Map to the Future (M2F):

Integrating soil mapping into cocoa farm development plans in Ghana

Version: 30 October 2020

GRAMEEN FOUNDATION
University of Colorado Boulder
USDA Agricultural Research Service

Acknowledgments

This project, *Map to the Future (M2F): Integrating soil mapping into cocoa farm development plans in Ghana*, submitted in response to the 2018 call for proposals by the World Bank's Development Economics Data Group (DECDG) and the Global Partnership for Sustainable Development Data (GPSDD), is supported by the World Bank's Trust Fund for Statistical Capacity Building III (TFSCB-III) with financing from the United Kingdom's Foreign, Commonwealth & Development Office, the Department of Foreign Affairs and Trade of Ireland, and the Governments of Canada and Korea.

The project team also acknowledges valuable inputs and cooperation from Peter van Grinsven (BrightLife Farming), Dr. Edward Yeboah and Stephen Owusu (Council for Scientific and Industrial Research [CSIR] - Soil Research Institute), Ernest Dwamena and Frank Mensah (Touton, Ghana), Selasse Gidiglio (Rainforest Alliance), and Odjobi Kwakye (Escape Poverty Africa). In addition, we would like to voice our appreciation for the following individuals who participated in the early consultations that helped inform the design of this project: Dr. Richard Asare (International Institute of Tropical Agriculture- IITA), Dr. Michael Opoku-Agyeman and Dr. Alfred Arthur (Cocoa Research Institute of Ghana- CRIG), Dr. Richard Ampadu-Ameyaw (Science and Technology Policy Research Institute-CSIR), Dr. Eunice Agyako Mintah (Ghana Atomic Energy Commission), Edward Kumah (Mondelez/Cocoa Life).

Project Team

The Grameen Foundation USA: Gigi Gatti (Technology for Development Director), Bobbi Gray (Research Director), Mona McCord (Innovations in Agriculture Director), Alfred Yeboah (Regional Director, Africa), Julián Gomez (Technical Advisor, Agriculture and Technology), Hannah Rubio (Technology Project Manager)

The Sustainability Innovation Lab at the University of Colorado, Boulder USA: Dr. Jason Neff (Director), Dr. Jonathan Maynard (LandPKS Analytical Scientist), Meghan Mize and Carolyn Kerchof

United States Department of Agriculture (USDA), Agricultural Research Service: Dr. Jeff Herrick (Soil Scientist and LandPKS Primary Investigator)

Contents

Project Information

Project Title	Map to the Future (M2F): Integrating soil mapping into cocoa farm development plans in Ghana
Date of Report	30 October 2020
Lead Organization	Grameen Foundation USA
Lead Collaborators	The Sustainability Innovation Lab at the University of Colorado (SILC), Boulder USA, USDA Agricultural Research Unit, Touton Ghana
Sustainable Development Goals (SDGs) Covered	1 (No Poverty), 2 (Zero Hunger), 8 (Decent Work and Economic Growth), 13 (Climate Action), 17 (Partnerships)
Country	Ghana
Data Types	Site specific farm and farmer data, Site based soil data, Soil maps, Crop Specific Soil Recommendations
Technologies	Land-Potential Knowledge System (LandPKS) is an open-sourced global suite of mobile phone applications, initially funded by the USAID, and implemented by the USDA. LandPKS is used to collect (1) basic soil and topographic information necessary to determine land potential and (2) soil and vegetation cover data necessary to assess and monitor major changes in plant community composition and wind and water erosion risk. The LandPKS team is led by the United States Department of Agriculture - Jornada Research Unit in partnership with the Sustainability Innovation Lab at the University of Colorado, Boulder, and the New Mexico State University. FarmGrow is an android-based decision making tool used by cocoa extension coaches that combines agronomy and economics. Using the tool, cocoa farmers are presented with recommendations and specific agronomic practices that they can adopt to increase their productivity to 1.5MT to 2 MT per hectare. It also generates a profit and loss statement that helps the farmers understand the financial investments needed to reach this level of productivity over a multi-year timeline. FarmGrow was developed by Grameen Foundation, USA in collaboration with the Mars and Rainforest Alliance.

Project Objective	This research project will integrate site-specific soil data with traditional survey data to explore the feasibility of supporting Bank projects which aim to improve the productivity of small-holder Cocoa farmers. The particular focus of this research project is to discern whether integrating new and traditional sources of available data may generate more insights into soil quality, to inform current (and future) Bank projects seeking to strengthen yield potential for smallholder cocoa farmers in the country. The activities conducted in this project have the potential to enable relevant Bank operational teams to improve the feedback loop between local data collection and the algorithms used to produce satellite based global soil mapping measurement products.
Risks	Aside from Covid-19 related risks which caused a delay in soil sampling and the cancellation of two in-person workshops, there were no other project risks. Originally, the Grameen and LandPKS team members were going to travel and engage directly in the data collection activities; however, due to travel restrictions, the team re-oriented funds to support the work of local soil scientists and a local data collection firm. This required frequent virtual meetings with both the local data collection firm (Escape Poverty Africa) and the soil scientists (Edward Yeboah and Stephen Owusu from the Council for Scientific and Industrial Research [CSIR] - Soil Research Institute) to both adapt to local travel and COVID restrictions and ensure safe data collection processes. Budgets were also slightly realigned to ensure hygiene practices would be followed, such as use of masks, hand sanitizer, and use of social distancing practices during in-door trainings. In lieu of the one-day, in-person event planned for Ghana in August and a brown-bag meeting scheduled for World Bank staff in September, one official webinar is now scheduled for January 2021 to share results with those directly engaged in the project. This medium will also allow a much wider audience to join and engage in the lessons learned by this project. Two informal webinars have already occurred with the World Bank and project team members and stakeholders to discuss and interpret the results of the study.
Gender Data	This research project was not set out to produce or generate gender statistics or data. Out of the 75 primary farmers under the study, 13 were female (17.3%).
Data / Methods	This project leveraged two geo-enabled mobile technology applications (FarmGrow and LandPKS) to generate site-specific—and more granular and accurate—soil and agronomic data for 75 farms. Three sites (soil pits) were evaluated per farm resulting in 225 data collection points. While the collection of site-specific soil data is generally cost-prohibitive due to the cost of sample collection and analysis, these applications provide a streamlined approach for acquiring information needed to generate farm-scale agronomic recommendations by using onsite observations and measurements of soil properties. This project also integrated traditional and digital soil map data with our site-specific soil data to create a comprehensive site-specific dataset. In addition, we downscaled the Food and Agriculture Organization's Agro-Ecological Zones (AEZ) soil suitability

	modelling framework, allowing us to generate crop-specific soil suitability ratings at the farm-scale. This modeling framework and data integration has been developed within an API, which with future development could be integrated into existing geo-enabled mobile applications (e.g., FarmGrow or LandPKS). The M2F project also contributed to new site-specific soil classification data for cocoa growing areas in the Ashanti and Western Regions in Ghana. There were no further farmer profile disaggregations of data given the primary focus on comparing site-specific soil data with existing soil maps and with farmer performance data (adoption observations captured from FarmGrow).

Executive Summary

Average cocoa yields in Ghana by smallholder farmers are currently far below their production potential (400 kg/ha vs. over 3,000 kg/ha) (Aneani and Ofori-Frimpong, 2013). FarmGrow (https://www.farmgrow.org), a geo-data enabled precision agriculture service and technology platform, is designed to assist smallholder cocoa farmers in Ghana to increase cocoa yields from 400 kg/ha to 1500-2000 kg/ha over an 8- to 10-year period. Agronomists use FarmGrow with participating cocoa farmers to provide them with individualized support in adopting good agricultural practices (GAPs) and increasing on-farm investments to sustainably improve cocoa yields and cocoa income. This requires an accurate assessment of the agronomic constraints currently limiting cocoa yields. Dominant factors contributing to this current yield gap include climate, cultural practices, and the soil. Among these, long-term soil degradation and, in particular, soil infertility is recognized as one of the main factors limiting cocoa yields in Ghana. FarmGrow's current agronomic recommendations highlight the issues surrounding fertility of cocoa soils as 97% of the farms evaluated received some type of soil management recommendation. However, a major limitation in assessing farm-scale soil suitability and limitations lies is acquiring accurate soil property data.

The main objective of the research conducted under the *Map to the Future (M2F): Integrating soil mapping into cocoa farm development plans in Ghana* project was to evaluate the integration of site-specific soil data with traditional soil data products (e.g., regional-to-global soil maps) in an effort to provide location-specific soil information that smallholder cocoa farmers can use to improve soil health and overall farm productivity. M2F focused directly on the detailed characterization of soils on cocoa farms using the LandPKS app (https://www.landpotential.org), the relation between LandPKS soil property data and the soil maps in these settings, and the soil data accuracy requirements for cocoa soil management. Based on these analyses, a framework for how site-specific soil information may be used to help farmers better manage their soils is illustrated in Figure 1 and consists of three main steps:

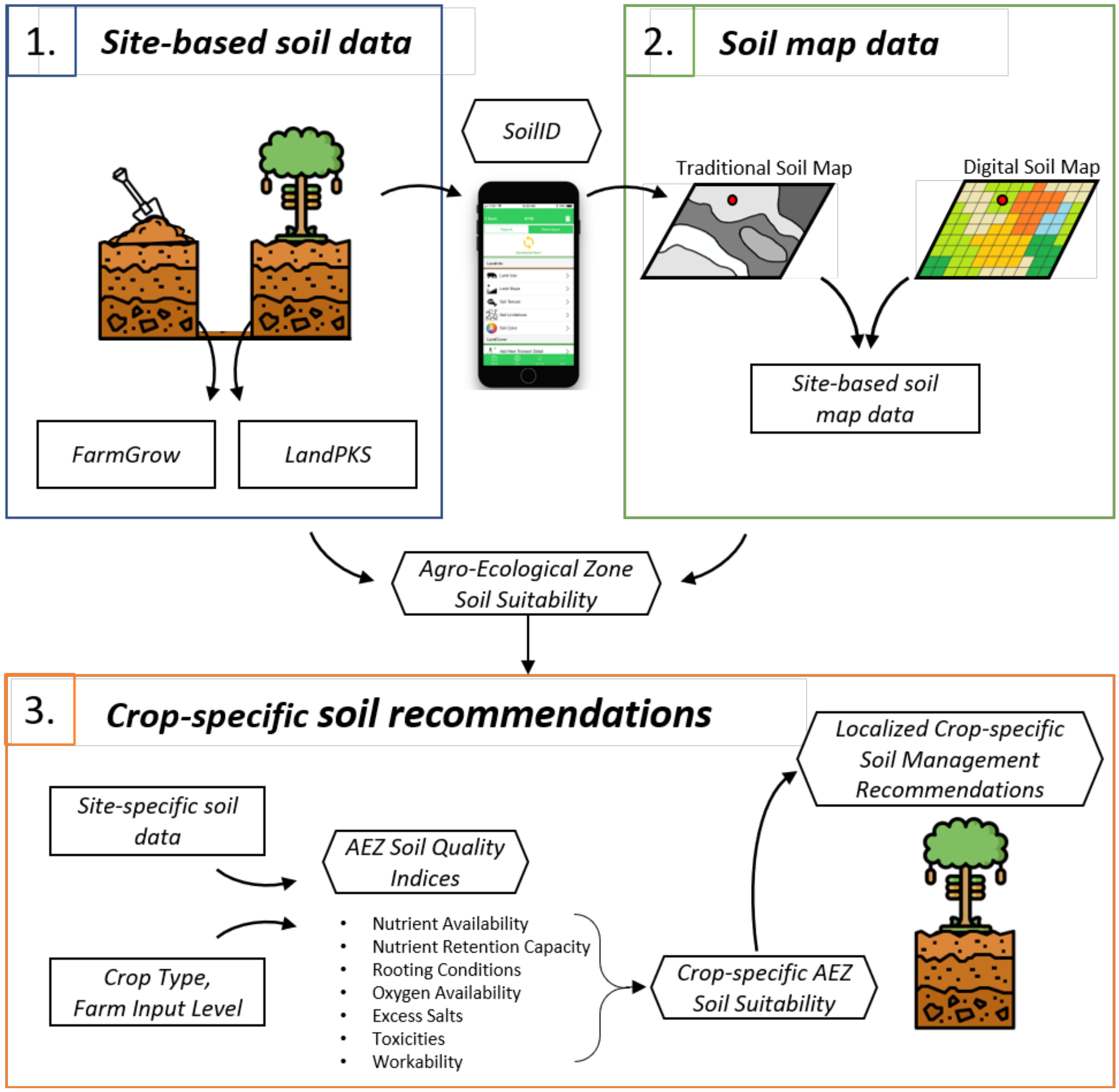
1. Acquire accurate site-specific soil data (FarmGrow + LandPKS). Seventy-five farms were assessed, with three data collection sites per farm, resulting in 225 soil sites.
2. Link existing soil map data at the site location.
3. Develop crop-specific soil recommendations based on site-specific + soil map data based on the Food and Agriculture Organization's (FAO) Agro-Ecological Zone (AEZ) soil suitability methodology.

The **first step** in this framework involves acquiring accurate site-specific soil data. The FarmGrow app currently collects a limited set of soil data that is used, in part, to generate individualized agronomic recommendations to smallholder cocoa farmers. This research study evaluated how the addition of more detailed soil characterization data collected using the LandPKS app could improve our understanding of soil limitations and be used to improve current agronomic recommendations. This study found that there was significant improvement in understanding soil limitations at cocoa farms with the addition of LandPKS soil data. For example, the FarmGrow assessment of soil physical condition only rated 1 farm as 'Bad' and 74 as 'Good', whereas the more detailed LandPKS soil assessment rated 47 farms as 'Bad' and 28 as 'Good'.

The **second step** in this framework involves linking soil map data at site locations. The relative accuracy of different soil map products (e.g., Harmonized World Soil Database [HWSD], World Inventory of Soil Emission Potential [WISE], SoilGrids) was evaluated relative to LandPKS soil data collected at each evaluation site.

Figure 1. M2F research framework

Map to the Future (M2F) page 5



In Land PKS' initial analysis of soil map data in Northern Ghana, SoilGrids was found to be the most similar to LandPKS soil data. However, HWSD/WISE was found to be the most accurate in Ghana's cocoa growing regions (Ashanti and Western regions) and SoilGrids the least accurate compared to LandPKS, which illustrates the variable accuracy of soil map products. The lower accuracy of SoilGrids in the cocoa growing regions is most likely due to limited soil observations in these regions to train the SoilGrids models.

Due to its higher accuracy in the cocoa growing regions, HWSD/WISE was used to link soil map data to the soil observation data generated from the LandPKS app. Two methods of linking soil map data to point locations were evaluated, (1) linking the dominant map unit component (area-specific linkage) and (2) using the LandPKS SoilID algorithm to match the most similar soil series based on the soil observation data at that site (site-specific linkage) (see Table 1 and Figure 8). It is important to note that traditional soil maps (HWSD and WISE) are designed to predict *groups* of soils occurring within a map unit and that most users of soil maps (and nearly all non-soil scientists) simply select the dominant component, which we evaluated in the area-specific linkage case. One of the principal values of the LandPKS SoilID is that it effectively gives non-soil scientists the ability to access *all* of the data in traditional soil maps, like HWSD and WISE, resulting in potentially higher accuracy than simply selecting the dominant map unit component.

The **third step** in this framework involves developing crop-specific soil recommendations based on site-specific data generated in Step 1, and the linked soil map data acquired from Step 2. We used the Agro-Ecological Zones (AEZ) soil suitability modelling framework, developed by the Food and Agriculture Organization of the United Nations (FAO) and the International Institute for Applied Systems Analysis (IIASA), to translate site-specific soil information into crop-specific soil suitability ratings (Figure 1). It should be noted that the AEZ modeling framework requires quantitative soil property input data and therefore only LandPKS soil data was used in the AEZ analysis. The AEZ methodology calculates seven crop-specific soil quality (SQ) indices for 54 different crop types at three different production input levels (low, intermediate, high). These indices include: (SQ1) nutrient availability, (SQ2) nutrient retention capacity, (SQ3) rooting conditions, (SQ4) oxygen availability, (SQ5) excess salts, (SQ6) toxicities (e.g., gypsum or calcium carbonate), and (SQ7) workability (Figure 1). The calculation of each soil quality index uses a unique set of soil property inputs, and based on the farm input level, these seven soil quality indices can be used to generate crop-specific suitability ratings (SR).

FAO AEZ soil quality indices are only available as spatial maps generalized at a 10 km grid resolution; a scale too large to reliably inform agronomic management at the farm scale. In this study, we have taken the AEZ methodology and applied it at the soil map unit component scale, effectively downscaling the crop-specific suitability ratings to a scale more appropriate for farm-scale decision making. This study localized the AEZ suitability calculations by leveraging site-specific soil property data (e.g., LandPKS texture, rock fragments, soil depth) and the LandPKS SoilID algorithm to identify the most likely soil component at each point location. Results from this downscaling effort showed that the localization of soil suitability ratings for cocoa often resulted in a different assessment relative to previous ratings based on more generic soil map data (e.g., dominant map unit component), particularly when results were compared across low-, intermediate-, and high-input farming systems (which vary regarding the amount of investment and mechanization that occurs on the farm). **The ability to calculate soil suitability as a function of farm input level makes it possible to conduct a cost-benefit analysis of crop intensification.** Soil infertility was found to be the dominant soil constraint for cocoa production, particularly for the low- and intermediate-input level farming systems. **Given all of the FarmGrow sites were found to fall into low- and intermediate-input farming systems, with low-input systems representing the majority, understanding the soil's response to fertilization will assist farmers in optimizing their return on investment, particularly for those farmers in low-input systems whose soils will not gain a lot of value from fertilization.** Two cases studies presented in this report demonstrate 1) that recommendations given to cocoa farmers through FarmGrow, such as replanting or increasing use of fertilizer, may not be to the real benefit of the farmer given existing soil constraints and 2) use of the AEZ framework can also be used to gauge soil suitability for other crops being grown on the land, such as rice, maize, or among 53 other crop types assessed by the AEZ soil quality indices.

Results from this analysis demonstrate the importance of site-specific soil data for understanding a soil's agronomic limitations and the feasibility of soil management interventions for improving crop yields. **Relying on soil map data alone may lead to an under or overestimation of land capability and thus fail to identify the soil management actions needed to improve cocoa yields**. When smallholder farmers have limited resources (financial, human, etc.), these differences could mean success or failure or limited impact of the investments they are making.

Further work is needed to refine and test this framework through acquiring more detailed soil maps (e.g., Ghana soil survey map) and to expand our site-specific soil dataset. **Future efforts to streamline the soil sampling process would also be needed in order to allow FarmGrow agronomists to integrate this data with their current data collection efforts. Future refinement of this work may also include integration of the site-specific AEZ soil suitability analysis into apps such as FarmGrow and LandPKS. This would allow end users in low resource settings to identify crops most suited for intercropping with cocoa based on their soil's condition, in an effort to improve soil fertility through increased biodiversity of cocoa farms.**

Acronyms

AEZ	Agro-Ecological Zones
AO	Adoption Observation
FAO	Food and Agriculture Organization of the United Nations
GAPs	Good Agricultural Practices
HWSD	Harmonized World Soil Database
IIASA	International Institute for Applied Systems Analysis
SMUs	Soil Map Units
SG	SoilGrids
SQ	Soil Quality
SQI	Soil Quality Index
WISE	World Inventory of Soil Emission Potential

Introduction

Ghana is the second largest cocoa producer in Africa, with the cocoa sector providing livelihoods for over 700,000 farmers in the country's southern tropical belt. Most cocoa farmers in Ghana operate small-area, low-input farms where cocoa is mostly grown under full sun or very low shade with little or no fertilizer applied. The continuous extraction of inherent fertility of cocoa soils without replenishment has resulted in the long-term depletion of soil fertility in cocoa farms and declining cocoa yields (Dossa et al., 2018). Recent increases in total cocoa production in Ghana have been largely due to agricultural expansion and conversion of remnant forests, particularly in the Western region where most soils are considered unsuitable for cocoa production (Appiah et al., 1997). Cocoa productivity is influenced by both inherent soil properties (such as clay content and pH) that determine a soil's potential productivity, and dynamic soil properties (such as organic matter and fertility) that determine soil health and that change over time as a result of management.

In order to help farmers improve soil health and cocoa productivity through intensification and more efficient farming practices using existing plantations, farmers and/or agronomists must first identify which soil factors are currently limiting and to what extent they may be minimized or corrected. However, the lack of available, actionable information on soil physical and chemical properties has complicated or limited opportunities for small-area farmers to both identify the best soils for optimal cocoa production and improve soil health through appropriate soil management practices. Acquiring accurate soil information is a critical first step needed to improve farm productivity, but information alone is not sufficient to ensure that appropriate management practices are implemented.

Interpretation of soil information within the context of specific crop requirements (i.e., nutrient requirements, texture requirements, drainage, etc.) is needed to identify and implement the most appropriate soil management practices given a soil's inherent characteristics and limitations (e.g., shallow soil depth). To generate agronomic recommendations relating to soil management, a farmer or agronomist must

(1) acquire accurate site-specific soil information;
(2) synthesize soil information into agronomic knowledge; and
(3) develop appropriate, actionable soil-specific agronomic recommendations.

The focus of this research project was to examine soil information sources (all in digital form) in Ghana's cocoa growing region, including traditional soil data sources (e.g., soil maps) and new sources of site-specific soil data (e.g., LandPKS, FarmGrow); and the ability to integrate all relevant soil data to inform farm management and improve overall farm productivity.

Specific research objectives were:

1. Conduct research that compares readily available site-specific soil information sources in Ghana's cocoa growing regions, including FarmGrow observations and LandPKS soil characterization, with soil property predictions from current soil maps; and
2. Explore how improved site-specific soil data may improve agronomic recommendations to smallholder farmers.

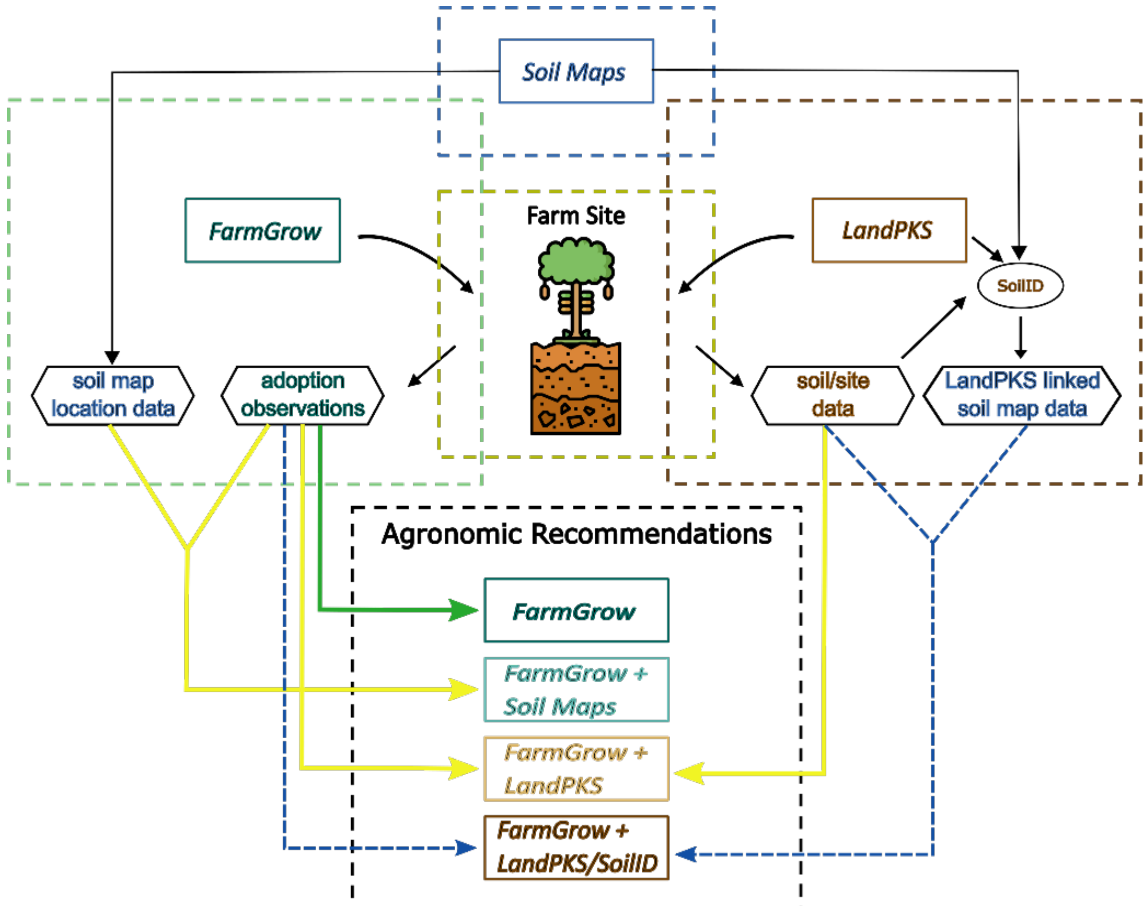
Site-specific Soil Information in Ghana’s Cocoa Growing Regions

The most accurate soil information at a site comes from direct observation and measurement. This involves digging a soil pit at a site and measuring various soil properties (e.g., texture, rock fragments). Some soil properties are relatively static and affect the long-term potential and consistent management approaches. These included factors such as soil texture and a number of soil properties that influence soil pH and fertilizer retention. Other soil properties, such as carbon or nitrogen content, can vary over shorter time-scales depending on soil fertility management practices. Site specific soil information can be used to improve long-term management recommendations but some decisions involving shorter term changes may require additional site-based measurements. While some soil properties can be directly measured at a site, many others that influence soil health (e.g., Cation Exchange Capacity[1]) require laboratory analysis. Limited access to soil laboratory analysis by farmers has restricted the soil data available to inform soil management. However, if a farmer or extension agent can identify the soil type (i.e., soil series) mapped at a location, it may allow them to infer some of these lab-measured soil properties and therefore gain improved insight into the soil’s agronomic potential and mitigate, when possible, known soil limitations. While global soil maps provide a complete set of physical and chemical soil data, these products are calibrated for national and global measures and not intended for accurate soil identification for small area estimates (field scale). This raises several important questions for end-users, including: How accurate are soil maps at my farm? Which soil map product is the most accurate? Can I use soil map information to inform my soil management decisions? Answers to these questions require an understanding of site-specific soil accuracy as it relates to both the relative accuracy of soil map information (i.e., compared to soil profile measurements) and the levels of soil accuracy required for different land management applications (e.g., crop-specific soil requirements).

Our evaluation framework for identifying the most accurate set of site-specific soil information for developing soil-specific agronomic recommendations is shown in Figure 2. In this framework we identify **three main sources of soil information** in Ghana’s cocoa growing regions: FarmGrow, LandPKS, and current soil maps. FarmGrow and LandPKS both provide site-specific soil data, while soil maps provide soil property and class predictions (often with high uncertainty) at point locations.

[1] **Cation exchange Capacity** (CEC) is a measure of the soil's ability to hold positively charged ions. It is a very important soil **property** influencing soil structure stability, nutrient availability, soil pH and the soil's reaction to fertilizers and other ameliorants (Hazleton and Murphy 2007).

Figure 2. Evaluation Framework for site-specific soil information in Ghana's cocoa growing regions



Based on these three soil information sources, **we identified four possible approaches for generating agronomic recommendations**, (i) FarmGrow, (ii) FarmGrow + Soil Maps, (iii) FarmGrow + LandPKS, and (iv) FarmGrow + LandPKS + SoilID (Figure 1, Table 1).

Table 1. Soil data sources evaluated

Soil Data Source	Description
1. FarmGrow	Limited set of soil property and soil management data collected at each farm. Includes soil physical condition, organic matter condition, fertilizer application and formulation, pH, and drainage.
2. LandPKS	Detailed set of soil property data collected at up to seven fixed-depth intervals. Includes soil texture class, rock fragment volume class, soil color, land slope, slope shape, and information on soil limitations and soil health.
3. Soil Map Data	
Soil Maps	Soil property data predicted at each location from both traditional (HWSD/WISE) and digital (SoilGrids) soil maps. Represents the dominant map unit component (traditional) or predicted (digital) soil property data at each location.
SoilID	LandPKS SoilID algorithm identifies the most likely soil map unit component at each point location from traditional soil maps by matching the soil map data to the measured soil data at each site.

FarmGrow, with its limited set of soil property observations (i.e., adoption observations), represents our baseline of soil information. The FarmGrow app is used to record soil properties (i.e., soil physical condition, organic matter condition, pH, and drainage) and related soil management observations (i.e. fertilizer application and formulation) from the top 30 cm of the soil at each farm site. The second soil dataset, FarmGrow + soil maps, evaluates the value-add of directly integrating soil map predictions with FarmGrow soil data. The third soil dataset, FarmGrow + LandPKS, evaluates the benefit of adding more detailed site-specific soil data. The LandPKS app is used to record soil property data (e.g., soil texture class, rock fragment volume class, and soil color) at up to seven fixed depth intervals (i.e., 0-1, 1-10, 10-20, 20-50, 50-70, 70-100, 100-120 cm). LandPKS is also used to document soil depth, land slope, slope shape, and information on soil limitations (e.g., soil cracking, surface salt accumulation, flooding, low pH, surface stoniness, high water table, and shallow soils) and soil health (e.g., biological activity, soil smell). The final soil dataset, FarmGrow + LandPKS + SoilID, evaluates the benefit of more detailed site-specific soil data plus a more accurate integration of soil map data by using the LandPKS SoilID algorithm. SoilID provides a modeling framework that can identify the most likely soil series at a sampling-location based on a simple set of soil property inputs measured at a site (e.g., soil texture class, rock fragment volume class, and soil color). The SoilID algorithm calculates a statistical similarity between the measured soil property values and the reported soil property values for each soil series mapped in that area. The soil series with the highest statistical similarity to the sampling site is identified as the most likely soil at that location.

The evaluation of these different soil data groups provides information necessary for a cost-benefit analysis of integrating more detailed site-specific soil data (i.e., LandPKS) and more detailed but less accurate soil information for available soil maps (e.g. SoilGrids). Our objective is to generate information that can be used to decide if, where and when collecting site-specific soil data can significantly improve management decisions, and whether the integration of soil map data would provide added value based on the accuracy of their predictions.

Evaluating and Comparing Four Approaches for Generating Agronomic Recommendations in Ghana's Cocoa Growing Regions

To evaluate the four approaches for generating agronomic recommendations, this study selected 75 cocoa farms in Ghana's southern cocoa growing regions (41 in Ashanti Region and 34 in the Western Region). These 75 farms are a subset of the ~3,000 farms participating in the FarmGrow program and selected to represent a diversity of farm conditions (e.g., age, size, yield, health). LandPKS soil and site characterization was performed at three sites per farm, with the location of each site selected to represent the variability of soil and site conditions at each farm. General information on the 75 farms is shown in Table 2. On average, the 75 profiled farms had 1 hectare of land under cocoa production with average productivity of 581 kg/ha and farm age of 25 years. Median production was a little under 500 kg/ha which is consistent with the production levels of the larger sample size (~3,000). A little more than half noted growing additional crops; approximately one third used hired labor. Among those growing additional crops, the majority were growing plantain and cassava in addition to some other crops such as maize, rice, cocoyam, yam, banana, oil palm, peppers and okra.

Table 2. Summary information for the 75 cocoa farms evaluated in this study

Cocoa farm area (ha)	Mean: 1.0; Min: 0.2; Max: 6.6
Productivity (kg/ha)*	Mean: 581; Min: 77; Max: 1931
Farm age (yrs.)	Mean: 25; Min: 6; Max: 60
Additional crops	Yes: 44; No: 30; Not specified: 1
Hired labor	Yes: 27; No: 3; Not specified: 45

*Note: There were two outliers in the data that have been removed; one with a minimum and one with a maximum level of production that did not realistically match their acreage. These questionable data points were removed to present a more accurate picture.

1. FarmGrow

FarmGrow is a digital agriculture advisory tool designed to assist smallholder cocoa farmers to increase cocoa yields from 400 kg / hectare to 1500 kg / hectare over an 8- to 10-year period. FarmGrow is used by agronomists to, (1) profile the farming household regarding basic demographics and farm data such as farm size, (2) assess the current condition of cocoa farms, such as determining tree age, density, health and the presence of diseases (3) identify which good agricultural practices (GAPs) and on-farm investments are needed to improve farm productivity, and (4) assist cocoa farmers in developing farm investment plans to support the agronomic and financial interventions needed to improve business performance and monitor success.

The FarmGrow assessment framework evaluates farm productivity in four main areas: plant genetics, farm condition, GAPs, and soil fertility. The standard GAPS include pruning; pest, disease and sanitation practices; weeding; harvesting conditions; shade management; soil fertility management which includes soil condition and health, fertilizer formulation and application. A total of 14 evaluation metrics (Adoption Observations or AOs) are used to assess farm condition and the four areas affecting farm productivity.

Table 3 shows the possible ratings for the four AO groups and 14 individual AOs. The AO groups are in bold (bold) and the 14 individual AOs are *italicized*. More in-depth descriptions of how the AOs are defined and scored are available Appendix 1.

Table 3. FarmGrow rating system

FarmGrow Adoption Observations	Possible Ratings
Genetics (Planting Material)	**Bad, Medium, Good**
GAPS	**Bad, Medium, Good**
Pruning	*Bad, Medium, Good*
Pest, Disease, Sanitation	*Bad, Medium, Good*
Weeding	*Bad, Good*
Harvesting	*Bad, Good*
Shade Management	*Bad, Good*
Soil Fertility	**Bad, Medium, Good**
Soil Condition	*Bad, Good*
Organic Matter	*Bad, Good*
Fertilizer Formulation	*Bad, Medium, Good*
Fertilizer Application	*Bad, Medium, Good*
Farm Condition	**Bad, Good**
Tree Age	*Bad, Good*
Tree Density	*Bad, Good*
Tree Health	*Bad, Good*

Debilitating Disease	*Bad, Good*

Table 4 shows the summary ratings for the four AO groups and 14 individual AOs. The scores for the individual AOs roll-up into an AO group score. In addition, two additional soil indicators assessing lime and drainage that were recorded in FarmGrow at the 75 farms were also evaluated in this study. The FarmGrow soil dataset represents our base-case (i.e., no additional soil information) for developing soil-specific agronomic recommendations. While most are scored as 'Bad' for Genetics (the planting material used), approximately half are either rated as 'Medium' or 'Good'. Forty-five were rated 'Bad' on Farm Condition, driven by Tree Age and Tree Density. Overall, the 75 farmers were rated as 'Bad' on GAPS; the scores are driven by the overwhelming poor practices related to pruning and shade management.

Regarding soil specifically (data in *italics*), while the majority (65) rated as 'Bad' for the Soil Fertility Group, all of the 75 farms were rated 'Good' for soil organic matter, all but one farm rated 'Good' for soil physical condition, none of the sites were in need of soil drainage improvements, and only one site required application of lime . The main soil limitations identified in FarmGrow were the application and formulation of fertilizer. Seventy-seven percent of farms were rated 'Bad' for both fertilizer application and formulation. Only one farm was rated 'Good' for both fertilizer application and formulation, with the remainder having varying combinations of 'Bad' and/or "Medium' for the two categories.

Table 4. Summary of FarmGrow AOs for the 75 cocoa farms evaluated in this study

Adoption Observations				
	Rating Categories	**Bad**	**Medium**	**Good**
Genetics		**39**	**11**	**25**
Planting Material	Bad, Medium, Good	39	11	25
GAPS		**75**	**0**	**0**
Pruning	Bad, Medium, Good	70	2	3
Pest, Disease, Sanitation	Bad, Medium, Good	36	17	22
Weeding	Bad, Good	27	--	48
Harvesting	Bad, Good	9	--	66
Shade Management	Bad, Good	59	--	16
Soil Fertility		***65***	***9***	***1***
Soil Condition	*Bad, Good*	*1*	*--*	*74*
Organic Matter	*Bad, Good*	*0*	*--*	*75*
Fertilizer Formulation	*Bad, Medium, Good*	*64*	*10*	*1*
Fertilizer Application	*Bad, Medium, Good*	*59*	*14*	*2*
Farm Condition		**45**	**--**	**30**
Tree Age	Bad, Good	30	--	45
Tree Density	Bad, Good	26	--	49
Tree Health	Bad, Good	13	--	62
Debilitating Disease	Bad, Good	13	--	62
Soil Indicator Summaries				
Lime	Bad, Good	1	--	74
Drainage	Bad, Good	0	--	75

As a result of the AO assessments, the investment plan is created for each farmer. The investment plan lays out an 8- to 10-year picture for the farmer of the potential yield and resulting income they could receive through their farm improvements, based on the recommendations they are provided.

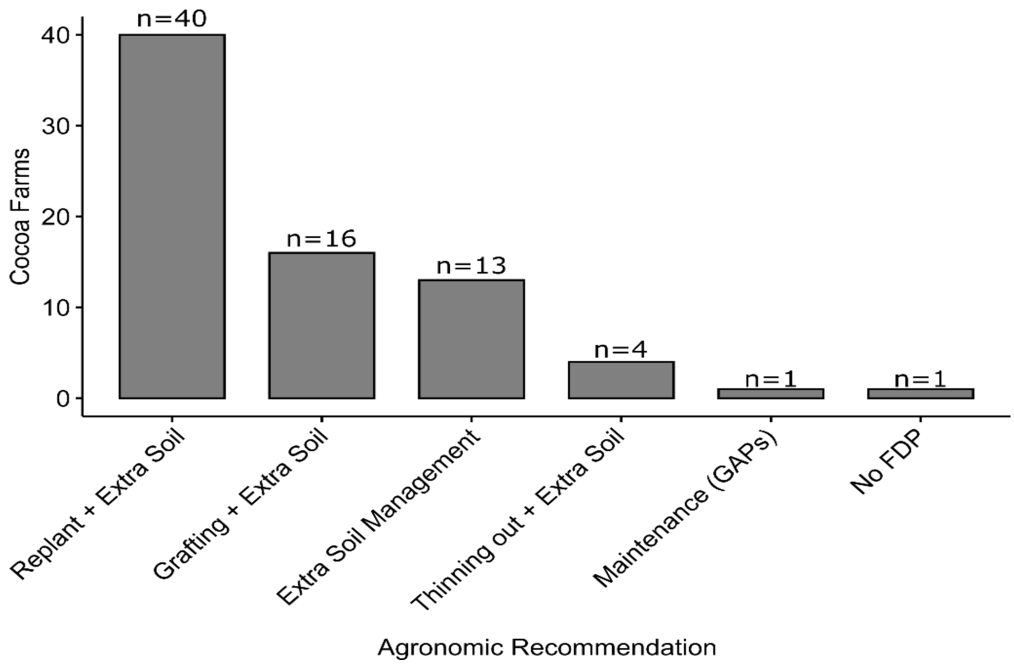
There are six primary categories of recommendations provided to farmers. These recommendations include:

- Replant (cut down old trees or diseased trees and replant with new planting material)
- Extra Soil Management (increase use of organic matter, proper application and formulation of fertilizer)
- Grafting (graft old trees with new planting material)
- Maintenance GAPs (follow basic GAPs)
- Thinning Out (remove some trees to meet recommended distance among trees)
- Filling In (plant new trees to maximize plot space and meet the recommended distances among trees)
- No Farm Development Plan (FDP; tree health and soil condition are both bad and it is not ideal for a farmer to plant cacao on the plot).

Any farmer can receive a combination of these recommendations, usually resulting in no more than two recommendations per plot. Extra Soil Management is the only recommendation that is coupled with other recommendations. Once recommendations are made and farmers agree upon a plan with the agronomist, they are then monitored at agreed-upon intervals with the agronomist. During a monitoring visit, farmers are assessed on the GAPs as well as their achievement of their targets per the recommendations provided by FarmGrow.

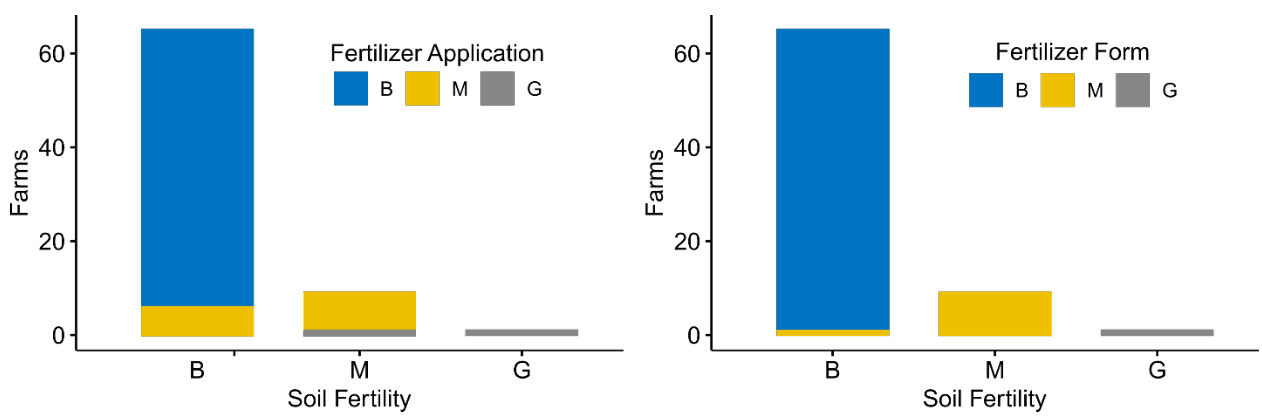
Based on the FarmGrow analysis, 97% of the 75 farms had agronomic recommendations that included some form of soil management (Figure 3). This is consistent with the majority of the farmers served by FarmGrow, whose results can be found elsewhere (Sarpong et al, 2020; Gray et al, 2020). Fertilizer application and formulation were identified as the main soil constraints affecting soil fertility, and therefore the agronomic recommendations of ‘Extra soil’ or ‘Extra Soil Management’ at 97% of farms can be almost exclusively attributed to these two factors (Figure 4).

Figure 3. FarmGrow agronomic recommendations for the 75 cocoa farms evaluated in this study



Notes: FDP, Farm Development Plan (which means they were not provided an investment plan given the unsuitability of the plot for increasing cocoa productivity); Extra Soil and Extra Soil Management are the same recommendation.

Figure 4. Evaluation of fertilizer application and fertilizer formulation ratings for the 75 cocoa farms evaluated in this study



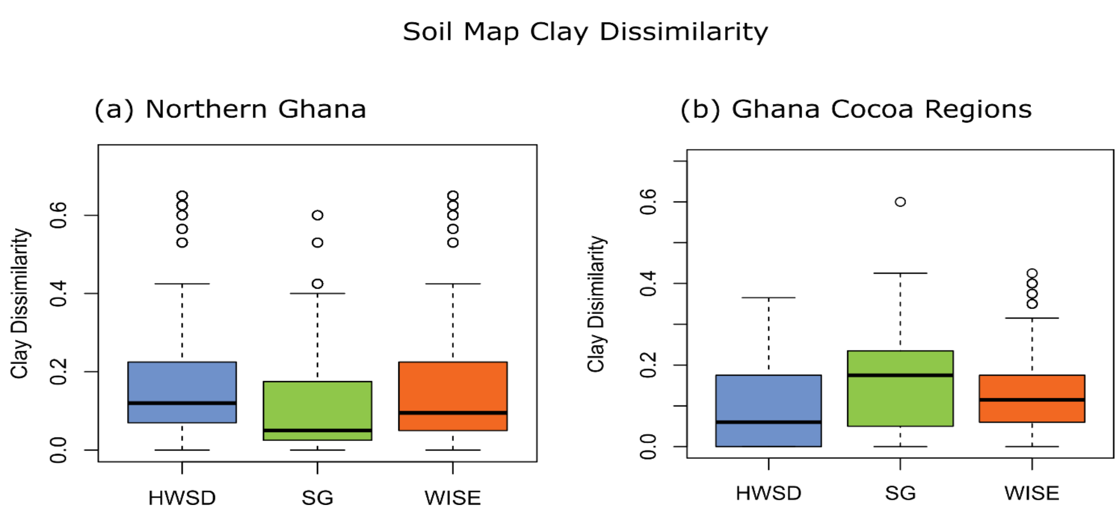
Notes: B, Bad; M, Medium; G, Good as assessed by FarmGrow agronomists for fertilizer application and fertilizer formulation

2. FarmGrow + Soil Maps

Soil maps provide a wealth of soil information essential to understanding land potential and informing GAPs that can optimize farm productivity. In this study we evaluated both traditional (HWSD/WISE) and digital (SoilGrids) soil maps. Traditional soil maps display soil map units (SMUs) representing distinct areas of a landscape composed of one or more soil series but do not show the exact location of each soil series. A common method for dealing with this spatial uncertainty is to assign any location within a SMU to its dominant soil series. SoilGrids and other digital soil mapping products offer an alternative to traditional soil maps by providing predictions of soil properties and classes at specific locations. These modeled products, however, can have high levels of uncertainty in their predictions at point locations. Soils in Ghana are generally mapped at large spatial scales (e.g., 1:5,000,000 map scale: HWSD, WISE), resulting in a disconnect between the spatial scales at which soil map information is generated (and its underlying spatial uncertainty) and the scale at which most land-use decisions are made (e.g., point locations on a farm). In order to utilize soil map information for site-specific decision making, a better understanding of soil map accuracy is needed.

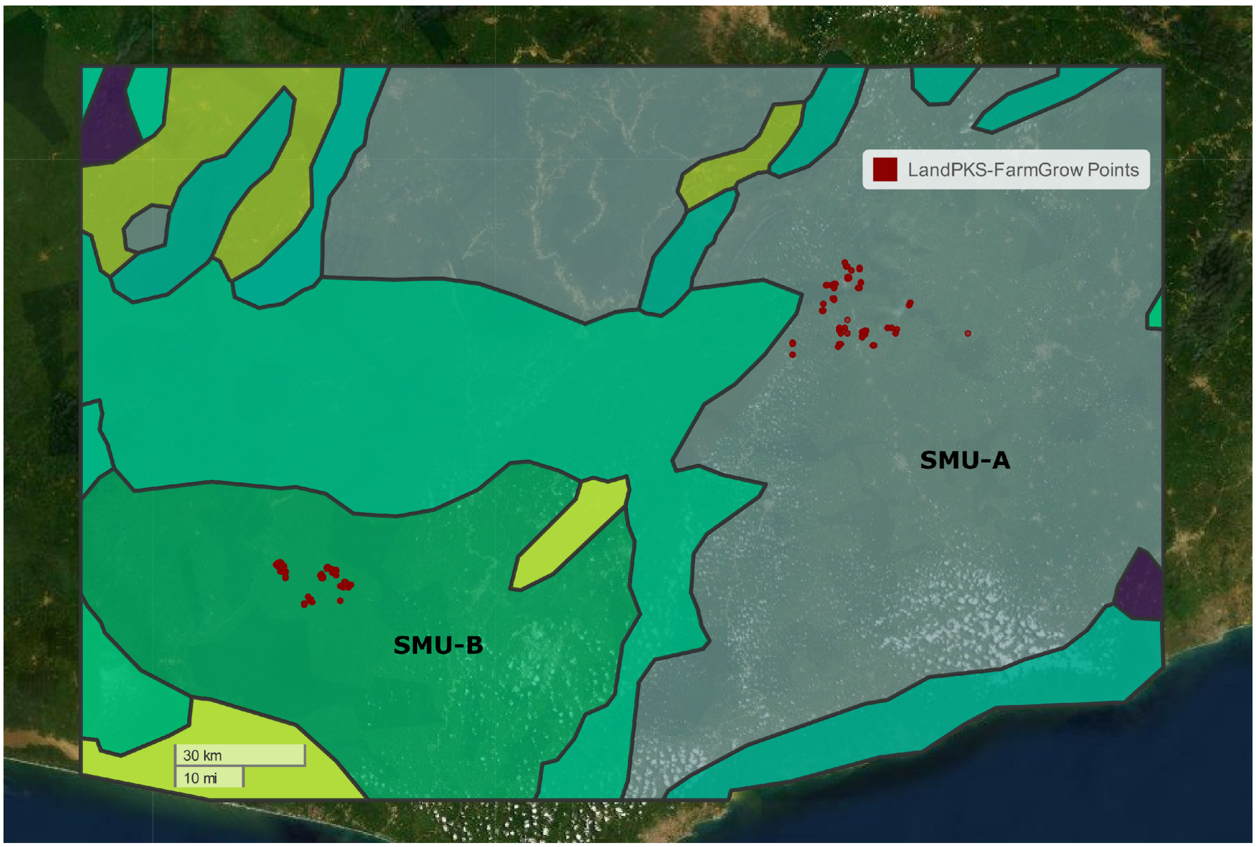
Among soil properties, soil texture is one of the most agronomically important due to its influence on soil water availability, infiltration rate, drainage, tillage conditions and capacity to retain nutrients (Sys et al, 1993). Our evaluation of soil map accuracy was based on a statistical comparison of soil texture (i.e., percent clay and sand) and rock fragment volume measured by LandPKS and predicted by the different soil maps (see Appendix 3 for details on methodology). In our initial draft report (Appendix 2), we evaluated the relative accuracy of soil map information within the Northern Regions of Ghana due to the large number of existing LandPKS sites (~6K). In this initial evaluation of soil map accuracy, we found that the soil property estimates of SoilGrids were more accurate than either HWSD or WISE when compared to LandPKS soil data (Appendix 3, Figure 5a). However, our subsequent analysis of the 225 LandPKS sites (75 farms, 3 sites per farm) in the Ashanti and Western Regions of Ghana revealed that SoilGrids was the least accurate soil map, and that HWSD and WISE had higher accuracies when compared to LandPKS soil data in this area (Figure 5). Figure 5 shows boxplots of the dissimilarity (statistical distance) between LandPKS and soil mapped data for representative percent clay values (see Appendix 3 for more details). The middle of each boxplot indicates the median value. The upper and lower edges of each boxplot indicate the 75th and 25th percentiles, respectively. The ends of the vertical lines indicate the minimum and maximum data values. Circles above the box plot are outliers and the distance of the circles from the maximum measure (top line) represent the magnitude of the difference between the LandPKS data and the soil map data. In the Northern Ghana graph, Soil Grids data had fewer outliers and those were located closer to the LandPKS data; however, in the Cocoa Regions, the outliers were more extreme with the SoilGrids data. Clay dissimilarity is on an absolute scale, so a value of 0.2 would equal a 20% difference in clay between the LandPKS value and the predicted map value. SoilGrids is a digital soil mapping product and its accuracy is dependent upon the availability of training data in an area. The lower accuracy of SoilGrids in the cocoa growing regions is likely due to limited soil observations in these regions to train the SoilGrids models.

Figure 5. Boxplots of clay dissimilarity (statistical distance) between LandPKS and soil mapped data for (a) Northern Ghana, and (b) Ghana's cocoa regions.



The spatial distribution of cocoa farms sampled in this study are shown in Figure 6. A total of 75 cocoa farms (41 in Ashanti Region and 34 in the Western Region) were evaluated with three LandPKS sites sampled per farm, totaling 225 sites. Figure 6 also shows the spatial delineation of HWSD/WISE soil map units, with all the farms in both the Ashanti and Western Regions lying within a single map unit. Each of these map units are attributed with only one soil component: Haplic Acrisols in the Ashanti Region (SMU-A) and Xanthic Ferralsols in the Western Region (SMU-B). **This lack of spatial resolution in the delineation of soil types severely limits their potential utility in helping inform soil management at the farm scale. Consequently, we would caution any direct integration of soil data from these map products with FarmGrow Adoption Observations.**

Figure 6. Location of the 75 cocoa farms sampled in this study. Polygons show the spatial location of HWSD/WISE soil map units (SMU).



3. FarmGrow + LandPKS

The FarmGrow soil fertility assessment is based on four different Adoption Observations, including, soil physical condition, organic matter, fertilizer formulation, and fertilizer application. The soil physical condition and organic matter content of the soil are assessed by a trained field technician, while the fertilizer formulation and application are self-reported by the farmers. FarmGrow uses a binary rating system (i.e., Good vs Bad) to assess both soil physical condition and soil organic matter content at each farm. LandPKS evaluates similar soil criteria used in FarmGrow but instead of a single binary rating, records quantitative evaluations of each soil property. We converted LandPKS soil property values to a binary based on FarmGrow soil evaluation criteria to allow for a direct comparison. Specific rating criteria for FarmGrow and LandPKS are shown in Table 5. In addition to the soil Adoption Observations, FarmGrow also records two additional soil indicators: soil drainage and soil lime requirement. Similar indicators are also recorded in the LandPKS app.

Table 5. FarmGrow and LandPKS rating criteria for organic matter condition and soil physical condition to achieve a 'Good' rating.

Soil Property	Criteria for a 'Good' rating
Organic Matter Condition	
FarmGrow	● Clear signs of microbial activity, **AND** ● Multiple layers of decaying organic material, **AND** ● Worms, worm castings, insect activity, soil pores, **AND** ● Organic material left in the farm
LandPKS	● Many signs of organisms in the soil, **OR** ● Fresh, sweet, earthy smell
Soil Physical Condition	
FarmGrow	● No signs of erosion, no roots visible on the surface **AND** ● Few rocks or gravel on farm surface or in the ground as measured by 3 holes of 30 cm deep per plot **AND** ● Soil is neither too sandy or clayey as measured by touch/roll test on soil from 3 holes of 30 cm deep per plot **AND** ● Well drained either naturally or through drainage canals **AND** ● Slope < 15%
LandPKS	● No signs of erosion: water flow patterns = None or Rare; Rills= None or Rare; Gullies = None or Present/Inactive **AND** ● Surface stoniness < 15%, surface rock fragment < 35% **AND** ● Surface texture is not Sand, Loamy Sand, Clay, or Silty Clay **AND** ● Flooding = None **AND** ● Slope < 15%

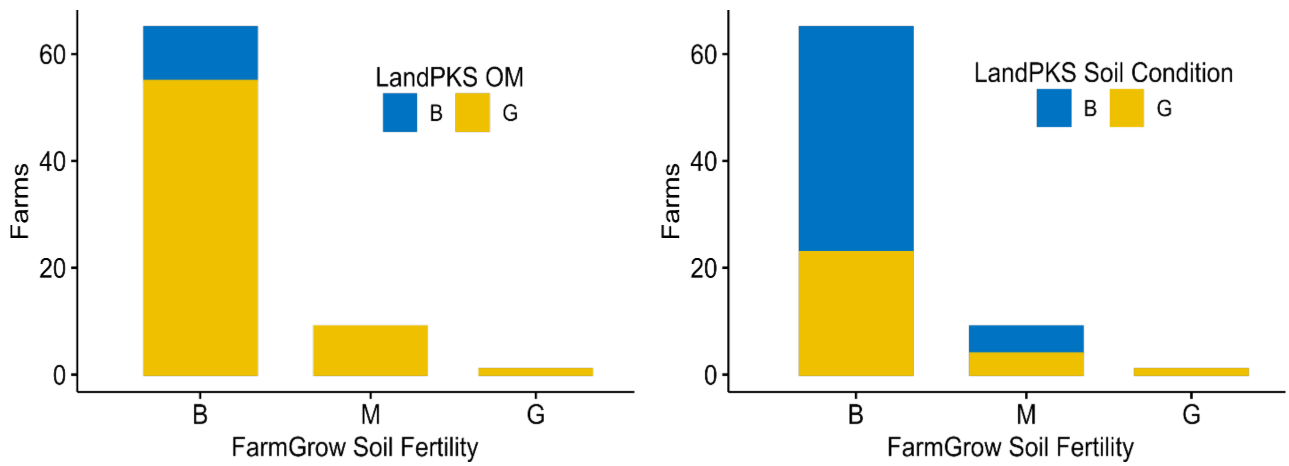
A comparison of the soil Adoption Observations and indicators for FarmGrow and LandPKS are shown in Table 6. For FarmGrow, all participating farms were rated as 'Good' for organic matter and drainage, and only one farm rated as 'Bad' for soil physical condition and lime requirement. In contrast, 10 farms were rated as 'Bad' for organic matter in LandPKS and 47 farms rated as 'Bad' for soil physical condition. No farms were identified as requiring lime and only four farms were rated as 'Bad' for drainage by LandPKS. The breakdown of individual soil properties contributing to the soil physical condition assessment for LandPKS are also shown in Table 6. Land slope and surface texture were rated as 'Bad' for roughly a third of all farms, followed by surface stoniness and soil erosion. Figure 7 illustrates the breakdown of LandPKS organic matter ratings relative to the FarmGrow soil fertility ratings. While all the FarmGrow farm assessments were rated as 'Good' for organic matter, the 10 farms rated as 'Bad' for organic matter by LandPKS fell in the 'Bad' category for FarmGrow soil fertility. For the LandPKS assessment of soil condition, most farms rated 'Bad' were also in the 'Bad' category for FarmGrow soil fertility (Figure 7). This indicates a general agreement between the two evaluation systems in terms of overall soil condition/potential, even though the factors contributing to those ratings differ. These results also suggest that the more detailed soil characterization performed by LandPKS was able to more accurately identify farms with suboptimal (i.e., 'Bad' or 'Medium') soil condition/potential. Consequently, **there is clear benefit in the additional soil/site data collected by LandPKS in identifying factors that limit soil potential (i.e., static soil properties -- texture, rock fragments, depth) vs. those that limit crop productivity as a result of the soil condition (i.e., dynamic soil properties – organic matter, drainage, erosion) that can be improved over time.**

Table 6. Comparison of FarmGrow and LandPKS soil property ratings

	FarmGrow		LandPKS	
FarmGrow Soil Properties	Bad	Good	Bad	Good
organic matter*	0	75	10	65
soil physical condition	1	74	47	28
-erosion*	--	--	8	67
-surface stoniness	--	--	13	62
-surface texture	--	--	20	55
-drainage*	--	--	4	71
-slope	--	--	25	50
Lime requirement*	1	74	0	75
Drainage*	0	75	4	71

Notes: *Relatively dynamic soil properties indicate condition. Other soil properties (static) are a function of land type.

Figure 7. Comparison of LandPKS organic matter (OM) and soil physical condition ratings in relation to the FarmGrow soil fertility ratings.



Notes: B=Bad; M=Medium; G=Good

4. FarmGrow + LandPKS + SoilID

Integrating LandPKS soil/site characterization data with FarmGrow's soil related Adoption Observations provides a significant improvement in our understanding of the soil limitations at a cocoa farm. The ability to accurately link soil observation data to soil map data at a site would significantly improve our understanding of soil limitations and potential. The accuracy of the SoilID matching is dependent upon the quality or accuracy of the input data, both the user-entered soil property data and the soil map data. In this study, the LandPKS soil property data was collected by field crews that were provided with 2 days of training. Additionally, at a subset of study sites (n=10), the LandPKS data was validated by professional soil scientists. The quality or spatial resolution of the soil map data, however, was the main limitation for obtaining accurate site-specific soil data. Future work incorporating more detailed soil maps (Ghana soil survey map) would likely improve the accuracy of our site-specific soil data set.

Translating Site-specific Soil Information into Agronomic Knowledge

Information on inherent soil properties, like those measured by LandPKS (e.g., soil texture), can be used to inform farmer decisions on a variety of management practices such as irrigation frequency, need for organic amendments, likelihood of erosion, and so on. If soil type can be determined, soil management decisions can be further tailored to a farmer's specific soil needs for improving soil health, including (for example) the use of modifiers (e.g. lime) of soil pH and the need for nutrient specific fertilization to address soil deficiencies.

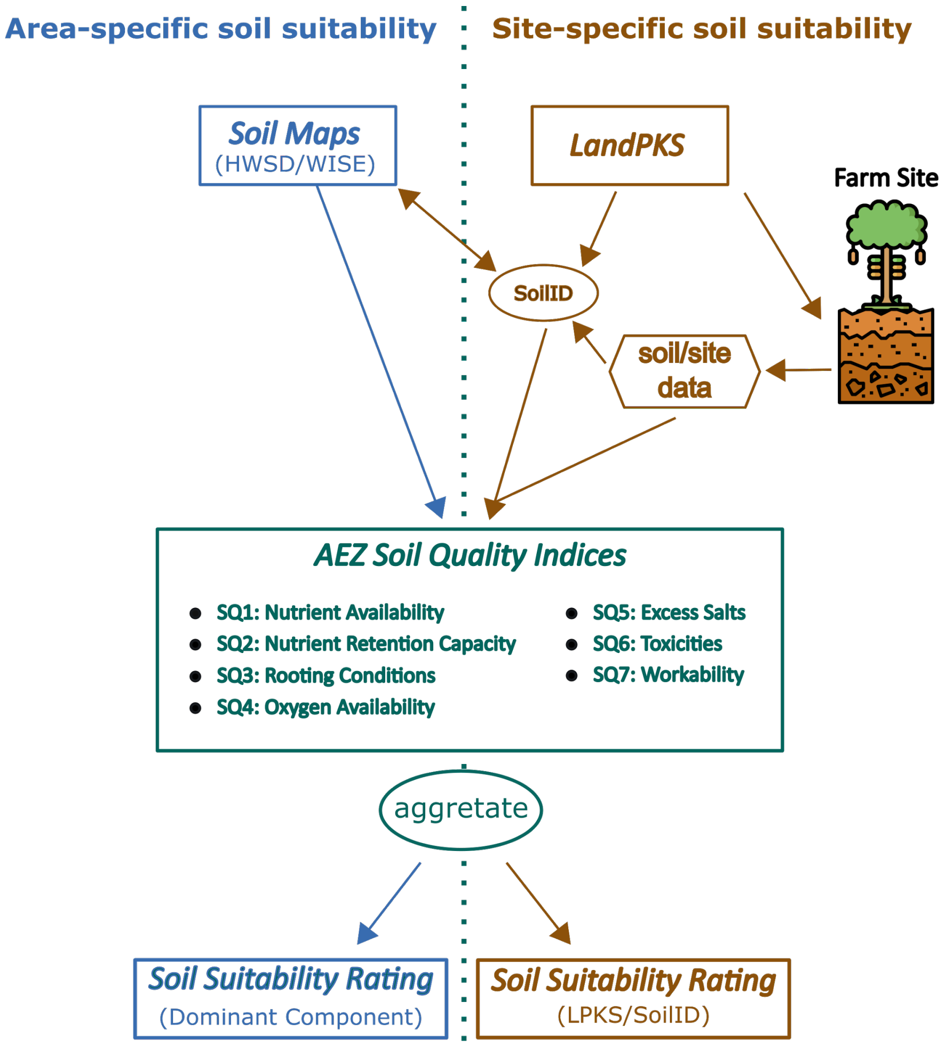
Translating site-specific soil information into agronomic recommendations, however, requires an understanding of how the soil property values at a given site affect the productivity of a particular crop. The Agro-Ecological Zones (AEZ) soil suitability modelling framework, developed by the Food and Agriculture Organization of the United Nations (FAO) and the International Institute for Applied Systems Analysis (IIASA), uses soil data and detailed agronomic knowledge to quantify land productivity and crop-specific agronomic potential (user guide). We used the AEZ modeling framework to translate site-specific soil information into crop-specific soil suitability ratings.

Downscaling Global Agro-Ecological Zones (AEZ) Soil Suitability

The current AEZ methodology uses HWSD soil property data to calculate crop-specific soil quality ratings for 54 different crop types at three different production input levels (low, intermediate, high). The AEZ modeling framework includes the calculation of seven crop-specific (in this case, for cocoa) soil quality indices. These indices include: (SQ1) nutrient availability, (SQ2) nutrient retention capacity, (SQ3) rooting conditions, (SQ4) oxygen availability, (SQ5) excess salts, (SQ6) toxicities (e.g., gypsum or calcium carbonate), and (SQ7) workability. The calculation of each soil quality index uses a unique set of soil property inputs. Based on the input level, these seven soil quality indices can be used to generate crop-specific suitability ratings (SR). In its current implementation, FAO AEZ soil quality indices are only available as spatial maps generalized at a 10 km grid resolution, a scale too large to reliably inform agronomic management at the farm scale. We have taken the AEZ methodology and applied it at the soil map unit component scale, effectively downscaling the crop-specific suitability ratings to a scale more appropriate for farm-scale decision making. Additionally, we have further localized the suitability calculations by leveraging site-specific soil property data (e.g., LandPKS texture, rock fragments, soil depth) and the LandPKS SoilID algorithm to identify the most likely soil component at a point location. In Ghana's cocoa growing regions, excess salts (SQ5) and toxicities relating to high gypsum or calcium carbonate (SQ6) are not relevant given the prevalence of highly weathered soils with neutral to low pH. Thus, crop-specific SRs in these regions will vary as a function of SQ1-SQ4, and SQ7. All calculations in this report were generated for the rain-fed irrigation type, since irrigation of cocoa in Ghana is rare. Figure 8 illustrates the general modeling framework for calculating area-specific (i.e., based on soil maps and the dominant soil type) and site-specific (i.e., soil maps + site data) soil suitability ratings for different crops.

The ability to localize crop-specific soil suitability ratings requires substituting comparable soil property values measured at a site. Table 7 shows the standard HWSD/WISE soil properties used to calculate each soil quality index and the equivalent or related soil properties from SoilGrids, FarmGrow, and LandPKS. The LandPKS soil properties listed in bold in Table 7 (i.e., texture, rock fragments, soil depth) were used to substitute HWSD/WISE soil properties in the site-specific soil suitability calculations. The AEZ soil suitability algorithm uses quantitative soil property input data. The additional site-specific soil properties in Table 7 (italicized) are related to the HWSD/WISE input properties but require additional work to formalize these relationships through statistical or logic-based approaches.

Figure 8. Modeling framework for area-specific and site-specific soil suitability ratings



Farm Input/Management Levels

Three generic levels of input/management are evaluated in the AEZ modeling framework: low, intermediate, and high input levels. Each input level, with its distinct set of cultivation practices and farm inputs, varies in its response to different soil conditions. Farming systems with low input levels (i.e., traditional management) are largely subsistence based. These types of farms use traditional cultivars, labor intensive management practices, no application of nutrients, herbicides, or pesticides, and minimum conservation measures. At the intermediate input level (improved management) the farming operation is partially market oriented where some portion of production is for commercial sale in addition to production for subsistence. These farming systems use improved varieties, some use of animal traction and/or mechanization, some use of fertilizers, herbicides, and pesticides, adequate fallows, and some conservation measures. Farming systems at the high input level (advanced management) are mainly market oriented and managed for commercial production. These systems use high yielding varieties, are fully mechanized with low labor requirements, and use optimal applications of fertilizers, herbicides, and pesticides. Using FarmGrow AO scores on fertilization for the 75 farms assessed, the M2F team would classify **59 farms as low-input farms, 16 as intermediate input-farms and none as high-input farms.** Despite some farmers receiving a “good” assessment on fertilization practices, they are not classified as high-input systems due to the fact that an underlying assumption of FarmGrow GAPs is that fertilization recommendations are typically being made to farmers with older trees that do not fully benefit from fertilizer. Fertilization recommendations focus more on nutrient replacement and not on maximizing yield potentials, due the cost implications.

Table 7. Relevant soil properties used to calculate AEZ soil quality indices in Ghana's cocoa growing regions

Soil Quality Indices	Soil Maps		Site-Specific Soil Data†	
	HWSD/WISE	SoilGrids	FarmGrow	LandPKS
SQ1: Nutrient availability	texture, organic carbon, pH, total exchangeable bases	texture, organic carbon, pH	organic matter, pH	**texture**, *soil smell, biological activity, lime requirement*
SQ2: Nutrient retention capacity	texture, base saturation, pH, cation exchange capacity (soil/clay)	texture, pH, cation exchange capacity	pH, *lime requirement*	**texture**, *lime requirement*
SQ3: Rooting conditions	soil depth, texture, rock fragments, soil phases	soil depth, texture, rock fragments	*soil physical condition, drainage*	**soil depth, texture, rock fragments,** *compaction layer, vertical cracks, water table depth*
SQ4: Oxygen availability	soil drainage class, soil phases	texture	*soil physical condition, drainage*	*water table depth, flooding,* **texture**
SQ7: Workability	soil depth, texture, rock fragments, soil phases	soil depth, texture, rock fragments	*soil physical condition*	**soil depth, texture, rock fragments,** *compaction layer, vertical cracks*

Notes: † Soil properties in *italics* indicate properties that are related to the original HWSD/WISE inputs but require additional work to formalize these relationships (e.g., pedotransfer functions).

The relative importance of the different soil quality indices varies for each farm input level due to differences in agronomic inputs and management practices. For example, **at the low input level, the inherent fertility of a soil, as measured by its nutrient availability (SQ1), is a dominant factor controlling crop productivity**. **Nutrient retention capacity (SQ2) on the other hand, is less important since there is no fertilizer application in low input systems. In high input systems the opposite relationships are found. Nutrient availability is of no consequence in high input systems due to optimal fertilizer application, while nutrient retention capacity or the ability of the soil to retain applied nutrients is of utmost importance.** In Ghana's cocoa growing regions we used the five soil quality indices listed in Table 7 to calculate the soil SR for cocoa (Fig. 8). The soil property ratings used to calculate each of the soil quality indices differs across the three farm input levels due to differences in agronomic inputs and cultural practices. Furthermore, the calculation of the soil SR at each input level differs based on the changing importance of the individual soil quality indices as discussed above. **This ability to calculate soil suitability as a function of farm input level makes it possible to conduct a cost-benefit analysis of crop intensification. Since soil infertility is recognized as the dominant soil constraint for cocoa production, understanding the soil's response to fertilization will assist farmers in optimizing their return on investment.**

Area-specific vs. Site-specific Soil Suitability in Ghana's Cocoa Growing Region

Farmers need accurate information on crop-specific soil suitability. AEZ soil suitability provides a methodology for achieving this but requires accurate soil information at the farm scale. This study evaluated two downscaling approaches: the first represents an area-specific soil suitability using soil data from the HWSD/WISE dominant map unit component at a location (FarmGrow + Soil Map test case) and the second represents a site-specific soil suitability using the HWSD/WISE SoilID map unit component data combined with LandPKS soil property data (FarmGrow + LandPKS + SoilID test case) (Fig. 8). While both downscaling approaches are an improvement over the current regional implementation of AEZ, **the site-specific soil suitability provides the most accurate estimation of soil limitations due to its incorporation of localized soil information.** Figure 8 illustrates these two downscaling approaches where soil data is used to calculate the seven soil quality indices, and these soil quality indices are then used to calculate a final soil suitability rating for the site. **A comparison of area- vs. site-specific soil suitability ratings showed they have the highest agreement at the intermediate farming input level** (i.e., 94% match = 211 sites), followed by the low (66% match = 150 sites) and high (45% match = 102 sites) input levels (Table 8). **At the low input level, site-specific soil suitability ratings were higher (i.e., less constrained) than area-specific suitability for the majority of no-match cases.** In contrast, at the high input level site-specific soil suitability ratings were lower (i.e., more constrained) than area-specific suitability for the majority of no-match cases (Tables 8 and 9).

Results from this analysis demonstrate the importance of site-specific soil data for understanding a soil's agronomic limitations and the feasibility of soil management interventions for improving crop yields. Relying on area-specific map data may lead to an under or overestimation of land capability and thus fail to identify the soil management actions needed to improve cocoa yields. When smallholder farmers have limited resources (financial, human, etc.), these differences could mean success or failure or limited impact of the investments they are making. For example, recommending costly fertilizer could be a waste of money if the soil is not able to provide long-term nutrient retention, which may be detrimental if households are making tradeoffs between paying for fertilizer vs school fees or health care.

Table 8. Comparison of site-specific and area-specific soil suitability ratings by farm input levels

Suitability*	Low Input	Intermediate Input	High Input
Match	150	211	102
Lower	23	14	96
Higher	52	--	27

Notes: *Soil suitability is evaluated as: match = site-specific and area-specific soil suitability are the same; lower = site-specific is lower than area-specific soil suitability; higher = site-specific is higher than area-specific soil suitability.

Table 9. Area-specific and site-specific cocoa soil suitability ratings for low, intermediate, and high input farming systems.

	Low Input		Intermediate Input		High Input	
Suitability*	**Area-specific**	**Site-specific**	**Area-specific**	**Site-specific**	**Area-specific**	**Site-specific**
S0	--	--	--	--	123	61
S1	--	--	--	--	102	131
S2	123	156	225	211	--	20
S3	102	58	--	4	--	3
S4	--	1	--	--	--	--
N	--	10	--	10	--	10

Notes: *S0: No constraint (100%-95%); S1: Slight constraint (95%-85%); S2: Moderate constraint (85%-60%); S3: Severe constraint (60%-40%); S4: Very severe constraint (40%-10%); N: Not suitable (<10%)

Case Study 1: FarmGrow-LandPKS Soil Suitability for a Cocoa Farm

Case Study 1 illustrates how the crop-specific soil suitability framework may be used to develop appropriate, localized soil management recommendations that farmers can implement to improve cocoa yields. We demonstrate the application of this approach on a smallholder cocoa farm in Ghana's Ashanti Region. Our example farm is 23 years old, has an average yield of 320 kg/ha, and has a current FarmGrow recommendation to replant plus extra soil management (Table 10).

Table 10. Location and general characteristics of the case study cocoa farm

Region	Ashanti
District	New Edubiase A
Village	Adansi Praso
Plot Name	P-00002454-2
Productivity	320 kg/ha
Farm Age	23
Additional Crops	Rice
FarmGrow Recommendation	Replant + Extra Soil

This FarmGrow recommendation is based on the 14 Adoption Observations shown in Table 11, which shows all four of the Adoption Observation groups (shown in bold) rated 'Bad' for this site. Among the soil related Adoption Observations, we see that both soil physical condition and organic matter were rated 'Good', while both fertilizer formulation and application were rated 'Bad'. From the LandPKS assessment of this site, organic matter was rated 'Good', while physical soil condition was rated 'Bad' due to the sandy surface texture (Table 12). Figure 9 shows photos of the sampling site, soil profile and soil samples. The soil profile photo (Figure 9, center) shows a uniform layer of sand for 0-50 cm, underlain by a redder clay loam layer high in rock fragments (Table 12, LandPKS).

Figure 9. Photos of sampling site P-00002454-2

Table 11. FarmGrow AO ratings at case study farm

FarmGrow Adoption Observations	Rating
Genetics	**Bad**
GAPS	**Bad**
Pruning	*Bad*
Pest, Disease, Sanitation	*Medium*
Weeding	*Bad*
Harvesting	*Good*
Shade Management	*Bad*

Map to the Future (M2F) page 26

Soil Fertility	Bad
Soil Condition	*Good*
Organic Matter	*Good*
Fertilizer Formulation	*Bad*
Fertilizer Application	*Bad*
Farm Condition	**Bad**
Tree Age	*Good*
Tree Density	*Bad*
Tree Health	*Good*
Debilitating Disease	*Good*

A comparison between the LandPKS measured soil property values and the soil property values mapped at this location showed a similar trend, with a uniform coarse textured, low rock fragment layer from 0-50 cm, followed by a finer-textured layer (50-70 cm) with high rock fragments (Table 12). However, the surface soil texture determined on-site by the observers using the texture key in LandPKS is considerably coarser than the mapped surface texture (~92% sand vs. 63% sand).

Table 12. Comparison of texture and rock fragment classes between HWSD/WISE dominant soil map unit component data and LandPKS measured data

	Soil Maps		LandPKS	
Depth (cm)	**Texture**	**Rock Fragments**	**Texture**	**Rock Fragments**
0-1	sandy clay loam	1-15%	sand	1-15%
1-10	sandy clay loam	1-15%	sand	1-15%
10-20	sandy clay loam	1-15%	sand	1-15%
20-50	sandy clay loam	1-15%	sand	1-15%
50-70	sandy clay	15-35%	clay loam	35-60%

This difference in surface soil texture influences both the intrinsic soil nutrient availability, as well as the ability of the soil to retain nutrients added through fertilization. This is illustrated in Table 13 where the nutrient availability (SQ1) was rated as a severe constraint (S3) based on the measured soil property data (LandPKS + SoilID), compared to a moderate constraint (S2) for the soil map data. For nutrient retention capacity (SQ2), both area- and site-specific calculations were rated as having a slight limitation (S1) for the low and intermediate levels where little-to-no fertilizer is applied. At the high input level where optimal levels of fertilizer are applied, the SQ2 rating increases to no constraint (S0) for the area-specific case and drops to a moderate constraint (S2) for the site-specific case. **This indicates that the sandy surface texture measured at the site limits the ability of the soil to retain added nutrients.**

Based only on soil map data, optimal fertilizer application (high input) would increase the soil suitability for cocoa production from being moderately constrained (S2) at the low and intermediate farming levels, to having no constraints (S0). However, **based on** our measured soil property data at this site, the suitability of the site would increase slightly from a low to intermediate input system (S3 to S2 with minimal nutrient additions), but would not be improved by optimal fertilizer application due to its sandy texture and limited ability to retain the added nutrients (Table 12). **Therefore, the specific recommendation to replant and provide extra soil management or any blanket recommendations to this particular farmer to apply fertilizer may not represent the best return on investment, particularly for a resource-constrained and low-income farmer.** Alternative management strategies for this soil would include the application of organic forms of nutrients that are slowly made available to crops over longer periods of time but this would likely not be sufficient for newly-planted cocoa trees to achieve the goal of 1500 kg / hectare given the overall low nutrient and moisture retention that would be needed for newly-established trees. Alternative crops may be needed for this soil but farmers or the agronomists that support them would need to calculate and balance the return on investment compared to cocoa and the land tenure issues that encourage continued use of cocoa or other long-term tree-based crops.[2]

Table 13. Cocoa soil suitability at case study farm based on HWSD/WISE soil map and LandPKS + SoilID data

Soil Quality Index	Area-Specific (Soil Map)			Site-Specific (LandPKS + SoilID)		
Farm Management Level	**Low**	**Int.**	**High**	**Low**	**Int.**	**High**
	--------------- soil quality ratings (%) ---------------					
SQ1: Nutrient availability	71%	71%	--	53%	53%	--
SQ2: Nutrient retention capacity	89%	89%	97%	89%	89%	65%
SQ3: Rooting conditions	100%	100%	100%	100%	100%	100%
SQ4: Oxygen availability	100%	100%	100%	100%	100%	100%
SQ7: Workability	100%	100%	100%	100%	100%	100%
Suitability Rating (SR)	71%	80%	97%	52%	71%	65%
Suitability Class*	S2	S2	S0	S3	S2	S2

Notes: *S0: No constraint (100%-95%); S1: Slight constraint (95%-85%); S2: Moderate constraint (85%-60%); S3: Severe constraint (60%-40%); S4: Very severe constraint (40%-10%); N: Not suitable (<10%)

Since fertilization is a common recommendation to improve soil quality, gaining a deeper understanding of how soils will respond to nutrient additions is essential if we are to optimize a farmer's return-on-investment. With growing interest in sustainable intensification, this approach provides a framework to evaluate how soils will likely respond to changes in management practices, and therefore allows farmers to refine and improve their management activities over general blanket recommendations.

Case Study 2: FarmGrow-LandPKS Soil Suitability for Farm Diversification

As mentioned earlier, the AEZ framework can calculate crop-specific soil quality and soil suitability ratings for 54 different agricultural crops. A farmer interested in diversifying their farm either through agroforestry practices or through converting from cocoa to another crop can evaluate the suitability of their soil for that crop. In our test case, the farmer described above in Case 1 also grows dryland rice in addition to cocoa. Table 14 shows the soil quality and soil suitability ratings for dryland rice at our example farm.

Soil quality results for dryland rice were similar to those of cocoa. **The sandy soil texture in the top 50 cm of the soil affects the soil's intrinsic fertility and ability to retain nutrients, both of which are limitations for most agricultural crops.** For dryland rice, the nutrient availability rating (SQ1) was slightly higher and nutrient retention capacity (SQ2) slightly lower than cocoa for the low and intermediate input systems at this site. In high input farming systems the SQ2 ratings were similar for the two crops, with area-specific ratings over estimating crop suitability relative to site-specific estimates (Table 14). For the overall soil SR, growing dryland rice in a low-input system is predicted to face a severe constraint (S3) based on the measured soil property data (LandPKS + SoilID) compared

---

[2]Under traditional land tenure agreements in Ghana, a farmer's land rights are tied to the trees growing on the land. Cutting down cocoa trees without replanting new ones can result in a farmer losing their lease of the land, particularly given the limited use of documented land agreements. This results in few incentives to rehabilitate farms or grow alternative crops.

to a moderate constraint (S2) for the soil map data. In a high input system, dryland rice would face moderate constraints based on soil property data (LandPKS + SoilID) compared to slight constraints based on soil map data (Table 14). These results indicate the need for soil management recommendations similar to cocoa, where soil nutrient additions need to be judiciously applied in order to prevent nutrient losses due to the soil's low nutrient retention capacity.

Table 14. Dryland rice suitability at our case study farm based on soil map and LandPKS data applied to the FAO's AEZ soil suitability modelling framework

AEZ Soil Quality Index	Area-Specific (Soil Map)			Site-Specific (LandPKS + SoilID)		
Farm Management Level	**Low**	**Int.**	**High**	**Low**	**Int.**	**High**
		--------------- soil quality rating (%) ---------------				
SQ1: Nutrient availability	84%	84%	--	60%	60%	--
SQ2: Nutrient retention capacity	77%	77%	94%	79%	79%	64%
SQ3: Rooting conditions	100%	100%	100%	100%	100%	100%
SQ4: Oxygen availability	100%	100%	100%	100%	100%	100%
SQ7: Workability	100%	100%	100%	100%	100%	100%
Suitability Rating (SR)	84%	81%	94%	59%	69%	64%
Suitability Class*	S2	S2	S1	S3	S2	S2

Notes: *S0: No constraint (100%-95%); S1: Slight constraint (95%-85%); S2: Moderate constraint (85%-60%); S3: Severe constraint (60%-40%); S4: Very severe constraint (40%-10%); N: Not suitable (<10%)

Lessons Learned

Technical Challenges and Organizational Experience

The M2F project started in late February 2020 and shortly after, the COVID-19 pandemic caused the lockdown of countries globally, including Ghana where the project was executed and the United States where technical resources were coming from. The rapidly changing and challenging situation posed by the COVID-19 pandemic created challenges for this project, but we were able to move forward without substantial changes to our research plan. Through the local Grameen office in Ghana, we were able to work effectively through video and phone calls to identify local resources that could support the research activities remotely. Whatsapp was used to monitor field activities closely and get instantaneous feedback. Through these early efforts we were able to enlist the help of Dr. Edward Yeboah of CSIR - Soil Research Institute. Dr. Yeboah's team led the soil training activities for the contractor, Escape Poverty Africa, who carried out all the field data collection. Establishing a strong collaboration with the CSIR - Soil Research Institute early in the project was a critical factor allowing us to carry out all of our planned research activities. The LandPKS soil sampling activities occurred during Ghana's rainy season which significantly slowed the data collection efforts and required continued adjustments to the sampling protocols (Appendix 3). Despite these obstacles, we were able to sample 225 cocoa sites. In addition, the team planned to have two in-person workshops, one in Ghana and one with World Bank staff in the United States, to share results as well as engage stakeholders in fruitful conversations regarding the implications of the research. At the time of writing this report, these in-person events were changed to a virtual event that would be held in early 2021. While in-person engagements would have been preferable, holding a virtual event will allow a wider range of actors to participate and engage with the research.

Potential for Replicability and Scalability

A main goal of this study was to investigate methods of integrating site-specific soil data with traditional soil map data in an effort to improve smallholder agronomic recommendations. An important consideration in this work was the ability to replicate and scale our data collection and modeling approach. Gathering site-specific soil data is often labor intensive and costly to implement at a large scale. The LandPKS app simplifies the soil data collection process, making it more feasible for agronomists and farmers to collect soil data and gain benefit. This study conducted a complete LandPKS soil and site characterization at the 225 cocoa sites which limited the number of sites we could visit within the sampling time-frame. The LandPKS sampling protocol could be simplified in future efforts through adjusting the sampling method. For example, sampling soil depths with an auger, minimizing the number of sampling depths (e.g., 0-30, 30-100 cm), measuring fewer soil properties (e.g., only texture and rock fragments). These types of adjustments can significantly streamline the soil sampling process and make it more feasible for larger-scale sampling efforts. A simplified LandPKS sampling methodology could still be incorporated into the AEZ soil suitability modelling framework, allowing for its broad-scale application. In the case of cocoa farming systems, we know that more than 80% of the cocoa root system is in the top 30 cm of the soil. Therefore, future LandPKS sampling efforts in cocoa systems could limit the soil sampling depth to 30 cm and limit the measured soil properties to texture and rock fragments. These changes to the sampling protocol would significantly shorten the time required to sample each farm site. Similar modifications can be made in other projects based on the known characteristics and requirements of different crops.

The direct implication of this research to FarmGrow is that 1) there is limited benefit of integrating soil map data without the benefit of site-specific soil data supplied through an application like LandPKS and given the inaccuracies of the soil map data predicting soil conditions for cocoa farmers and 2) the long-term investment plan could benefit from integration with applications such as LandPKS that provide site-specific soil recommendations. As part of the FarmGrow product roadmap, such site-specific soil recommendations that support a modified investment plan that includes crop diversification and soil health management practices to improve the farmer's return on investment, beyond optimization of fertilizer use, would be more responsive to the income needs of farming households to make investments, such as farm rehabilitation or renovation. While it is not the purpose of this study to delve into the cost/benefit analysis of including a more rigorous site-specific soil data collection process into the FarmGrow processes, we believe that it merits further study and that approaches on data collection, especially during initial farmer profiling, can be streamlined to keep the service costs at a reasonable level. Some farmer observation data could be replaced with data collected through the LandPKS app (e.g., LandPKS's LandInfo and SoilHealth modules), therefore reducing agronomist or farmer error.

Opportunities for Other Farmer Decision Support Tools

Crop-specific decision support tools, like FarmGrow for cocoa smallholder farmers, provide a robust framework for understanding agronomic limitations. LandPKS, in contrast, is crop-agnostic, providing farmers and agronomists a systematic framework for measuring site-specific soil property information. Most farmer decision support tools are deficient in soil information and therefore the integration of these crop-specific tools with more generalized soil tools like LandPKS can provide significant benefit. While there was a clear benefit to integrating FarmGrow and LandPKS, future applications of the modeling approach presented in this study do not require crop-specific application like FarmGrow. Current and future World Bank projects only interested in soil specific crop limitations could use the LandPKS and AEZ modeling framework, allowing for broader scale application.

Conclusion and Future Refinement

This research project integrated site-specific soil data with traditional survey data to explore the feasibility of supporting World Bank projects which aim to improve the productivity of small-holder cocoa farmers. It focused on whether integrating new and traditional sources of available data could generate more insights into soil quality to inform current (and future) World Bank projects seeking to strengthen yield potential for smallholder cocoa farmers in Ghana. Healthy soils are essential for the productivity of cocoa farms to produce pods and repair damaged parts of the trees. Gaining a deeper understanding of how soils will respond to nutrient additions is essential if World Bank projects are to help farmers identify efficiencies that lead to higher productivity and profitability, lower input costs, and optimized fertilizer use.

Long-term soil degradation, and in particular soil infertility, is recognized as one of the main factors limiting cocoa yields in Ghana. The research team sought to leverage data from FarmGrow, LandPKS and existing soil maps to provide a more accurate assessment of the inherent soil properties that may affect crop yields. However, a major limitation in assessing farm scale soil suitability and limitations lies is acquiring accurate soil property data. In our initial analysis of soil map data in Northern Ghana, we found SoilGrids was the most similar to LandPKS soil data. However, our more recent analysis in the Ashanti and Western regions revealed that SoilGrids was the least similar to LandPKS. This illustrates the variable accuracy of soil map products. Additional analysis is needed to further evaluate factors affecting soil property similarity (e.g., land-use, climate, topography) so that we can better understand which soil map areas provide accurate predictions and where they are less reliable. We are hopeful that the use of higher resolution mapping products such as the national Ghana Soil Map, with a map scale of 1:400,000, will provide even more accurate predictions of soil properties.

For FarmGrow cocoa farmers, location-specific soil data generated by the LandPKS app has the potential to provide tailored information for a farm's specific need, particularly around optimizing fertilizer use for improved soil health. However, as a result of this research we recognize the need to be mindful of inherent soil properties and fertilizer efficiency. Given that most cocoa farmers in Ghana are practicing low-input farming on depleted cocoa soils, any blanket-application of fertilizer without knowledge of the site-specific soil condition may fail to provide the expected benefits. While site-specific soil data is essential for making informed soil management decisions, acquiring the data is often cost-prohibitive.

Therefore, future efforts to streamline the soil sampling process would be needed in order to allow FarmGrow agronomists to integrate this data with their current data collection efforts. Future refinement of this work may also include integration of the site-specific AEZ soil suitability analysis into apps such as FarmGrow and LandPKS. This would allow end users in low resource settings to identify crops most suited for intercropping with cocoa based on their soil's condition, in an effort to improve soil fertility through increased biodiversity on cocoa farms.

Map to the Future (M2F) page 31

Appendix 1: FarmGrow Cocoa Adoption Observations

Adoption Observations and Rationale

Adoption Observations		Mechanism to monitor	Rationale
Plant Material	1. Planting Material - Genetic Potential	Interview and Observation	Plant material determines maximum yield - it must produce 1.5 MT/ha or more
Farm Condition	2. Tree Age	Interview and Observation	Trees over 25 yrs. old must be replaced as they are or will soon be in decline
	3. Tree Density	Observation	We need maximum production per ha and need between 800 - 1350 trees/ha
	4. Tree Health	Observation	If many trees are in poor health, it is better to replace
	5. Debilitating Disease	Observation	If there is a disease such as CSSV, trees must be replaced
GAP	6. Pruning	Observation	Only good pruning will ensure both energy and nutrient sequestration to pods
	7. Pest and Disease (P&D) and Sanitation	Observation	Only good P&D management will protect high pod load
	8. Weeding	Observation	Good weeding allows fertilizer uptake by trees
	9. Harvesting	Observation	Good harvesting (leave nothing on the tree) to reach highest production
	10. Shade Management	Observation	Light shade is wanted to allow enough sunlight, but also some stress protection
Soil	11. Soil Condition (pH separately)	Observation	Only good soil condition (not too argillic, sandy, rocky etc.) allows high yield
	12. Organic Matter	Observation	Organic matter supports high microbial activity
	13. Fertilizer Formulation	Interview	We need all nutrients, and in the right ratios, whilst we avoid Urea and Ammonia
	14. Fertilizer Application	Interview	We need enough fertilizer, in the right place at the right time to support 1.5 MT/ha

Adoption Observations and Assessment Summary

Rating	Criteria
A. Plant Material Genetics	
Plant Material: What is the yield potential of planting material used at the farm?	
Good	<u>Interview:</u> • >80% of Plant Material sourced after 1990 from research station, extension service, accredited plant material distributor <u>OR</u> • if historical known Yield reached 1500kg/ha <u>Field observation:</u> • Identification of clone or hybrid OR • If in peak season: yield on tree

Map to the Future (M2F) page 32

Medium	<u>Interview:</u> • >80% of Plant Material sourced before 1990 from research station, extension service, accredited plant material distributor <u>OR</u> • if historical known Yield was between 900-1500kg/Ha <u>Field observation:</u> • If in peak season: yield on tree
Bad	<u>Interview:</u> • Plant Material source not known or taken from farms with unknown parentage <u>OR</u> • if historical known Yield never reached 900 kg/Ha <u>Field observation:</u> • If in peak season: yield on tree <u>OR</u> • other indicators of low yield potential i.e. 70/30 yield distribution
B. Farm Condition	
B1: Tree age: Are the trees above or below the theoretical maximum production threshold?	
Good	<u>Interview:</u> • <26 years <u>Observation:</u> • best judgement
Bad	<u>Interview:</u> • 26 years and older (age 25 - 30 graft or replant, > 30 only replant) <u>Observation</u> : • best judgement
B2. Tree density: Does the density of trees support targeted production per hectare? (i.e. spacing between trees as proxy to number of trees and average density)	
Good	<u>Observation:</u> • Farm has adequate density (800 – 1320 trees per ha)
Bad	<u>Observation:</u> • Farm has poor density (<800 trees per ha or more than 1320 tree/ha)
B3. Tree health: Are the trees on a farm healthy enough to support targeted yield?	
Good	<u>Observation:</u> • >80% trees are healthy and without physical damage
Bad	<u>Observation:</u> • >20% of trees look unhealthy with irreparable problems (i.e. cannot be fixed by GAP or soil management) <u>OR</u> •20% of trees with physical damage
B4. Debilitating disease: Is the farm free of any signs of major diseases that may imperil the farm?	
Good	<u>Observation:</u> No observable CSSV on the farm
Bad	<u>Observation:</u> Evidence of CSSV on the farm
C. Good Agricultural Practices	

<table>
<tr><th colspan="2">C1. Pruning</th></tr>
<tr><td>Good</td><td><u>Observation</u>:
Hybrid Trees, >90% of trees must have:
• Max height of the tree: < 4.5 m <u>AND</u>
• 3-5 main branches <u>AND</u>
• All main branches visible <u>AND</u>
• >50% of leaves capture direct light <u>AND</u>
• Good aeration under and in the tree canopy <u>AND</u>
• chupons on <10% of trees

<u>Other criteria to support positive judgement</u>
• Height of Jorquette: <1.5m <u>AND</u>
• Branches exhibit vertical growth habit <u>AND</u>
• Canopies of trees do not touch (CSSVD prevention) <u>AND</u>
• Mostly single stem trees

Clonal Trees, major criteria of all trees:
<u>Observation:</u>
• Height of tree < 3.5 m <u>AND</u>
• 2-3 main branches, in balance, clearly visible <u>AND</u>
• >75% of leaves capture direct or a lot of indirect light <u>AND</u>
• good aeration in the whole farm <u>AND</u>
• chupons on <10% of trees

<u>Other criteria (to support positive judgement)</u>
• Branches exhibit vertical growth habit <u>AND</u>
• Canopies of trees do not touch each other (CSSVD prevention)</td></tr>
<tr><td>Medium</td><td><u>Observations:</u>
Hybrid Trees, >90% of trees must have:
• Max height of the tree: < 5 m <u>AND</u>
• 2-5 main branches, in balance <u>AND</u>
• all main branches are visible <u>AND</u>
• 50% of leaves likely to capture direct and indirect light <u>AND</u>
• good aeration <u>AND</u>
• Chupons on <25% of trees

<u>Other criteria to support positive judgement</u>
• Height of Jorquette: 1.5-2m <u>AND</u>
• Branches exhibit at least some vertical growth habit <u>AND</u>
• <25% - 50% of canopies of trees touch each other <u>AND</u>
• Mostly single stem trees

<u>Observations:</u>
Clonal Trees, >90% must have:
• Height of tree < 4.5 m <u>AND</u>
• Max 4 main branches, in balance, clearly visible <u>AND</u>
• 50-75% of leaves likely to capture light <u>AND</u>
• Good aeration

<u>Other criteria to support positive judgement</u>
• Branches exhibit mostly vertical growth habit <u>AND</u>
• Some (<25%) canopies of trees touch each other <u>AND</u>
• Chupons on <10% of trees</td></tr>
</table>

Bad	Observations: **Hybrid Trees, most trees on the farm have** • Height of the tree: > 5m OR • Only one stem until crown or >5 main branches, poor balance, some or most main branches not visible OR • Most leaves are not likely to capture light and trees are not aerated well under or within the canopy • >25% chupons on the trees Other criteria (to support negative judgement) • Height of Jorquette: >2m OR • Most branches have horizontal growth habit OR • >25% of canopies of trees touch each other OR • many multiple-stem trees (>25%) Observations: **Clonal trees, most trees have** • Height tree > 4.5 m OR • >3 main branches, poor balance, most branches not visible OR • <50% of leaves do not capture enough light OR • poor aeration under or within canopy OR Other criteria (to support negative judgement) • Branches exhibit mostly horizontal growth habit OR • >50% of canopies of trees touch each other OR • >25% chupons on the trees
C2. Pest, Disease and Sanitation: What is the Pest and Disease (P&D) and Sanitation condition for supporting or limiting the yield potential of the planting material?	
Good	Observation: **P&D** • Spread of pest disease is low measured by few pods and branches affected on < 10% of the trees OR only in a few pockets on <10% of farm area) AND • the P&D presence causes little loss **Sanitation** • trees are nearly free of diseased, damaged, wilted, dead or mummified pods, epiphytes, or ant nests and tunnels AND • no diseased plant material on the ground near the tree
Medium	Observation: **P&D** • < 25% of trees have significant presence of non-debilitating diseases on pods, stems and branches leading to loss of <15% Sanitation • < 25% have diseased, damaged, wilted, dead or mummified pods, epiphytes, dead branches, or ant nests and tunnels AND • <25% of land have some diseased plant material on the ground near the tree

Bad	Observation: **P&D** • > 25%) have significant presence of non-debilitating diseases on pods, stems and branches leading to significant loss of >20% OR • The spread of diseases to many trees all over the farm Sanitation • > 25% of trees have diseased, damaged, wilted, dead or mummified pods, epiphytes, dead branches, or ant nests and tunnels OR • >25% of land has diseased plant material on the ground near the tree
C3. Weeding: What is the weeding condition for supporting or limiting the yield potential of the planting material?	
Good	Observation: • The ground under the canopy of trees is kept clean of undesired undergrowth and very little weed is visible
Bad	Observation: • Undesired undergrowth or weeds up to knee height on >10%) of the farm and outside canopy of cocoa trees OR • >10% of area under canopy of cocoa trees has weeds
C4. Harvesting: What is the harvest condition for supporting or limiting the yield potential of the planting material?	
Good	Observation: • Few over-ripe pods on maximum 10% of the trees AND • <10% under-ripe pods harvested (if this can be observed)
Bad	Observation: • >10% of trees have over-ripe pods OR • >10% of harvested pods are under-ripe (if this can be observed)
C5. Shade: What is the shade level for supporting or limiting the yield potential of the planting material?	
Good	Observation • Good shade is light shade which can be measured by 70 - 80% of sunlight reaching the canopy of most cocoa trees OR presence of 12 to 18 large shade trees of >20 m tall per ha AND • >75% cocoa trees receive shade during part of the day AND • Shade trees are compatible with cocoa i.e. no host of disease, no competition for root or canopy space, no breaking branches
Bad	Observation: Bad shade is insufficient shade or too much shade which is measured by <70% or more than 80% of sunlight reaching the canopies of most cocoa trees OR < 12 or > 18 large shade trees of > 20m tall per ha OR • <75% receive shade during part of the day OR • Shade trees that are not compatible with cocoa i.e. host of disease, competition for root or canopy space, no breaking branches
D. Soil Fertility Management	
D1. Physical condition of farm land (soil condition): What is the physical condition of the land and its limiting factors for cocoa cultivation?	

Good	Observation: • No signs of erosion, no roots visible on the surface AND • few rocks or gravel on farm surface or in the ground as measured by 3 holes of 30 cm deep per plot AND • soil is neither too sandy or argillic as measured by touch/roll test on soil from 3 holes of 30 cm deep per plot AND • well drained either naturally or through drainage canals AND • slope < 15%
Bad	Observation: • signs of erosion, roots visible on the surface OR • many rocks or gravel on farm surface or in the ground as measured by 3 holes of 30 cm deep per plot OR • soil is too sandy or too argillic measured by touch/roll test on soil from 3 holes of 30 cm deep per plot) OR • poorly drained (waterlogged) OR • slope > 15%
D2. Organic Matter (Soil Health): What is the volume and level of decomposition of organic matter on and in the soil and what are other indicators of soil health i.e. worm, insect activity and microbial life for supporting or limiting the yield potential of the planting material?	
Good	Observation: • Clear signs of microbial activity everywhere on the farm with multiple layers of decaying organic material covering the soil under the cocoa canopies of all trees, worms, worm castings, insect activity, soil pores AND • Organic material left in the farm and/or extra organic material (compost, manure) around cocoa trees or in ‘mulching rows or trenches’ evenly spread through the farm (note: pod husk left in the farm is a strong positive indicator)
Bad	Observation: • >10% of soil under the cocoa tree canopies is exposed without at least one layer of decaying organic material OR • Little or no signs of organic material in the farm or microbial activity in the soil
D3. Fertilizer Formulation: What kind (formulation) of fertilizer is used at the farm i.e. nutrient content, nutrient balance and non-acidifying and does it support or limit the yield potential of the planting material?	
Good	Interview: • Use of well-balanced NPK + Secondary + Micro nutrients fertilizer with N in CaNitrate AND • No use of Urea AND • If pH <5.7 apply mechanism to add Ca to soil i.e. use relevant dose of lime, higher doses of Nitrabor, more organic material etc.
Medium	Interview • Use of Ammonium based NPK + Secondary + Micro nutrients fertilizers with reasonable nutrient balance, if accompanied with significant doses of lime/kieserite/dolomite or Nitrabor AND • No use of Urea AND • If pH <5.8 apply mechanism to add Ca to soil i.e. use relevant dose of lime, higher doses of Nitrabor, more organic material etc.
Bad	Interview • Use of poorly balanced fertilizer OR • Use of Ammonium based N without additional lime, Kieserite or Dolomite OR • Use of Urea
D4. Fertilizer application: How is fertilizer used i.e. dosage, timing and application technique, and does it support or limit the yield potential of the planting material?	

Map to the Future (M2F) page 37

Good	Interview For details see manual For Mature trees and to sustain 1.5 mt/ha AND • > 700 kg/ha of all fertilizer combined excluding lime/dolomite AND • applied under the leaf litter or in the soil at the root system of the trees AND • applied at least once per year
Medium	Interview For details see manual For Mature trees and to sustain 1.5 mt/ha AND • > 400 - 700 kg/ha of all fertilizer combined excluding lime/dolomite AND • applied mostly under the leaf litter or in the soil at the root system of the trees AND • applied at least once per year
Bad	Interview For details see manual For Mature trees and to sustain 1.5 mt/ha OR • < 400 kg/ha of all fertilizer combined excluding lime/dolomite OR • mostly applied close to trunk or far from tree root system, applied on leaf litter OR • applied <1 time per year

Appendix 2. Comparing Existing Soil Map Data to Site-specific Soil Data

The ability to identify a soil's type or series provides a critical link to soil information a farmer can use to implement sustainable soil management practices. The main objective of this analysis was to evaluate the relative accuracy of soil maps available for Ghana using existing maps and soil data points. This appendix describes the results from two different research tasks designed to evaluate the relative accuracy of current soil maps.

The first task (Task 1) is focused on inherent soil properties. These include factors such as sand or clay content. Inherent soil properties can be used to inform farmer decisions on a variety of management practices such as irrigation frequency, need for organic amendments, likelihood of erosion, and so on. This task performed a comparison of inherent soil properties (e.g., texture, rock fragments) between soil maps (dominant mapped soil) and LandPKS study sites in northern Ghana. LandPKS data provides information on inherent soil properties (the long-term potential of the soil), including soil texture and rock fragment volume. These properties directly affect cocoa production. They also determine how susceptible soils are to declines in fertility, and how responsive they are likely to be to different types and amounts of fertilizer and organic amendments such as compost and manure. The second research task (Task 2) focused on soil type (taxonomy). Soil types can be used to inform decisions tailored to a farmer's specific needs on a range of management practices to improve soil health management decisions including (for example) the use of modifiers (e.g. lime) of soil pH and the need for nutrient specific fertilization to address soil deficiencies. Task 2 performed a comparison of the World Reference Base (WRB)[3] soil taxonomic classification (Reference Soil Group [RSG]) between soil maps (dominant mapped soil) and the Ghana Soil Profile Database (GSPD), a collection of historic soil profile descriptions collected from different soil mapping and/or research studies in Ghana. The GSPD soil profiles were collected by professional soil scientists and each profile is classified to a soil type using the World Reference Base (WRB) soil taxonomic system.

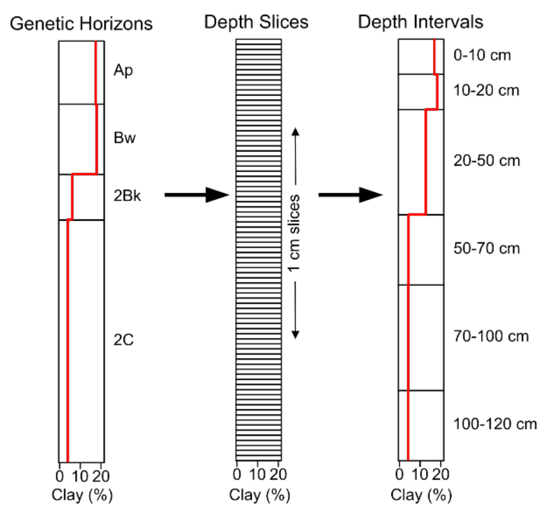
[3] WRB is an international soil classification system for naming soils and creating legends for soil maps.

These two research tasks provided insight into the relative accuracy of the different soil mapping products both in terms of their functional similarity (indicated by a similarity in inherent soil property values from (Task1) and a taxonomic similarity through matching RSGs (Tasks 2). Both of these types of properties can be obtained from geospatial soil map resources and from the LandPKS mobile application either directly (e.g. texture) or indirectly (e.g. soil type) following assessment of a site. Results from these analyses show that (1) on average all of the soil map products have some value in predicting soil texture in northern Ghana but high variability may limit their utility when applied at the farm-scale, and (2) that SoilGrids provides more accurate soil property and taxonomic class predictions in Northern Ghana relative to either HWSD or WISE. These analyses are a key step toward understanding differences in soil map products, the relation of these products to soil profile data (e.g., LandPKS, GSPD), and the next steps for building linkages between site specific soil assessment and potential farming recommendations.

Task 1: Relative accuracy of soil maps in Northern Ghana—Soil property similarity

To evaluate the relative accuracy of soil maps in Northern Ghana we compared the mapped soil property data at 5,815 LandPKS sites for each of the available soil maps to the measured (i.e., validation) LandPKS soil property data collected at each site. LandPKS soil profiles were sampled in the Northern, Upper West, and Upper East regions of Ghana (Figure 1) as part of a monitoring program of USAID Feed the Future projects. Data collection was performed by trained field crews following standard LandPKS sampling protocols. To account for differences in soil property depths between the LandPKS sites and the different soil maps, all soil map profiles were segmented at 1 cm increments and reaggregated at LandPKS standard depth intervals (i.e., 0-1, 1-10, 10-20, 20-50, 50-70, 70-100, 100-120 cm) (Figure 2). Quality control filtering was performed on LandPKS data to remove incomplete or questionable sites. This included excluding sites with less than three measured soil layers or sites that failed a simple logic filter based on incompatible data inputs (i.e., sandy textures and deep vertical cracking).

Figure A1.1. Soil profile slicing and aggregation method for converting contrasting soil sampling depths to the standard LandPKS sampling depths



While the availability of different soil mapping products has increased in Ghana over the past several decades, current soil maps are still only available at coarse spatial scales (i.e., 1:400,000 to 1:5,000,000). Table A1.1 lists the four soil mapping products available in Ghana. Traditional soil maps (i.e., HWSD, WISE, Ghana Soil Map) do not show the exact location of a soil series but instead display Soil Map Units (SMUs) representing distinct areas of a landscape composed of one or more soil series.

A common method for dealing with this spatial uncertainty is to assign any location within a SMU to its dominant soil series. In our comparisons of soil property values, we used the property values associated with the dominant SMU soil. In Ghana, the Harmonized World Soil Database (HWSD) is derived from the FAO-UNESCO Digital Soil Map of the World (DSMW) which has a map scale of 1:5,000,000. HWSD soil property data is derived using actual soil profile data from the World Inventory of Soil Emission Potential (WISE) soil profile database and pedotransfer rules, producing two aggregated soil depth intervals (0-30 and 30-100 cm). The WISE soil map is a recent improvement upon HWSD, where an expanded WISE soil profile database and new pedotransfer rules were used to derive soil profile data at 7 standardized depth intervals (0-20, 20-40, 40-60, 60-80, 80-100, 100-150, 150-200 cm). HWSD and WISE have identical spatial data but differ in their soil property data (2 vs 7 depths for HWSD and WISE, respectively). The Ghana Soil Map (not available for this report) is a traditional polygon-based soil map with a map scale of 1:400,000.

Digital soil mapping products (e.g., SoilGrids) offer an alternative to traditional soil maps by providing predictions of soil properties and classes at specific locations. SoilGrids is a global digital soil map that predicts soil properties at a 250 m spatial resolution at seven standard depths (0, 5, 15, 30, 60, 100 and 200 cm).

We evaluated the accuracy of each soil map by calculating the statistical distance between their mapped soil property values and the soil property values measured at each LandPKS site. Statistical distance or dissimilarity was calculated as:

Eq. 1
$$D = \frac{|X_O - X_P|}{100}$$

where D is the statistical distance or dissimilarity for property X, and X_O and X_P are the observed (i.e., LandPKS) and predicted values for property X. The following soil properties were evaluated, percent clay, percent sand, and rock fragment volume. Since LandPKS soil texture and rock fragment volume data are class-based measurements, we converted all numeric soil property values to their corresponding property class and then derived the representative numeric value associated with that property class. For example, if the SoilGrids sand and clay percentage values were 14 and 20, respectively, it would classify as a silt loam texture. Based on the representative sand and clay values for a silt loam, SoilGrids sand and clay values would be reassigned as 25 and 13.5, respectively (Table A1.8, A1.9). These generalization steps ensure that all data comparisons are made using the same level of generalization.

Table A1.1. Soil mapping products in Ghana

Soil Map	Map Extent and Scale	Map-unit	Depth support
HWSD	Global, 1:5,000,000	Polygon	2 layers: 0-30, 30-100 cm
WISE	Global, 1:5,000,000	Polygon	7 layers: 0-20, 20-40, 40-60, 60-80, 80-100, 100-150, 150-200 cm
SoilGrids	Global, 250 m	Raster	7 depths: 0, 5, 15, 30, 60, 100, 200 cm
Ghana Soil Map†	National, 1:400,000	Polygon	Genetic horizons

†Not evaluated in this draft report

Table A1.2. Georeferenced soil profile data in Ghana

Soil Profile Data	# points	Taxonomy	Soil properties†	Depth support
LandPKS	5,815	No	TXCL, RFV	Up to 7 layers: 0-1, 1-10, 10-20, 20-50, 50-70, 70-100, 100-120 cm
Ghana Soil Profile Database	53	Yes	TXCL, RFV	Genetic horizons

†TXCL, soil texture class; RFV, rock fragment volume

We evaluated 5,815 LandPKS study sites which comprised a total of 33,953 individual soil layers. The distribution of soil texture classes for the LandPKS sites and corresponding dominant mapped soils are shown in Table A1.3. Texture classes in Table A1.3 are ordered from low-to-high based on their available water holding capacity (AWC) (i.e., sand = low AWC, clay = high AWC). LandPKS sites have predominantly coarse textured soil with 72% of soil layers classified as sandy loam or coarser. In contrast, HWSD only has 13% of soil layers classified as sandy loam or coarser. WISE and SoilGrids are more similar with 40% and 64% of soil layers classified as sandy loam or coarser, respectively. LandPKS sites span a wider range of textures which is expected given the natural variability of soil texture in the field versus the representative texture values used to populate soil maps. Table A1.4 shows the distribution of soil rock fragment classes for the LandPKS sites and corresponding dominant mapped soils. LandPKS sites show a range of soil rock fragment classes, with an almost equal distribution among the first four classes. HWSD predicted low rock fragments (90% = 1-15% class) while WISE and SoilGrids had higher percentages in the higher rock fragment classes which more closely aligned with LandPKS values.

Table A1.3. Percent distribution of soil texture classes among the different soil data sources

	LandPKS	HWSD	WISE	SoilGrids
Texture Class†	--- % LandPKS Soil Layers (n=33,953) ---			
Sand	27	--	--	--
Loamy sand	26	--	--	--
Sandy loam	19	13	40	64
Loam	4	28	11	3
Silt loam	4	--	--	--
Silt	--	--	--	--
Sandy clay loam	7	29	26	31
Clay loam	3	21	15	2
Silty clay loam	--	--	--	--
Sandy clay	2	--	--	--
Silty clay	--	--	--	--
Clay	1	9	9	--

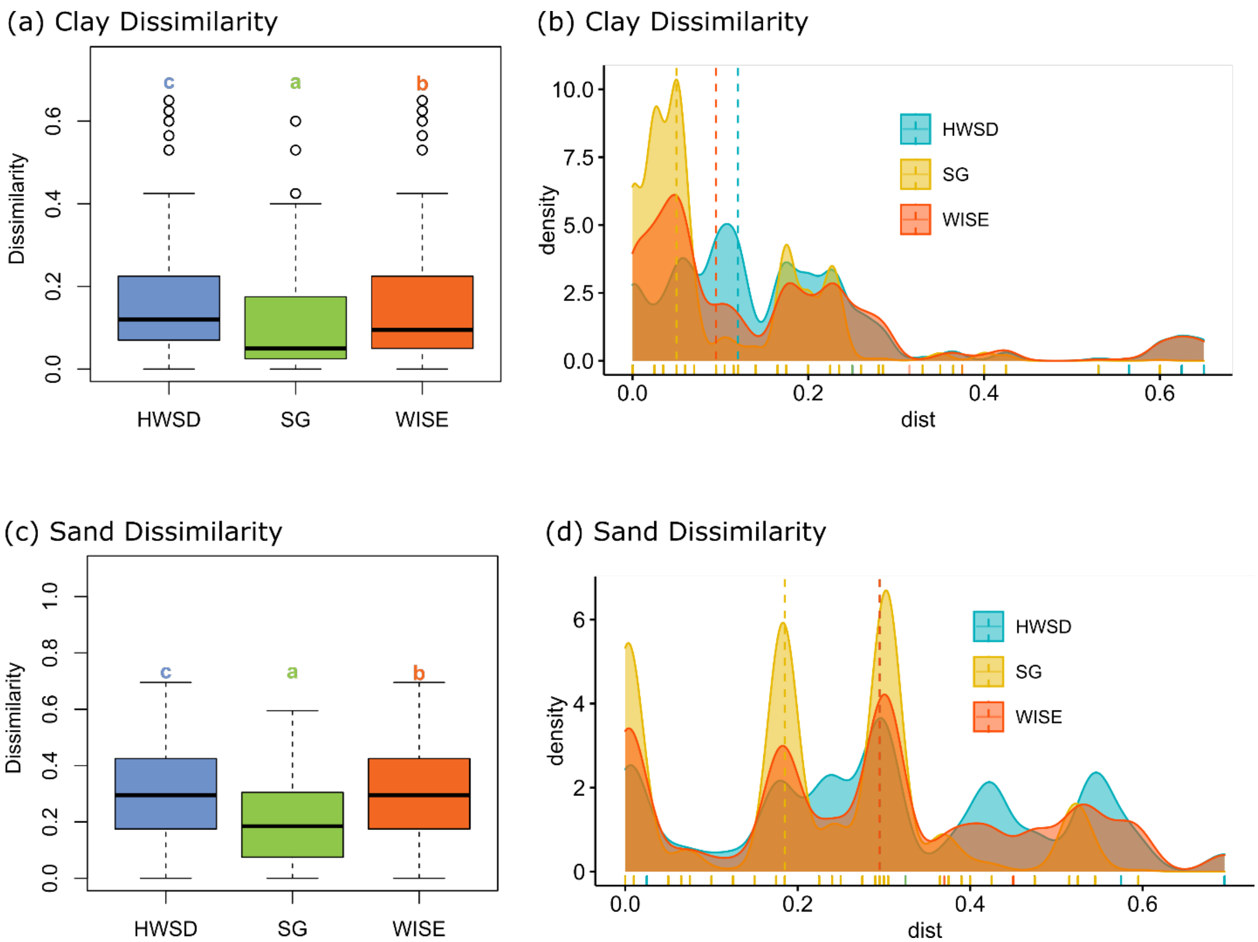
†Texture classes are ordered by their available water holding capacity (AWC), e.g., Sand = low AWC and Clay = high AWC.

Table A1.4. Percent distribution of rock fragment classes among the different soil data sources

	LandPKS	HWSD	WISE	SoilGrids
Rock Fragment Class	--- % LandPKS Soil Layers (n=33,872) ---			
0-1%	23	--	--	--
1-15%	25	90	61	64
15-35%	21	10	21	36
35-60%	20	--	18	--
>60%	1	--	--	--

Dissimilarity values were calculated between soil property values from LandPKS and the dominant soil from each soil map using Eq. 1. This resulted in a dissimilarity value for HWSD, WISE, and SoilGrids for each of the 33,953 soil layers. Analysis of variance (ANOVA) was performed to evaluate the pair-wise differences in dissimilarity among the soil maps. Posthoc comparisons between soil maps were made using Tukey's honestly significant difference (HSD). All statistical differences were tested at the $p = 0.05$ confidence level.

Figure A1.2 Boxplots of the dissimilarity between LandPKS and soil mapped data



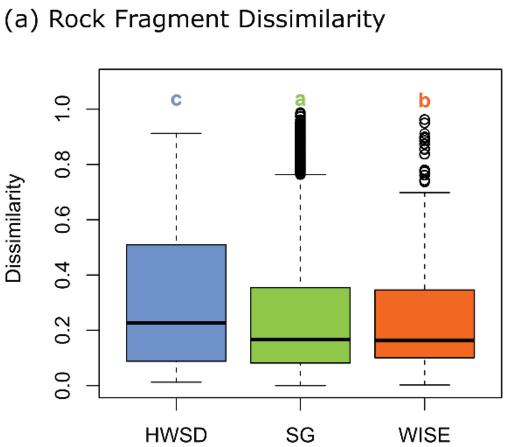
Notes: Boxplots (median ± 95%CI) of the dissimilarity (statistical distance) between LandPKS and soil mapped data for the representative values of (a) percent clay and (c) percent sand. Letters above each boxplot indicate statistically significant differences among the soil maps. Density plots of dissimilarity values for (b) percent clay and (d) percent sand between LandPKS and soil mapped data. Vertical dashed lines in panels (b) and (d) represent the median dissimilarity value for each soil map. The different letters (c, a, b) within each plot indicate significant differences between soil maps for that plot.[4]

Dissimilarity values for clay and sand were statistically different among the soil maps, with SoilGrids having the lowest dissimilarity, followed by WISE and then HWSD (Figure A1.3a). This indicates that, for this particular dataset, the soil texture accuracy for SoilGrids is significantly better (i.e., closer to LandPKS texture values) relative to either WISE or HWSD. Rock fragments showed a similar pattern to soil texture, with dissimilarity values statistically different among the soil maps and with SoilGrids having the lowest dissimilarity, followed by WISE and then HWSD (Figure 4a). On average SoilGrids

[4] The graphical display of significant differences uses a compact letter display where each lower-case letter symbolizes the significant difference between soil maps at the p=0.05 significance level. For example, in a pairwise comparison of clay values between the different soil maps, if we found that that 1) SoilGrids ↔ HWSD = significant difference, 2) SoilGrids ↔ WISE = no significant difference and 3) HWSD ↔ WISE = no significant difference, we would assign the following letter: SoilGrids=a, WISE=ab, HWSD=b. These letters tell us that SoilGrids and WISE are not significantly different because they both share the letter ‘a’. Similarly, WISE and HWSD are not significantly different because they both share the letter ‘b’. However, SoilGrids and HWSD are significantly different because they do not have any letters in common. Given each graphic is showing only one letter indicates that all three soil maps were significantly different.

had only a 9% difference in clay from the LandPKS measured clay, compared to 15 and 17% for HWSD and WISE, respectively (Table A1.5). Relative differences in dissimilarity values among the soil maps were less pronounced for sand with mean values ranging from 21 to 30% (Table A1.5). Density plots of dissimilarity values for clay, sand, and rock fragments showed similar patterns among the soil maps, with SoilGrids having a consistently higher density of low dissimilarity values relative to HWSD and WISE (Figs. A1.3b&d, A1.4b).

Figure A1.4. Boxplots (median ± 95%CI) of the dissimilarity (statistical distance) between LandPKS and soil mapped data for the representative values of the percent rock fragment volume (a). Letters above each boxplot indicate statistically significant differences among the soil maps. Density plots of dissimilarity values for (b) percent rock fragment volume between LandPKS and soil mapped data. Vertical dashed line in panel (b) represents the median dissimilarity value for each soil map.



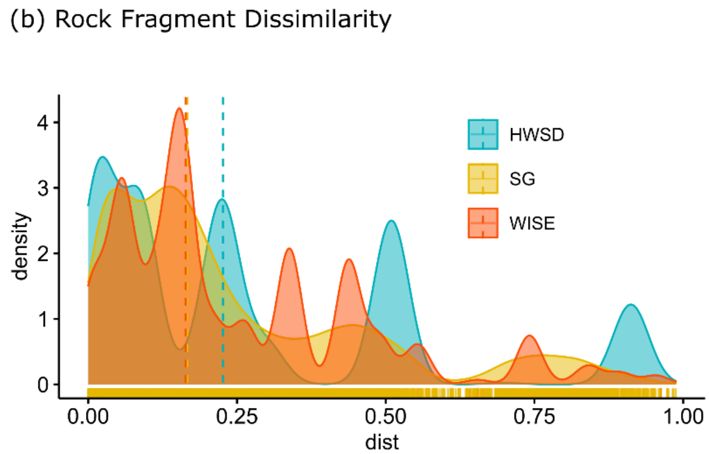


Table A1.5. Summary statistics for soil property dissimilarity values

	HWSD		WISE		SoilGrids	
Dissimilarity	mean	cv†	mean	cv	mean	cv
Clay	17	18	15	19	9	10
Sand	30	26	28	26	21	19
RFV	27	37	26	29	24	29

Notes: † cv=coefficient of variation

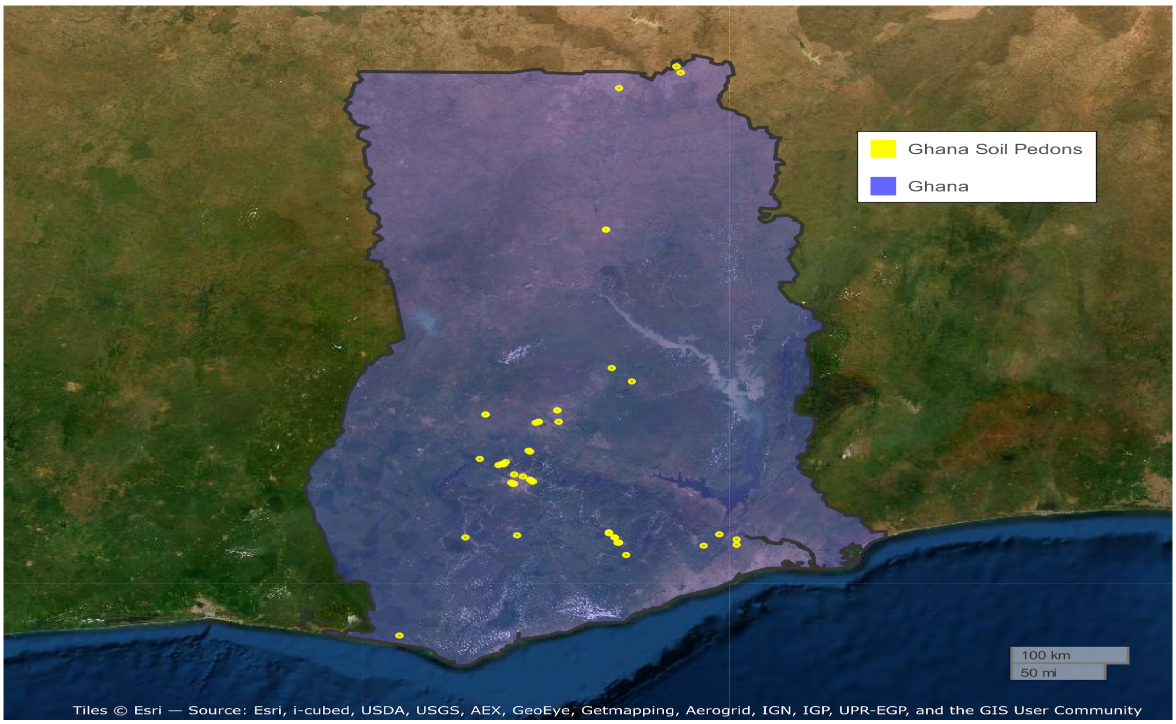
To help further illustrate accuracy differences between the soil maps, we calculated several measures of similarity between LandPKS soil property values and the predicted soil map values. The first measure is based on an exact match between LandPKS and soil map values. In other words, it is a measure of how often the soil map property value was the same as the LandPKS measured value. For soil texture, the percentage of LandPKS soil layers that had the same texture class as the soil map ranged from 6 (HWSD) to 15% (SoilGrids) (Table A1.6). Match rates for rock fragment classes were higher and very similar among the soil maps, ranging from 24-26%.

An additional method for evaluating the relative accuracy of the soil maps is to determine which soil map has the highest similarity (lowest dissimilarity) for each of the 33,953 LandPKS soil layers. We should note that at any given site more than one soil map can be assigned the rank of highest similarity for a given site when they share the same soil property value and that value matches or is closest to the LandPKS property value. Using this approach, SoilGrids property values had the highest similarity to the LandPKS property values for 83% of LandPKS clay values and at 73% of LandPKS sand values. This contrasts with WISE and HWSD which had the highest similarity to LandPKS values at far fewer sites for both clay and sand (HWSD: 34-43%; WISE: 50-59%). For rock fragments, the percentage of sites with the highest similarity was similar among soil maps, ranging from 33-36%.

Table A1.6. Measures of soil property similarity between LandPKS values and predicted soil map values

	HWSD	WISE	SoilGrids
Similarity Measure†	--- % LandPKS Soil Layers (n=33,953) ---		
Texture match	6	11	15
Rock Fragment match	24	25	26
Highest Site Similarity			
Sand	43	50	73
Clay	34	59	83
Rock Fragments	34	36	33

Figure A1.5. Soil pedons in Ghana with WRB taxonomic classifications



Notes: †Multiple soil maps can obtain the same similarity measure (e.g., highest similarity) for a given site when they share the same soil property value and that value matches or is closest to the LandPKS property value

Task 2: Relative accuracy of soil maps in Ghana—Taxonomic similarity

We acquired existing soil profile observations in Ghana from three sources: (1) the African Soil Profile Database (AfSP: 163 profiles), (2) the World Inventory of Soil Emissions potentials (WISE: 11 profiles), and (3) the World Soil Information Service (WoSIS: 432 profiles). These soil profile datasets were collected by trained soil scientists and thus represent high-quality reference data that can be used to evaluate the accuracy of soil maps in Ghana. The datasets were combined and filtered to exclude profiles with missing soil property input data (i.e., soil texture class, rock fragment volume class) and missing taxonomic classifications. These aggregation and filtering steps resulted in 53 soil profiles meeting these criteria and hereafter referred to as the Ghana Soil Profile Database (GSPD)(Table A1.2, Figure A1.5).

At each of these 53 soil profile locations, we queried the dominant soil series from HWSD and WISE soil maps and queried the top predicted soil series class from SoilGrids. We then compared the reference soil group (RSG) (e.g., Acrisol) from each of the soil maps relative to the RSG correlated to each of the GSPD profiles. Results from this analysis show that SoilGrids was twice as accurate in predicting the correct RSG when compared to HWSD and WISE, with a 66% correct match rate compared to 30% and 32% for HWSD and WISE, respectively (Table A1.7).

Table A1.7. Measures of soil property similarity between LandPKS values and predicted soil map values

	RSG match rate
HWSD	30%
WISE	32%
SoilGrids	66%

Initial conclusions on Ghana soil map accuracy

Based upon these initial results, SoilGrids provided more accurate soil property and RSG predictions than either HWSD or WISE soil maps. Results from Task 1 showed that SoilGrids had (1) significantly lower sand and clay dissimilarity values, (2) higher exact match rates for soil texture classes, and (3) a considerably higher occurrence of maximum site-wise similarity for sand and clay relative to HWSD or WISE. Since dissimilarity values were normalized on a 100 percent scale, values reflect the absolute difference in soil property values between the measured LandPKS data and the predicted soil map property values. These results showed that the mean difference (absolute value) in clay percentage for SoilGrids relative to LandPKS was only 9% (9% CV). This means that on average the clay percentage reported by LandPKS and SoilGrids at a site was different by only 9% (e.g., 10% clay vs 19% clay). Mean differences in sand and rock fragment volume percentages for SoilGrids were higher, 21% (19% CV) and 24% (29% CV), respectively. These results indicate that in many cases SoilGrids provides relatively accurate predictions of soil texture relative to LandPKS.

Results from Task 2, using GSPD profiles as reference, showed that SoilGrids was twice as accurate in predicting the RSG (66% accuracy) relative to HWSD and WISE (30 and 32% accuracy, respectively; Table A1.7). Task 2 applied an exact matching criterion where the RSG name correlated to each soil profile must match the RSG name from the dominant soil queried from each soil map. Depending on the between-class variability of soil properties that influence land potential, different soil series may function similarly with respect to specific land-use practices. Thus, evaluating the accuracy of soil maps based on an exact RSG name match may under represent the functional or management-relevant similarity. Additionally, both HWSD and WISE are traditional soil maps that often include multiple soil types mapped at a sampling location. Thus, the lower RSG match rates for HWSD and WISE may be due in part to our reliance on the dominant soil from each map unit.

Our reliance on the dominant soil from each map unit is an important point as HWSD and WISE are designed to predict *groups* of soils occurring within a map unit. However most users of soil maps (and nearly all non-soil scientists) simply select the dominant component, as we did here. One of the principal values of the LandPKS + SoilID approach (#4 in the main text) is that it effectively gives non-soil scientists the ability to access *all* of the data in traditional soil maps, like HWSD and WISE, resulting in potentially higher accuracy than any of the other 3 approaches.

Table A1.8. Representative sand and clay numeric values for the different soil texture classes

	Range		Representative Value	
Texture Class	Sand	Clay	Sand	Clay
Sand	85-100	0-10	92.5	5
Loamy sand	70-90	0-15	80	7.5
Sandy loam	43-80	0-20	61.5	10
Loam	23-52	7-27	37.5	17
Silt loam	0-50	0-27	25	13.5
Silt	0-20	0-12	10	6
Sandy clay loam	45-80	20-35	62.5	27.5
Clay loam	20-45	27-40	32.5	33.5
Silty clay loam	0-20	27-40	10	33.5
Sandy clay	45-65	35-55	55	45
Silty clay	0-20	40-60	10	50
Clay	0-45	40-100	22.5	70

Table A1.9. Representative rock fragment volume numeric values for the different rock fragment classes

	Range	Representative Value
Texture Class	RFV	RFV
0-1%	0-1	0
1-15%	2-15	8
15-35%	15-35	25
35-60%	36-60	48
>60%	61-100	80

Appendix 3. LandPKS Soil and Site Characterization Protocol

LandPKS Site Characterization

Purpose: Complete LandPKS site characterization per the detailed instructions in the protocol below

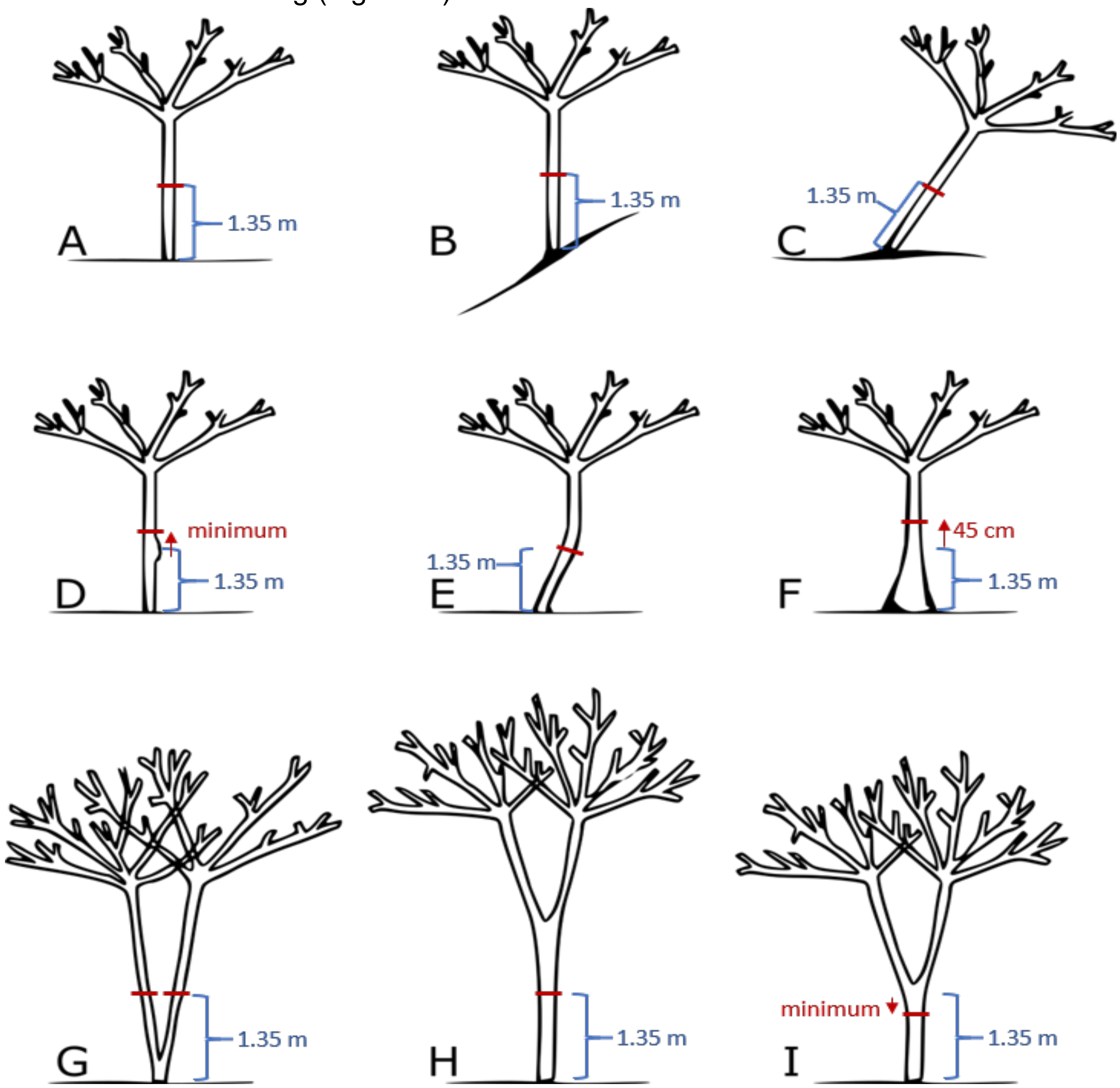
Materials:

- Smartphone/tablet with LandPKS App
- Measuring tape

Protocol:

1. Create a new sampling location
   a) Determine the location of the soil pit based on site selection guidance provided by LandPKS. Create a LandPKS plot at the location and assign it the correct sample name. Make sure that you give the phone or tablet enough time to get an accurate geolocation for the site with the highest possible positional accuracy (<15 m if possible).
2. Cocoa tree measurements
   a. Identify the location of the 10 closest cocoa trees to the location of the LandPKS soil pit.
   b. At each tree, locate the point on the cocoa trunk that is 1.35 m from the ground. If irregular tree growth results in non-representative trunk dimensions at breast height (1.35 m), then height adjustments will need to be made. Figure 1 illustrates the different tree shapes and where to measure the tree's circumference (circumference at breast height; CBH). Decide which tree class from Figure 1 each tree belongs and determine the mark along the trunk where the CBH measurement should be taken.
   c. At this height, wrap the tape around the trunk, ensuring the tape is perpendicular to the trunk and tight against the bark, and measure the circumference (i.e., CBH). Record this value for each tree in the notes section of the LandPKS app, also indicating its tree class from Figure 1 (i.e., Tree 1 = Class A, 42 cm CBH; Tree 2 = Class H, 30 cm CBH; etc.).
   d. Figure 1 illustrates the different cocoa tree shapes and the method uses to measure CBH.
      i. CBH should be measured 1.35 m from the ground (Figure 1A) on the uphill side of the tree trunk (Figure 1B).
      ii. Trees that lean should be marked and measured at 1.35 m along the trunk, and the measurement should be taken perpendicular to the tree's trunk (Figure 1C).
      iii. In cases where trees have scars, knots, burls, or other deformities at 1.35 m, the mark and measurement should instead be made at the minimum distance necessary to clear the obstacle, above or below 1.35 m (Figure 1D).
      iv. Trees that curve, or "sweep" up from the ground are treated differently from ones that lean; curved stems are marked and measured at 1.35 vertical meters from the ground, making sure the measurement is made perpendicular to the stem at the point of measurement (Figure 1E).
      v. For trees that have swollen-butted trunks where the swelling affects trunk diameter at 1.35 m, the mark and measurement are 0.45 m above the top of the swelling (Fig 1F).
      vi. Forked trees are treated as two separate trees, with two separate marks and measurements if they fork below 1.35 m. In this case, marks and measurements should be made 1.35 m from the ground above the fork (Figure 1G).
      vii. If a fork occurs above 1.35 m, then a single measurement should be made 1.35 m (Figure 1H).

viii. Forks just above 1.35 m may cause a distortion of the trunk shape and size at 1.35 m; in this case, mark and measure at the minimum distance below 1.35 m necessary to obtain an undistorted reading (Figure 1I).



3. Site characteristics
   a. In the LandPKS app, go to the Data Input screen and select Land Slope in the LandInfo module. Record the site's slope using either the LandPKS slope meter or by selecting the picture representing the site's slope class.
   b. Select the Slope Shape tab and record the site's slope shape. Slope shape describes the shape of the land both in the direction of the slope (i.e., down-slope from the site, left column of pictures in LandPKS app), and the shape perpendicular to the slope direction (i.e., across-slope from the site, right column of pictures). Slope shape is assessed by looking down and across the slope within a 20-50m diameter area around the site.
   c. On the Data Input screen, scroll down the list of input modules and select Photos. On the next screen, select landscape photos. Capture landscape photos of the site in the north, south, east, and western directions.

LandPKS Soil Profile Characterization

Map to the Future (M2F) page 48

Purpose: Complete LandPKS soil profile characterization per the detailed instructions in the protocol below

Materials:
- Smartphone/tablet with LandPKS App
- Digging tools (shovel, hoe, etc.)
- Tarp, with marked areas labeled with LandPKS soil sample depths
- ~2mm sieve(s) - 20cm or larger diameter.
- Measuring tape
- Water bottle with water
- Trowel or knife
- Tray, board, or small bin; ~30x30cm for collection soil samples from the pit face

Protocol:
1. Dig the soil pit
    a. Before digging the soil pit, determine the location of the pit face, and what time of day the profile description will occur. The pit face refers to the portion of the soil pit that will be described. Ideally, the pit face should be situated so that it is in full sunlight during the description and should be as vertical as possible. If the soil pit is located under tree canopy, determine which side of the soil pit will receive the most sunlight and make that side the pit face.
    b. Dig a large hole and clear a 70 x 100 cm area on one edge of the hole where the labeled tarp will be placed. Place all soil excavated during digging on the opposite side.
    c. The soil pit should be as deep as possible (>120 cm); but at least 70 cm or to the depth of bedrock. Dig a hole at least 15cm *deeper* than the bottom of the layer to be sampled.
    d. If bedrock was reached, record depth value in the *Soil Texture* section of the LandPKS app
    e. Measure the depth of any surface plant litter layer and record in notes. Scrape away surface litter layer down to the start of the mineral soil.
2. Prepare the Pit Face and Locate Soil Layers
    a. Once the pit has been excavated, the face needs to be cleaned of loose material and smearing before sampling. This is done by picking the loose material off with a knife, rock hammer, spade, or similar instrument. It is best to clean from the top down so that loose material does not fall onto areas that have already been cleaned.
    b. After the cleaning process is complete, use a measuring stick or tape to determine the top and bottom boundary for each LandPKS sampling layer based on the depths in the soil texture screen in LandInfo. Use a small stick to mark each boundary depth.
3. Photograph the Soil Pit
    a. After marking all sampling depths, take a photograph of the pit face using the Photos sections of the app on the Data Input screen. Select Photos, and then Pit Photo. It is helpful to place the measuring stick or tape against the pit face to show the depth scale in the photo.
    b. Try to take the photo as vertical to the pit face as possible and with the most even lighting conditions possible (e.g., not partly direct sun and shadow).
4. Collect and Prepare Soil Samples (for field evaluation – *no samples will be taken from the farm*)
    a. Place the sampling tarp on the 70 x 100 cm cleared area adjacent to the pit. There should be a separate space for each soil depth marked on the tarp and it should be labeled.
    b. Starting at the bottom sampling depth, place the flat collection surface (tray, board, small bin, shovel) perpendicular to the pit face at the bottom of each sampling depth.
    c. Using a knife or trowel, excavate approximately 0.5-1 liter of soil from each layer and place it on its designated location of the sampling tarp. Ensure that the sample for each layer is representative of that layer. For example, if collecting soil from the 20 to 50 cm layer, make sure to collect some soil at each depth from 20 to 50. This can be easily done by scraping the side of

the pit with a knife or some other semi-sharp object, so that some of the soil falls onto the flat collection surface from every depth between 20 and 50.
   d. Manually break apart the soil with your hands so that the soil can dry and will be easier to pass through the 2mm sieve.
   e. Collect a photo of the tarp containing the prepared soil samples (i.e., LandPKS app → Photos → Samples Photo).
5. Estimate Soil Coarse Fragments
   a. For each soil layer, pass several handfuls of soil through the 2mm sieve. Separate all material >2mm and place in a pile. Once the soil has been sieved for a soil layer, place the sieved soil in a pile next to the coarse fragment pile for that layer.
   b. By visually comparing the relative volume of the sieved soil pile and coarse fragment pile, estimate the proportion of the *total soil volume* (soil + coarse fragments) that the coarse fragments (all >2mm) occupy. Remember to account for any larger rocks that were not sampled from each layer of the pit face or larger rocks from the sampled soil that were not included in the sieved sample. Additionally, use the pictures of coarse fragment classes displayed in the LandPKS app's *Rock Fragment Volume* section on the *Soil Texture* screen in *LandInfo* to assist in selecting the correct coarse fragment class.
6. Measure Soil Color
   a. Using the 2mm sieved soil, take a subsample (about a handful) and flatten the soil pile on the tarp. Make sure the flattened pile is at least 1 cm deep.
   b. Place a color reference card next to the soil pile.
   c. Open the Soil Color screen of the LandPKS app under LandInfo. In uniform lighting conditions (not direct sunlight), take a photo of the color card and soil about 10 to 20 cm above the sample following the color measurement instructions within the LandPKS app.
7. Determine Soil Texture
   a. For each depth, start with a small handful of soil, about the size of a golf ball, and slowly add water a drop at a time, mixing as you go, until you have a ball of soil that has the consistency of putty. If too much water is added and the soil appears saturated, add additional dry soil until a putty consistency is achieved.
   b. Using the LandPKS soil texture guide, follow each step to determine the correct soil texture class and record this in the LandPKS app on the soil texture screen for each depth
8. Soil Health Assessment

In the SoilHealth module, record all information on soil health indicators for:
   a. Erosion
   b. Compaction (include any hard layers not caused by compaction)
   c. Aggregate stability
   d. Soil smell
   e. Biological activity

10. Soil Limitations

In the LandInfo module, record information on soil limitations in the app, including:
   a. Surface/Vertical soil cracks
   b. Surface stoniness
   c. Soil depth
   d. Water table depth (if encountered)
   e. Flooding (if visible signs or information from farmer)
   f. Lime requirement (information from farmer if available)

Appendix 4. Useful Links and References

Useful Links

About FarmGrow - www.farmgrow.org

About LandPKS - www.landpotential.org

References

Aneani, F., and K. Ofori-Frimpong. 2013. An Analysis of Yield Gap and Some Factors of Cocoa (Theobroma cacao) Yields in Ghana. Sustain. Agric. Res. 2(4): 117. doi: 10.5539/sar.v2n4p117.

Appiah, M., S. Sackey, K. Ofori-Frimpong, and A. Afrifa. 1997. The consequences of cocoa production on soil fertility in Ghana : A review. Ghana J. Agric. Sci. 30(2). doi: 10.4314/gjas.v30i2.1970.

Dossa, E.L., A. Arthur, W. Dogbe, A. Mando, D. Snoeck, et al. 2018. Improving fertilizer recommendations for Cocoa in Ghana based on inherent soil fertility characteristics. Improving the Profitability, Sustainability and Efficiency of Nutrients Through Site Specific Fertilizer Recommendations in West Africa Agro-Ecosystems. p. 287–299

Gray, B., Gomez, J., Arthur, F., Rubio, H. 2020. FarmGrow: Farm Investment Plans for Smallholder Cocoa Farmers in Ghana Midline Assessment. Grameen Foundation. https://grameenfoundation.org/partners/resources/farmgrow-farm-investment-plans-for-smallholder-cocoa-farmers-in-ghana-midline-assessment

Sarpong, D., Osei-Asare, Y., Gray, B., Tsekpo, E. 2020. FarmGrow: Farm Development Plans for Smallholder Cocoa Farmers in Ghana Baseline Report. Rainforest Alliance, Grameen Foundation and University of Ghana, Legon. https://grameenfoundation.org/partners/resources/farmgrow-baseline-report-2-18-2020.

Sys, C., Van Ranst, E., Debaveye, J. and Beernaert, F. 1993. Land Evaluation Part 1. Principles in land evaluation and crop production calculations. International Training Centre for Post Graduate Soil Scientists, University of Ghent. Agricultural Publications No. 7, General Administration for Development Co-operation, Brussels, Belgium - Page 86.