

# Collecting the Dirt on Soils

## Advancements in Plot-Level Soil Testing and Implications for Agricultural Statistics

*Sydney Gourlay*  
*Ermias Aynekulu*  
*Keith Shepherd*  
*Calogero Carletto*



**WORLD BANK GROUP**

Development Data Group

May 2017

## Abstract

Much of the current analysis on agricultural productivity is hampered by the lack of consistent, high quality data on soil health and how it is changing under past and current management. Historically, plot-level statistics derived from household surveys have relied on subjective farmer assessments of soil quality or, more recently, publicly available geospatial data. The Living Standards Measurement Study of the World Bank implemented a methodological study in Ethiopia, which resulted in an unprecedented data set encompassing a series of subjective indicators of soil quality as well as spectral soil analysis results on plot-specific soil samples for 1,677 households. The goals of the study, which was completed in partnership with the World Agroforestry Centre and the Central Statistical Agency of Ethiopia, were twofold: (1) evaluate the feasibility of integrating a soil survey into household socioeconomic data collection operations, and (2) evaluate local knowledge of farmers in assessing their soil quality. Although a costlier method than subjective assessment, the integration of spectral soil analysis in household surveys has potential

for scale-up. In this study, the first large scale study of its kind, enumerators spent approximately 40 minutes per plot collecting soil samples, not a particularly prohibitive figure given the proper timeline and budget. The correlation between subjective indicators of soil quality and key soil properties, such as organic carbon, is weak at best. Evidence suggests that farmers are better able to distinguish between soil qualities in areas with greater variation in soil properties. Descriptive analysis shows that geospatial data, while positively correlated with laboratory results and offering significant improvements over subject assessment, fail to capture the level of variation observed on the ground. The results of this study give promise that soil spectroscopy could be introduced into household panel surveys in smallholder agricultural contexts, such as Ethiopia, as a rapid and cost-effective soil analysis technique with valuable outcomes. Reductions in uncertainties in assessing soil quality and, hence, improvements in smallholder agricultural statistics, enable better decision-making.

---

This paper is a product of the Development Data Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The authors may be contacted at [gcarletto@worldbank.org](mailto:gcarletto@worldbank.org).

*The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.*

**Collecting the Dirt on Soils: Advancements in Plot-Level Soil Testing and Implications for  
Agricultural Statistics**

*Sydney Gourlay<sup>1</sup>, Ermias Aynekulu<sup>2</sup>, Keith Shepherd<sup>2</sup>, Calogero Carletto<sup>3</sup>*

<sup>1</sup> The World Bank, Washington, D.C., USA

<sup>2</sup> World Agroforestry Centre (ICRAF), Nairobi, Kenya,

<sup>3</sup> The World Bank, Rome, Italy

**Key words:** Land productivity; household survey; soil spectroscopy; soil fertility; local knowledge

**JEL classification:** Q10; Q12; Q24; C81

## 1. Introduction

“Noting that soils constitute the foundation for agricultural development, essential ecosystem functions and food security and hence are key to sustaining life on Earth,” the UN General Assembly declared 2015 the International Year of Soils (A/RES/68/232).<sup>1</sup> The recent increased attention afforded to soil health is for naught, however, if soil health measurements are inaccurate or of inadequate resolution. This is especially critical in the face of increased variability in weather conditions brought on by climate change. Renewed interest in increasing agricultural productivity to meet food security needs and increasing resilience of agricultural systems in developing countries, especially in Sub-Saharan Africa, makes understanding soil fertility constraints and trends ever more important.

Much of the current analysis on agricultural productivity is hampered by the lack of consistent, high quality data on soil health and how it is changing under past and current management. This is beginning to change, however. As soil testing methods become increasingly rapid and affordable, data constraints lessen. In Ethiopia, for example, an innovative national-scale soil mapping operation is underway. The Ethiopia Soil Information Service (EthioSIS) project, supported by the World Bank-funded Agricultural Growth Program and implemented by the Ethiopian Agricultural Transformation Agency, has begun to reveal its value (World Bank, 2016).<sup>2</sup> Although the project has yet to be completed at full scale, knowledge acquired through EthioSIS and disseminated by extension agents has already led to the reformulation of critical inputs and substantial increases in wheat yields (Sawa, 2016). The early successes of EthioSIS illustrate the potential agricultural gains that can be unlocked by improving the detail and geographical scope of soil data.

With an ever-expanding population and finite land resources, soils will become more and more taxed as we strive to produce sufficient food to meet the needs of the world population. Not only will there be a need to increase food production to accommodate the growing population, but at present roughly 795 million people are estimated to be undernourished, 98 percent of whom are in developing regions (FAO et al., 2015). Land, although of finite quantity, can be used more productively, as evidenced by startling yield gaps observed across the world. The magnitude of yield gaps varies significantly across crops and context. For example, Lobell et al. (2009) clearly illustrate the variation in maize yield gaps, as average tropical lowland maize yields in Africa are less than 20 percent of yield potential, while tropical lowland maize yields reach approximately 40 percent of yield potential on average in East and Southeast Asia. Rice exhibits consistently smaller yield gaps, with average rice yields exceeding 80 percent of yield

---

<sup>1</sup> A/RES/68/232: [http://www.un.org/en/ga/search/view\\_doc.asp?symbol=A/RES/68/232&Lang=E](http://www.un.org/en/ga/search/view_doc.asp?symbol=A/RES/68/232&Lang=E).

<sup>2</sup> To learn more about EthioSIS, please visit: <http://www.ata.gov.et/highlighted-deliverables/ethiosis/>.

potential in Bangladesh, Indonesia, and Nepal, among others (Lobell et al., 2009). Insufficient soil health is commonly used to explain, at least partially, said yield gaps (Cassman, 1999; Lobell et al., 2009; Tittonell et al., 2008).

Several methods for closing yield gaps have been identified by the scientific community. According to the FAO, the use of sustainable soil management techniques, such as zero tillage and agroforestry, could boost food production by as much as 58 percent (FAO, 2015). The use of improved crop varieties and chemical input use have been shown to improve productivity and/or resilience exponentially (Cassman, 1999; Duflo et al., 2008). Additionally, Kumar and Quisumbing (2011) draw positive linkages between improved varieties and nutritional status. However, the uptake of such practices has been unenthusiastic, particularly in Sub-Saharan Africa. Marenya and Barrett (2009) suggest that farmer demand for fertilizer use is variable on soil carbon level, with higher carbon content plots achieving greater marginal product of fertilizer, suggesting that soil quality has implications for adoption of fertilizer use. As will be illustrated in this paper, relying on subjective farmer assessments of soil quality as a proxy for carbon content may provide data unsuitable for use in targeting of fertilizer adoption programs.

Productivity has also been observed to vary with farm size. Soil quality has long been argued to explain the inverse farm-size productivity puzzle, which suggests that small farms are more productive than larger farms (Bhalla and Roy, 1988; Barrett et al., 2010; Carletto et al., 2013; Carletto et al., 2015; Lamb, 2003; Tatwangire and Holden, 2013). Despite the results from Barrett et al.'s (2010) experimental study which concluded that soil properties did not explain away the inverse productivity relationship, much research suggests that omitted high quality data on soil properties is at least partially responsible for the inverse relationship (for example, Bhalla and Roy (1988)).

Yield gaps and the *quantity* of crop production are not the only concerns related to soils, food security, and nutrition. The *quality* of food produced can vary, and lack of micronutrients can lead to hidden hunger (Cakmak, 2002; FAO, 2015). With a direct link between the micronutrient content found in crops and the soils from which they grew, soil health measurement and monitoring could lead to the identification and, ideally, prevention of micronutrient malnutrition.

Agricultural analysis is multidimensional. Knowing the quantity of production alone, or even productivity, is not sufficient to analyze determinants of strong yields, estimate adoption of sustainable or improved farm management practices, or establish causal links between agriculture and nutrition. These data, however, are most readily available in household-level surveys with a focus on agriculture, such as

the Living Standards Measurement Study – Integrated Surveys on Agriculture (LSMS-ISA; [www.worldbank.org/lsms](http://www.worldbank.org/lsms)). Historically, plot-level soil statistics derived from household surveys have relied on subjective farmer assessments of soil quality or on linking with soil raster data (when plots are geo-referenced). Direct systematic measurement of soil fertility as part of a large-scale household-level data collection operation has rarely been attempted due to the high costs of soil sampling and analysis. Recently developed rapid low-cost technology for assessing soil characteristics using infrared spectroscopy, however, has increased the potential for direct soil fertility characterization in large studies.

The value of soil data is unquestionable, but the sources, quality, and resolution of such data vary widely. And while platforms like EthioSIS provide invaluable information on soil from an agronomic perspective, having soil data integrated with household-level or plot-level data on input use, farm management practices, agricultural labor, agricultural production, and household socioeconomic characteristics holds extensive analytical value. Soil data from household and farm surveys also provide great opportunities for validation of information obtained through other means. However, the quality of the subjective soil data that are most often found with such inclusive agricultural household surveys has rarely been validated. In this paper, we seek to compare subjective farmer assessment of plot-level soil quality against objective laboratory analyses, by utilizing the data purposively collected for methodological validation under the LSMS Methodological Validation Program.

Using a unique plot-level data set collected by the Living Standards Measurement Study (LSMS) of the World Bank in collaboration with the World Agroforestry Centre (ICRAF) and the Central Statistical Agency of Ethiopia, and with funding from UK Aid, which consists of a menu of subjective farmer-estimated indicators of soil quality and results from objective conventional and spectral soil tests, this paper analyzes the impacts of relying on subjective farmer estimates of soil quality for policy-based decision making through comparison of subjective and objective measures of soil properties. Results from the methodological experiment data suggest that smallholder farmers are unable to clearly discriminate between soil fertility levels, which we hypothesize may partially explain the slow adoption of improved agricultural practices and inputs often observed in Africa.

Building on the few previously existing studies, such as those by Dawoe et al. (2012), Desbiez et al. (2004), Odendo et al. (2010), and Gray and Morant (2003), we aim to validate the use of subjective soil quality indicators against objective measures. Specifically, we compare a multidimensional farmer assessment of soils with plot-level soil analysis conducted using conventional and spectral testing, similar to the data used by Marenya and Barrett (2009).

The remainder of the paper is organized as follows. Section 2 details the specific subjective and objective soil data collected in the Ethiopia Land and Soil Experimental Research (LASER) project and provides descriptive statistics on each. Analytical comparison of the measurement methods is explored in Section 3, with an emphasis on the ability of respondents with various characteristics to more or less accurately assess the quality of their soils against the objective benchmark. Section 4 concludes.

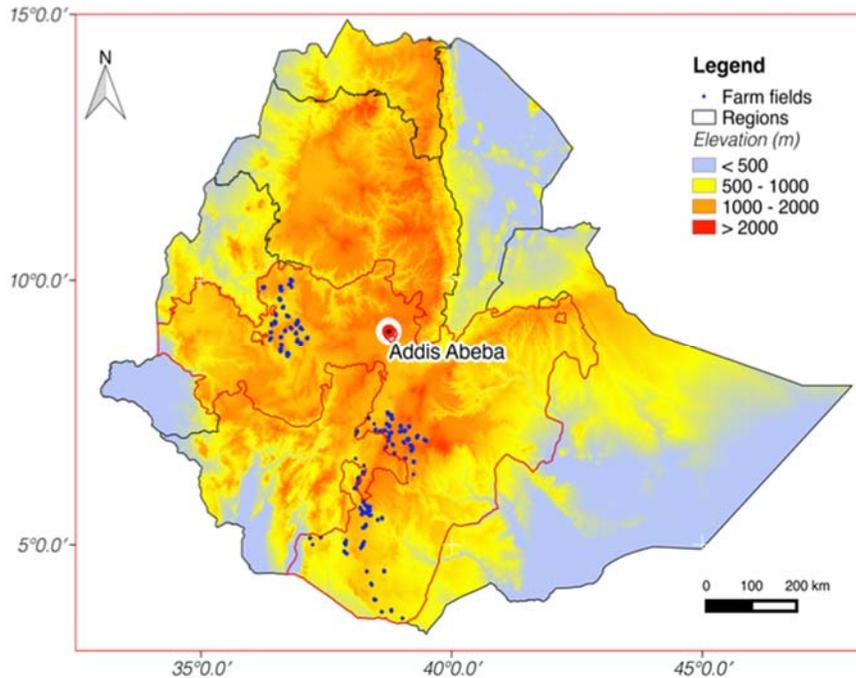
## **2. Data**

### *2.1 LASER Study*

In an effort to collect the highest quality data possible in a large-scale household survey context, the Living Standards Measurement Study has prioritized methodological research in recent years through implementation of the LSMS Methodological Validation Program. With the aim of identifying the magnitude and (potential) systematic nature of measurement error associated with various measurement methods, and with financial support from UK Aid, the LSMS has designed several methodological experiments focused on key aspects of agricultural analysis, including soil fertility. Such methodological experiments strive to find balance between quality and scalability, and ultimately implement the most appropriate methods in future surveys.

Nationally representative LSMS-ISA surveys commonly include basic subjective questions on soil fertility, often asked to the head of household or plot-manager. Additionally, when plots are geo-referenced, indicators of soil health such as nutrient availability, toxicity, and salinity are derived from outside sources including the Harmonized World Soil Database and provided as supplementary data along with the full LSMS-ISA data set. However, in order to know how well the subjective assessments of soil quality correlate with true soil fertility measures, and whether there are any systematic measurement biases based on topography or respondent characteristics, the subjective measures must be taken alongside objective, plot-level measures. This was the motivation behind the Land and Soil Experimental Research (LASER) project.

Data collection for the LASER study was conducted in 3 zones of the Oromia region in Ethiopia (refer to Figure 1). Oromia region was selected because it represents a large area of Ethiopia and encompasses areas with great variation in rainfall, elevation, and agroecological zones. In total, 85 enumeration areas (EAs) were randomly selected using the Central Statistical Agency of Ethiopia's Agricultural Sample Survey (AgSS) as the sampling frame. Within each EA, 12 households were randomly selected from the AgSS household listing completed in September 2013.



**Figure 1.** Location of the study area: markers indicate fields where soil samples are collected from interviewed households in Borena, East Wellega, and West Arsi zones of the Oromia region, Ethiopia.

Fieldwork was conducted in multiple waves. Post-planting activities were conducted during September – December 2013. Post-harvest activities were conducted from January to March 2014. Crop-cutting was conducted at any point during this period when the maize was deemed ready for harvest by the respondent. The post-planting, crop-cutting, and post-harvest questionnaires were administered using computer-assisted personal interviewing.

## 2.2 Farmer Subjective Assessment

Prior to the collection of physical soil samples, a series of subjective plot-level questions was administered to the self-identified ‘best informed’ household member on each plot. These questions ranged from a categorical coded-response “what is the soil quality of [field]?” to questions on soil color, texture, and type (clay, sand, loam, etc.). It is worth noting that the subjective questions were administered at the dwelling, not upon direct respondent observation of the soils, as the study was aimed at assessing farmer knowledge for larger-scale surveys that may not allow for visitation of each plot. Refer to Annex I for the relevant portion of the questionnaire instrument.

**Table 1. Subjective Assessment Summary**

	Plots with Objective Soil Analysis					Full Sample	
	East Wellega	Borena	West Arsi	Total		Total	
N	589	496	592	1677	100%	4149	100%
<i>Soil Quality</i>							
Good	128	276	304	708	42%	1458	35%
Fair	419	193	274	887	53%	2445	59%
Poor	42	27	14	83	5%	246	6%
<i>Soil Color*</i>							
Black	97	128	413	638	38%	1362	33%
Red	403	270	87	761	45%	2170	52%
White/Light	88	84	92	264	16%	592	14%
Yellow	1	14	0	14	1%	24	1%
<i>Soil Type</i>							
Sandy	143	123	93	360	21%	803	19%
Clay	291	250	360	901	54%	2292	55%
Mixture of Sand/Clay	92	120	139	351	21%	785	19%
Other	63	3	0	65	4%	269	6%
<i>Soil Texture°</i>							
Very Fine	4	26	26	56	3%	87	2%
Fine	240	225	404	870	52%	2070	50%
Between Coarse and Fine	259	187	140	586	35%	1584	38%
Coarse	81	56	21	158	9%	390	9%
Very Coarse	5	2	1	8	0%	18	0%

\* Categories "White/Light" and "Yellow" combined for analysis

° Categories "Very Fine" and "Fine" were combined for analysis, as were "Coarse" and "Very Coarse"

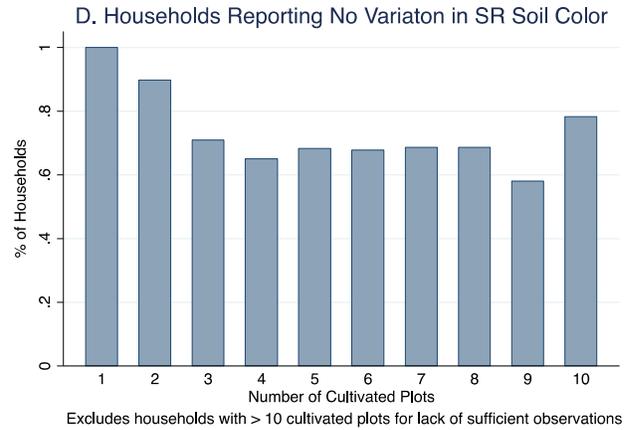
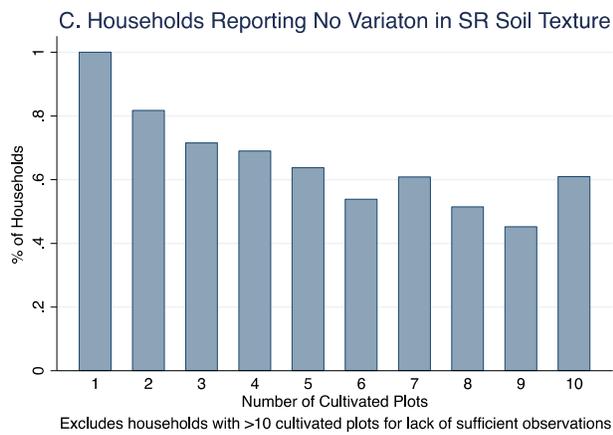
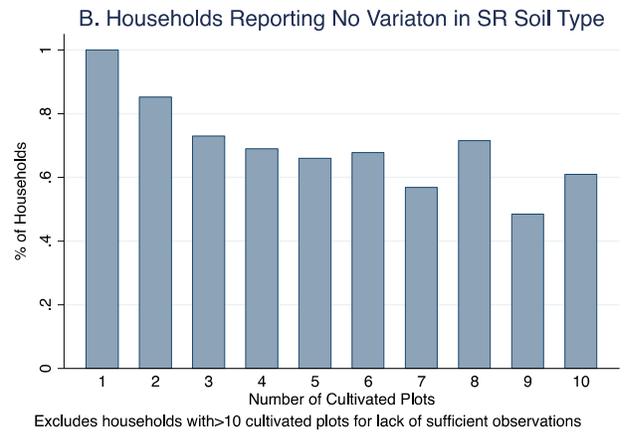
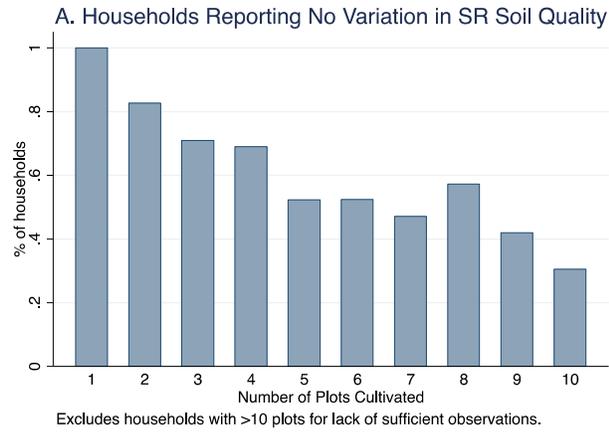
While subjective assessments of soil quality are both cost- and time-efficient, the quality of results may be questionable. Summary statistics of the subjective questions included in the LASER study, found in Table 1, reveal little discrimination by respondents.<sup>3</sup> The table focuses on the sample of plots for which spectral soil analysis was completed, as this is the sample that will be compared to the laboratory results in subsequent sections, but there is also a column for the full sample of plots. When asked about the quality of soil on a particular plot, 94 percent of all plots were reported to have either *good* or *fair* soil (35 percent *good*, 59 percent *fair*). On the whole, only 6 percent of plots were reported to have *poor* soils. This heavy-tailed distribution holds across administrative zone and agroecological zone, with no more than 8 percent of plots in a single agroecological or administrative zone reported as *poor*. This finding is not unique to the LASER data set. In nationally representative LSMS-ISA surveys from Uganda (2013-14), Malawi (2013), and Tanzania (2012-13), only 3 percent, 12 percent, and 6 percent of plots were reported as having poor soil, respectively (UBOS, 2013; Malawi NBS, 2013; Tanzania NBS, 2012). In

<sup>3</sup> The sample is limited to plots in which a topsoil sample was tested in the laboratory. Due to mislabeling of soil samples and/or transportation between the field and the laboratories, 120 plots with subjective measurements do not have matching objective measurements. These observations have been dropped. The number of samples lost to mislabeling was significantly reduced by the use of barcoded labels in a replication study.

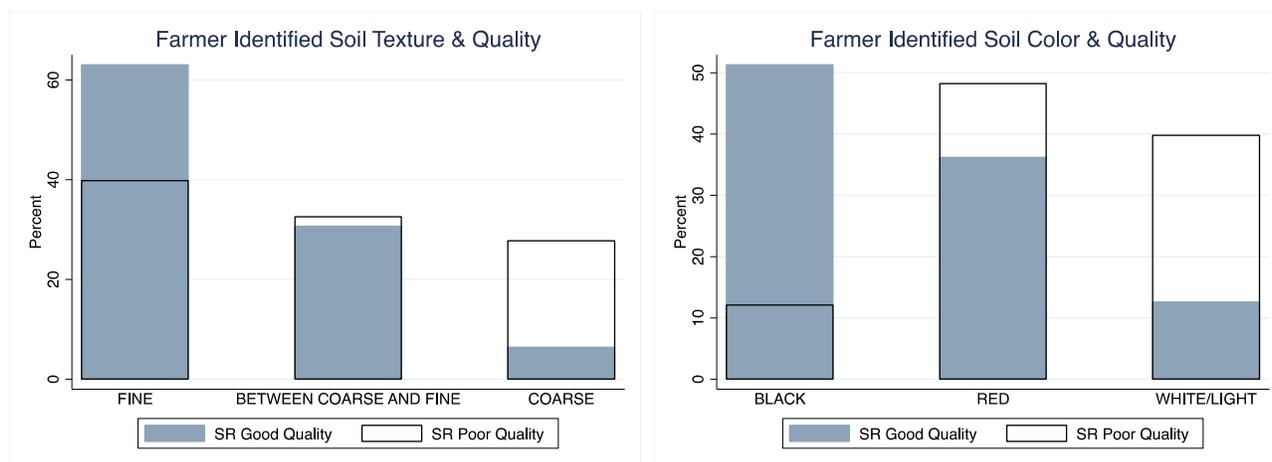
the most recent nationally representative LSMS-ISA survey in Ethiopia (2013-14), nearly 21 percent of *parcels* (not fields, as measured in LASER) were reported as having poor soil quality (CSA, 2013).

Subjective assessments of soil fertility also suffer from a lack of intra-household variation. Of households with more than one cultivated field in the sample, 63 percent reported the same soil quality on all plots. Similarly, 71 percent reported the same soil type, 73 percent reported the same soil color, and 68 percent the same soil texture. This is striking, especially given the high number of fields cultivated per household. Figure 2 (A–D) illustrates the percentage of households reporting no variation in the abovementioned indicators, by number of plots cultivated.

Descriptive analysis suggests that farmers use soil color and texture as indicators of soil quality. As observed in Figure 3, self-reported dark and fine textured soils were categorized as *good* soils while red and course textured soils were more frequently categorized as *poor* soils. While the more specific subjective questions, such as texture and color, appear to be correlated with the overall quality assessments, the value of these questions in terms of correlation with objectively measured soil properties, believed to be the truest measure, remains to be analyzed. Section 3 will explore these correlations.



**Figure 2.** Percent of households reporting no variation in subjective soil quality questions, by number of plots cultivated per household.



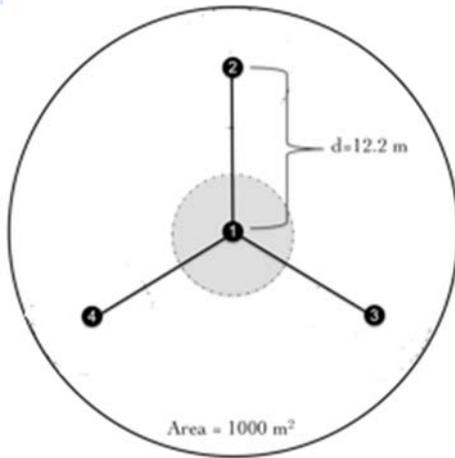
**Figure 3.** Farmers use soil texture (left) and color (right) as indicators of soil quality.

### 2.3 Objective Data

Soil samples were collected from up to two randomly selected plots per household (where applicable, one pure-stand maize plot was selected for crop-cutting). The in-field sampling protocol was designed with ICRAF, adapting the Land Degradation Surveillance Framework of the African Soil Information Service to fit the smallholder farm structure.<sup>4</sup> From each selected plot, two samples were tested: (1) a composite sample collected from four points within the plot at 0-20 cm depth following the layout in Figure 4 (referred to as *topsoil*), and (2) a single sample from the center of the plot at 20-50 cm depth (referred to as *subsoil*). Field staff were trained by ICRAF personnel to ensure comparability of field protocols. A thorough explanation of the soil collection, processing, and analysis protocols followed in LASER are found in the guidebook by Aynekulu et al. (2016).

Soil samples were delivered to local processing laboratories within five days of collection to prevent decomposition of organic matter. Local laboratories, which were also trained on ICRAF protocols, were responsible for drying, grinding, sieving, and weighing the samples. After processing, samples were shipped to ICRAF laboratories in Nairobi, Kenya for analysis. All analyses completed by ICRAF were done following African Soil Information Service (AfSIS) protocols so as to ensure comparability of results with separate pre-existing and ongoing research in the region. On average, soil sample collection took approximately 40 minutes per field. In a replication study, also by the LSMS, this time was reduced by incorporating implementation lessons from LASER, such as using barcoded labels rather than handwritten labels (see the guidebook by Aynekulu et al. (forthcoming) for details).

<sup>4</sup> For more information on the Land Degradation Surveillance Framework see <http://www.africasoils.net/data/lds-description>.



**Figure 4.** Sample plot layout on agricultural plots, with four points (dotted circles). The distance along the radial arms between the center point and the other three points is 12.2 m. Point 1 is the center of the plot. The composite topsoil sample is composed of samples from points 1, 2, 3, and 4.

Two objective measures were employed by ICRAF laboratories. Conventional soil analysis (CSA), which includes traditional wet chemistry methods for soil nutrient extraction and some basic soil physical analyses, was conducted on 10 percent of samples (n=361). Conventional analysis, while often regarded as the gold standard in soil analysis, is expensive and destructive in nature. Spectral soil analysis (SSA), or soil infrared spectroscopy, the second set of tests conducted under the LASER study, is significantly less expensive and non-destructive, allowing for multiple tests over time.

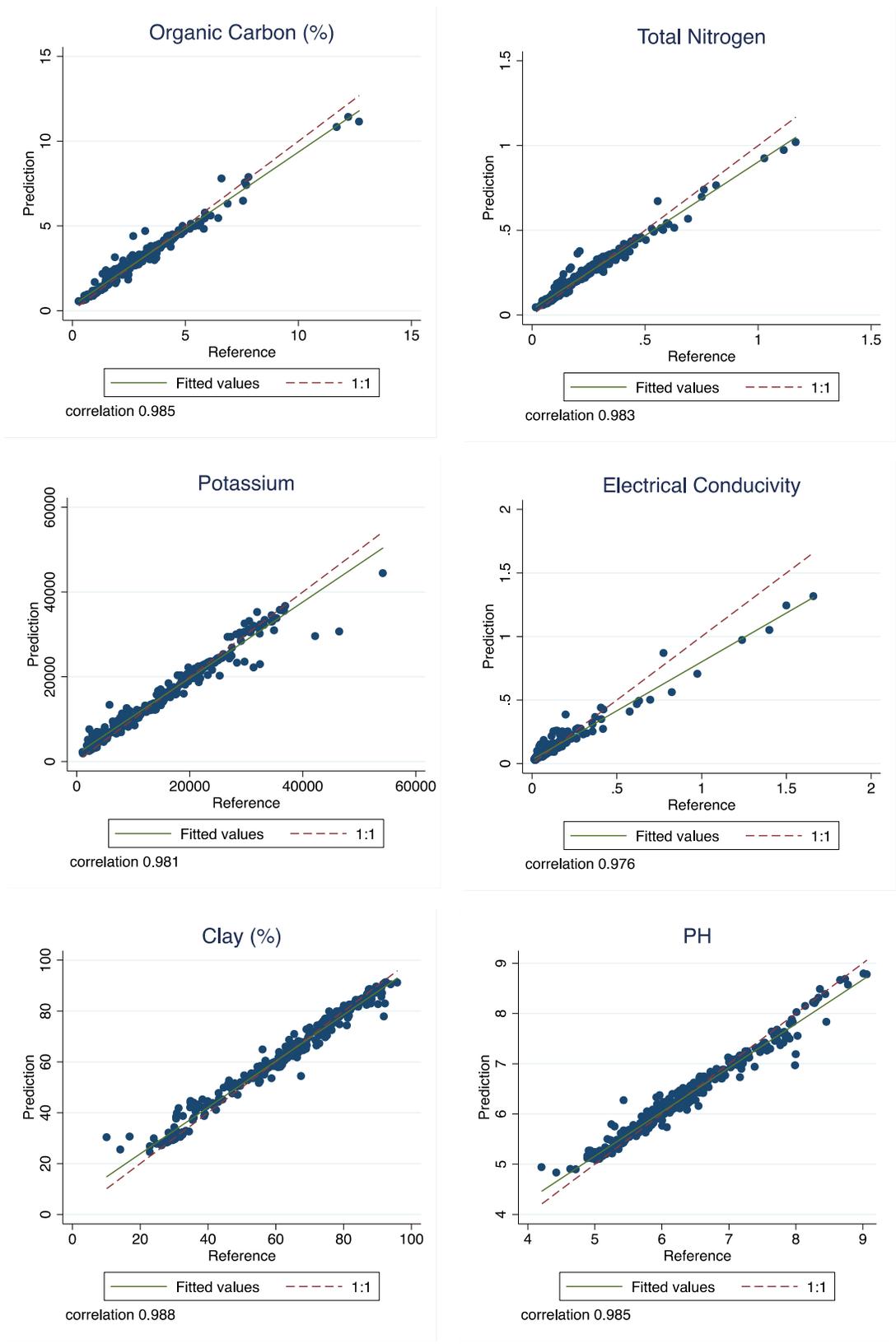
Soil infrared spectroscopy (IR) is an emerging technology that makes large area sampling and analysis of soil health feasible (AfSIS, 2014; Shepherd and Walsh, 2007) and overcomes the current impediments of high spatial variability of soil properties and high analytical costs, which are key challenges in monitoring soil health at a large scale (Conant et al., 2011). A review by Bellon-Maurel and McBratney (2011) showed an exponential increase in the use of near infrared (NIR) and mid-infrared (MIR) reflectance spectroscopy for soil analysis. Because spectral analysis is rapid, it greatly increases the quantity of soil samples that can be processed while also expanding the number of fundamental soil properties that can be simultaneously predicted with little increase in analytical costs. This reduces errors in quantifying soil carbon and other key properties that are often caused by spatial heterogeneity of soils. Infrared data can be integrated with geostatistic data (Cobo et al., 2010), remote sensing data and topographic information for digital soil mapping at the landscape level (Croft et al., 2012). Rossel et al. (2014), for instance, used infrared data to develop a soil carbon map of Australia.

The suite of spectral analyses includes the following tests: mid-infrared diffuse reflectance spectroscopy (MIR), laser diffraction particle size distribution analysis (LDPSA), x-ray methods for soil mineralogy (XRD), and total element analysis (TXRF). MIR and LDPSA spectral tests were conducted on all top- and sub-soil samples (n=3,611), while the x-ray tests, XRD and TXRF, were conducted on the same 10 percent on which conventional testing was executed. Ultimately, approximately 50 variables were predicted for each top and subsoil sample, containing both chemical and physical soil properties.

### *2.3.1 Predictions of Soil Properties from Spectra*

Following the methods designed by Shepherd and Walsh (2002) the results of the CSA were used to predict soil properties onto the full sample based on the spectral signatures, an example of which is found in Annex II. Figure 5 illustrates the predictive power of the mid-infrared spectroscopy on key soil properties, while Table 2 summarizes selected predicted properties, disaggregated by top- and sub-soil. The predictive models are successful in that, of the variables predicted, the lowest correlation between predicted value and actual value (using the reference sample upon which CSA was conducted) was 0.946 (prediction of zinc concentration using Mehlich 3 method). The highest correlation was in the prediction of aluminum concentration by TXRF, with a rho of 0.989. Key soil properties such as total carbon (percent), total nitrogen (percent), clay (percent), and pH were strongly predicted with correlation coefficients of 0.984, 0.983, 0.988, and 0.985, respectively. The near-perfect predictions lend confidence to our assumption that laboratory results obtained through spectral analysis are strong proxies for true measures.

Table 2 illustrates that significant differences are observed between the top- and sub-soil samples, motivating the need to analyze both separately if study objectives and resources allow. The rooting depth of the crop(s) of interest should be considered when determining if top and/or sub-soils should be tested, as it is preferable to test the soil properties at the level at which the plant absorbs the majority of its nutrients (Lorenz and Lal, 2005). Levels of all presented properties are significantly different between top- and sub-soil at the 1 percent level, with the exception of sand percentage, which is significant only at the 10 percent level. In addition to variation across soil depths, levels of key soil properties vary across administrative zone. Figure 6 illustrates that distribution of total carbon and pH by administrative zone. Carbon levels are highest in the West Arsi zone, followed by East Wellega and Borena (means across zones significantly different at the 1 percent level). High carbon and pH variability is observed in Borena, likely due to the great variation in agroecological zones enclosed within its borders. East Wellega has more acidic soils, which could be suitable for maize production (FAO, 1983).



**Figure 5.** Mid-infrared spectroscopy strongly predicts multiple soil properties

**Table 2. Selected Predicted Properties Summary**

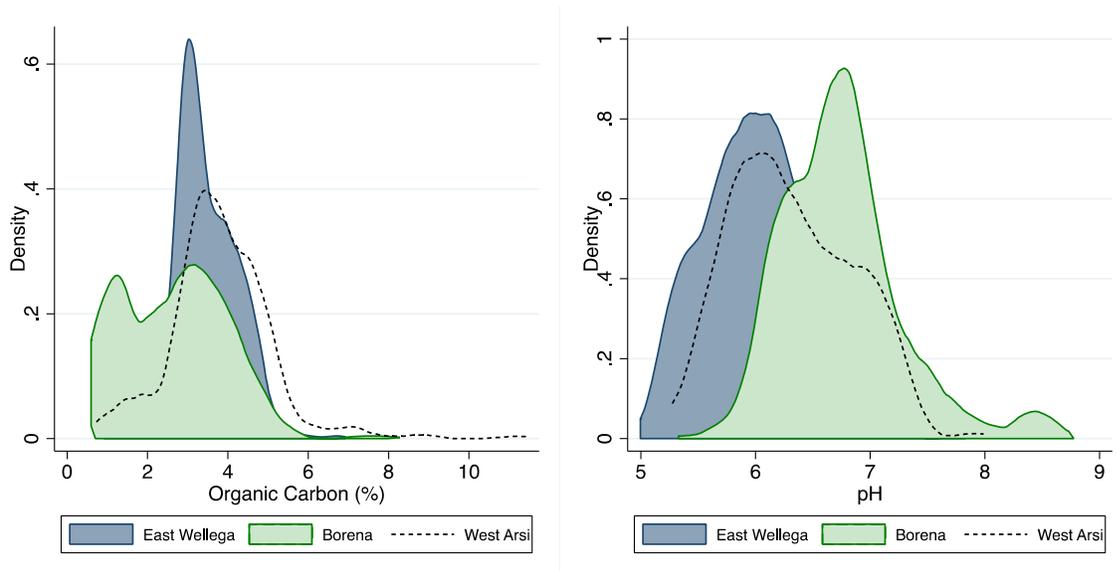
Soil properties	Top Soil (0-20 cm)		Sub Soil (20-50 cm)		Difference in means
	Mean	SD	Mean	SD	
<i>Physical</i>					
% Sand	12.2	7.0	11.8	7.4	0.4*
% Clay	65.0	12.8	67.3	13.3	-2.3***
% Silt	22.6	7.4	20.8	7.5	1.8***
<i>Chemical</i>					
pH	6.3	0.6	6.3	0.6	0.0***
<i>Macronutrients:</i>					
Total Carbon (%)	3.4	1.2	2.9	1.0	0.5***
Total Nitrogen (%)	0.3	0.1	0.2	0.1	0.1***
Exchangeable Calcium (mg kg <sup>-1</sup> ) <sup>+</sup>	3445	1826	3193	1933	252***
Potassium (mg kg <sup>-1</sup> ) <sup>+</sup>	742	297	663	259	79***
Exchangeable Magnesium (mg kg <sup>-1</sup> )	540	192	510	198	30***
<i>Micronutrients:</i>					
Iron (mg kg <sup>-1</sup> ) <sup>+</sup>	160	62	148	55	12***
Zinc (mg kg <sup>-1</sup> ) <sup>+</sup>	5.6	3	5.11	3	0.49***
Exchangeable Manganese (mg kg <sup>-1</sup> )	182	52	173	56	9***

+ Extracted with Mehlich 3 method

\* Extracted with wet method

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Note: Data limited to plots with both top and subsoil samples (n=1599)



**Figure 6.** West Arsi has the highest organic carbon content while soils from East Wellega are more acidic than Borena and West Arsi (top-soil levels reported).

### 3. Comparison of Methods

Given the complexity of soil and the varying needs of different crops and agricultural systems, assessing the overall quality of soil at an objective level can be difficult in itself. Comparing categorical subjective questions to the array of objective measurements and evaluating how well those subjective data reflect the true soil quality is even more challenging. To simplify the process, we first analyze the ability of subjective questions to predict soil carbon levels, a proxy for overall soil health. We attempt to explain which respondent and plot characteristics improve the ability of subjective questions to accurately (or relatively more accurately) assess soil quality. Subsequently, in order to incorporate more of the rich laboratory data and better capture the complex nature of the soil, we construct two variations of soil quality indices. Basic OLS regression is then used to identify which subjective questions, if any, significantly predict changes in the soil quality indicators. Finally, to reinforce the value of plot-level spectral analysis, the spectral results are briefly compared with publicly available geospatial data. All analyses are conducted using top soil (0-20 cm depth) measurements unless otherwise specified.<sup>5</sup>

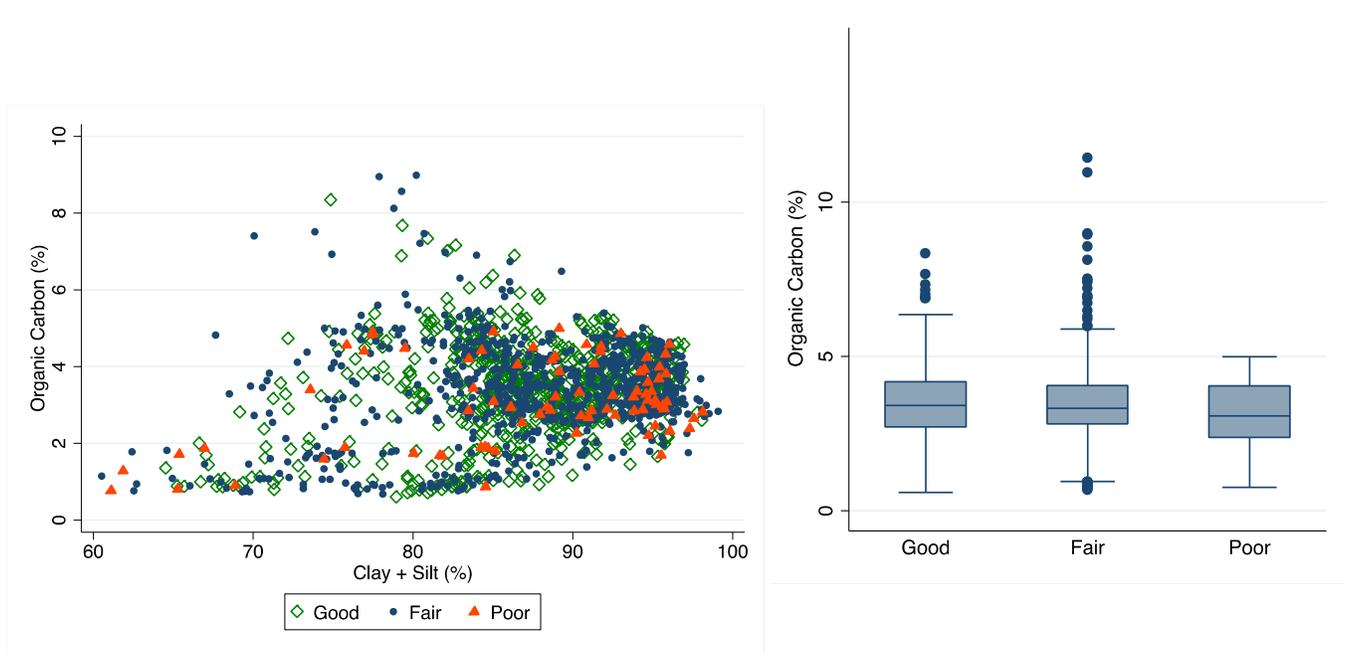
#### 3.1 Carbon as a proxy for overall soil quality

Carbon content is often considered to be the best single indicator of soil quality (IIASA/FAO, 2012). Higher levels of organic carbon indicate greater soil fertility and more optimal soil structure (IIASA/FAO, 2012). Carbon is also highly correlated with other key properties such as total nitrogen (with rho of 0.974 in this data set). Do farmer assessments of overall soil quality reflect carbon levels? Descriptive analysis reveals little relation between organic carbon content (percent) and the respondent's assessment of the soil as *poor*, *fair*, or *good*. As seen above, 42 percent, 53 percent, and 5 percent of the household respondents classified the status of their soil as *good*, *fair*, and *poor*, respectively. T-test results provide weak evidence of distinction between organic (or acidified) carbon content.<sup>6</sup> In plots with reportedly *good* or *fair* soils, there is a greater organic carbon content than in plots reported with *poor* soils (difference is statistically significant at the 10 percent level). There is no significant difference in organic carbon content on plots with *good* and *fair* soils. The significant difference in organic carbon content on *good* and *poor* soils (3.36 percent and 3.10 percent, respectively) is consistent with other

---

<sup>5</sup> The regression analysis found in Section 3.2 was also conducted using sub-soil results. For brevity, the results are not reported here. The findings using sub-soils are largely consistent with those using top soils, however subjective indicators appear to be a slightly better predictor of top soil soil quality indices. Sub soil results available from the authors.

<sup>6</sup> No significant difference is found between total carbon content in plots reported as good, fair, and poor. However, correlation between total carbon and acidified carbon among top soil samples in the LASER data is very high (0.9851).



**Figure 7.** Scatter plot of organic carbon and clay/silt (%) by self-reported soil quality (left), and box plots of organic carbon by self-reported soil quality (right).

studies, such as Desbiez et al. (2004) and Mtambanengwe and Mapfumo (2005), but with those studies finding a greater divergence in organic carbon content between categories.<sup>7</sup>

To better illustrate the distribution of carbon levels across self-reported soil quality categories, Figure 7 presents a scatter plot relating organic carbon, clay and silt content, and self-reported quality category (left) and box plots of carbon levels disaggregated by self-reported quality category (right). The scatter plot reveals that the soils reported as *poor* are not concentrated in areas with low carbon levels, but rather seemingly randomly distributed. This suggests that the local assessment on *overall* soil quality may not be a robust method for mapping soil quality and making decisions on potential interventions like fertilizer recommendations to improve land productivity.

Disaggregating the self-reported indicators by respondent, geographic, and plot characteristics reveals slightly more explanation. Splitting the data into two age categories above and below 40 years (excluding the 68 observations in which the plot manager was not the respondent) shows that the younger respondents were able to differentiate between *poor* and *good* soils ( $p < 0.05$ ), and between *fair* and *poor* soils ( $p < 0.01$ ), but not between *good* and *fair* soils, where we define successful differentiation as a relative measure (higher carbon levels in reportedly better soils). There is no significant difference in total organic carbon levels across the three self-reported soil quality categories for the respondent age group of greater than 40 years. One might expect farmer age to be inversely correlated with education and literacy,

<sup>7</sup> Both Desbiez et al. (2004) and Mtambanengwe and Mapfumo (2005) used a binary classification of plots, rather than ‘good’, ‘fair’, and ‘poor’.

but when disaggregating by manager literacy, there is no significant difference in organic carbon levels between subjective soil quality categories. Disaggregation by manager (and respondent) sex yields less insight. Neither male nor female manager assessments of overall soil quality discriminate by carbon level.

Geographic characteristics, particularly the variation in soils in the immediate vicinity of the household, may play a role in the correlation between subjective assessments of overall soil quality and objectively measured indicators. Overall quality is a highly subjective and relative measure and thus, it is likely to vary with the reference set available to the farmer. That is, in areas with greater variation in soil properties, a farmer may be better able to distinguish between plots that have good, fair, and poor soil because they have a wider range of soils against which they can make comparisons. This theory is supported by the results in Table 3.

Limiting the sample to the enumeration areas with the highest and lowest quartile of variance in organic carbon content indeed suggests that subjective assessment of overall soil quality better approximates organic carbon content in areas with greater variation. In the enumeration areas with the highest quartile

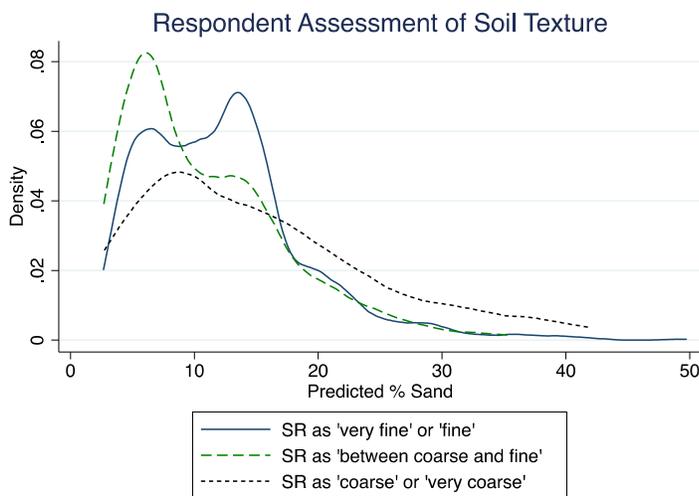
**Table 3.** Subjective Overall Soil Quality and Organic Carbon Content, by Geographic Area (top-soils reported)

	N	Mean Organic Carbon Content (%)	Difference of Means		
			Good	Fair	Poor
All					
Good	708	3.36		-	*
Fair	886	3.37	-		*
Poor	83	3.10	*	*	
EAs with lowest 25% variance					
Good	151	2.66		-	*
Fair	243	2.69	-		*
Poor	23	3.13	*	*	
EAs with highest 25% variance					
Good	184	3.65		*	**
Fair	211	3.94	*		***
Poor	35	3.13	**	***	
West Arsi					
Good	304	3.83		-	-
Fair	274	3.79	-		-
Poor	14	3.53	-	-	
East Wellega					
Good	128	3.56		-	-
Fair	419	3.47	-		-
Poor	42	3.42	-	-	
Borena					
Good	276	2.76		*	-
Fair	193	2.54	*		-
Poor	27	2.38	-	-	

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

of variance in organic carbon content, statistically significant differences are observed in the carbon content of soils reported as *good* and *poor*, *fair* and *poor*, and, to a lesser degree, *good* and *fair*. In enumeration areas with the lowest variance, not only are the differences only marginally significant, but reportedly *poor* soils have a higher mean carbon content than soils reported as *good* and *fair*. Breaking down the sample by administrative zone reveals some support to the idea of variation affecting the ability of farmers to rate the overall quality of their plots, as Borena, the zone with the highest variance, is the only zone in which any significant difference is observed between the three categories, albeit with weak statistical significance.

Farm management practices and property rights may have implications on the ability of respondents to assess overall soil quality. Although there is no significant difference in organic carbon levels between plots that received and did not receive fertilizer (organic or inorganic), there is a difference in the relationship between subjective quality assessments and carbon content. On plots on which fertilizer was *not* used, respondents are better able to distinguish between lower and higher organic carbon levels. On these plots there is a significant difference in carbon levels between plots identified as *good* and *poor*, and *fair* and *poor*, but not between *good* and *fair*. No significant difference is found between plots of different classifications on which fertilizer was used. In a similar trend, plots for which the household holds a title are assessed more appropriately, again with a significant distinction between *good* and *poor*, and *fair* and *poor*, but not between *good* and *fair*. There was no significant distinction on plots without a title, which is potentially explained by reduced knowledge of plots that are not owned and perhaps have not been farmed by the respondent over multiple growing periods.



**Figure 8.** Sand content (%) disaggregated by farmer assessment of soil texture.

Descriptive analysis suggests that on the whole, farmers do not do well at assessing overall soil quality, at least in terms of carbon content. Above, Figure 3 provided evidence that suggested farmers use texture as an indicator of overall soil quality. In fact, there does appear to be a relationship between farmer-reported soil texture and percent sand. Figure 8 plots the distribution of sand concentration in soils reported as fine, coarse, and between coarse and fine, with coarser soils expected to have a higher

concentration of sand as opposed to silt and clay.

The difference in sand concentration is significantly different than zero between all three categories, but the levels are in an unexpected direction as reportedly fine soils have 12.4 percent sand while soils reported between coarse and fine have 11.0 percent sand (reportedly coarse soil has 15.1 percent sand on average). Theoretically, soil texture does have an impact on objective soil quality, with sandy soils having less nutrient holding capacity. The impact of soil texture on soil quality indices is explored in the next section.

### *3.2 Soil quality indices*

While carbon is commonly used as a proxy for soil fertility, it may not be the primary limiting factor of soils in the sample. To achieve a more dynamic measure of soil quality two indices are created. The indices were constructed following the guidance set forth by Mukherjee and Lal (2014) in their comparison of three approaches to soil quality indices. A simple additive and a weighted additive approach were utilized.<sup>8</sup> The indices proposed by Mukherjee and Lal include three components: root development capacity, water storage capacity, and nutrient storage capacity. Data are only available for the construction of the nutrient storage component, which is 40 percent of the complete weighted additive SQI. Therefore, results presented here only indicate constraints related to nutrient storage capacity.

Mukherjee and Lal use their expertise and existing literature to assign linear scores to relevant soil properties ranging from 0 to 3 based on the constraint posed by the level of the specific property (Mukherjee and Lal, 2014). These linear scores are summed to create the simple additive soil quality index (SA SQI). While Mukherjee and Lal assign scores for a multitude of soil properties, data in the LASER study allow for the inclusion of pH, organic carbon content (percent), total nitrogen content (percent), and electrical conductivity, the properties that together make up the nutrient storage capacity component. Unlike the weighted additive index (discussed below), the SA SQI is not normalized on the sample and therefore provides an indicator of overall soil quality that is not relative to the study sample. The SA SQI ranges from 0 to 7, with a mean of 4.61.

The weighted additive index, referred to henceforth as the WA SQI, was constructed by assigning linear scores to the relevant soil properties (pH, soil electrical conductivity, organic carbon (percent), and total nitrogen (percent)), normalizing the scores for each individual property over the sample, and then

---

<sup>8</sup> Additionally, a principal component analysis was conducted following Mukherjee and Lal (2014). This was not the preferred soil quality index method and is therefore not reported. Results available upon request.

applying the indicated weights and summing the scores.<sup>9</sup> The linear scores for each included property ranged from 0 to 1 and were determined by dividing all observations by the highest value in the sample for soil properties in which a higher value is more beneficial (carbon

**Table 4.** Correlation of Soil Quality Indices

	SA SQI	WA SQI	Organic Carbon (%)
SA SQI	1		
WA SQI	0.659	1	
Organic Carbon (%)	0.830	0.800	1

*Note: All significant at the 1% level.*

and nitrogen) and dividing the all observations by the lowest value in the sample for properties in which a lower value is preferred. Soil electrical conductivity and pH have an optimal range, and these were treated as such.<sup>10</sup> This method follows Mukherjee and Lal (2014), who learn from Karlen and Stott (1994) and Fernandes et al. (2011). WA SQI scores range from 0.32 to 0.86, with a mean of 0.46.

Table 4 presents the correlation matrix of the three abovementioned soil quality indices, as well as the organic carbon content. All correlation coefficients are significant at the 1 percent level.

### 3.2.1 Soil quality indices: Regression analysis

In order to determine how well subjective soil indicators are correlated with objective measures, including the soil quality indices and organic carbon content, basic ordinary least squares regression analysis is conducted. The primary objective of the regression analysis is to determine how well subjective soil assessments predict soil quality index measures, and which subjective questions perform best. The following model is executed:

$$SQI_i = \alpha + \beta_1 X_i + \epsilon_i \quad (2)$$

where  $SQI$  is one of the two soil quality indices defined above,  $\alpha$  is a constant,  $X$  is a matrix of subjective soil indicators, and  $\epsilon$  is a random error term with the usual desirable characteristics. Organic carbon content is also run as a dependent variable for robustness. While several factors, such as plot slope and various agricultural practices, may influence the soil quality on the plot, those covariates are excluded

<sup>9</sup> Scores for each of the four soil properties normalized as (observation score – sample min)/(sample max – sample min). Weights were applied as follows: pH (0.3); electrical conductivity (0.3); organic carbon (0.2); total nitrogen (0.2). Scores and weights taken from Mukherjee and Lal (2014).

<sup>10</sup> For properties that have an optimal range, the observations were split into those above and below the critical thresholds (as defined by Mukherjee and Lal, 2014), with those below the threshold treated as though a higher value is preferred and those above the threshold treated as though a lower value is preferred.

from the simple model presented here. The objective is not to analyze the determinants of soil quality but rather to determine how well subjective measures of soil quality predict true measures (as proxied by laboratory results). The models are first run with individual subjective indicators in order to identify how well each variable predicts the index score independently, then with all subjective variables, in order to analyze the predictive power of the subjective indicators as a whole. Note that subjective soil texture was included rather than type, as the “other, specify” category of the soil type question consisted primarily of soil colors and as such, correlation between soil type and color was a concern.

Results of the regression analysis are presented in Table 5. Immediately evident is the low explanatory power of the subjective indicators of soil quality, as expressed by the  $R^2$ , which ranges from 0.002 to 0.060. Specifications (1), (2), and (3), which look at the individual subjective indicators separately, suggest that soil color explains more of the variation in the soil quality indices and carbon content than do the subjective questions on overall quality and texture. The direction of the coefficients on *red* and *white/light* soils are as expected – they have a lower soil quality index score or organic carbon content than black soils. Coarse soils would be expected to have lower levels of nutrient availability, and therefore greater soil fertility, and this is reflected in the results, albeit with limited magnitude in the WA SQI model.

The descriptive analysis on the subjective assessment of overall soil quality invoked little confidence in its relationship with objective soil quality, at least in this particular sample. This sentiment is reflected in the regression analysis. The subjective assessment of overall soil quality had no significant relationship with the WA SQI when self-reported soil color and texture were controlled for. In the model on the SA

**Table 5. SQI Regression Analysis, No Fixed Effects**

Dependent Variable: Specification:	SA SQI				WA SQI				Organic Carbon (%)			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
self-reported soil quality												
fair	-0.024			0.138*	-0.009**			-0.002	0.003			0.143**
poor	-0.375*			-0.011	-0.017**			-0.005	-0.265**			0.023
self-reported color (collapsed, 'black' omitted)												
red		-0.623***		-0.640***		-0.033***		-0.032***		-0.570***		-0.588***
white/light		-0.889***		-0.866***		-0.027***		-0.025***		-0.609***		-0.586***
self-reported soil texture (collapsed, 'fine' omitted)												
between coarse and fine			-0.067	-0.02			-0.008*	-0.005			-0.054	-0.022
coarse			-0.428***	-0.239*			-0.015**	-0.009			-0.379***	-0.265**
constant	4.640***	5.039***	4.675***	5.001***	0.461***	0.475***	0.460***	0.478***	3.364***	3.712***	3.409***	3.674***
N	1677	1677	1677	1677	1677	1677	1677	1677	1677	1677	1677	1677
R <sup>2</sup>	0.002	0.048	0.006	0.052	0.005	0.041	0.004	0.043	0.002	0.053	0.008	0.060
Independent Var Mean:			4.61				0.46				3.35	
Independent Var. Std Dev:			1.59				0.08				1.22	

Robust standard errors.

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

SQI and organic carbon, the results in specification (4) suggest that soils reported as *fair* are of greater quality than those reported as *good*.

The results presented in Table 5 do not control for inter-household differences. Within the full sample, and not controlling for differences across households, the subjective indicators of soil quality do not exhibit strong predictive power of the soil quality indices or organic carbon content, but there is *some* relationship. Looking strictly at intra-household effects by including household fixed effects (and limiting the sample to households which had top soil samples for two plots), suggests that *within* household, subjective indicators have even less relationship with soil quality indices (refer to Table 6). After controlling for household fixed effects the strength of the model is reduced, as evidenced by the lower R<sup>2</sup> values. This is potentially associated with the lack of intra-household variation of subjective indicators illustrated previously.

Although the amount of variation in the soil quality indices and carbon content explained by the subjective indicators falls with the inclusion of fixed effects, there is one positive outcome. The coefficients on subjective overall soil quality gain statistical significance and move in the right direction, with self-reported *poor* soil possessing a negative coefficient (marginally significant in the WA SQI and carbon models, not significant in the SA SQI model), suggesting that plots may be ranked appropriately within households. Overly positive conclusions on the ability of subjective questions to reflect soil quality should not be drawn from this result, however, given the lack of intra-household variation observed and the low magnitude of the coefficients.

Several differences are observed in relationship between subjective indicators and the various soil quality

**Table 6.** SQI Regression Analysis, Household Fixed Effects

Dependent Variable:	SA SQI				WA SQI				Organic Carbon (%)			
Specification:	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
self-reported soil quality												
fair	-0.188			-0.172	-0.015**			-0.012*	-0.166*			-0.138
poor	-0.109			-0.096	-0.033***			-0.030**	-0.404**			-0.364*
self-reported color (collapsed, 'black' omitted)												
red		-0.049		-0.041		-0.014		-0.013		-0.151		-0.138
white/light		0.165		0.191		0.004		0.006		-0.192		-0.171
self-reported soil texture (collapsed, 'fine' omitted)												
between coarse and fine			-0.252*	-0.238			-0.021**	-0.020**			-0.219**	-0.201*
coarse			0.047	0.096			-0.027*	-0.023			-0.222	-0.200
constant	4.841***	4.732***	4.822***	4.899***	0.467***	0.463***	0.467***	0.479***	3.548***	3.538***	3.540***	3.711***
N	1384	1384	1384	1384	1384	1384	1384	1384	1384	1384	1384	1384
R <sup>2</sup>	0.003	0.002	0.006	0.011	0.010	0.005	0.010	0.023	0.009	0.002	0.006	0.016

Robust standard errors.

All specifications include HH Fixed Effects.

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

indices. The SA SQI, which is not normalized on the sample, exhibits a weaker relationship with subjective indicators than the WA SQI, particularly when household fixed effects are included. This is likely explained by the fact that the WA SQI is normalized on the sample and, therefore, is more a measure of relative soil quality.

### 3.2.2 Spectral Analysis & Geospatial Data

To provide further confidence in the value of conducting spectral soil analysis on plot-level soil samples, a brief comparison is made with publicly available geospatial data. Admittedly, comparison may be made with more than one source of geospatial data. However, the AfSIS data have among the highest resolutions currently available in public data sets (250m) for Ethiopia (for details see Hengl et al, 2015). They also may be the most comparable to the LASER data in a methodological sense considering both are conducted by the ICRAF. For these reasons, the comparisons made here may present an upper bound of comparability, at least in this particular context. Values were extracted from the AfSIS geospatial data set using the GPS coordinates of the specific plots. Comparison is made between organic carbon content (percent) as measured by plot-level spectral testing and that indicated in the AfSIS map.

Table 7 summarizes the mean organic carbon content observed in LASER and AfSIS. In the full sample, the difference in means between the two data sets is statistically different from zero at the 1 percent level. Although the magnitude of the difference may be immaterial depending on the research question of

interest, it is important to note that the correlation between the two measures is only 0.586.

Concerns with the use of geospatial data are often related to their (in)ability to capture variation in soil properties within small areas. Indeed, a closer look at the correlation between the spectral analysis and the geospatial data reveals that the correlation falls when limiting the sample to the EAs with the highest quartile of variance in carbon content (as measured by spectral analysis). In EAs with the highest variance, correlation is only 0.387, while in EAs with the lowest

**Table 7. Soil Organic Carbon (%) in LASER and**

	Organic Carbon (%)		Difference in Means
	LASER	AFSIS	
<b>All Top-Soil</b>			
Mean	3.350	3.476	***
Standard Dev.	1.216	1.192	
Correlation		0.586	
N		1674	
<b>EAs with lowest 25% variance</b>			
Mean	2.775	2.884	**
Standard Dev.	1.200	1.045	
Correlation		0.715	
N		412	
<b>EAs with highest 25% variance</b>			
Mean	3.739	3.767	-
Standard Dev.	1.418	1.189	
Correlation		0.387	
N		429	

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Difference in means: T-test

variance, the correlation is 0.715 (refer to Table 7).

#### **4. Conclusions**

Knowledge of soil quality indicators and overall health is becoming increasingly important as food security issues become more pressing and climate change threatens to change the face of agriculture. Soil health, both perceived and actual, can have impacts on the targeting and uptake of improved agricultural practices, which can improve both the quality and quantity of food produced. For instance, Marenya and Barrett (2009) show that fertilizer effectiveness, and in turn, demand, is dependent on soil carbon content. However, results of the LASER study suggest that subjective soil quality indicators fail to effectively reflect true levels of organic carbon, thereby limiting the value of subjective assessments of soil quality in policy making.

Certainly, asking a farmer to categorically rate overall soil quality has limited benefit. This particular subjective soil quality question does not successfully distinguish between soil carbon levels or predict soil quality index scores, at least within this sample in Ethiopia. Subjective questions on soil color and texture were more effective in predicting soil quality index scores and organic carbon content, although the low explanatory power of these variables leaves much to be desired. The value of subjective soil quality indicators is further questioned by the severe lack of intra-household variation observed. Further research validating different subjective questions, potentially formulated with soil scientists, may yield more optimistic results. However, the questions included in the LASER study are those that have been historically included in LSMS-ISA surveys in multiple countries.

From a fieldwork implementation standpoint, the experience of the LASER study gives promise that the integration of soil spectroscopy into socioeconomic household panel surveys is feasible. The methodology is a relatively rapid and cost-effective soil measurement technique that could unlock further understanding of the effects of farm management practices and changes in soil health over time. Detailed guidance on implementation strategies and protocols implemented in the LASER study can be found in Aynekulu et al. (2016).

Despite the weak correlation observed here between laboratory analysis and subjective assessment, several studies find subjective assessments of soil quality to be a significant determinant of plot-level productivity (for example, Carletto et al., 2013). This suggests that if subjective soil quality assessments are not capturing true soil properties, they must be capturing something else relevant to agricultural production. As a potential explanation for this, we echo the sentiments of Tittonell et al. (2008) and others, who suggest that farmers have a ‘holistic’ view of soils, and that rather than assessing the soil properties explicitly, they often incorporate other components such as overall agricultural productivity

and likelihood of crop theft, for example. This finding would indeed render subjective assessments of soil quality significant predictors of agricultural productivity, but largely leaving true soil quality omitted.

Additional research is needed (and ongoing) to determine the effects of including these objectively measured soil properties in productivity analysis as opposed to, or in addition to, subjective assessments. Additionally, while a brief comparison of plot-level spectral analysis and AfSIS geospatial data was included to illustrate the ability of plot-level analysis to capture a greater degree of variation within small areas, further research in this arena would be valuable, including, for instance, a comparison of LASER results with the EthioSIS national soil map. Geospatial data on soil quality have been recently compared with subjective data by Kelly and Anderson (2016), who find a similar pattern in that farmers are often over-optimistic about the fertility of their soils with respect to the Harmonized World Soil Database. This line of work could be extended to include plot-level soil analysis and further validate the need for objective plot-level analysis.

Ethiopia is poised to benefit greatly from advancements in soil testing, particularly with the rollout of projects like EthioSIS combined with the upscaling of data collection efforts at the farm household level. The results of the LASER study, which bring subjective estimates of soil quality under scrutiny and point to the need for more direct, yet practical, soil measurements, show the potential value of the complementarities between platforms like EthioSIS, and household-level data collection, based on which accurate soil information can be made available as part of rich data sets on the socioeconomic condition and farming practices of farming units. Soil data collection through household and farm surveys may also provide a much needed vehicle to groundtruth remote sensing information and calibrate soil models. In this vein, fostering stronger linkages between national EthioSIS soil data and surveys like the Ethiopian Rural Socioeconomic Survey, a household panel survey supported by the LSMS-ISA, offers great opportunities from the research and operational perspectives.

Evidence from the Ethiopia LASER study suggests that subjective farmer assessments of soil quality poorly explain objective laboratory results and lack intra-household variation. Spectral analysis has been proven to near-perfectly predict key soil parameters as measured by conventional wet chemistry methods while providing highly detailed data that can be useful in policy aimed at increasing agricultural output, such as fertilizer input programs and identifying optimal crop selection, as well as agricultural productivity analysis. Improving agricultural statistics by reducing the uncertainties in soil quality assessment via objective measurement can enable better decision-making, both at micro and macro levels.

## References

- AfSIS. 2014. The Africa soil information service. <<http://www.africasoils.net>> (Accessed 2015.07.10)
- Aynekulu, E., Carletto, C., Gourlay, S., Shepherd, K. (forthcoming). *Spectral Soil Analysis & Household Surveys: A Guidebook for Integration*. The World Bank, Washington, D.C., and World Agroforestry Centre, Nairobi.
- Aynekulu, E., Carletto, C., Gourlay, S., Shepherd, K. (2016). *Soil Sampling in Household Surveys: Experience from Ethiopia*. The World Bank, Washington, D.C., and World Agroforestry Centre, Nairobi. Available at: <http://go.worldbank.org/PYIZRV60K0>
- Barrett, C.B., Bellemare, M.F., and Hou, J.Y. (2010). “Reconsidering Conventional Explanations of the Inverse Productivity-Size Relationship”, *World Development*, 38(1), pp. 88-97.
- Bellon-Maurel V., McBratney A. 2011. Near-infrared (NIR) and mid-infrared (MIR) spectroscopic techniques for assessing the amount of carbon stock in soils e Critical review and research perspectives. *Soil Biology & Biochemistry* 43:1398-1410.
- Bhalla, S. and P. Roy (1988). ‘Mis-Specification in Farm Productivity Analysis: The Role of Land Quality’, *Oxford Economic Papers, New Series* 40(1): 55-73.
- Cakmak, I. (2002). “Plant nutrition research: Priorities to meet human needs for food in sustainable ways,” *Plant and Soil*, 247, pp. 3-24.
- Carletto, C., S. Savastano and A. Zezza (2013). ‘Fact or Artefact: The Impact of Measurement Errors on the Farm Size-Productivity Relationship’, *Journal of Development Economics*, 103: 254–26.
- Carletto, G., Gourlay, S., and Winters, P. (2015). “From Guesstimates to GPStimates: Land Area Measurement and Implications for Agricultural Analysis,” *Journal of African Economies*, first published online May 29, 2015 doi:10.1093/jae/ejv011.
- Cassman, K.G. (1999). “Ecological intensification of cereal production systems: Yield potential, soil quality, and precision agriculture,” *Proceedings of the National Academy of Sciences of the United States of America*, 96(11), pp: 5952-5959; doi:10.1073/pnas.96.11.5952.
- Cobo JG, Dercon G, Yekeye T, Chapungu L, Kadzere C, Murwira A, Delve R, Cadisch J. (2010). Integration of mid-infrared spectroscopy and geostatistics in the assessment of soil spatial variability at landscape level. *Geoderma* 158:398–411.
- Conant RT, Ogle S, Paul EA, Paustian K. (2011). Measuring and monitoring soil organic carbon stocks in agricultural lands for climate mitigation. *Frontiers in Ecology and the Environment* 9:169–173.
- Croft H, Kuhn NJ, Anderson, K. (2012). On the use of remote sensing techniques for monitoring spatio-temporal soil organic carbon dynamics in agricultural systems. *Catena* 94: 64–74
- CSA (Central Statistical Agency of Ethiopia). (2013). Dataset from Ethiopia Socioeconomic Survey (2013-14) / Living Standards Measurement Study – Integrated Surveys on Agriculture. Retrieved from: <http://go.worldbank.org/ZK2ZDZYDD0>

- Dawoe, E.K., Quashie-Sam, J., Isaac, M.E. and Oppong, S.K. (2012). “Exploring farmers’ local knowledge and perceptions of soil fertility and management in the Ashanti Region of Ghana”, *Geoderma*, Elsevier B.V., Vol. 179-180, pp. 96–103.
- Desbiez, A., Matthews, R., Tripathi, B. and Ellis-Jones, J. (2004). “Perceptions and assessment of soil fertility by farmers in the mid-hills of Nepal”, *Agriculture, Ecosystems & Environment*, 103(1), pp. 191–206.
- Duflo, E., Kremer, M. and Robinson, J. (2008). “How High Are Rates of Return to Fertilizer? Evidence from Field Experiments in Kenya,” *American Economic Review: Papers & Proceedings*, 98(2), pp. 482-488.
- FAO. (1983). *Guidelines: Land Evaluation for Rain-fed Agriculture*. Soil Bulletin No 52, Food and Agriculture Organization, Rome, pp. 237.
- FAO. (2015). “Healthy soils are the basis for healthy food production,” I4405E/1/02.15. Available from: <http://www.fao.org/3/a-i4405e.pdf> (Accessed 07/21/2015).
- FAO, IFAD and WFP. (2015). *The State of Food Insecurity in the World 2015. Meeting the 2015 international hunger targets: taking stock of uneven progress*. Rome, FAO.
- Fernandes, J.C., Gamero, C.A., Rodrigues, J.G.L., and Miras-Avalos, J.M.. (2011). Determination of the quality index of a Paleudult under sunflower culture and different management systems. *Soil and Tillage Research*, 112: pp. 167–174.
- Gray, L.C. and Morant, P. (2003). “Reconciling indigenous knowledge with scientific assessment of soil fertility changes in southwestern Burkina Faso”, *Geoderma*, 111(3-4), pp. 425–437.
- Hengl T, Heuvelink GBM, Kempen B, Leenaars JGB, Walsh MG, Shepherd KD, et al. (2015) Mapping Soil Properties of Africa at 250 m Resolution: Random Forests Significantly Improve Current Predictions. *PLoS ONE* 10(6): e0125814. doi:10.1371/journal.pone.0125814
- IIASA/FAO. (2012). *Global Agro-ecological Zones (GAEZ v3.0)*. IIASA, Laxenburg, Austria and FAO, Rome, Italy.
- Karlen, D.L., and Stott, D.E. (1994). A framework for evaluating physical and chemical indicators of soil quality. In: Doran JW, Coleman DC, Bezdicek DF, Stewart BA, editors. *Defining soil quality for a sustainable environment*. Madison, WI: Soil Science Society of America. pp. 53–72.
- Kelly, Allison C. and C. Leigh Anderson. (2016). “Comparing farmer and measured assessments of soil quality in Tanzania: Do they align?”, *Journal of Natural Resources and Development*, 06, pp. 55-65.
- Kumar, N., and A. R. Quisumbing. (2011). “Access, Adoption, and Diffusion: Understanding the Long-Term Impacts of Improved Vegetable and Fish Technologies in Bangladesh.” *Journal of Development Effectiveness* 3 (2): 193–219.
- Lamb, R.L. (2003). ‘Inverse Productivity: Land Quality, Labor Markets, and Measurement Error’, *Journal of Development Economics*, 71(1): 71-95.

- Lobell, D.B., Cassman, K.G., and Field, C.B. (2009). “Crop Yield Gaps: Their Importance, Magnitudes, and Causes,” *Annual Review of Environment and Resources*, 34, pp: 179-204.
- Lorenz, K., & Lal, R. (2005). The depth distribution of soil organic carbon in relation to land use and management and the potential of carbon sequestration in subsoil horizons. *Advances in agronomy*, 88, 35-66.
- Malawi NBS (National Bureau of Statistics). (2013). Dataset from Integrated Household Panel Survey (2013) / Living Standards Measurement Study – Integrated Surveys on Agriculture. Retrieved from: <http://go.worldbank.org/NOXNI9YDS0>
- Marenya, P., & Barrett, C. B. (2009). “Soil quality and fertility use rates among smallholder farmers in western Kenya”, *Agricultural Economics*, 40(5), pp. 561-572.
- Mtambanengwe, F. and P. Mapfumo. (2005). “Organic matter management as an underlying cause for soil fertility gradients on smallholder farms in Zimbabwe”, *Nutrient Cycling in Agroecosystems*, 73, pp. 227-243.
- Mukherjee A., and R. Lal. (2014). Comparison of Soil Quality Index Using Three Methods. PLoS ONE 9(8): e105981. doi:10.1371/journal.pone.0105981
- Odendo, M., Obare, G. and Salasya, B. (2010). “Farmers’ perceptions and knowledge of soil fertility degradation in two contrasting sites in western Kenya”, *Land Degradation & Development*, 21(6), pp. 557–564.
- Rossel RAV, Webster R, Bui EN, Baldoc JA. (2014). Baseline map of organic carbon in Australian soil to support national carbon accounting and monitoring under climate change. *Global Change Biology* 20:2953–2970
- Sawa, P. (2016, September 26). Ethiopia soil map arms farmers with new fertilisers in climate fight. *Reuters*. Retrieved from <http://www.reuters.com/article/us-ethiopia-climatechange-agriculture-idUSKCN11Z197>.
- Shepherd, K.D., and Walsh, M.G. (2002). Development of reflectance spectral libraries for characterization of soil properties. *Soil Sci. Soc. Am. J.* 66:988–998.
- Shepherd, K.D., and Walsh, M.G. (2007). Infrared spectroscopy-enabling an evidence-based diagnostic surveillance approach to agricultural and environmental management in developing countries. *Journal of Near Infrared Spectroscopy* 15:1-19.
- Tanzania NBS (National Bureau of Statistics). (2012). Dataset from Tanzania National Panel Survey (2012-13) / Living Standards Measurement Study – Integrated Surveys on Agriculture. Retrieved from: <http://go.worldbank.org/EJMAC1YDY0>
- Tatwangire, A. and Holden, S. (2013). Land Tenure Reforms, Land Market Participation and the Farm Size – Productivity Relationship in Uganda. In *Land Tenure Reform in Asia and Africa*, pp.187–210.
- Tittonell, P., Shepherd, K.D., Vanlauwe, B., and K.E. Giller. (2008). “Unravelling the effects of soil and crop management on maize productivity in smallholder agricultural systems of western Kenya - An

application of classification and regression tree analysis”, *Agriculture, Ecosystems and Environment*, 123, pp. 137-150.

UBOS (Uganda Bureau of Statistics). (2013). Dataset from Uganda National Panel Survey (2012-13) / Living Standards Measurement Study – Integrated Surveys on Agriculture. Retrieved from: <http://go.worldbank.org/FS2M7AYE00>

World Bank. (2016, November 8). Scaling Climate-Smart Agriculture in Ethiopia, from the Ground Up. Retrieved from <http://www.worldbank.org/en/news/feature/2016/11/08/scaling-climate-smart-agriculture-in-ethiopia-from-the-ground-up>.

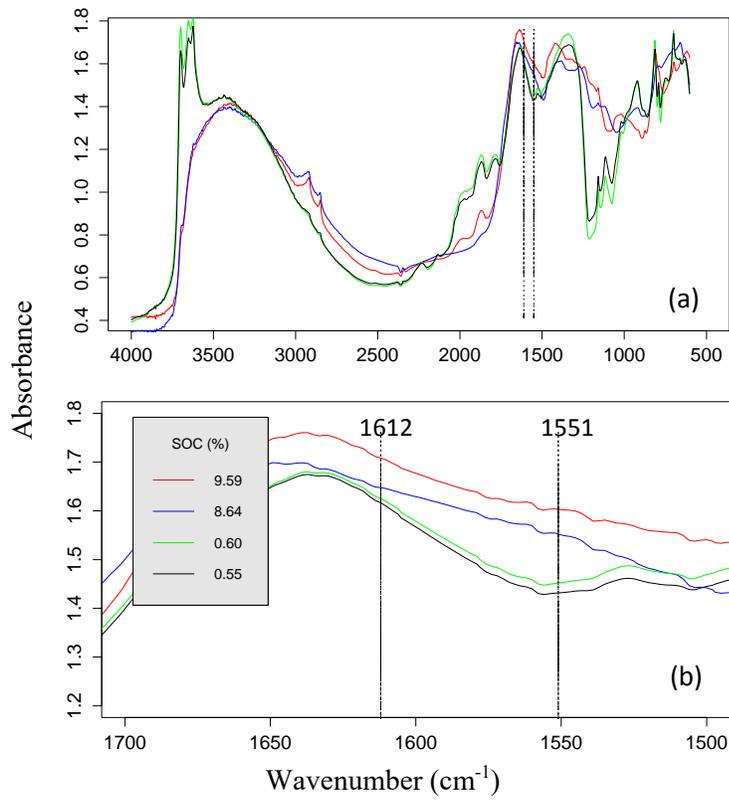
Annexes

Annex I. Subjective Soil Questionnaire Excerpt

PP SECTION 6: FIELD DETAILS  
THIS SECTION TO BE COMPLETED FOR ALL CULTIVATED FIELDS. THE RESPONDENT SHOULD BE THE BEST INFORMED.

1. FIELD DESCRIPTION		2. RESPONDENT ID		SLOPE		SOIL							
3. What is the slope of [FIELD]?		4. In Figure A, which diagram best represents the slope of the [FIELD]?		5. What is the predominant soil type of this [FIELD]?		6. What is the color of the soil?		7. What is the soil quality of this [FIELD]?		8. How do you know the soil quality of the [FIELD]?		9. What is the texture of the soil on this [FIELD]?	
READ ANSWERS		READ ANSWERS		READ ANSWERS		READ ANSWERS		READ ANSWERS		READ ANSWERS		READ ANSWERS	
P	A	F	I	FLAT.....1	ENUMERATOR: SHOW RESPONSENT FIGURE A AND RECORD THE NUMBER SELECTED	SANDY.....1	SLACK.....1	GOOD...1	SCIENTIFICALLY TESTED.....1	VERY FINE...1			
R	E	I	E	MODERATE SLOPE..2		CLAY.....2	RED.....2	FAIR...2	OWN EXPERIENCE...2	FINE.....2			
C	L	E	L	STEEP SLOPE.....3		MAINLY MIXTURE OF SAND AND CLAY.....3	WHITE/LIGHT...3	POOR...3	OTHER PERSON..3	BETWEEN COARSE AND FINE.....4			
E	L	D	I			OTHER (SPECIFY).....4	YELLOW.....4		OTHER PERSON..4	COARSE.....4			
L	I	I	I				OTHER (SPECIFY).....5		SPECIFY.....4	VERY COARSE..5			
I	I	D	D										

## Annex II. Example of Spectral Soil Signatures



*Mid-infrared spectra signature of soils with different level of soil organic carbon content collected from the study area.*