The use of GPS technology in the Living Standards Measurement Study

LSMS

Sydney Gourlay
Senior Economist
Living Standards Measurement Unit
World Bank

THE WORLD BANK
IBRD • IDA | WORLD BANK GROUP
Development Economics • Data

CEPAL/ECLAC
Seminario regional sobre innovaciones y desafíos en encuestas de hogares
November 21, 2024

Overview

- What is the LSMS?
- How, what, and where have we used GPS technology?
- Work on integrating georeferenced survey data with satellites
- Recent and ongoing methodological validation of alternative tools

What is the Living Standards Measurement Study (LSMS)?

The **LSMS** is the World Bank's flagship household survey program focused on:

- Strengthening household **survey systems** in countries
- Improving the **quality of microdata** to better inform development policies

DATA PRODUCTION

Designing, implementing, and disseminating large-scale surveys on households, farms, firms, and facilities

METHODS AND TOOLS

Improving methods and tools for survey data collection and analysis, via field experiments and rigorous research

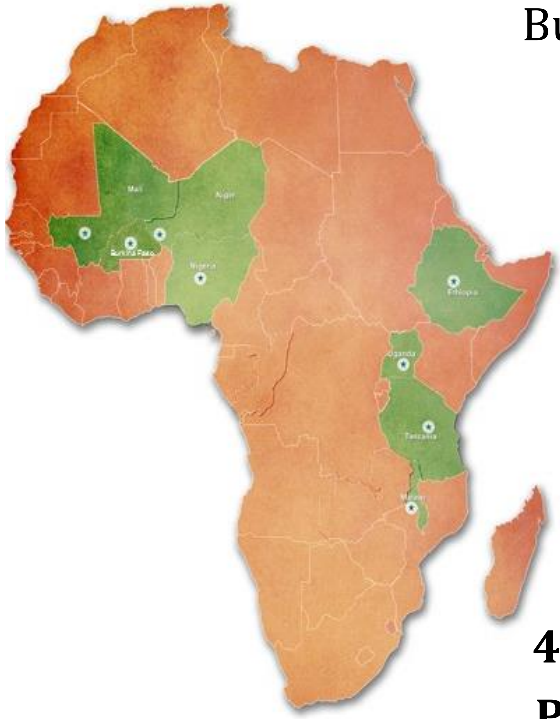
POLICY RESEARCH

Conducting and promoting research to inform development policies

LSMS-Integrated Surveys on Agriculture (2008-2024)

LSMS-ISA is a **unique system of longitudinal surveys** designed to improve the understanding of wellbeing, livelihoods and smallholder agriculture in Africa.

8 Partner National Statistical Offices



Burkina Faso
Ethiopia
Malawi
Mali
Niger
Nigeria
Tanzania
Uganda

45% of SSA Population

LSMS-Integrated Surveys on Agriculture (2008-2024)

Integrated into the **NSSs**, implemented by NSOs

Individual- & plot-level data collection

Representative at the **national & regional-levels**

Georeferencing of household & plots, GPS-based plot area measures

Tracking of **households & individuals**

Computer-assisted personal and telephone interviewing (Survey Solutions)

Integrating agricultural data collection into a multi-topic framework

Open access survey & derived geospatial data policy

Geolocations captured at multiple levels: household and agricultural plot and/or parcel

+

Measurement of plot area using GPS

Collecting Geolocations: Coverage & Tools

Household-level coordinates:

typically collected directly in computer assisted personal interviewing (CAPI) tablet, nearly full coverage of all sampled households

Plot-level

GPS coordinates

Typically collected directly in **CAPI tablet;** following recent guidance, collected from plot center

GPS area measurement

Area measurement and plot boundary collected using **handheld GPS device** such as Garmin eTrex 30

Collecting Geolocations: Coverage & Tools

Example coverage rates: GPS Plot Area Measurement

Country	Plots with GPS Area	Farms with GPS on all cultivated plots
Malawi 2010/11	96%	93%
Uganda 2009/10	58%	62%
Tanzania 2010/11	78%	80%
Niger 2011	46%	51%

Variation in coverage partially due to survey guidelines, for example:

- "all parcels within the enumeration area"
- "plots within 1 hour from household by any means of transportation"

Plot-level

GPS area measurement

Area measurement and plot boundary collected using **handheld GPS device** such as Garmin eTrex 30

Anonymization & dissemination of geovariables

- Responsibilty to protect the anonymity of survey respondents
- LSMS-ISA datasets are typically released with **offset (anonymized) GPS coordinates** at community level to integrate other geo-localized data.
  - Following the DHS, random offset 0-2km in urban EAs, 0-5km rural EAs 0-10 km for 1% of EAs
- LSMS-ISA are also disseminated with a set of ready-to-use **geovariables** for each community/EA, household, plot

Theme	Geovariables
Climatology	Temperature, precipitation
Landscape Typology	landcover class, agriculture, agro-ecological zones, population density
Soil & terrain	Roughness, Nutrient availability, Nutrient retention capacity, Rooting conditions, Oxygen availability to roots, Excess salts, Toxicity, Workability
Crop season parameters	Total rainfall, start of wettest quarter, total rainfall in wettest quarter, change in greenness (EVI), average EVI, timing of greenness onset, peak EVI
Distance	Household distance to nearest major road, Household distance to nearest town of >20,000, Household distance to nearest major market, border control posts, distance to regional capital
Plot geovars	Field distance to household, elevation, wetness index, slope

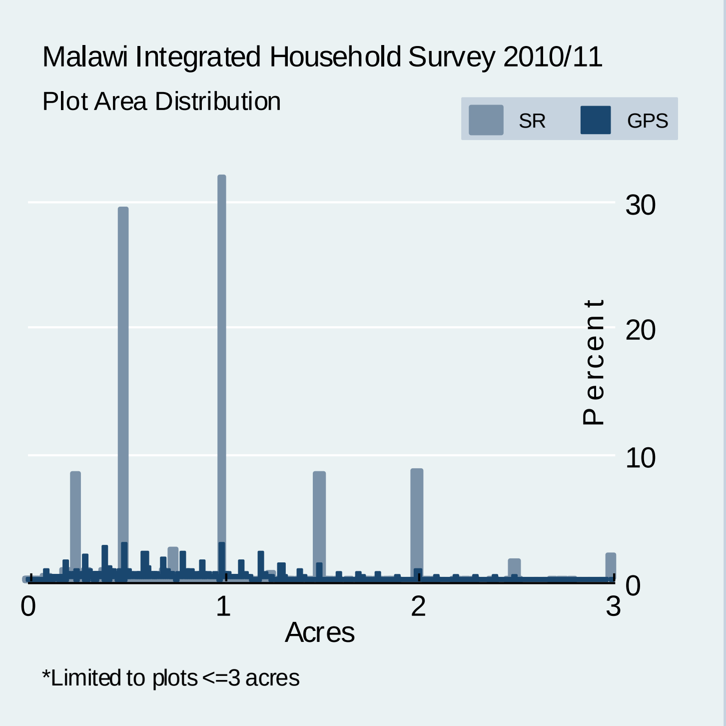
LSMS
Living Standards Measurement Study

Some motivation for using GPS to measure plot areas and capture locations?

Motivation #1 – Evidence of systematic bias in respondent estimates of plot area

Self-reported estimates are potentially sensitive to:

- Respondent characteristics
- Perceived use of the data
  - Property taxes
  - Access to a program
- Traditional/local units of measurement
- **Rounding**

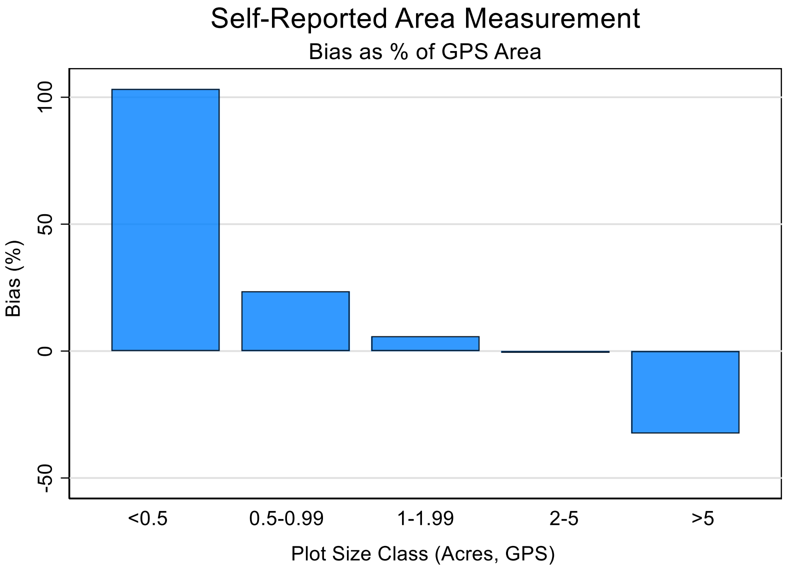


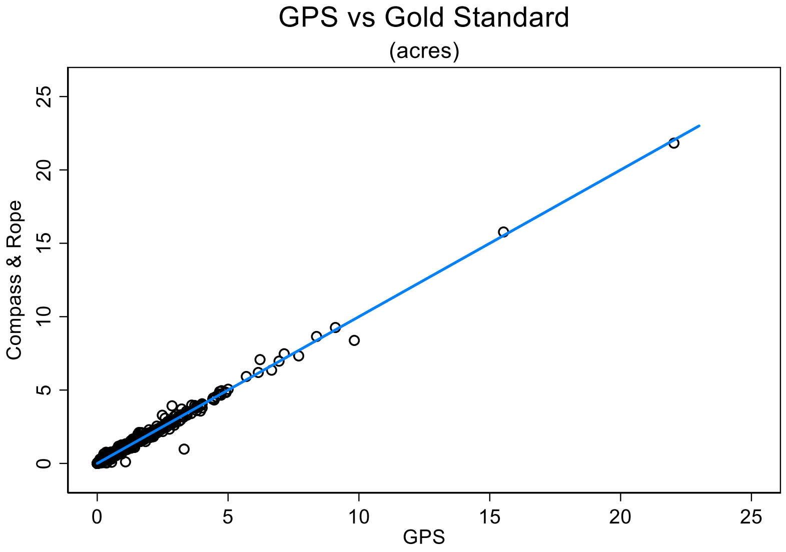
Some motivation for using GPS to measure plot areas and capture locations?

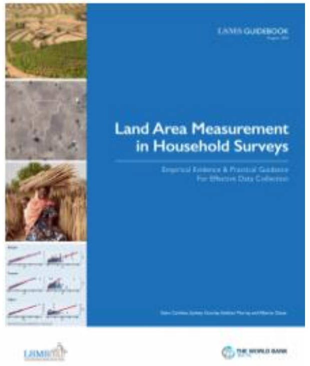
Motivation #1 – Evidence of systematic bias in respondent estimates of plot area

- cross-country evidence of **systematic bias in farmer estimates**
- cross-country evidence of **consistency of GPS against gold standard***

See the LSMS guidebook on area measurement







Some motivation for using GPS to measure plot areas and capture locations?

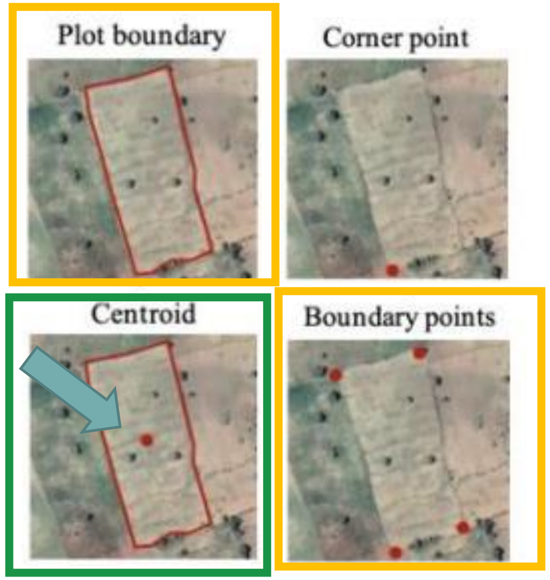
Motivation #2 – Value of integrating georeferenced survey data with satellite imagery and machine learning methods

Integration of Surveys & Satellites
Research on optimal georeferencing protocols

- Methodological work conducted by the LSMS and partners with support of the 50x2030 Initiative (Azzari et al. 2021), used georeferenced survey data to, among other objectives:

→ explore the implications of different georeferencing protocols on the ability to train and validate remote sensing models for crop-type mapping

Georeferenced plot-level survey data was integrated with publicly available Sentinel-2 imagery and other geospatial data to assess, via machine learning methods, how crop type prediction accuracy changes based on georeferencing protocols.



Strongly preferred options!

Third-best option, used in the case that only a single point can be collected

Integration of Surveys & Satellites

Development of Crop Type & Yield Maps

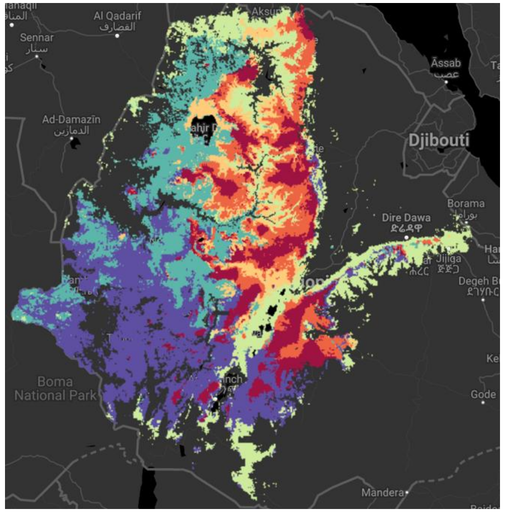
- LSMS and partners, including Atlas AI, working to answer high-priority questions in microdata collection, analysis and dissemination to support scaling up accurate integrated satellite-survey applications
  - Focus on an attainable set of agricultural outcomes – production, cultivated area, and yield – for **a minimum of four cereal crops** - namely maize, rice, sorghum, and wheat

- Production and publication of **high-resolution crop and maize area maps** for Ethiopia and Malawi

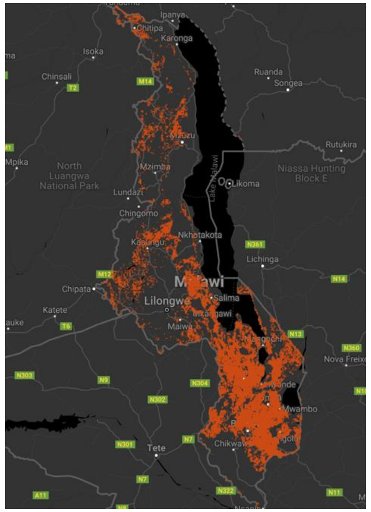
- Ongoing: Extension of research with a focus on crop area mapping and crop yield estimation for an expanded set of cereal crops in Malawi, Ethiopia and Mali.

- **Forthcoming:** ***Guidelines*** for georeferenced microdata collection and processing for calibrating and validating remote sensing models.

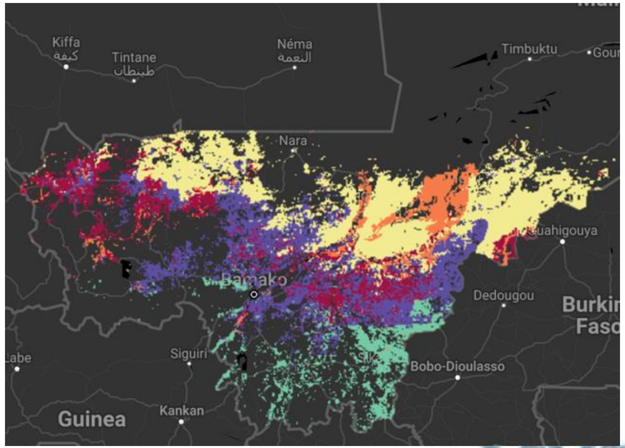
10-meter crop type maps



Ethiopia 2018



Malawi 2019



Mali 2018

Developing and testing alternative tools for georeferencing and area measurement

Handheld GPS now integrated into many national survey programs, delivering significant improvements in land area measurement accuracy.

Despite its advantages, challenges remain:

- **Cost**: Devices cost approximately $300 each.
- **Data Integration**: Manual linking of survey data with GPS measurements is error-prone and requires substantial cleaning.
- **Time**: requires travel to the plot

Testing/developing alternative approaches using Survey Solutions, the World Bank's CAPI software

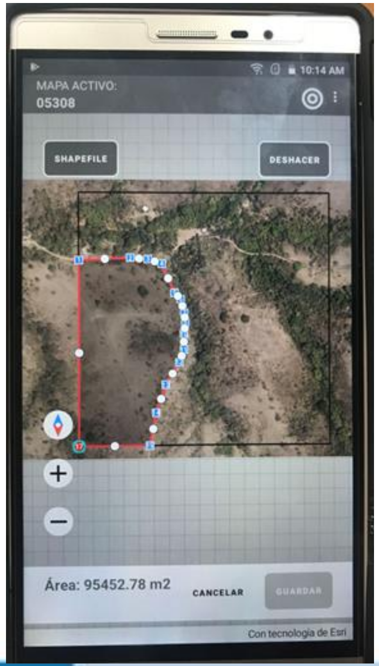
Boundary delineation on (offline) satellite image

Tablet-based GPS boundary delineation

Testing the feasibility & accuracy of a satellite-image based measurement and georeferencing approach

Two efforts aimed at testing the feasibility of use of **the Survey Solutions satellite image-based area measurement** against current methods, including handheld GPS:

- small-scale pilot conducted in Ecuador implemented by INEC (NSO), in partnership with FAO and the World Bank.
- larger methodological research study implemented in Armenia, joint with the Statistical Committee of the Republic of Armenia and the ICARE Foundation



Testing the feasibility & accuracy of a satellite-image based measurement and georeferencing approach

Select take-aways from the Armenia study:

Duration

Avg duration for parcel measurement:

- Satellite: 3.4 mins
- handheld GPS*: 7.3 mins

Difficulty

Satellite-image approach deemed impossible on 17% of parcels

→ Primarily a function of respondents not able to find/locate the parcel on the image (54% of cases)

Accuracy

Challenges in correctly locating parcels on a satellite image (22% of satellite-based parcel boundaries having no intersection with the handheld-GPS parcel boundaries)

Mean area measurements do not statistically differ across the two methods, but absolute value of the measurement error is ~23% (among parcels whose GPS and satellite boundaries intersect)

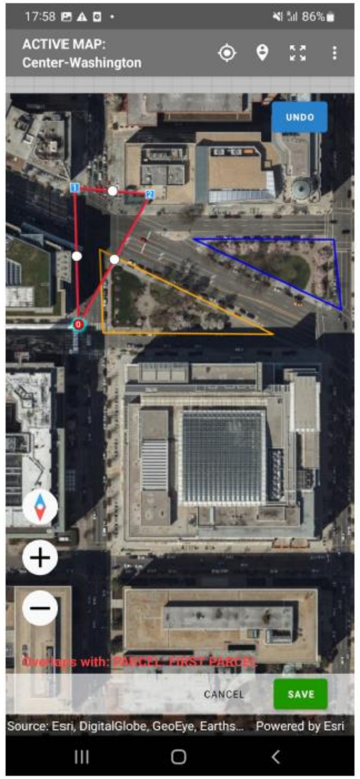
*does not include time to travel to parcel

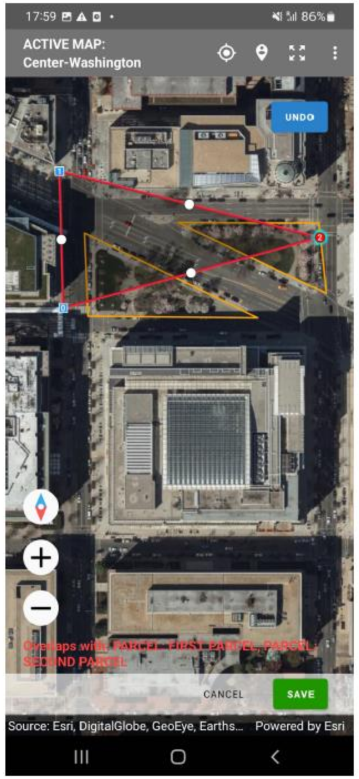
Testing the feasibility & accuracy of a satellite-image based measurement and georeferencing approach

Following the experiences in Armenia and Ecuador, as well as experience in other countries with handheld GPS measurement, 50x2030 Initiative & the LSMS team have supported enhancements to this Survey Solutions feature:

Real-time boundary overlap detection!

Color-coded boundaries and warning messages flag overlapping boundaries

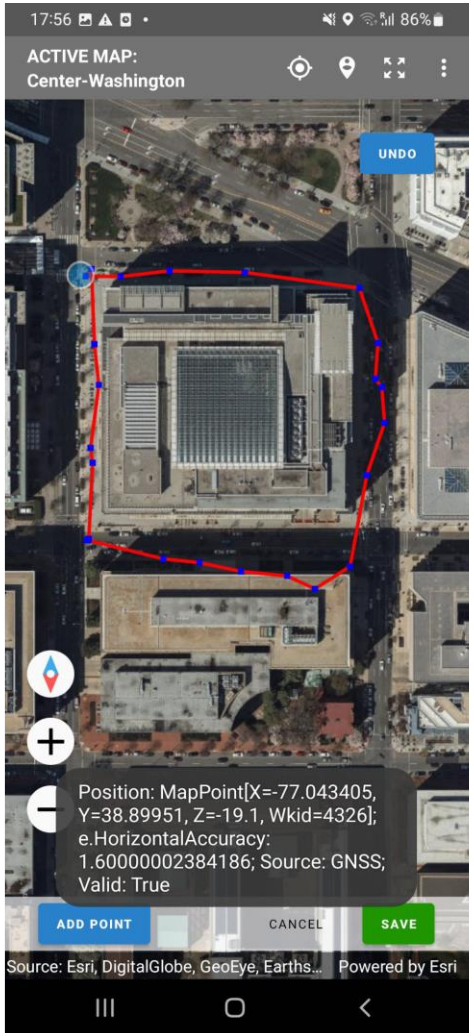


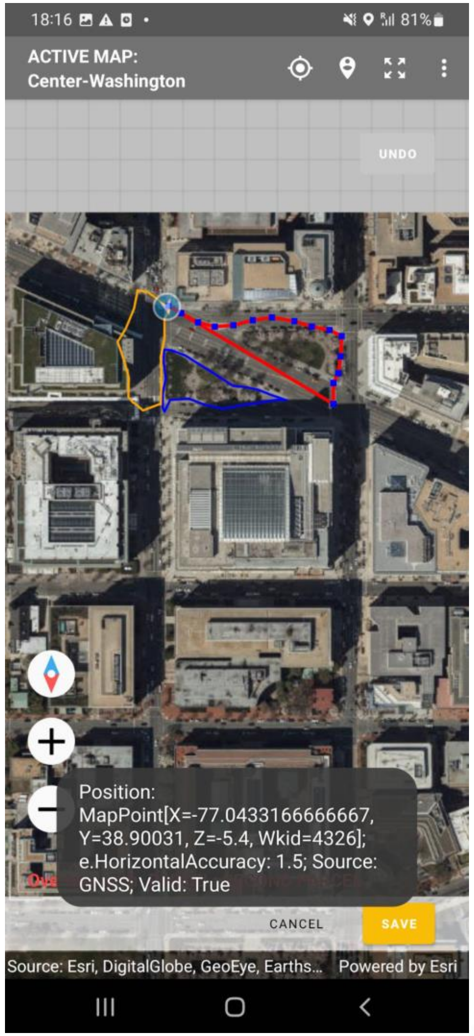


Tablet-based GPS boundary delineation and area measurement

To address challenges in use of handheld GPS-based measurement, 50x2030 Initiative supported development of new feature designed in World Bank's Survey Solutions CAPI software to **create and store polygons**, and **measure area**, by walking around the perimeter of parcel/plot.

- Coordinates are automatically collected at set time interval (e.g., 5 secs)
- String of ordered coordinates exported and can be easily converted to polygon
- New tool requires validation against tools with validated accuracy





LSMS

Tablet-based GPS boundary delineation and area measurement

Feasibility of implementation and data quality of this feature assessed in a methodological study in Uganda (Uganda Climate, Land Area, and Soil Study 2023-24).

One randomly selected maize plot per household was measured with each of the following:

- Self-Reported Estimate
- Handheld GPS (Garmin eTrex 30)
- Survey Solutions (Tablet*)
- Survey Solutions (Tablet*) + GPS booster

All objective measurements conducted simultaneously to ensure the same boundaries were walked for each

Garmin Glo 2.0: Palm-sized GPS receiver that overrides embedded tablet GPS, which may be insufficient for accurate measurement

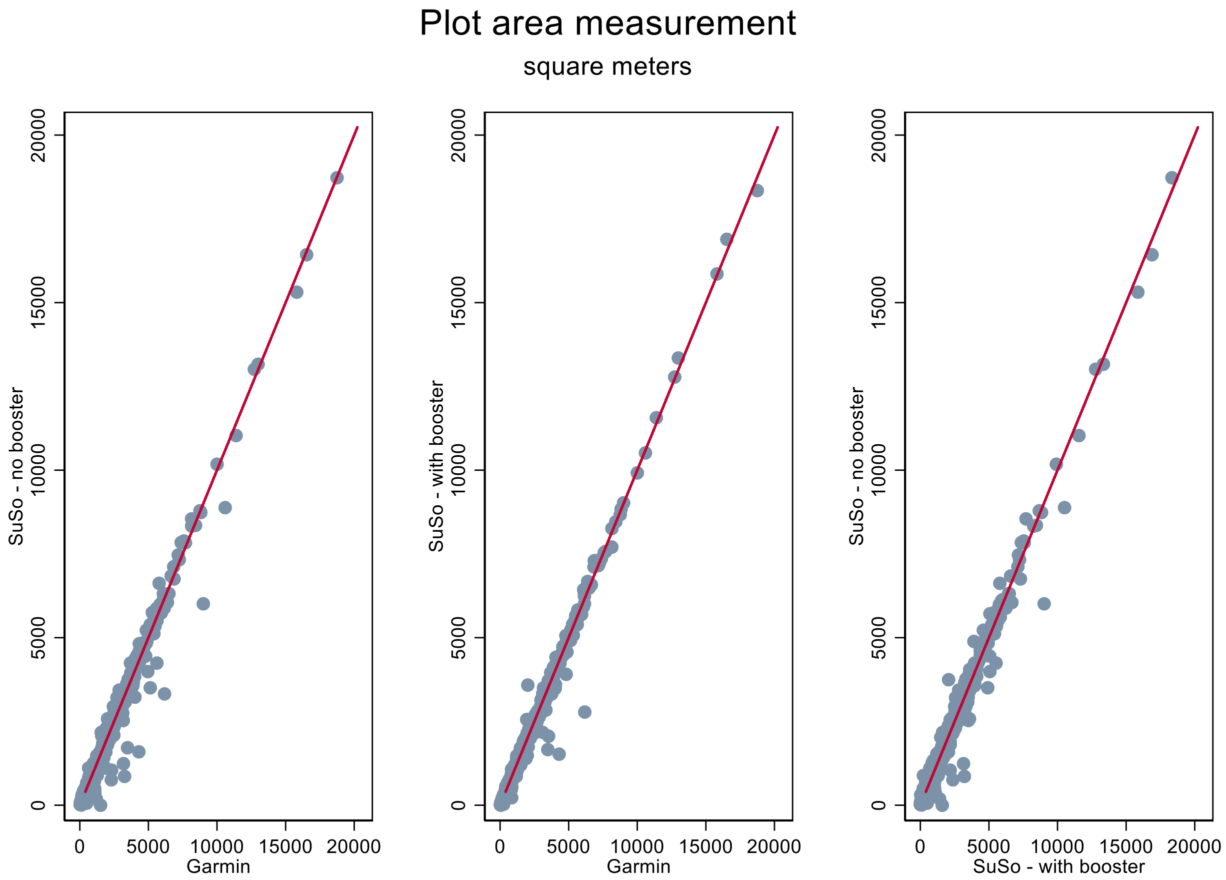
*Tablet model: Samsung Galaxy T580

Tablet-based GPS boundary delineation and area measurement

Comparison of Garmin (handheld GPS) & Survey Solutions with and without GPS booster reveals similar (but statistically different) area estimates

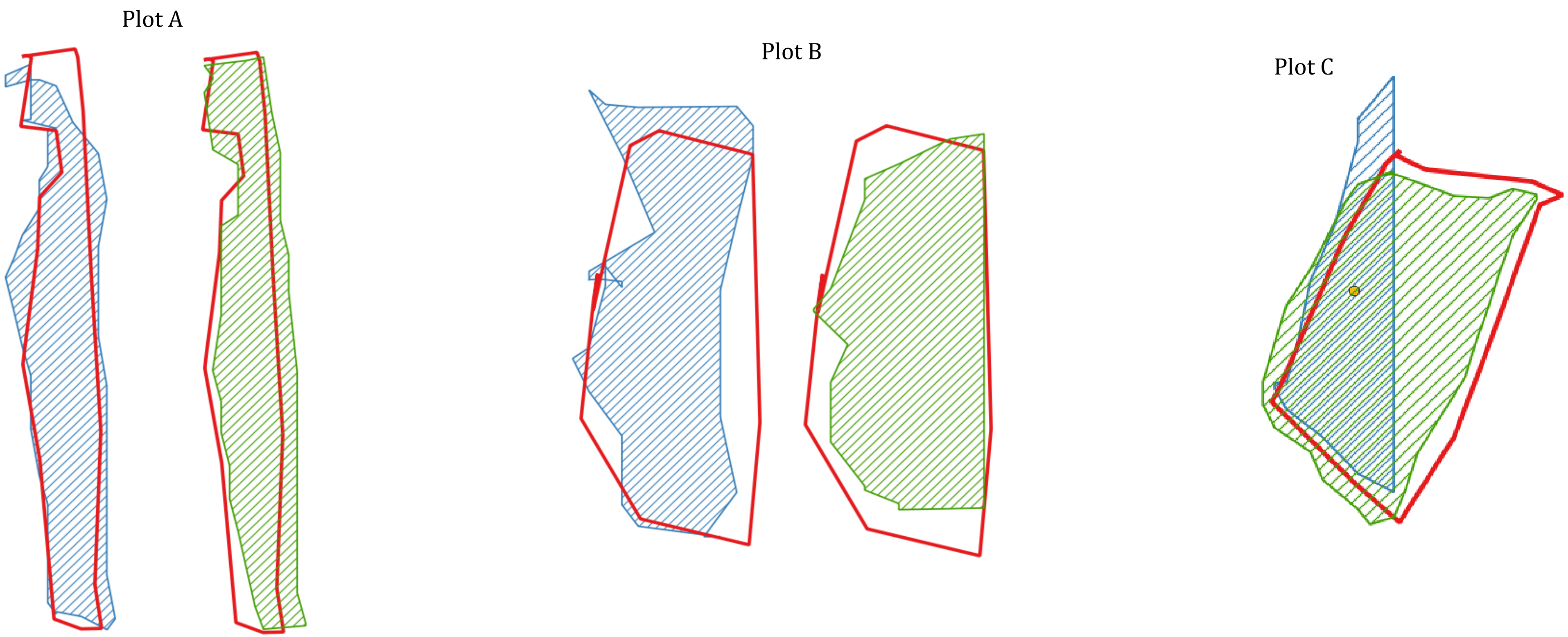
	Mean (m^2)	Bias (SuSo-Garmin)	| Bias |	| % Bias |
Garmin	1670.27	-	-	-
SuSo with booster	1628.25	-42.02	94.73	10.53
SuSo no booster	1641.06	-29.21	126.71	12.05

N=869



Tablet-based GPS boundary delineation and area measurement

Observation of plot boundaries recorded with different methods reveals differences that may have implications for remote-sensing based applications



Red = Garmin GPS; Blue = SuSo without booster; Green = SuSo with booster

Takeaways

- Following validation of the **use of handheld GPS devices** for area measurement, LSMS-ISA surveys (and others) have adopted this practice at scale – and if you haven't already, you can too!

- Collection of geolocations of household and agricultural land has added significant value to the survey data by allowing **integration with geospatial data** sources
  - Dissemination of constructed geovariables to facilitate data use while maintaining respondent confidentiality
  - Integrating with satellite imagery to generate crop type (and soon, crop yield) maps at scale

- Work is underway to validate and provide practical guidance on **alternative tools** to enable wider, more efficient geolocation and GPS-based area measurement in surveys

Stay tuned for upcoming publications on: (i) georeferenced microdata collection and processing for calibrating and validating remote sensing models, and the (ii) validation of the use of offline satellite imagery and (iii) tablet-based GPS for boundary delineation

LSMS
Living Standards Measurement Study

Thank you!

Stay connected:

Worldbank.org/lsms

lsms@worldbank.org

www.facebook.com/WorldBankLSMS/

GET UPDATES:

Annex:
Experience & Preliminary Findings
Practical Implementation

[preliminary – not for citation]

Survey Solutions tool and booster implemented with **minimal technical challenges**

Did you face any technical challenges with the…

	Non-Boosted Tablet		Boosted Tablet	
	N	%	N	%
No	842	94.1%	870	97.2%
Yes	53	5.9%	25	2.8%

If yes, what were the challenges you faced?

	Non-Boosted Tablet		Boosted Tablet	
	N	%	N	%
Accuracy too low to allow measurement	16	1.8%	7	0.8%
Took a long time to reach required accuracy	13	1.5%	9	1.0%
Shape does not look like plot shape	24	2.7%	3	0.3%
Booster would not connect	-	-	1	0.1%
Other	7	0.8%	5	0.6%

(multiple responses possible)

UGANDA CLIMATE, LAND AREA, AND SOIL STUDY (CLASS)

Main objective: testing & validating **innovative** methods for improved measurement of **climate variability & extreme weather**, **soil health** and **land area** in small-scale agriculture:

- Evaluating accuracy & performance against **benchmark** measures (and also against each other)
- Assessing **trade-offs** of different methods in terms of quality, costs and scalability

Study elements:

- **Household survey** conducted in multiple visits with inclusion of objective measurements of soil health and agricultural land area through different methods, and improved sets of subjective questions for soil, weather and climate
- **Community-level weather data collection** via weather sensors installed at survey locations

Sample:

- 900 maize-growing households in 75 rural enumeration areas (across 5 districts)
- **All methods of measuring land area and soil health implemented on one randomly selected maize plot per household, allowing for direct comparison**

- Timeline: second agricultural season of 2023 (July – December)

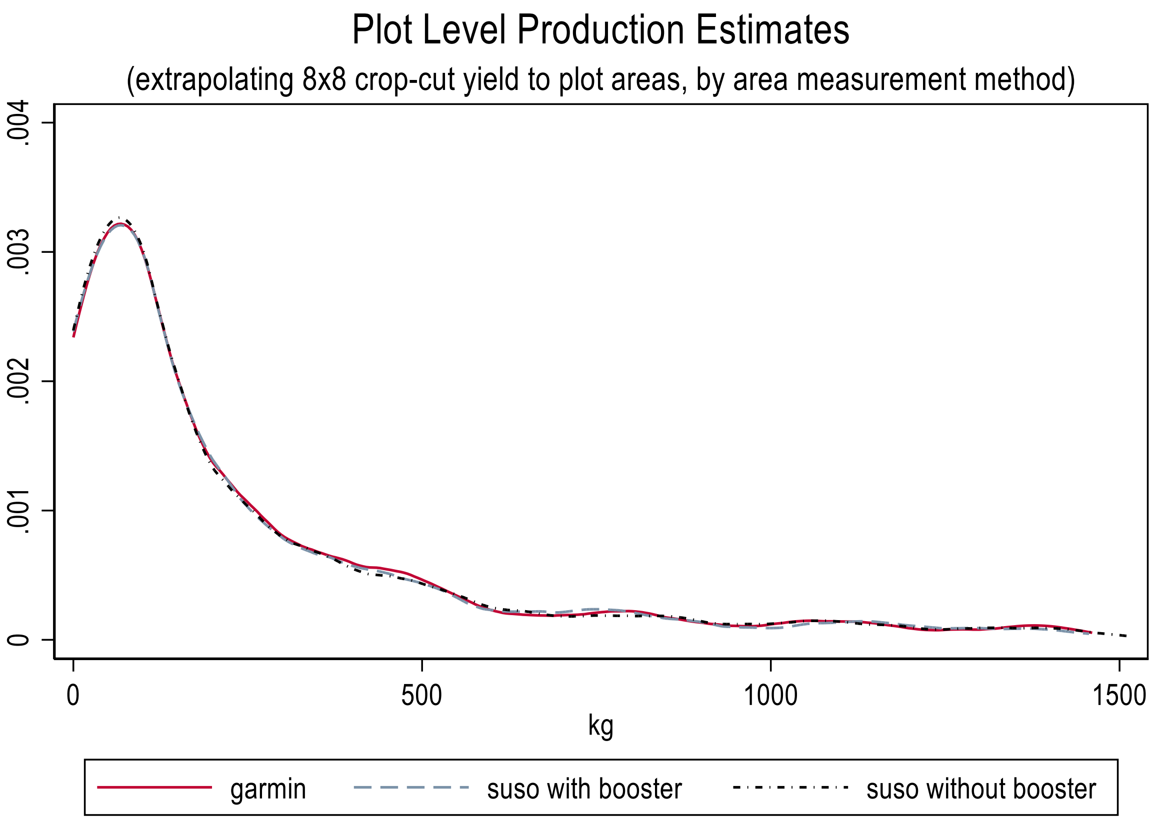
LSMS
Living Standards Measurement Study

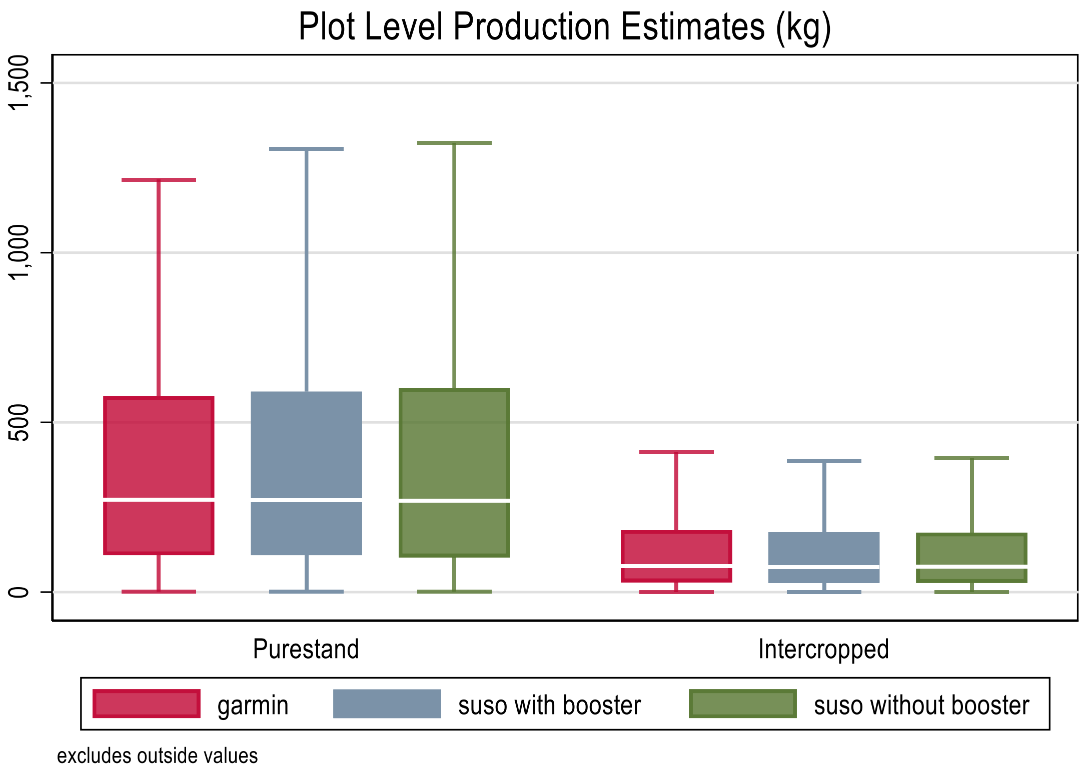
Experience & Preliminary Findings

Implications for Plot-Level Production Estimation

(extrapolation of CC yields to full plot area)

[preliminary – not for citation]





Note: Subsample with 8x8 crop-cutting subplot only (N=557); excludes households that pre-harvested 100% of maize prior to CC and households with zero harvest.

Experience & Preliminary Findings

[preliminary – not for citation]