

OPEN

DATA DESCRIPTOR

Survey data on factors that influence the adoption of soil carbon enhancing practices in Western Kenya

George Magambo Kanyenji^{1*}, Stanley Karanja Ng'ang'a² & Evan Hartunian Girvetz³

The data described in this paper were collected in Western Kenya, specifically in Kakamega and Vihiga Counties. The data were collected from 334 households with the aim of assessing factors that facilitate or constrain the adoption of practices that enhance the sequestration of soil carbon. The data were collected through a structured questionnaire that was designed in SurveyCTO. The data were later downloaded from SurveyCTO servers and exported to STATA version 14 for cleaning and analysis. This data can be used by researchers to assess the probability and extent of adoption of specific soil carbon enhancing practices in the two counties of Western Kenya. Additionally, it can be utilized to access the impact of adopting soil carbon enhancing practices on maize and beans yield at both the plot and the farm level.

Background & Summary

Enhancing food security and reducing the poverty rate have been and continue to be vital policy targets for most countries in sub-Saharan Africa (SSA). These challenges are of great concern considering that in SSA the population is expected to double¹ by 2050 and hence, the demand for food will continue rising sharply². However, overcoming these challenges has been elusive due to weak policy formulation in an ever-changing environment. Additionally, most agricultural land in SSA is faced by low soil fertility due to poor agricultural techniques such as continued mono-cropping resulting in soil degradation^{3,4}. At the same time, frequent dry spell and rainfall variability have worsened the situation leading to low agricultural productivity⁵. As a result, in most SSA countries, food security policies are geared towards the adoption of sustainable agricultural practices that can aid in enhancing soil fertility and increasing household income^{6,7}. Some of the sustainable practices that have been advocated for in SSA include climate-smart agricultural technologies (CSAs) and soil carbon enhancing practices (SCEPs). SCEPs are preferred as they help improve soil fertility and at the same time sequestering atmospheric carbon^{8,9}. Additionally, the implementation of certain practices such as the use of farmyard manure, compost manure, intercropping with legumes is encouraged, as they are considered as low-cost practices¹⁰ suitable for a majority of the small-scale farmers in SSA who are largely dependent on agriculture and live below the poverty line¹¹.

Adoption of SCEPs in Kenya is being advocated for as it can enhance soil fertility over time; however, the adoption rates are still low¹²⁻¹⁵. Published literature indicates that several factors (such as socio-economic factors, external support, and wealth category)^{13,14} can influence the adoption of these technologies. Nevertheless, the influence of plot-specific characteristics has been ignored by most researchers in Kenya. At the same time plot specific characteristics have been established to be critical in influencing the adoption rate of agricultural technologies¹⁶⁻¹⁸. Motivated by the aforementioned, a survey was conducted in Western Kenya, a high agricultural potential area faced with low soil fertility, soil degradation, land fragmentation due to high population density and relatively high poverty rate^{19,20}. In addition, several programs have been implemented in this area to enhance the adoption of these practices. The data contained herein were used to evaluate factors that may facilitate or constrain the adoption of SCEPs. The data contains socioeconomic factors, plot-specific characteristics, external support institution factors, wealth category and information relating to access to specific infrastructures.

¹Department of Agricultural Economics, University of Nairobi, P.O. Box 29053-00625, Kangemi, Nairobi, Kenya.

²International Center for Tropical Agriculture, P.O Box 6247, Kampala, Uganda. ³International Center for Tropical Agriculture, P.O Box 823-00621, Nairobi, Kenya. *email: geomagambo@gmail.com

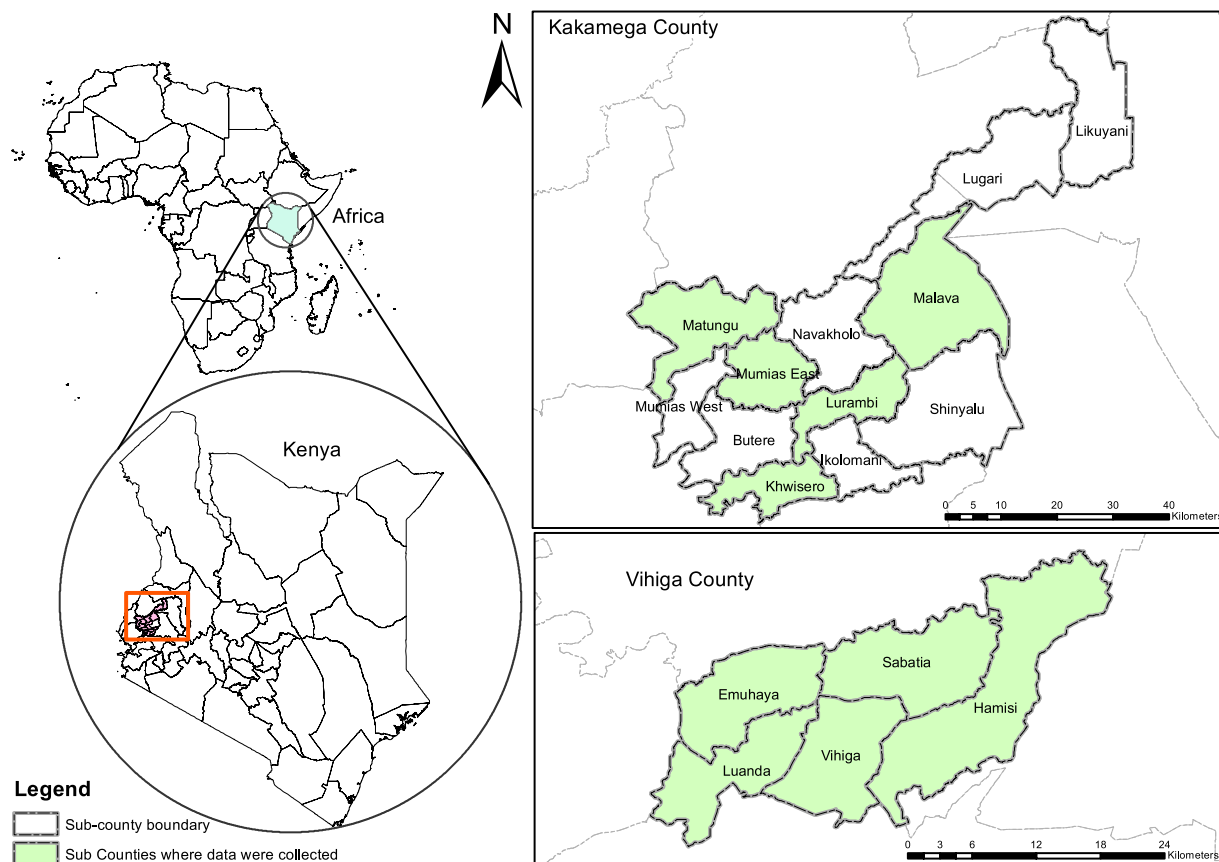


Fig. 1 Map of the study area.

Additionally, the data takes into account maize and beans yield at the plot level under the different combinations of practices implemented by farmers. The survey tool that was used to collect the data was reviewed and approved by the Internal Review Board (IRB) at the International Centre for Tropical Agriculture (CIAT) before the study was commenced. Additionally, the farmers signed a consent form that indicated that they were willing to participate in the study and would terminate the interview at any point.

The data can be utilized to investigate the probability and extent of adoption of specific SCEPs, reasons behind farmers' decision to implement particular SCEPs, the challenges they encounter while implementing the practices and the impact of adopting practices – singly or in combination – on maize and bean yield at both plot level and household level.

Methods

Data collection. The data were collected from Vihiga and Kakamega counties in Western Kenya (Fig. 1). The study area was selected as it is classified as a high agricultural potential area but faced with low soil fertility, soil erosion, soil degradation, and low agricultural productivity. A four-stage multistage sampling technique was utilized to generate the sample. In the first stage, five sub-counties (i.e., Khwisero, Matungu, Malava, Lurambi, and Mumias East) were randomly selected in Kakamega county, while in Vihiga county Vihiga (Emuhaya, Hamisi, Sabatia, and Luanda sub-county) were selected. In the second stage, two wards were selected from each sub-county with the help of the county extension officer. This involved checking the probability of finding farmers that had adopted different SCEPs. In the third stage, from four wards one village and from 6 wards 2 villages were randomly selected in each county. In total 16 villages from each county were selected. In the fourth stage, ten farmers from each village were interviewed, by first picking random farmers from the extension officer list and then snowballing to find the next farmers. The final sample was determined using Eqs. 1 and 2, which resulted in 320 farmers (i.e., 160 farmers from each county). However, to cater to data problems such as missing observation and incompletely filled questionnaires, 14 additional respondents were interviewed, leading to a final sample size of 334 farmers operating 710 plots.

$$n_0 = \frac{Z^2 pq}{e^2} \quad (1)$$

$$n_0 = \frac{1.96^2(0.5 * 0.5)}{0.055^2} = 317(\sim 320) \quad (2)$$

where n_0 is the sample size, e is the desired level of precision, Z^2 is standard normal deviation at 95% confidence interval, p is the estimated proportion of an attribute that is present in the population, and q is $1-p$.

The data were collected with the help of five trained enumerators who were fluent in both English and Kiswahili. The enumerators undertook a three-day intensive training on how to conduct the interview, understanding the research question and how to key in the data into the tablets since the questionnaire had been digitized into Survey CTO. A pre-test was conducted to test whether the enumerators had mastered the survey tool and whether there were errors that needed to be rectified in the tool before its deployment. Survey CTO was utilized because it helps reduce errors by controlling for missing values and therefore makes it easy to control the quality of data obtained.

Data Records

The data is available as Stata Data Format (.dta) and across the entire data, missing values are identified with a dot (.), the standard way of representing missing values under the Stata Data Format. All data are stored in the Harvard Dataverse Repository²¹ and are accessible through the Harvard Dataverse Repository online portal. The data are arranged as per the questionnaire that was used to collect the data. The questionnaire was divided into eight sections as shown in Table 1. Section one collected the general information of the study area. Section two collected information relating to the respondents – their name, years in farming and years in farming for the household head). Section three collected data on the household demographic characteristics (i.e., that is the number of people in the household, their age, gender, relation to the household head, occupation, and if they participated in farming activities). Additionally, the section also collected information that helped in determining the household's wealth category as based on a simple wealth scorecard as well as information relating to household access to different infrastructures such as motorable road, tarmac road, local, livestock and urban markets, electricity, and clinic. Section four contains data relating to farmer's plot characteristics, perception towards soil erosion, soil type and soil fertility, practices implemented under each plot, main crops grown in the plots, output over the last two growing seasons, inputs utilized, source of labour, livestock, and their participation in crop and livestock markets. Section five contains data associated with farmer's social capital, while section six and seven established access to credit and access to extension services respectively. Lastly, section eight contains data pertaining to a farmer's different sources of income. In the survey the household head was targeted; however, in their absence, other household members that were above 18 years were interviewed, provided they had more than five years of farming experience. In this dataset, a household head was defined as the key decision-maker as far as farming was concerned and participated in farming. The different sections of the questionnaire helped us generate the specific variable such as socio-economic factors, plot-specific characteristics, external support, wealth category and access to specific infrastructures factors that may influence the adoption of practices.

Technical Validation

The entire data set described is cross-sectional and was obtained through interviews with farmers. It is common for this type of data to have several problems such as missing information or under or over-reporting. To correct for this SurveyCTO was utilized whereby constraints were embedded in the survey questionnaire to ensure responses to key questions were obtained. The survey employed three techniques to enhance the reliability of data collected. Firstly, before the start of an interview, the farmers were given an overview of questions they were to be asked and in case they objected to giving any sensitive information they had the right to terminate the interview. Secondly, key experts (extension officers) were utilized to validate the values the farmers were giving. Additionally, a focus group discussion with farmers was conducted to get rough estimates of the inputs utilized, land size, and yield. As such their estimated married with the survey results. Additionally, the values were counterchecked with estimates from the literature review, key experts (extension officers) and focus group discussion results.

Among the data described in this paper, there are some variables we would not ascertain fully as they required a farmer's recollection capability and honesty. However, measures were put in place to enhance the reliability of the variables. With reference to Table 1, variables under demographic, infrastructure, wealth index, SCEPs, labour, livestock, livestock market participation, social capital, access to credit and extension, and source of income are highly reliable and certain. However, other variables that required quantification would be considered as highly certain as described below. Under plot specific information, we are certain of all the variables apart from plot size. Nevertheless, to ensure the reliability of the variable, enumerators would first enquire if the farmer had a title deed and this would be utilized as a measure of the plot size. However, if a farmer had subdivided their land, they would be asked if they know the size of the plot either in meters or acres. The different measurements were then standardized to acres. The study avoided utilizing hectare since most farmers were not familiar with hectare standardization as they were more familiar with acre and meters.

Under crop yield subcategory, the quantity of crop harvested is the only uncertain variable. However, to enhance its reliability, a farmer would quantify their harvest either in bags either 90 kg bag, 70 kg bag, 50 kg bag, in kilograms or other local measurement units that were then converted into kilograms during data cleaning. Under the input section, the only variable the study is not certain of was manure usage, unlike fertilizer usage which is accurate since farmers would recall the quantity they purchased. However, to enhance reliability on the quantity of manure farmers utilized, farmers would specify quantities utilized in local measurements or in kilograms that were later converted into kilograms. Lastly, some farmers would over report selling price for some of their products mainly cash crops such as tea and sugarcane and to correct for this, we replaced the outliers with the mean selling price during the time of the interview as obtained from the Kenya National Bureau of Statistics (KNBS).

Variable	Theme	Description	Duration Considered
Demographic	SOC	Gender, age, education level, and occupation of the household members.	12 Months
Infrastructure	INF	Distance in walking minutes to the motorable road, tarmac road, local market, nearest livestock market, the urban market, electricity, and clinic.	Current
Wealth Index	WEA	The probability of a household not being poor as guided by 10 questions.	Current
Plot specific Information	PLI	Describes the plot size, the tenure system, manager of the plot, distance of the plot from the homestead, plot's slope, soil type, and farmers' perception towards their soil fertility and soil erosion.	Current
SCEPS	AGR	Practice implemented in each plot and the reason behind their implementation and challenges associated with each practice.	Current
Crop Yield	AGR	The three main crops grown by the farmers and their yield in the last two seasons.	Last two cropping season
Crop Market Participation	AGR	Indicates if farmers sold their crop produce and the revenue received.	Last 12 months
Input	AGR	Indicates if farmers utilized chemical fertilizer and inorganic fertilizer. It takes into account the amount utilized, source, and price per unit if purchased.	Last Season
Labour	AGR	Labour sources for farming activities.	Last two cropping season*
Livestock	AGR	The livestock owned and purpose of owning the livestock.	Current
Livestock Market Participation	AGR	Indicates if farmers sold their livestock and the revenue received.	Last 12 months
Social Capital	ESS	If the farmers belonged to any farmer group or organization and the purpose of the group.	Last 12 months
Access to Credit	ESS	If a farmer accessed credit, source of the credit, the amount received, and purpose of the credit.	Last 12 months
Access to Extension	ESS	If farmers accessed extension service, source, nature, and terms of provision of the information.	Last 12 months
Sources of Income	WEA	The different sources of income that the household head has.	Last 12 months

Table 1. Description of the dataset as per the questionnaire. The questionnaire generated six themes: Socio-economic factors (SOC), Access to infrastructure (INF), Wealth information (WEA), Plot specific information (PLI), Agricultural practices and activities (AGR), and Access to external support services (ESS). *Included the short rain season and long rains season of 2017. The survey was conducted in August 2018.

Usage Notes

The data is available in Stata Data Format and can be opened by any Stata program Version 13 and above. However, if using an older version of Stata, Stata provides means of converting the data to be compatible with the previous version. For those who do not have access to Stata software, Statistical Package for Social Sciences (SPSS) software can also be utilized to open the data, under the import tab and specifying that the data is in Stata Data Format (.dta). The questionnaire that was used to collect the data is provided together with the data and will be key to the understanding of the data.

The observations do vary across the files since some households did not participate in the said activity. Additionally, the data in the sub-files is in the long format since the presiding question in the main files was a select multiple. For instance, if a farmer did not access extension, they could not be included in the Access_Extension file. Additionally, if a farmer did access extension but from three sources, they have three entries describing the three sources of extension. Additionally, the data use quite a complex key, computer generated by the Survey CTO data collection software, which was found to be give unique labels preferred to the enumerator pre-assigned key. Lastly, all names of the respondents and their household members were replaced with a 1 to hide their identity and protect their privacy as indicated in the consent form that pseudonyms would be utilized.

Received: 2 September 2019; Accepted: 7 January 2020;

Published online: 10 February 2020

References

1. FAO. *The Future of Food and Agriculture: Trends and Challenges*. (Food and Agriculture Organization of the United Nations, 2017).
2. Van Ittersum, M. K. *et al.* Can sub-Saharan Africa feed itself? *Proceedings of the National Academy of Sciences*. **113**, 14964–14969 (2016).
3. Cavanagh, C. J., Chemarum, A. K., Vedeld, P. O. & Petursson, J. G. Old wine, new bottles? Investigating the differential adoption of 'climate-smart' agricultural practices in western Kenya. *J. Rural. Stud.* **56**, 114–123 (2017).
4. Kihara, J. *et al.* Application of secondary nutrients and micronutrients increases crop yields in sub-Saharan Africa. *Agron. Sustain. Dev.* **37**, 25 (2017).
5. Ngwira, A. R., Aune, J. B. & Mkwinda, S. On-farm evaluation of yield and economic benefit of short term maize legume intercropping systems under conservation agriculture in Malawi. *Field Crop Res.* **132**, 149–157 (2012).
6. Place, F., Barrett, C. B., Freeman, H. A., Ramisch, J. J. & Vanlauwe, B. Prospects for integrated soil fertility management using organic and inorganic inputs: evidence from smallholder African agricultural systems. *Food Policy*. **28**, 365–378 (2003).
7. Godfray, H. C. J. *et al.* Food security: the challenge of feeding 9 billion people. *Science*. **327**, 812–818 (2010).
8. Stockmann, U. *et al.* The knowns, known unknowns and unknowns of sequestration of soil organic carbon. *Agric. Ecosyst. Environ.* **164**, 80–99 (2013).
9. Lal, R. Beyond COP 21: Potential and challenges of the "4 per thousand" initiative. *J. Soil. Water. Conserv.* **71**, 20A–25A (2016).

10. Li, Y., Shibusawa, S. & Kodaira, M. Carbon sequestration potential and farming income: Identifying the optimal carbon farming practices in Japanese paddy fields. *Engineering Agr. Environ. Food*. **6**, 68–76 (2013).
11. Holden, S. T. Fertilizer and sustainable intensification in sub-Saharan Africa. *Glob. Food. Sec.* **18**, 20–26 (2018).
12. Dallimer, M. *et al.* Who uses sustainable land management practices and what are the costs and benefits? Insights from Kenya. *Land Degrad. Dev.* **29**, 2822–2835 (2018).
13. De Vente, J., Reed, M. S., Stringer, L. C., Valente, S. & Newig, J. How does the context and design of participatory decision-making processes affect their outcomes? Evidence from sustainable land management in global drylands. *Ecol Soc.* **21**, 24 (2016).
14. Mutoko, M. C., Hein, L. & Bartholomeus, H. Integrated analysis of land use changes and their impacts on agrarian livelihoods in the western highlands of Kenya. *Agr. Syst.* **128**, 1–12 (2014).
15. Marenya, P. P. & Barrett, C. B. Household-level determinants of adoption of improved natural resources management practices among smallholder farmers in western Kenya. *Food Policy*. **32**, 515–536 (2007).
16. Manda, J., Alene, A., Gardebroek, C., Kassie, M. & Tembo, G. Adoption and Impacts of Sustainable Agricultural Practices on Maize Yields and Incomes: Evidence from Rural Zambia. *J. Agr. Econ.* **67**, 130–153 (2015).
17. Kassie, M., Jaleta, M., Shiferaw, B., Mmbando, F. & Mekuria, M. Adoption of interrelated sustainable agricultural practices in smallholder systems: Evidence from rural Tanzania. *Technol. Forecast. Soc. Change*. **80**, 525–540 (2013).
18. Teklewold, H., Kassie, M. & Shiferaw, B. Adoption of multiple sustainable agricultural practices in rural Ethiopia. *J. Agr. Econ.* **64**, 597–623 (2013).
19. Jaetzold, R., Schmidt, H., Hornetz, B. & Shisanya, C. *Farm Management Handbook of Kenya—Natural Conditions and Farm Management Information-Subpart A1 Western Province – Kakamega and Vihiga Counties 2nd edn.* (Ministry of Agriculture Kenya: Government printers, 2010).
20. Odendo, M., Obare, G. & Salasya, B. Farmers’ perceptions and knowledge of soil fertility degradation in two contrasting sites in western Kenya. *Land Degrad. Dev.* **21**, 557–564 (2010).
21. Ng’ang’a, S. K., Kanyenji, G. M., Jalango, D. A., Nguru, W. M. & Girvetz, E. An integrated approach for understanding the factors that facilitate or constrain the adoption of soil carbon enhancing practices in East Africa, specifically Western Kenya. *Harvard Dataverse*, <https://doi.org/10.7910/DVN/HE6CEM> (2019).

Acknowledgements

The authors are grateful for the support and perseverance of the data collection team; Kevin Maina, Robert Otieno, Brian Cheye, Reuben Omega, Felix Shiundu, and Richard Njenga. Thanks also go to the farmers who willingly accepted to provide information to this study. The authors would like to thank Wilson Nguru and Dorcas Jalango for their contribution during the data collection procedure. This work was conducted as part of the Federal Ministry for Economic Cooperation and Development, Germany (BMZ) funded project on “Scaling up soil carbon enhancement interventions for food security and climate across complex landscapes in Kenya and Ethiopia (contract: 81206681. Project no.: 16.7860.6-001.00.). This project was undertaken as part of the CGIAR Research Programs on Water, Land and Ecosystems (WLE). We thank all donors that globally support our work through their contributions to the CGIAR system.

Author contributions

G.M.K. organized the data collection, processing, and quality control of the data sets documented here. G.M.K. also organized and wrote the data paper. S.K.N. assisted with the review, overall guidance to the development of this data paper, and editing of the manuscript, while G.E. assisted with editing and acquisition of resources that facilitated the collection of the data and the development of this manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to G.M.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

The Creative Commons Public Domain Dedication waiver <http://creativecommons.org/publicdomain/zero/1.0/> applies to the metadata files associated with this article.

© The Author(s) 2020