010). Applicability of Green-Red Vegetation Index for Remote Sensing of Vegetation Phenology. *Remote Sensing*, *2*(10), 2369–2387. https://doi.org/10.3390/rs2102369

Mpeketula, P. M. G. (2016). *Soil organic carbon dynamics and mycorrhizal fungal diversity in contrasting agroecosytems* [Michigan State University]. https://d.lib.msu.edu/etd/3907/datastream/OBJ/download/Soil_Organic_Carbon_Dynamics_and_Mycorrhizal_Fungal_Diversity_in_Contrasting_Agroecosystems.pdf

Mponela, P., Kassie, G. T., & Tamene, L. D. (2018). Simultaneous adoption of integrated soil fertility management technologies in the Chinyanja Triangle, Southern Africa. *Natural Resources Forum*, *42*(3), 172–184. https://doi.org/10.1111/1477-8947.12155

Mponela, P., Tamene, L. D., Ndengu, G., Magreta, R., Kihara, J., & Mango, N. (2016). Determinants of integrated soil fertility management technologies adoption by smallholder farmers in the Chinyanja Triangle of Southern Africa. *Land Use Policy*, *59*. https://doi.org/10.1016/j.landusepol.2016.08.029

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

Mungai, L. M., Snapp, S. S., Messina, J. P., Chikowo, R., Smith, A., Anders, E., Richardson, R. B., & Li, G. (2016). Smallholder Farms and the Potential for Sustainable Intensification. *Frontiers in Plant Science*, *7*. https://doi.org/10.3389/fpls.2016.01720

Mupangwa, W. T., & Tagwira, F. (2007). Calcium, Mg, and K interaction on groundnut (Arachis hypogaea L.) yield, nutrient uptake, and soil nutrient levels of an acid sandy soil. *Tropical Agriculture*, *84*(1–4), 21–28.

Murray, U., Gebremedhin, Z., Brychkova, G., & Spillane, C. (2016). Smallholder Farmers and Climate Smart Agriculture: Technology and Labor-productivity Constraints amongst Women Smallholders in Malawi. *Gender, Technology and Development*, *20*(2), 117–148. https://doi.org/10.1177/0971852416640639

Musinguzi, P., EBANYAT, P., TENYWA, J. S., BASAMBA, T. A., TENYWA, M. M., & Mubiru, D. N. (2016). CRITICAL SOIL ORGANIC CARBON RANGE FOR OPTIMAL CROP RESPONSE TO MINERAL FERTILISER NITROGEN ON A FERRALSOL. *Experimental Agriculture*, *52*(4), 635–653. https://doi.org/10.1017/S0014479715000307

Musinguzi, P., Tenywa, J. S., Ebanyat, P., Tenywa, M. M., Mubiru, D. N., Basamba, T. A., Leip, A., Patrick, M., Tenywa, J. S., Ebanyat, P., Tenywa, M. M., Mubiru, D. N., Basamba, T. A., & Leip, A. (2013). Soil Organic Carbon Thresholds and Nitrogen

Management in Tropical Agroecosystems: Concepts and Prospects. *Journal of Sustainable Development*, *6*(12). https://doi.org/10.5539/jsd.v6n12p31

Mustafa-Msukwa, A. K., Mutimba, J. k., Masangano, C., & Edriss, A. K. (2011). An assessment of the adoption of compost manure by smallholder farmers in Balaka District, Malawi. *South African Journal of Agricultural Extension*, *39*(1), 17–25. http://www.scielo.org.za/pdf/sajae/v39n1/02.pdf

Mutegi, J., Kabambe, V., Zingore, S., Harawa, R., & Wairegi, L. (2015). *The Status of Fertilizer Recommendation in Malawi: Gaps, Challenges and Opportunities*. Soil Health Consortium of Malawi. http://ssa.ipni.net/beagle/AFR-3091&f=Malawi soil fertility recommendations.pdf.

Muyanga, M., & Jayne, T. S. (2019). Revisiting the Farm Size-Productivity Relationship Based on a Relatively Wide Range of Farm Sizes: Evidence from Kenya. *American Journal of Agricultural Economics*, *101*(4), 1140–1163. https://doi.org/10.1093/ajae/aaz003

Naes, T., & Mevik, B.-H. (2001). Understanding the collinearity problem in regression and discriminant analysis: COLLINEARITY PROBLEM IN REGRESSION AND DISCRIMINANT ANALYSIS. *Journal of Chemometrics*, *15*(4), 413–426. https://doi.org/10.1002/cem.676

Nakhumwa, T. O. (2004). *Dynamic costs of soil degradation and determinants of adoption of soil conservation technologies by smallholder farmers in Malawi* (Issue July) [University of Pretoria]. https://repository.up.ac.za/bitstream/handle/2263/25032/Complete.pdf?sequence=5

Naohiro, M., Koji, N., Chisambi, C., & Moyo, M. (2016). Chemical properties and fermentation behavior of the composts prepared by three composting methods in Malawi. *African Journal of Agricultural Research*, *11*(25), 2213–2223. https://doi.org/10.5897/ajar2016.10919

Nautiyal, P. C. (2002). *GROUNDNUT: Post-harvest Operations*. http://www.fao.org/fileadmin/user_upload/inpho/docs/Post_Harvest_Compendium_-_Groundnut.pdf

Nearing, M. A., Foster, G. R., Lane, L. J., & Finkner, S. C. (1989). A Process-Based Soil Erosion Model for USDA-Water Erosion Prediction Project Technology. *Transactions of the ASAE*, *32*(5), 1587–1593. https://doi.org/10.13031/2013.31195

Ngwira, R. A., Aune, J. B., & Thierfelder, C. (2014). ON-FARM EVALUATION OF THE EFFECTS OF THE PRINCIPLES AND COMPONENTS OF CONSERVATION AGRICULTURE ON MAIZE YIELD AND WEED BIOMASS IN MALAWI. *Experimental Agriculture*, *50*(04), 591–610. https://doi.org/10.1017/S001447971400009X

Ngwira, R. A., Johnsen, F. H., Aune, J. B., Mekuria, M., & Thierfelder, C. (2014). Adoption and extent of conservation agriculture practices among smallholder farmers in Malawi. *Journal of Soil and Water Conservation*, *69*(2), 107–119. https://doi.org/10.2489/jswc.69.2.107

Njira, K. O. W. (2013). Microbial Contributions in Alleviating Decline in Soil Fertility. *British Microbiology Research Journal*, *3*(4), 724–742. https://doi.org/10.9734/BMRJ/2013/5577

Njira, K. O. W., Semu, E., Mrema, J. P., & Nalivata, P. C. (2017). Biological nitrogen fixation by pigeon pea and cowpea in the doubled-up and other cropping systems on the Luvisols of Central Malawi. *African Journal of Agricultural Research*, *12*(15), 1341–1352. https://doi.org/10.5897/AJAR2017.12167

Njoloma, J. P., Sileshi, W. G., Sosola, B. G., Nalivata, P. C., & Nyoka, B. I. (2016). Soil fertility status under smallholder farmers fields in malawi. *African Journal of Agricultural Research*, *11*(19), 1679–1687. https://doi.org/10.5897/AJAR2015.10018

Nkonya, E., Johnson, T., Kwon, H. Y., & Kato, E. (2016). Economics of Land Degradation in Sub-Saharan Africa. In E. Nkonya, A. Mirzabaev, & J. von Braun (Eds.), *Economics of Land Degradation and Improvement – A Global Assessment for Sustainable Development* (pp. 215–259). Springer International Publishing. http://link.springer.com/10.1007/978-3-319-19168-3_9

O'Hara, R. B., & Kotze, D. J. (2010). Do not log-transform count data. *Methods in Ecology and Evolution*, *1*(2), 118–122. https://doi.org/10.1111/j.2041-210X.2010.00021.x

Ortega, D. L., Waldman, K. B., Richardson, R. B., Clay, D. C., & Snapp, S. (2016). Sustainable Intensification and Farmer Preferences for Crop System Attributes: Evidence from Malawi's Central and Southern Regions. *World Development*, *87*(June), 139–151. https://doi.org/10.1016/j.worlddev.2016.06.007

Palm, C., Gachengo, C. N., Delve, R., Cadisch, G., & Giller, K. E. (2001). Organic inputs for soil fertility management in tropical agroecosystems: application of an organic resource database. *Agriculture, Ecosystems and Environment*, *83*, 27–42. https://doi.org/PII_ S0167-8809(00)00267-X

Palm, C., Giller, K. E., Mafongoya, L. P., & Swift, J. M. (2001). Management of organic matter in the tropics: translating theory into practice. *Nutrient Cycling in Agroecosystems*, *61*(1–2), 63–75. https://doi.org/10.1023/A:1013318210809

Panagos, P., Standardi, G., Borrelli, P., Lugato, E., Montanarella, L., & Bosello, F. (2018). Cost of agricultural productivity loss due to soil erosion in the European Union: From direct cost evaluation approaches to the use of macroeconomic models. *Land Degradation and Development*, *29*(3), 471–484. https://doi.org/10.1002/ldr.2879

Partey, S. T., Thevathasan, N. V., Zougmoré, R. B., & Preziosi, R. F. (2016). Improving maize production through nitrogen supply from ten rarely-used organic resources in Ghana. *Agroforestry Systems*. https://doi.org/10.1007/s10457-016-0035-8

Pattanayak, K. S., Mercer D., E., Sills O., E., Yang, J.-C., & Cassingham, K. (2002). Taking stock of agroforestry adoption studies. In *Environment and Natural Resource Economics*. Research Triangle Institute. http://www.rti.org/sites/default/files/resources/rtipaper_02_04.pdf

Peters, P. E. (2006). Rural income and poverty in a time of radical change in Malawi. *Journal of Development Studies*, *42*(2), 322–345. https://doi.org/10.1080/00220380500405568

Peterson, R. A. (2000). A Meta-Analysis of Variance Accounted for and Factor Loadings in Exploratory Factor Analysis. *Marketing Letters*, *11*(3), 261–275. https://doi.org/10.1023/A:1008191211004

Petschel-Held, G., Block, A., Cassel-Gintz, M., Kropp, J., Ludeke, M. K. B., Moldenhauer, O., Reusswig, F., & Schellnhuber, H. J. (1999). Syndromes of Global

Change: a qualitative modelling approach to assist global environmental management. *Environmental Modeling and Assessment*, *4*, 295–314.

Pieri, C. (1992). Fertility of Soils: A Future for Farming in the West African Savannah. In I. Douglas, F. Joly, M. Marcus, & B. Messerli (Eds.), *Springer Series in Physical Environment* (Vol. 10). Springer Berlin Heidelberg. http://link.springer.com/10.1007/978-3-642-84320-4

Pieri, C. (1995). Long term soil management experiments in semiarid francophone Africa. In R. Lal, A. Stewart, B., & R. Boca (Eds.), *Soil management. Experimental basis for substainability and environmental quality* (pp. 225–266). CRC press, Inc.

Pircher, T., Almekinders, C. J. M. M., & Kamanga, B. C. G. G. (2013). Participatory trials and farmers' social realities: understanding the adoption of legume technologies in a Malawian farmer community. *International Journal of Agricultural Sustainability*, *11*(3), 252–263. https://doi.org/10.1080/14735903.2012.738872

Polhill, G. J., Parker, D., Brown, D., & Grimm, V. (2008). Using the ODD Protocol for Describing Three Agent-Based Social Simulation Models of Land-Use Change. *Journal of Artificial Societies and Social Simulation*, *11*(2), 3. http://jasss.soc.surrey.ac.uk/11/2/3.html

Potter, P., Ramankutty, N., Bennett, E. M., & Donner, S. D. (2010). Characterizing the Spatial Patterns of Global Fertilizer Application and Manure Production. *Earth Interactions*, *14*(2), 1–22. https://doi.org/10.1175/2009EI288.1

Pretty, J. N. (1997). The sustainable intensification of agriculture. *Natural Resources Forum*, *21*(4), 247–256. https://doi.org/10.1111/j.1477-8947.1997.tb00699.x

Pretty, J. N., Toulmin, C., & Williams, S. (2011). Sustainable intensification in African agriculture. *International Journal of Agricultural Sustainability*, *9*(1), 5–24. https://doi.org/10.3763/ijas.2010.0583

Quang, D. V., Schreinemachers, P., & Berger, T. (2014). Ex-ante assessment of soil conservation methods in the uplands of Vietnam: An agent-based modeling approach. *Agricultural Systems*, *123*, 108–119. https://doi.org/10.1016/j.agsy.2013.10.002

R Core Team. (2018). *R: A language and environment for statistical computing.* (3.5.1). R Foundation for Statistical Computing. https://www.r-project.org/

Railsback, S. F., & Grimm, V. (2012). *Agent-Based and Individual-Based Modeling: A Practical Introduction*. Princeton University Press.

Raverkar, K. P., & Konde, B. K. (1988). Effect of Rhizobium andAzospirillum lipoferum inoculation on the nodulation, yield and nitrogen uptake of peanut cultivars. *Plant and Soil*, *106*(2), 249–252. https://doi.org/10.1007/BF02371220

Ray S., S., Singh P., J., Das, G., & Panigrahy, S. (2004). Use of high resolution remote sensing data for generating site specific soil management plan. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *35*, 127–132. https://pdfs.semanticscholar.org/0556/36c2da45c46f0d6f096f0f2063121e7af35d.pdf

Renard, K. G., Foster, G. R., Weesies, G. A., Mccool, D. K., & Yoder, D. C. (1997). Predicting soil erosion by water: a guide to conservation planning with the Revised Universal Soil Loss Equation (RUSLE). In *Agricultural Hand Book* (No.

703; Issue No. 703). U.S. Department of Agriculture. https://naldc.nal.usda.gov/download/CAT10827029/PDF

Ricker-Gilbert, J., & Jayne, T. S. (2017). Estimating the Enduring Effects of Fertiliser Subsidies on Commercial Fertiliser Demand and Maize Production: Panel Data Evidence from Malawi. *Journal of Agricultural Economics*, *68*(1), 70–97. https://doi.org/10.1111/1477-9552.12161

Ricker-Gilbert, J., Mason, N. N., Jayne, T. S., Darko, F., & Tembo, S. (2013). *What are the Effects of Input Subsidies on Maize Prices? Evidence from Malawi and Zambia*. Indaba Agricultural Policy Research Institute (IAPRI). http://fsg.afre.msu.edu/zambia/wp75.pdf

Riveiro-Valiño, J. A., Álvarez-López, C. J., & Marey-Pérez, M. F. (2009). The use of discriminant analysis to validate a methodology for classifying farms based on a combinatorial algorithm. *Computers and Electronics in Agriculture*, *66*(2), 113–120. https://doi.org/10.1016/j.compag.2008.12.001

Roose, E. (1977). *Erosion et ruissellement en Afrique de l'Ouest. Vingt années de mesures en petites parcelles expérimentales.* (No. 78). http://horizon.documentation.ird.fr/exl-doc/pleins_textes/pleins_textes_6/Tra_d_cm/09011.pdf

Rosegrant, M. W., Koo, J., Cenacchi, N., Ringler, C., Robertson, R., Fisher, M., Cox, C., Garrett, K., Perez, N. D., & Sabbagh, P. (2014). *Food security in a world of natural resource scarcity The role of agricultural technologies*. International Food Policy Research Institute. http://ebrary.ifpri.org/cdm/ref/collection/p15738coll2/id/128022

Rware, H., Wairegi, L., Oduor, G., Macharia, M., Romney, D., Tarfa, B. D., Maria, R. de, Ley, G., Tetteh, F., Makumba, W., Dicko, M., Nouri, M., Leon, N. N., Gondwe, B., Kaizzi, K., Negash, D., Kibunja, C., Alawy, A., Wartmann, C., & Rutsimba, E. (2014). Assessing the Potential to Change Stakeholders Knowledge and Practices on Fertilizer Recommendations in Africa. *Agricultural Sciences*, *05*(14), 1384–1391. https://doi.org/10.4236/as.2014.514149

Sachs, J. D., McArthur, J. W., Schmidt-Traub, G., Kruk, M., Bahadur, C., Faye, M., & McCord, G. (2004). *Ending Africa's Poverty Trap*. The Millenium Project. http://www.social-exclusion-housing.com/ending-africas-poverty-trap.pdf

Sakonnakhon, S. P. N., Toomsan, B., Cadisch, G., Baggs, E. M., Vityakon, P., Limpinuntana, V., Jogloy, S., & Patanothai, A. (2005). Dry season groundnut stover management practices determine nitrogen cycling efficiency and subsequent maize yields. *Plant and Soil*, *272*(1–2), 183–199. https://doi.org/10.1007/s11104-004-4780-5

Salvador, S., & Chan, P. (2004). Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms. *16th IEEE International Conference on Tools with Artificial Intelligence*, 576–584. https://doi.org/10.1109/ICTAI.2004.50

Sambo, P. T., Haywood, C., Wardell, D. A., Kibugi, R., & Cordonnier-Segger, M. C. (2015). *Enabling legal frameworks for sustainable land use investments in Zambia: Legal assessment report*. Center for International Forestry Research (CIFOR). http://www.cifor.org/library/5777/enabling-legal-frameworks-for-sustainable-

land-use-investments-in-zambia-legal-assessment-report/
Sanchez, P. A. (2002). Soil Fertility and Hunger in Africa. *Science*, *295*(5562), 2019–2020. https://doi.org/10.1126/science.1065256
Sandram, I. (2018). *Soil Erosion Hotspot Mapping and Climate Change Sensitivity Analysis of Rivirivi Watershed* (Issue February) [LILONGWE UNIVERSITY OF AGRICULTURE AND NATURAL RESOURCES]. http://rgdoi.net/10.13140/RG.2.2.20329.54885
Sardans, J., & Peñuelas, J. (2015). Potassium: a neglected nutrient in global change. *Global Ecology and Biogeography*, *24*(3), 261–275. https://doi.org/10.1111/geb.12259
Sau, K. R. (1971). Agricultural Revolution by Production Function. *Economic and Political Weekly*, *6*(26), A50–A58.
Sauer, J., & Tchale, H. (2009). The Economics of Soil Fertility Management in Malawi. *Review of Agricultural Economics*, *31*(3), 535–560. https://doi.org/10.1111/j.1467-9353.2009.01452.x
Schillaci, C., Acutis, M., Lombardo, L., Lipani, A., Fantappiè, M., Märker, M., & Saia, S. (2017). Spatio-temporal topsoil organic carbon mapping of a semi-arid Mediterranean region: The role of land use, soil texture, topographic indices and the influence of remote sensing data to modelling. *Science of The Total Environment*, *601–602*, 821–832. https://doi.org/10.1016/j.scitotenv.2017.05.239
Schindler, J., Vlek L, G, P., & Ehlers, E. (2009). *A multi-agent system for simulating land-use and land-cover change in the Atankwidi catchment of Upper East Ghana.* [Univeristy of Bonn]. http://hss.ulb.uni-bonn.de/2009/1954/1954.pdf
Schindler, F. V, Woodard, H. J., & Doolittle, J. J. (2005). Assessment of Soil Potassium Sufficiency as Related to Quantity-Intensity in Montmorillonitic Soils. *Communications in Soil Science and Plant Analysis*, *36*(15–16), 2255–2270. https://doi.org/10.1080/00103620500196630
Scholz, W. R., Binder, R. C., & Lang, J. D. (2011). The HES framework. In R. Scholz W. (Ed.), *Environmental literacy in science and society: from knowledge to decisions* (pp. 453–462). Cambridge University Press.
Schreinemachers, P., & Berger, T. (2011). An agent-based simulation model of human–environment interactions in agricultural systems. *Environmental Modelling & Software*, *26*(7), 845–859. https://doi.org/10.1016/j.envsoft.2011.02.004
Scoones, I. (1998). SUSTAINABLE RURAL LIVELIHOODS: A FRAMEWORK FOR ANALYSIS. In *IDS*.
Scoones, I., & Toulmin, C. (1999). Soil Nutrient Budgets And Balances: What Use For Policy? In *Managing Africa's Soils*. Institute of Development Studies (IDS). http://pubs.iied.org/pdfs/7399IIED.pdf
Shannon, R. E. (1998). Introduction to the Art and Science of Simulation. In D. J. Medeiros, E. F. Watson, J. S. Carson, & M. S. Manivannan (Eds.), *Winter Simulation Conference* (pp. 7–14).
Sheahan, M., & Barrett, C. B. (2017). Ten striking facts about agricultural input use in Sub-Saharan Africa. *Food Policy*, *67*, 12–25. https://doi.org/10.1016/j.foodpol.2016.09.010
Shefrin, H., & Statman, M. (2000). Behavioral Portfolio Theory. *JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSI*, *35*(2), 127–151.

http://www.jstor.org/stable/2676187
Shen, F., Wu, J., Fan, H., Liu, W., Guo, X., Duan, H., Hu, L., Lei, X., & Wei, X. (2019). Soil N/P and C/P ratio regulate the responses of soil microbial community composition and enzyme activities in a long-term nitrogen loaded Chinese fir forest. *Plant and Soil*, *436*(1–2), 91–107. https://doi.org/10.1007/s11104-018-03912-y
Signorelli, S., Azzarri, C., & Haile, B. (2016). *Typology characterization of farmers in Africa RISING sites in Malawi*. International Institute of Tropical Agriculture. https://cgspace.cgiar.org/handle/10568/83137
Silberg, T. R., Richardson, R. B., Hockett, M., & Snapp, S. S. (2017). Maize-legume intercropping in central Malawi: determinants of practice. *International Journal of Agricultural Sustainability*, 1–19. https://doi.org/10.1080/14735903.2017.1375070
Sileshi, G. W., & Akinnifesi, F. K. (2017). COMMENTS ON COE ET AL. (2016)– 'LOADING THE DICE IN FAVOUR OF THE FARMER . . .' *Experimental Agriculture*, 1–6. https://doi.org/10.1017/S0014479717000060
Simler, K. (1994). Agricultural Policy and Technology Options in Malawi: Modelling Responses and Outcomes in the Smallholder Subsector. In *Cornell Food and Nutrition Policy Program*. Cornel 1 University. http://www.cfnpp.cornell.edu/images/wp49.pdf
Simpson, D. R. (2011). *The "Ecosystem Service Framework": A Critical Assessment*. The United Nations Environment Programme. http://www.bioecon-network.org/pages/UNEP_publications/05 Ecosystem Service Framework.pdf
Sinha, P., Hobbs, P. V., Yokelson, R. J., Blake, D. R., Gao, S., & Kirchstetter, T. W. (2003). Distributions of trace gases and aerosols during the dry biomass burning season in southern Africa. *Journal of Geophysical Research D: Atmospheres*, *108*(17), 1–23. https://doi.org/10.1029/2003jd003691
Sinsabaugh, R. L., Hill, B. H., & Follstad Shah, J. J. (2009). Ecoenzymatic stoichiometry of microbial organic nutrient acquisition in soil and sediment. *Nature*, *462*(7274), 795–798. https://doi.org/10.1038/nature08632
Smajgl, A., Brown, D. G., Valbuena, D., & Huigen, M. G. A. (2011). Empirical characterisation of agent behaviours in socio-ecological systems. *Environmental Modelling & Software*, *26*(7), 837–844. https://doi.org/10.1016/j.envsoft.2011.02.011
Smaling, E. M. A., & Fresco, L. O. (1993). A decision-support model for monitoring nutrient balances under agricultural land use (NUTMON). *Geoderma*, *60*(1–4), 235–256. https://doi.org/10.1016/0016-7061(93)90029-K
Smaling, E. M. A., Nandwa, S. M., & Janssen, B. H. (1997). Soil fertility in Africa is at sake. *Replenishing Soil Fertility in Africa.*, *SSSA Spec.*(51), 47–62.
Smartt, J. (Ed.). (1994). *The Groundnut Crop*. Springer Netherlands. https://doi.org/10.1007/978-94-011-0733-4
Smith, A., Snapp, S. S., Chikowo, R., Thorne, P., Bekunda, M., & Glover, J. (2017). Measuring sustainable intensification in smallholder agroecosystems : A review. *Global Food Security*, *12*(November 2016), 127–138. https://doi.org/10.1016/j.gfs.2016.11.002
Smith, A., Thorne, P., & Snapp, S. (2015). *Measuring sustainable intensification in smallholder agroecosystems: A review*. International Livestock Research Institute. https://cgspace.cgiar.org/bitstream/handle/10568/67181/AR_SI_metrics_mar2015.

pdf?sequence=1
Snapp, S. S. (1998). Soil nutrient status of smallholder farms in Malawi. *Communications in Soil Science and Plant Analysis*, *29*(17–18), 2571–2588. https://doi.org/10.1080/00103629809370135
Snapp, S. S., Blackie, M. J., & Donovan, C. (2003). Realigning research and extension to focus on farmers' constraints and opportunities. *Food Policy*, *28*(4), 349–363. https://doi.org/10.1016/j.foodpol.2003.08.002
Snapp, S. S., Blackie, M. J., Gilbert, R. A., Bezner-Kerr, R., & Kanyama-Phiri, G. Y. (2010). Biodiversity can support a greener revolution in Africa. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(48), 20840–20845. https://doi.org/10.1073/pnas.1007199107
Snapp, S. S., Grabowski, P., Chikowo, R., Smith, A., Anders, E., Sirrine, D., Chimonyo, V., & Bekunda, M. (2018). Maize yield and profitability tradeoffs with social, human and environmental performance: Is sustainable intensification feasible? *Agricultural Systems*, *162*, 77–88. https://doi.org/10.1016/j.agsy.2018.01.012
Snapp, S. S., Jayne, T. S., Mhango, W., Benson, T. D., & Ricker-Gilbert, J. (2014). Maize-nitrogen response in Malawi's smallholder production systems. *Eight Years of FISP – Impact and What Next?*
Snapp, S. S., Mafongoya, P. ., & Waddington, S. (1998). Organic matter technologies for integrated nutrient management in smallholder cropping systems of southern Africa. *Agriculture, Ecosystems & Environment*, *71*(1–3), 185–200. https://doi.org/10.1016/S0167-8809(98)00140-6
Snapp, S. S., & Pound, B. (2008). *Agricultural Systems: Agroecology and Rural Innovation for Development* (1st ed.). Academic Press. https://www.elsevier.com/books/agricultural-systems-agroecology-and-rural-innovation-for-development/snapp/978-0-12-372517-2
Snapp, S. S., Rohrbach, D. D., Simtowe, F., & Freeman, H. A. (2002). Sustainable soil management options for Malawi: can smallholder farmers grow more legumes? *Agriculture, Ecosystems & Environment*, *91*(1–3), 159–174. https://doi.org/10.1016/S0167-8809(01)00238-9
Sosola, B., Gudeta, S., Akinnifesi, F., & Ajayi Clifford, O. (2011). Conservation Agriculture Practices in Malawi: Opportunities and Challenges. *FAO Regional Symposium on Conservation Agriculture, 7-10 February 2011*. http://evergreenagriculture.net/sites/default/files/Conservation Agriculture Practices in Malawi - Opportunities and Challenges.pdf
Sperfeld, E., Martin-Creuzburg, D., & Wacker, A. (2012). Multiple resource limitation theory applied to herbivorous consumers: Liebig's minimum rule vs. interactive co-limitation: Co-limitation theory applied to herbivores. *Ecology Letters*, *15*(2), 142–150. https://doi.org/10.1111/j.1461-0248.2011.01719.x
StataCorp. (2017). *Stata Statistical Software: Release 15*. StataCorp LLC.
Stefano, C. D., Ferro, V., Porto, P., & Rizzo, S. (2005). Testing a spatially distributed sediment delivery model (SEDD) in a forested basin by cesium-137 technique. *Journal of Soil and Water Conservation*, *60*(3), 148–157.
Stephens, W., & Middleton, T. (2002). Why has the uptake of decision support systems been so poor? In R. Matthews & W. Stephens (Eds.), *Crop-soil simulation models:*

applications in developing countries (pp. 129–147). CABI. http://www.cabi.org/cabebooks/ebook/20083014516

Stevenson, F. J., & Cole, M. A. (1999). *Cycles of soils: Carbon, nitrogen phosphorus, sulfur, micronutrients* (2nd ed.). Wiley.

Stoorvogel, J. J., & Smaling, E. M. A. (1990). *Assessment of soil nutrient depletion in sub-Saharan Africa, 1983-2000* (No. 28, Vol 1 of 4). https://library.wur.nl/WebQuery/wurpubs/fulltext/305170

Sud, C. U., Ahmad, T., Gupta, K. V, Chandra, H., Sahoo, M. P., Aditya, K., Singh, M., & Biswas, A. (2015). Synthesis of Literature and Framework: Research on Improving Methods for Estimating Crop Area, Yield and Production under Mixed, Repeated and Continuous Cropping. In *Global Strategy*. ICAR-Indian Agricultural Statistics Research Institute. http://gsars.org/wp-content/uploads/2016/01/WP_Synthesis-of-Literature-and-Framework_Improving-Methods-for-Estimation-of-Crop-Area-190116.pdf

Tallis, H., Kareiva, P., Marvier, M., & Chang, A. (2008). An ecosystem services framework to support both practical conservation and economic development. *Proceedings of the National Academy of Sciences*, *105*(28), 9457–9464. https://doi.org/10.1073/pnas.0705797105

Tamene, L. D. (2005). *Reservoir siltation in Ethiopia: Causes, source areas, and management options* (P. L. G. Vlek, M. Denich, C. Martius, C. Rodgers, & N. van de Giesen (Eds.); Issue 30). Cuvillier Verlag Göttingen.

Tamene, L. D., Adimassu, Z., Aynekulu, E., & Yaekob, T. (2017a). Estimating landscape susceptibility to soil erosion using a GIS-based approach in Northern Ethiopia. *International Soil and Water Conservation Research*, *5*(3), 221–230. https://doi.org/10.1016/j.iswcr.2017.05.002

Tamene, L. D., Adimassu, Z., Aynekulu, E., & Yaekob, T. (2017b). Estimating landscape susceptibility to soil erosion using a GIS-based approach in Northern Ethiopia. In *International Soil and Water Conservation Research* (Vol. 5, Issue 3, pp. 221–230). https://doi.org/10.1016/j.iswcr.2017.05.002

Tamene, L. D., Le, Q. B., & Vlek, P. L. G. (2014). A Landscape Planning and Management Tool for Land and Water Resources Management: An Example Application in Northern Ethiopia. *Water Resources Management*, *28*(2), 407–424. https://doi.org/10.1007/s11269-013-0490-1

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

Tamene, L. D., Mponela, P., Sileshi, G. W., Chen, J., & Tondoh, J. E. (2016). Spatial variation in tree density and estimated aboveground carbon stocks in Southern Africa. *Forests*, *7*(3). https://doi.org/10.3390/f7030057

Tamene, L. D., Sileshi, G. W., Ndengu, G., Mponela, P., Kihara, J., Sila, A., & Tondoh, J. (2019). Soil structural degradation and nutrient limitations across land use categories and climatic zones in Southern Africa. *Land Degradation and Development*. https://doi.org/10.1002/ldr.3302

Tamene, L. D., & Vlek, P. L. G. (2007). Assessing the potential of changing land use for reducing soil erosion and sediment yield of catchments: a case study in the highlands of northern Ethiopia. *Soil Use and Management*, *23*(1), 82–91. https://doi.org/10.1111/j.1475-2743.2006.00066.x

Tan, M. L., Ficklin, D. L., Dixon, B., Ibrahim, A. L., Yusop, Z., & Chaplot, V. (2015). Impacts of DEM resolution, source, and resampling technique on SWAT-simulated streamflow. *Applied Geography*, *63*, 357–368. https://doi.org/10.1016/j.apgeog.2015.07.014

Team, Q. D. (2016). *QGIS Geographic Information System* (2.18 Las P). Open Source Geospatial Foundation. http://qgis.osgeo.org

Telles, T. S., Dechen, S. C. F., de Souza, L. G. A., & Guimarães, M. de F. (2013). Valuation and assessment of soil erosion costs. *Scientia Agricola*, *70*(3), 209–216. https://doi.org/10.1590/S0103-90162013000300010

The Montpellier Panel. (2013). *Sustainable Intensification: A New Paradigm for African Agriculture*. http://ag4impact.org/publications/montpellier-panel-report2013/

Thiele, J. C., Kurth, W., & Grimm, V. (2015). Jan C . Thiele , Winfried Kurth and Volker Grimm (2014) Facilitating Parameter Estimation and Sensitivity Analysis of Agent-Based Models : A Cookbook Using NetLogo and R. *Journal of Artificial Societies and Social Simulation*, *17*(3), 1–45. https://doi.org/10.18564/jasss.2503

Thirtle, C., Lin, L., & Piesse, J. (2003). The impact of research led agricultural productivity growth on poverty reduction in Africa, Asia and Latin America. *International Association of Agricultural Economists (IAAE)*, *25*. http://ecsocman.hse.ru/data/609/661/1219/118.pdf

Timler, C., Michalscheck, M., Klapwijk, C., Mashingaidze, N., Ollenburger, M., Falconnier, G., Kuivanen, K., Descheemaeker, K., & Groot, J. (2014). *Characterization of farming systems in Africa RISING intervention sites in Malawi, Tanzania, Ghana and Mali*. International Institute of Tropical Agriculture. https://www.researchgate.net/profile/Jeroen_Groot/publication/280737196_Characterization_of_farming_systems_in_Africa_RISING_intervention_sites_in_Malawi_Tanzania_Ghana_and_Mali/links/55ddf52e08aeaa26af0f1f4b/Characterization-of-farming-systems-in-Africa-

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. https://doi.org/10.1016/j.fcr.2012.10.007

Tittonell, P., Muriuki, A., Shepherd, K. D., Mugendi, D., Kaizzi, K. C., Okeyo, J., Verchot, L., Coe, R., & Vanlauwe, B. (2010). The diversity of rural livelihoods and their influence on soil fertility in agricultural systems of East Africa – A typology of smallholder farms. *Agricultural Systems*, *103*(2), 83–97. https://doi.org/10.1016/j.agsy.2009.10.001

Tittonell, P., Vanlauwe, B., Leffelaar, P. A., Shepherd, K. D., & Giller, K. E. (2005). Exploring diversity in soil fertility management of smallholder farms in western Kenya: II. Within-farm variability in resource allocation, nutrient flows and soil fertility status. *Agriculture, Ecosystems & Environment*, *110*(3–4), 166–184. https://doi.org/10.1016/j.agee.2005.04.003

Tochombe, K. (2002). Agricultural production function, labor productivity, and wage embedded in Meiji peasant productive behaviors: a quantitative analysis of Nohi

Ryogoku Choson Ryakushi, ca.1881. *XIIth Congress of the International 700 Economic History Association*, *12*.

Towett, E. K., Shepherd, K. D., Tondoh, J. E., Winowiecki, L. A., Lulseged, T., Nyambura, M., Sila, A., Vågen, T.-G., & Cadisch, G. (2015). Total elemental composition of soils in Sub-Saharan Africa and relationship with soil forming factors. *Geoderma Regional*, *5*, 157–168. https://doi.org/10.1016/j.geodrs.2015.06.002

Tsai, Y., Zia, A., Koliba, C., Bucini, G., Guilbert, J., & Beckage, B. (2015). An interactive land use transition agent-based model (ILUTABM): Endogenizing human-environment interactions in the Western Missisquoi Watershed. *Land Use Policy*, *49*, 161–176. https://doi.org/10.1016/j.landusepol.2015.07.008

Tully, K., Sullivan, C., Weil, R., & Sanchez, P. (2015). The State of Soil Segradation in Sub-Saharan Africa: Baselines, Trajectories, and Solutions. *Sustainability*, *7*(6), 6523–6552. https://doi.org/10.3390/su7066523

United Nations. (1987). Report of the World Commission on Environment and Development: Our Common Future. In *A/42*. United Nations. http://www.un-documents.net/wced-ocf.htm

United Nations. (2000). United Nations Millennium Declaration. In *A/RES*. United Nations. http://www.unmillenniumproject.org/documents/ares552e.pdf

United Nations. (2015a). *The millennium development goals report*. United Nations. http://www.un.org/millenniumgoals/2015_MDG_Report/pdf/MDG 2015 rev (July 1).pdf

United Nations. (2015b). Transforming our world: the 2030 Agenda for Sustainable Development. In *A/RES*. United Nations.

USGS. (2018). *SRTM DEM data*. https://earthexplorer.usgs.gov/

Vågen, T.-G., & Winowiecki, L. A. (2013). Mapping of soil organic carbon stocks for spatially explicit assessments of climate change mitigation potential. *Environmental Research Letters*, *8*(1), 15011. https://doi.org/10.1088/1748-9326/8/1/015011

Vågen, T.-G., Winowiecki, L. A., Tondoh, J. E., Desta, L. T., & Gumbricht, T. (2016). Mapping of soil properties and land degradation risk in Africa using MODIS reflectance. *Geoderma*, *263*, 216–225. https://doi.org/10.1016/j.geoderma.2015.06.023

van Biljon, J. J., Fouche, D. S., & Botha, A. D. P. (2008). Threshold values and sufficiency levels for potassium in maize producing sandy soils of South Africa. *South African Journal of Plant and Soil*, *25*(2), 65–70. https://doi.org/10.1080/02571862.2008.10639897

Van den Bosch, H., Gitari, J. ., Ogaro, V. ., Maobe, S., & Vlaming, J. (1998). Monitoring nutrient flows and economic performance in African farming systems (NUTMON). *Agriculture, Ecosystems & Environment*, *71*(1–3), 63–80. https://doi.org/10.1016/S0167-8809(98)00132-7

van der Ploeg, J. D., & Long, A. (1990). *Labor, Markets, and Agricultural Production*. Westview Press. http://edepot.wur.nl/358315

van der Poel, P. (1980). *Rainfall erosivity and its use for soil loss estimation*. National government publication, Soil Conservation Section, Division of Land Utilization, Dept. of Agricultural Field Services, Ministry of Agriculture.

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

van Ittersum, M. K., van Bussel, L. G. J., Wolf, J., Grassini, P., van Wart, J., Guilpart, N., Claessens, L., de Groot, H., Wiebe, K., Mason-D'Croz, D., Yang, H., Boogaard, H., van Oort, P. A. J., van Loon, M. P., Saito, K., Adimo, O., Adjei-Nsiah, S., Agali, A., Bala, A., … Cassman, K. G. (2016). Can sub-Saharan Africa feed itself? *Proceedings of the National Academy of Sciences*, *113*(52), 14964–14969. https://doi.org/10.1073/pnas.1610359113

van Noordwijk, M., & Brussaard, L. (2014). Minimizing the ecological footprint of food: closing yield and efficiency gaps simultaneously? *Current Opinion in Environmental Sustainability*, *8*, 62–70. https://doi.org/10.1016/j.cosust.2014.08.008

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., … 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*, *15*(6). https://doi.org/10.1080/14735903.2017.1393038

Vanlauwe, B., Coe, R. I. C., & Giller, K. E. (2016). Beyond averages: new approaches to understand heterogeneity and risk of technology success or failure in smallholder farming. *Experimental Agriculture*, *55*(S1), 1–23. https://doi.org/10.1017/S0014479716000193

Vanlauwe, B., Six, J., Sanginga, N., & Adesina, A. A. (2015). Soil fertility decline at the base of rural poverty in sub-Saharan Africa. *Nature Plants*, *1*(7), 15101. https://doi.org/10.1038/nplants.2015.101

Vanlauwe, B., Tittonell, P., & Mukalama, J. (2006). Within-farm soil fertility gradients affect response of maize to fertiliser application in western Kenya. *Nutrient Cycling in Agroecosystems*, *76*(2–3), 171–182. https://doi.org/10.1007/s10705-005-8314-1

Vargas, R., & Omuto, C. (2016). *Soil loss assessment in Malawi*. Food and Agriculture Organization of the United Nations, UNDP-UNEP Poverty-Environment Initiative & Ministry of Agriculture, Irrigation and Water Development, Malawi. http://www.fao.org/3/a-i6387e.pdf.

Vayssières, J., Vigne, M., Alary, V., & Lecomte, P. (2011). Integrated participatory modelling of actual farms to support policy making on sustainable intensification. *Agricultural Systems*, *104*(2), 146–161. https://doi.org/10.1016/j.agsy.2010.05.008

Villamor, G. B. (2012). Flexibility of multi-agent system models for rubber agroforest landscapes and social response to emerging reward mechanism for ecosystems services in Sumatra, Indonesia [Rheinischen Friedrich-Wilhelms-Universität Bonn]. In *Centre for development research(ZEF)* (Issue 88). https://hss.ulb.uni-bonn.de/2012/2865/2865.pdf

Villamor, G. B., Le, Q. B., Djanibekov, U., van Noordwijk, M., & Vlek, P. L. G. (2014). Biodiversity in rubber agroforests, carbon emissions, and rural livelihoods: An agent-based model of land-use dynamics in lowland Sumatra. *Environmental*

Modelling & Software, *61*, 151–165. https://doi.org/10.1016/j.envsoft.2014.07.013
Vlek, P. L. G., Bao, Q., & Tamene, L. D. (2008). *Land decline in Land-Rich Africa – A creeping disaster in the making.* CGIAR Science Council Secretariat. http://library.wur.nl/isric/fulltext/isricu_i34313_001.pdf
Vrieze, S. I. (2012). Model selection and psychological theory: A discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychological Methods*, *17*(2), 228–243. https://doi.org/10.1037/a0027127
Waldman, K. B., Ortega, D. L., Richardson, R. B., Clay, D. C., & Snapp, S. (2016). Preferences for legume attributes in maize-legume cropping systems in Malawi. *Food Security*, *8*(6), 1087–1099. https://doi.org/10.1007/s12571-016-0616-4
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*(8), 1–17. https://doi.org/10.1371/journal.pone.0219296
Wang, J., Wang, H., & Tan, C. (2008). Cryptanalysis and improvement of an "efficient remote mutual authentication and key agreement." In D. J. Medeiros, E. F. Watson, J. S. Carson, & M. S. Manivannan (Eds.), *Proceedings of the 3rd IEEE Asia-Pacific Services Computing Conference, APSCC 2008* (pp. 835–840). https://doi.org/10.1109/APSCC.2008.173
Wang, X., Hoogmoed, W. B., Cai, D., Perdok, U. D., & Oenema, O. (2007). Crop residue, manure and fertilizer in dryland maize under reduced tillage in northern China: II nutrient balances and soil fertility. *Nutrient Cycling in Agroecosystems*, *79*(1), 17–34. https://doi.org/10.1007/s10705-006-9070-6
Weltin, M., Zasada, I., Piorr, A., Debolini, M., Geniaux, G., Moreno Perez, O., Scherer, L., Tudela Marco, L., & Schulp, C. J. E. (2018). Conceptualising fields of action for sustainable intensification – A systematic literature review and application to regional case studies. *Agriculture, Ecosystems & Environment*, *257*, 68–80. https://doi.org/10.1016/j.agee.2018.01.023
Wilensky, U. (1999). *NetLogo* (6.0.3). Center for Connected Learning and Computer-Based Modeling, Northwestern University. http://ccl.northwestern.edu/netlogo/
Wischmeier, W. H., & Smith, D. D. (1978). *Predicting rainfall erosion losses: A guide to conservation planning* (No. 537). https://naldc.nal.usda.gov/download/CAT79706928/PDF
Wooldridge, J. M. (2012). *Introductory Econometrics: A Modern Approach* (Fifth). South-Western CENGAGE Learning.
World Bank. (1992). *Malawi Economic Report on Environmental Policy* (No. 9888). http://documents.worldbank.org/curated/en/807441468774613140/pdf/multi-page.pdf
World Bank. (2019). *World Development Indicators*. https://datacatalog.worldbank.org/dataset/world-development-indicators
Xiao, J., Shen, Y., Tateishi, R., & Bayaer, W. (2006). Development of topsoil grain size index for monitoring desertification in arid land using remote sensing. *International Journal of Remote Sensing*, *27*(12), 2411–2422. https://doi.org/10.1080/01431160600554363

Xu, X., Thornton, P. E., & Post, W. M. (2013). A global analysis of soil microbial biomass carbon, nitrogen and phosphorus in terrestrial ecosystems. *Global Ecology and Biogeography*, *22*(6), 737–749. https://doi.org/10.1111/geb.12029

Xue, J., & Su, B. (2017). Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications. *Journal of Sensors*, *2017*, 1–17. https://doi.org/10.1155/2017/1353691

Zambonelli, F., Jennings, R. N., & Woordridge, M. (2003). Developing Multiagent Systems: The Gaia Methodology. *ACM Transactions on Software Engineering and Methodology*, *12*(3), 317–370. https://doi.org/1049-331X/03/0700-0317

Zhang, L., Ding, X., Peng, Y., George, T. S., & Feng, G. (2018). Closing the Loop on Phosphorus Loss from Intensive Agricultural Soil: A Microbial Immobilization Solution? *Frontiers in Microbiology*, *9*. https://doi.org/10.3389/fmicb.2018.00104

Zingore, S., Tittonell, P., Corbeels, M., van Wijk, M. T., & Giller, K. E. (2011). Managing soil fertility diversity to enhance resource use efficiencies in smallholder farming systems: a case from Murewa District, Zimbabwe. *Nutrient Cycling in Agroecosystems*, *90*(1), 87–103. https://doi.org/10.1007/s10705-010-9414-0

APPENDICES

S1. Conformity to existing maps and soil mapping covariance importance

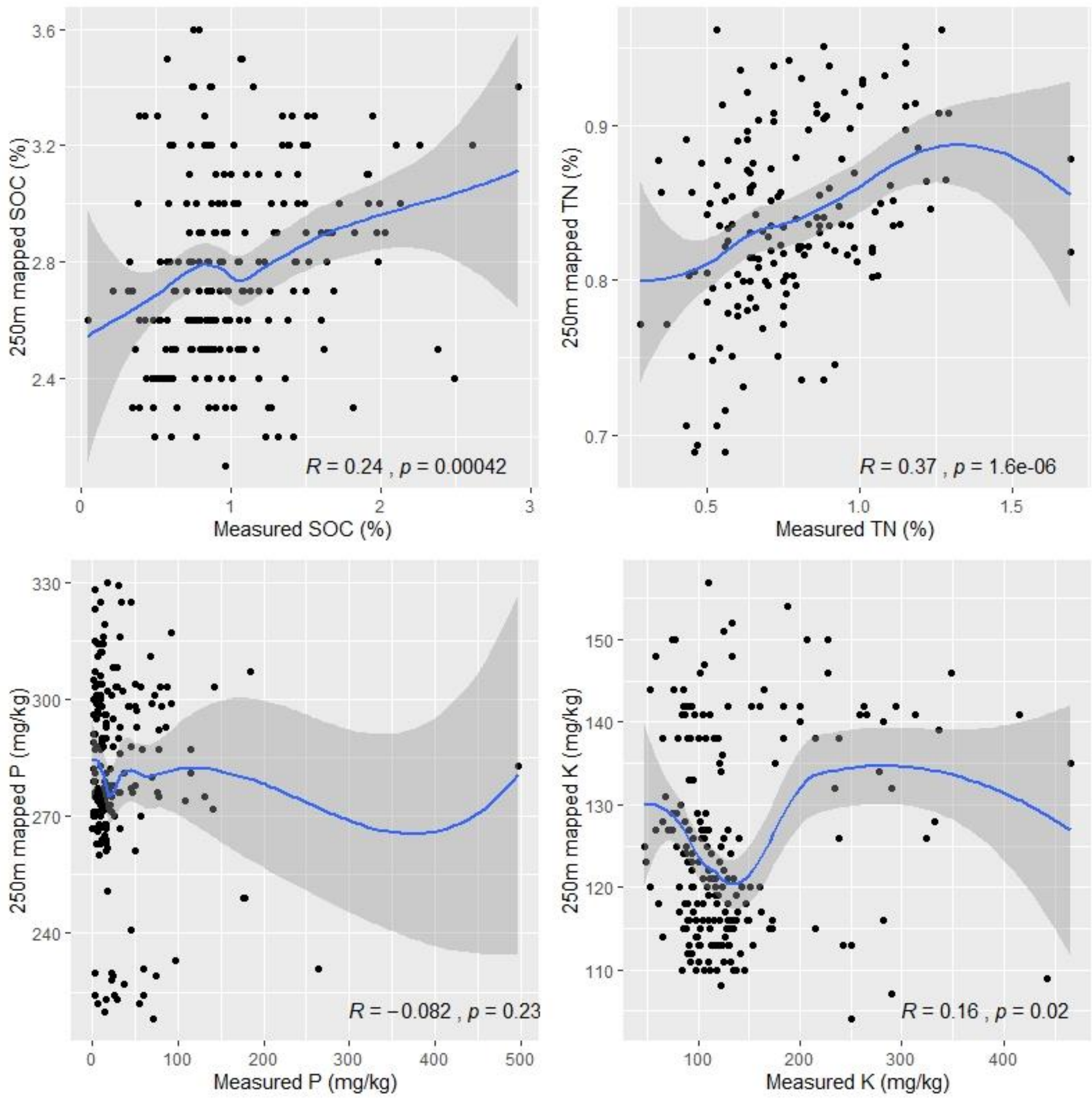
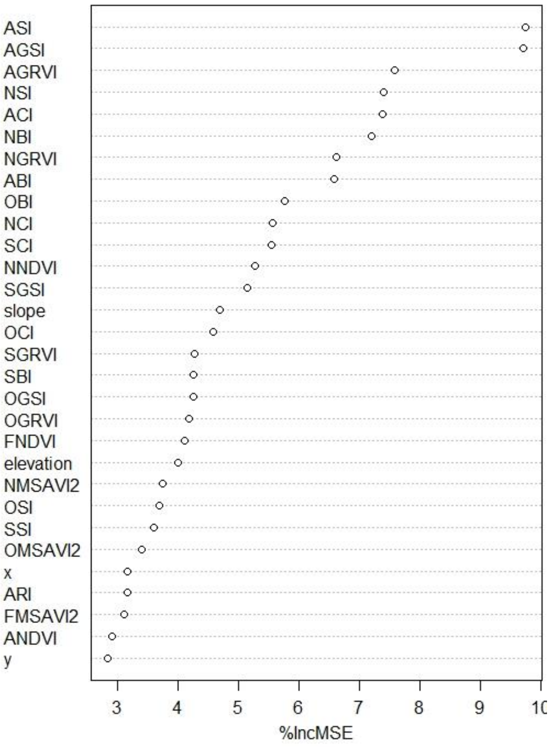
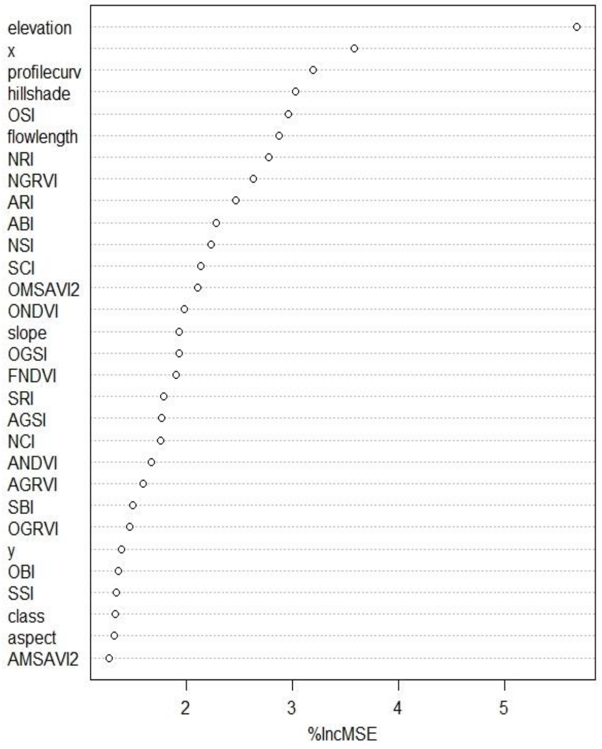


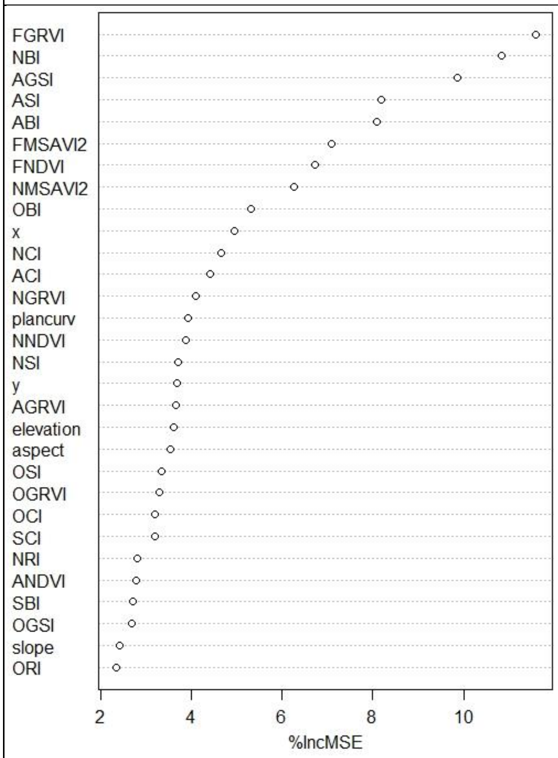
Figure S1: The conformity among measured and the 250m resolution soil prediction by Hengl (2015).



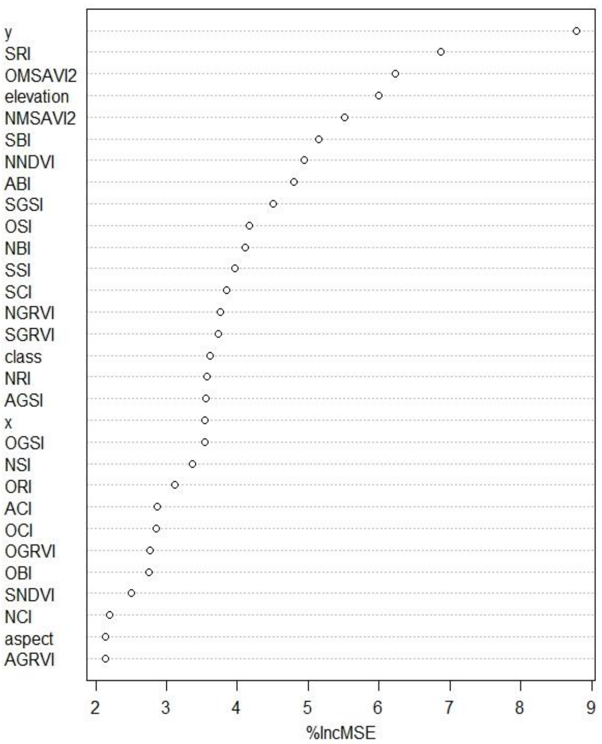
(a) SOC



(c) Phosphorus



(b) Total nitrogen



(d) Potassium

Figure S2 (a-f) co-variate importance for the distribution of NPK and SOC indicated by magnitude of mean square error (%IncMSE)

Table S 1 baseline nutrient and SOC balances estimated using commonly used stoichiometric and biomass conversion factors

	Nitrogen					Phosphorus					Potassium					Soil organic carbon				
	$\bar{x}$	sd	min	max	0>%	$\bar{x}$	sd	min	max	0>%	$\bar{x}$	sd	min	max	0>%	$\bar{x}$	sd	min	max	0>%
Inputs (population)																				
IN1$_{\text{fertilizer*}}$	56	82	0	640		6.2	10.5	0.0	85											
IN2$_{\text{organic*}}$	3	6	0	32		1.0	2.1	0.0	11		3	8	0	39		29	63	0	353	
IN2$_{\text{residues}}$																475	583	0	5304	
IN3$_{\text{legume-bnf*}}$	27	70	0	455																
IN4$_{\text{sediment}}$	4	13	0	175		0.2	1.3	0.0	36		1	3	0	38		53	176	0	2285	
IN5$_{\text{atm-depo}}$	14	0	14	15		1.4	0.0	1.4	1		4	0	4	4						
Total-in	104	111	14	838		8.8	10.8	1.4	86		8	8	4	62		556	625	0	5559	
Outputs (population)																				
OUT1$_{\text{product*}}$	34	40	0	341		5.0	6.3	0.0	52		26	32	0	266						
OUT2$_{\text{residues*}}$	27	34	0	250		2.3	2.8	0.0	22		24	28	0	222						
OUT3$_{\text{erosion}}$	31	44	0	432		2.6	8.9	0.0	134		6	10	0	104		411	610	1	5930	
OUT4$_{\text{leaching}}$	4	1	2	7							54	12	34	74						
OUT5$_{\text{gaseous}}$	32	0	32	32												720‡	75	465	973	
Total-out	128	84	36	665		9.9	12.5	0.0	150		110	59	36	566		1131	591	554	6595	
Full balance																				
Population	-25	115	-588	542	31	-1.14	16.20	-132.68	77.29	42	-102	58	-534	-4	0.00	-575	858	-6173	4700	13
Farmtype1	-40	96	-474	400	17	-2.96	14.25	-107.65	48.19	33	-103	59	-488	-33	0.00	-480	757	-4318	4700	11
Farmtype2	-11	116	-459	532	38	1.41	16.88	-132.68	76.52	53	-96	55	-534	-20	0.00	-622	840	-6173	4540	9
Farmtype3	-33	122	-588	542	31	-3.35	15.95	-131.72	77.29	32	-108	61	-490	-4	0.00	-571	931	-5884	3991	18
*Partial Balance**																				
Population	24	106	-448	575	58	-0.2	13.7	-66.0	76.1	39	-46	56	-460	36	2.5	503®	599	0	5554	94
Farmtype1	1	92	-437	422	47	-2.9	12.3	-65.6	46.8	30	-49	58	-419	8	2.0	508®	593	10	4968	100
Farmtype2	36	104	-391	553	66	2.3	14.0	-65.9	75.1	47	-41	53	-460	4	2.5	437®	593	0	5554	92
Farmtype3	21	115	-448	575	55	-1.6	13.5	-66.0	76.1	35	-52	59	-420	36	2.8	587®	602	0	4802	94

$\bar{x}$ = mean; sd = standard deviation; 0>% = percentage of plots with positive balances; *partial balance estimated from flows influenced by human action; ‡SOM degradation, ®inputs only.

S3. MASSAI Netlogo interface and codes

The interface (Figure S3) has two main codes for users. First, with the **SET-UP** button, the user loads both spatial and household-plot data, generates the remaining population for the unsampled landscape and estimates the initial nutrient inputs, outputs and balances. Then with the **go** button, the user simulates changes in input and output processes and estimates the associated nutrient balances. The simulated changes are based on the assumption that some factors that drive inputs and outputs progressively and dynamically change each year which is set **on** using the switch '**Progressive**'. This could be switched **off** if the user assumes the principle of *ceteris paribus* to set the baseline. The interface also features the sliders for adjusting the major nutrient input strategy: **fertilizer subsidy**. Currently set 20% (reduced), 0% (reduced to zero) 70% (universal) and as described in section 3.9.2.

The user can also set the simulation period using the chooser **stop-when**. Currently set at **20** years counting from the growing season 2016/17 to 2036/37. The user should be mindful that adjusting this downwards will steepen the rate of change for policy variables. The user can also view the years of simulation through the reporter **elapsed-years**. On the right the user can view data and statistics after leading data and the summary statistics can be used to check estimates for the baseline sample and the most recent scenario. The interface also has a graphical results section which are used for trend reporting and exported for moving average analysis in a statistical software. The procedures for the main codes are presented in Box S1. Detailed sub-procedures are saved separately in __includes files.

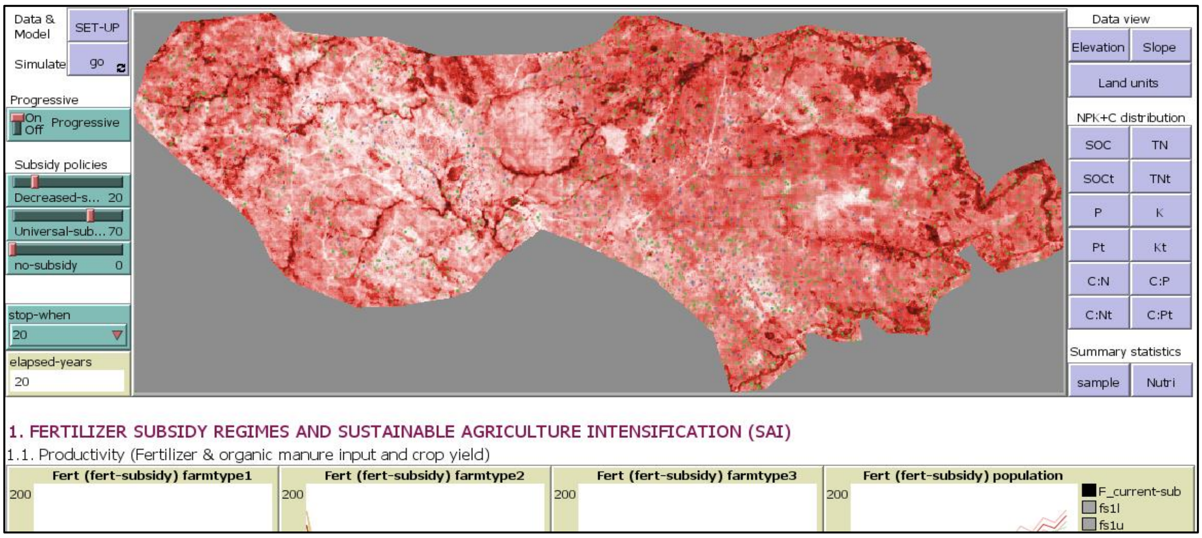


Figure S3 MASSAI Netlogo interface

Box S1 The MASSAI main procedures

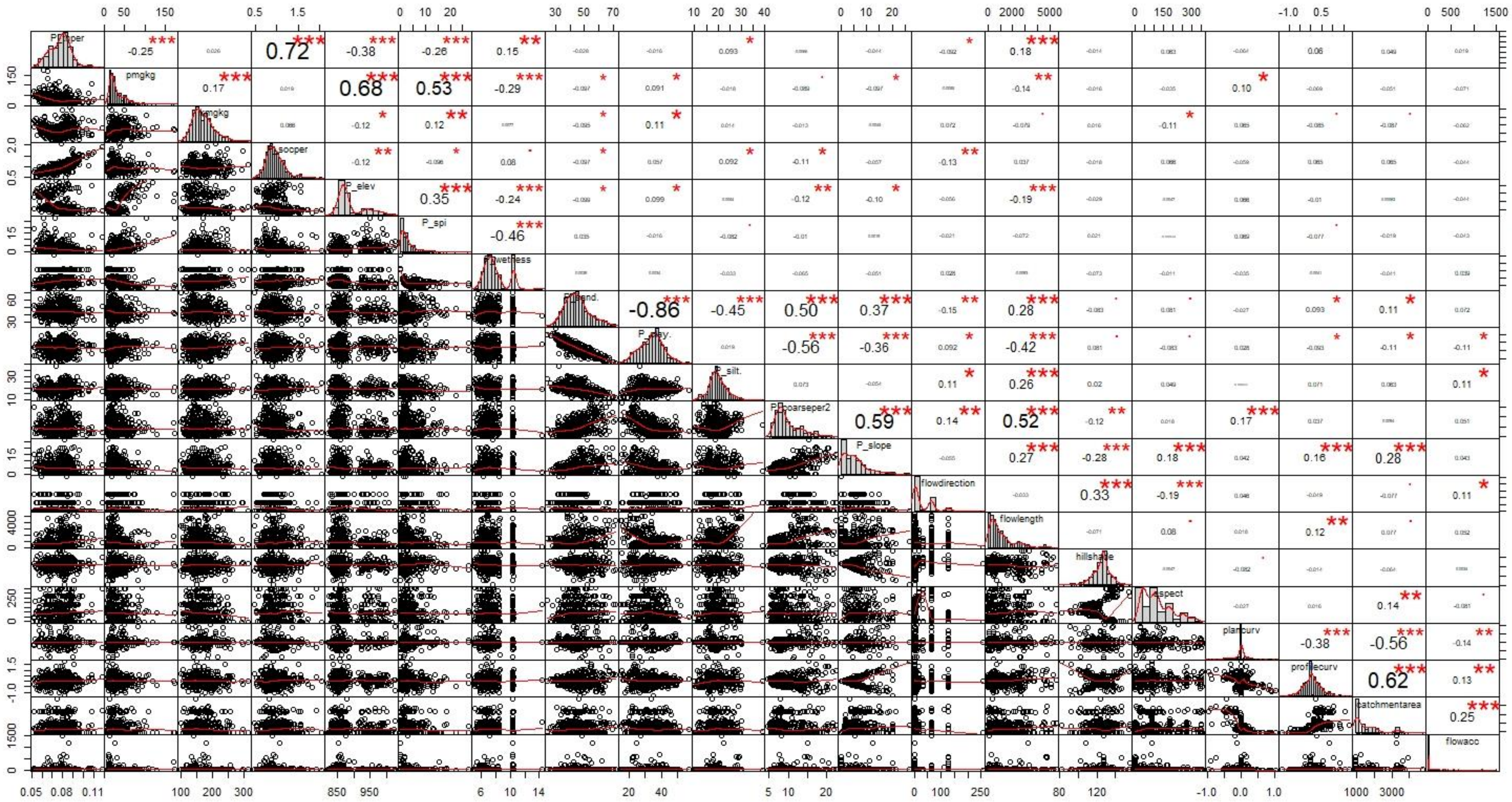
```
; codes for sub-procedures, created by __includes["file.nls"] then
saved first
__includes[
  "P0.Patch&Turtle-Attributes.nls" "P11.Import-Spatial-Data.nls"
  "P12.Import-Sample-Data.nls"  "P123.Generate-coefficients.nls"
  "P22.Nutrient-Stocks&Flows.nls" "P3.Update-Attributes&Processes.nls"
  "P4&5.Scenarios&Iterations.nls" "P6.Draw-graphs.nls"
"P61.Maps&Statistics.nls"]

globals [elapsed-years]
;LANDSCAPE AND HOUSEHOLDS (in file: patch-turtle-attributes.nls )

To SET-UP
   __clear-all-and-reset-ticks
  set elapsed-years 0
  P11.Import-Spatial-Data
  P12.Import-Household-Plot-Data
  P13.Generate-The-Remaining-Population
  P14.Working-policy-variables
  P15.erosion-parameters
  P22.Nutrient-Stocks&Flows
  P226.Calculate-productivity-Index           ;
  P6.Draw-graphs
  show "INITIALIZATION has been done."
End

To GO
tick
set elapsed-years elapsed-years + 1
P311.Update-policy&dynamic-factors
P51.subsidy-scenarios
show (word "Year " ticks " subsidy scenarios done")
;P52.gender&labour-scenarios
;show (word "Year " ticks " gender&labour scenarios done")
P6.Draw-graphs
show (word "Year " ticks  " has elapsed! Beeindrucked!")
if ticks >= stop-when [stop]
End
To P51.subsidy-scenarios
P511.subsidy-current P512.subsidy-reduced
P513.subsidy-universal P514.subsidy-zero
End
To P52.gender&labour-scenarios
P521.baseline
P522.weai-increase
P523.labour-decrease
P524.dependecy-decrease
End
To P4.iterations ;run repetitions based on the random number, set at 3
P41.iteration1 show (word "1st iteration done")
P42….P49
P410.iteration10 show (word "10th iteration done")
P411.sum-iterations
End
```

S4. Correlations among continuous vaiables



S5. Simulated nutrient states and distributions for the centroids of cultivated plots among alternative fertilizer regimes

Figure S4 Simulated fertilizer inputs under three subsidy scenarios (SC = current (28%), SU = universal increased to 70%, SZ= reduced to zero) at year 01, 10 and 20.

Figure S5 Simulated fertilizer input under three subsidy scenarios (SC = current (28%), SU = universal increased to 70%, SZ= reduced to zero) at year 01, 10 and 20.

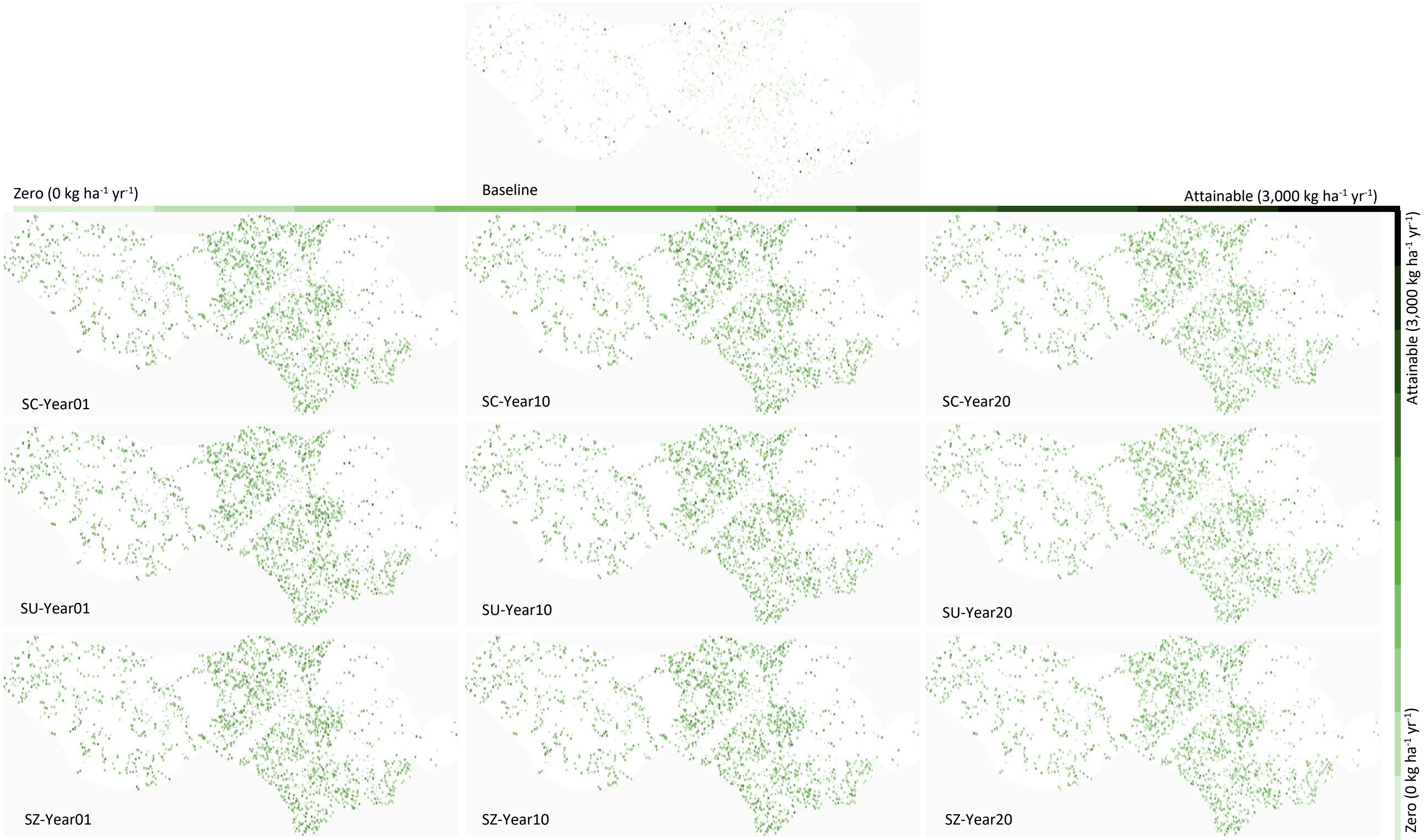


Figure S6 Simulated maize yield under three subsidy scenarios (SC = current (28%), SU = universal increased to 70%, SZ= reduced to zero) at year 01, 10 and 20.

Figure S7 Simulated carbon balance under three subsidy scenarios (SC = current (28%), SU = universal increased to 70%, SZ= reduced to zero) at year 01, 10 and 20.

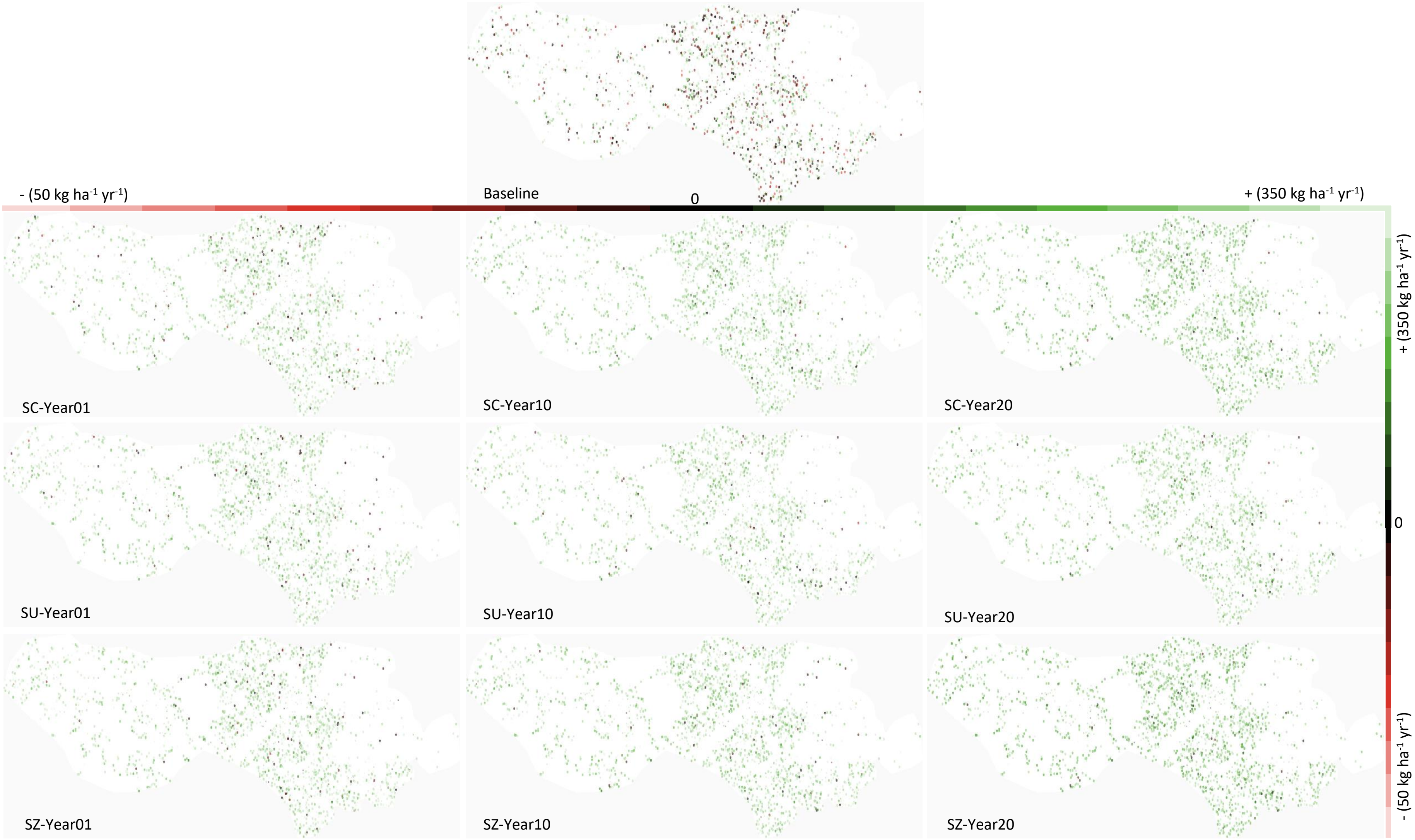


Figure S8 Simulated nitrogen balance under three subsidy scenarios (SC = current (28%), SU = universal increased to 70%, SZ= reduced to zero) at year 01, 10 and 20.

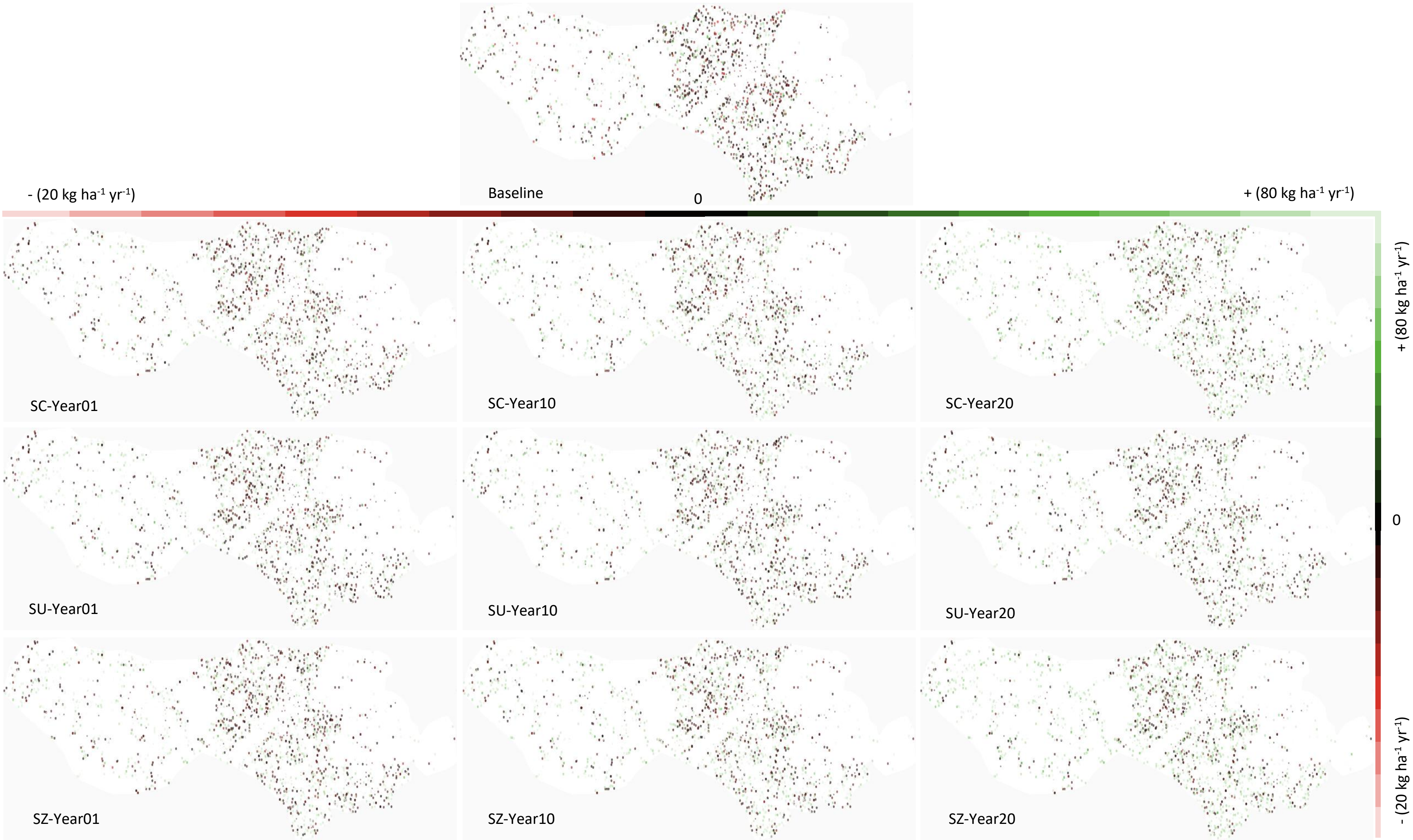


Figure S9 Simulated phosphorus balance under three subsidy scenarios (SC = current (28%), SU = universal increased to 70%, SZ= reduced to zero) at year 01, 10 and 20.

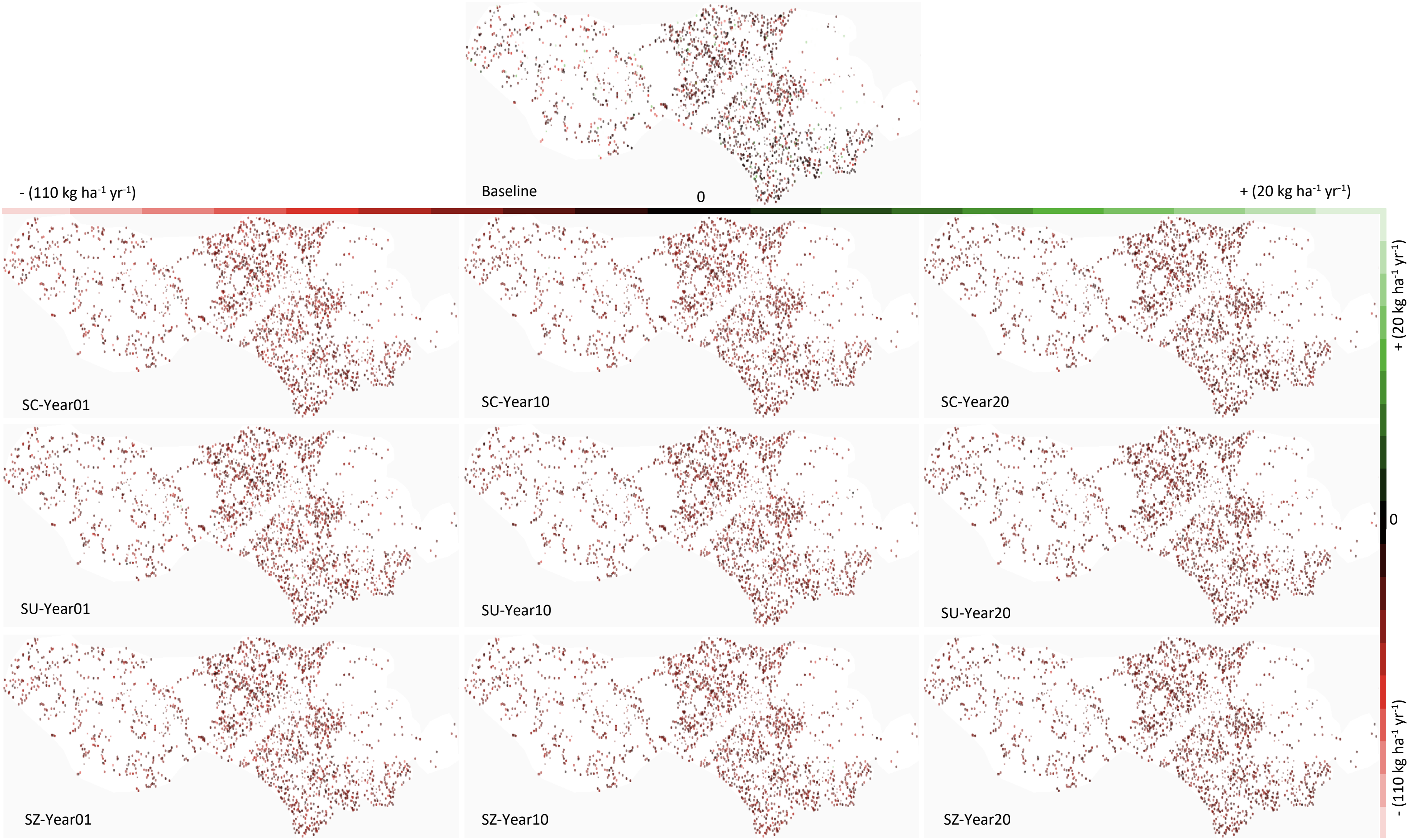


Figure S10 Simulated potassium balance under three subsidy scenarios (SC = current (28%), SU = universal increased to 70%, SZ= reduced to zero) at year 01, 10 and 20.