50x2030
DATA-SMART AGRICULTURE

TECHNICAL NOTE ON GPS COORDINATE COLLECTION IN 50x2030 SURVEYS

May 2025

50x2030 Technical Notes for Country Teams

ACKNOWLEDGMENTS

This document is a product of the 50x2030 Initiative to Close the Agricultural Data Gap and was developed by Josefine Durazo (World Bank), Asmelash H Tsegay (World Bank), Sydney Gourlay (World Bank), and Kevin McGee (World Bank), with contributions from other members of the 50x2030 Methods & Tools coordination team. Valuable comments were received on an earlier draft from FAO colleagues working on 50x2030 Data Production activities. This document draws heavily from various research outputs produced by the LSMS Team[1,2], Survey Solution CAPI manuals[3], and the DHS field guide for GPS data collection[4].

The 50x2030 Initiative to Close the Agricultural Data Gap is a multi-agency effort aimed at supporting 50 low- and lower-middle-income countries to produce fundamental agricultural and rural data through the use of integrated agricultural and rural surveys. For more on the Initiative, please visit www.50x2030.org.

This publication is part of a series of 50x2030 Technical Notes for Country Teams that will provide digestible, implementation-focused guidance for data producers and survey practitioners. Each note offers a brief summary of the motivation for specific survey design decisions followed by detailed, practical guidance that can be directly translated into survey design or training efforts. These notes are part of the existing 50x2030 Technical Note series.

[1] Azzari, G., Jain, S., Jeffries, G., Kilic, T., & Murray, S. (2021). Understanding the requirements for surveys to support satellite-based crop type mapping: Evidence from sub-Saharan Africa. *Remote Sensing*, *13*(23), 4749. (https://www.mdpi.com/2072-4292/13/23/4749)

[2] 50x2030 Guidelines for Data Collection in Large-Scale Surveys to Support Machine Learning Analysis (forthcoming)

[3] https://docs.mysurvey.solutions/questionnaire-designer/questions/geography-question/

[4] https://dhsprogram.com/publications/publication-dhsm9-dhs-questionnaires-and-manuals.cfm

Contents

1. INTRODUCTION & BACKGROUND

This technical note is intended to be a reference for survey practitioners seeking guidance on the collection of GPS coordinates in household and agricultural surveys. This note provides a brief motivation for collecting GPS coordinates as part of agricultural and household surveys, referred to as georeferenced survey data, followed by practical guidance for implementing this type of data collection. The focus is on recommendations for georeferencing in surveys supported by the 50x2030 Initiative, including *where* GPS coordinates should be collected to maximize analytical value and to support data quality control (summarized in Box 1), as well as guidance on *how* to collect them, equipment considerations, training needs, and questionnaire design related to the collection of GPS coordinates. This technical note can be complemented by the 50x2030 Technical Note on Land Area Measurement, which provides guidance on the use of GPS technology to measure the area of agricultural parcels and plots and store referenced outlines.

Georeferenced data refers to data that can be ascribed to (or 'tagged with') a particular location. This type of data – also referred to as geospatial data, spatial data, or geographical information – captures location by referencing a satellite navigation system, such as the Global Positioning System (GPS). Note that while georeferenced location data can use GPS as well as other navigations systems (such as GLONASS) as their source, herein for shorthand we refer to all georeferenced location data, regardless of system source, as GPS data.

The advancement of georeferencing technologies over the last two decades, coupled with the significant decrease in their cost, have created greater opportunity for researchers and policymakers to integrate georeferenced data into their analysis. While some surveys collect georeferenced location data (for households, farm plots, enterprises, community centers, etc.), most household and agricultural surveys in developing countries still do not capture such data, due in part to a lack of information about the benefits of doing so. This note intends to bridge this gap by providing the motivation and practical guidance to support survey practitioners in collecting GPS coordinates of both household dwellings and agriculture land, as part of 50x2030 surveys. This work builds largely on methodological research implemented by the World Bank's Living Standards Measurement Study team under the LSMS-ISA project and 50x2030 Initiative, as well as the broader research community – and on lessons learned from the framework of the Global Strategy to Improve Agricultural and Rural Statistics. This document does not cover in detail the recommended protocols for disseminating georeferenced data but includes key considerations for the handling of such data and indicates future work in this domain.

Value of georeferenced survey data

The collection of GPS coordinates in household and agricultural survey operations is valuable in achieving several objectives, including enhancing the analytical potential of the survey data, enabling greater data quality control, and providing the opportunity for survey data to train and validate remote sensing models, among others.

Research and policy

Integrating the collection of georeferenced location information with overall survey collection allows for a better understanding of the dynamics between and within the economy, environment, and society. Linking data about people, businesses, buildings, infrastructures, agriculture, and natural resources, can result in an improved understanding of social, economic, and environmental issues.[5] Given this wealth of additional analysis that georeferenced household and agricultural survey data

[5] PARIS21. 2021. Guide on Geospatial Data Integration in Official Statistics. Available at: https://www.paris21.org/geospatial

unlocks, it also leads to improved economic and policy analysis.[6] The collection of household, farm, and plot coordinates can provide vast insights useful for policy, such as the linkages between various climatic conditions and agricultural production (through the integration of GPS coordinates with geospatial data on rainfall, temperature, etc.), and the impact of market access on agricultural production and incomes (through the integration of GPS coordinates with geospatial data on market locations and road density), among others.

Improving data collection

GPS data can be used to improve both the quality of household and agricultural survey data and the cost-effectiveness of collecting it. This can occur at several phases of data collection, from the development of a sample frame to quality control protocols and use for follow-up surveys. More accurate and cost-effective surveying enables researchers to carry out better analysis and provide better evidence-based advice to policymakers.[7]

Box 1. Recommendations for 50x2030 Surveys: Where to Collect GPS Coordinates

In 50x2030 surveys, GPS coordinates should be collected at two different levels:

- **Farm Location:** For the household sector, this should be the **household dwelling**. This should be collected standing directly outside the entrance to the dwelling to avoid issues of GPS accuracy or extended satellite acquisition time inside the dwelling. For the non-household sector, this can be the main agricultural building, main agricultural parcel, or, when those are not possible, the administrative office of the holding. These coordinates are collected in all 50x2030 instruments.

- **Cultivated plots:** Coordinates should be collected from the **center of cultivated plots**. While some surveys have historically collected coordinates from a single corner of the plot, 50x2030-supported research has shown that the collection of a center point is significantly more valuable than a single corner point in satellite-based crop mapping applications (Azzari et al., 2021). These center points should be collected regardless of if GPS-based area measurement is conducted. However, if plot outlines that are collected in an external GPS device are saved, collection of the center point becomes optional (though still highly recommended). The different 50x2030 instruments have different levels of coverage for plot-level GPS coordinates. These are summarized below:
    - CORE & MEA: All cultivated plots, excluding kitchen gardens
    - ILP & PME: All cultivated plots, including kitchen gardens

  Note that coordinates at the plot level are only collected for the household sector by default. These can be added to the non-household sector instruments depending on the country context.

  I**n the case of 1-visit surveys** that are implemented in the post-harvest visit, coordinates should be collected from the center of *parcels* instead of plots as plot boundaries may no longer be visible or may have changed since the seasonal harvest.

For 50x2030 surveys, the recommendation is to **collect the above GPS coordinates directly in the CAPI tablet** using the embedded tablet GPS rather than a handheld GPS device. However, for area measurement, the recommendation is to use handheld GPS devices due to the superior accuracy.

---

[6] Gibson, John; McKenzie, David. 2007. *Using the Global Positioning System in Household Surveys for Better Economics and Better Policy*. Policy Research Working Paper; No. 4195. World Bank, Washington, DC. © World Bank. https://openknowledge.worldbank.org/handle/10986/7003 License: CC BY 3.0 IGO."

[7] Gibson, John; McKenzie, David. 2007. *Using the Global Positioning System in Household Surveys for Better Economics and Better Policy*. Policy Research Working Paper; No. 4195. World Bank, Washington, DC. © World Bank. https://openknowledge.worldbank.org/handle/10986/7003 License: CC BY 3.0 IGO."

Filling data gaps

By integrating survey data with satellite imagery and other remote sensing data, georeferencing survey data helps researchers and policymakers fill data gaps that face-to-face surveys cannot capture. This integration provides comprehensive insights into land use, climate information, accessibility distance to public facilities and infrastructures, and environmental changes that face-to-face surveys alone cannot capture.[8]

Machine learning models

Georeferenced multi-topic household and farm survey data can also be used to train and validate remote sensing models that aim to estimate, for example, crop-covered areas and their yields using satellite data. Accurate models for estimating agriculture outcomes based on remote sensing and satellite imagery have the potential to provide timely insights into conditions on the ground and can fill gaps in agricultural monitoring and statistics by enabling the production of large scale crop type and yield maps (see Box 2).[9] For detailed guidance on integrating satellite and survey data for high-resolution crop area mapping and crop yield prediction in low- and lower-middle income countries, refer to the 50x2030 *Guidebook for Designing and Using Surveys to Inform Earth Observation Applications on Crop Area Mapping and Crop Yield Estimation in Smallholder Contexts* (forthcoming).

Box 2. Georeferenced Survey Data for Machine Learning

The value of georeferencing in surveys is well-established. It enables the integration of data from a multitude of auxiliary sources that can capture general landscape characteristics as well as highly localized conditions such as roof material for dwellings or crop type and vigor at the field level. Collecting GPS information as part of agricultural data in household and farm surveys not only provides valuable data on its own, but this data can also serve as interoperable complements and validation instruments for other data sources, such as for the ground-truthing of remote sensing data (Carletto et al., 2021). The forthcoming *Guidebook for Designing and Using Surveys to Inform Earth Observation Applications on Crop Area Mapping and Crop Yield Estimation in Smallholder Contexts* focus primarily on the needs of specialized remote-sensing applications such as machine learning models for crop type and crop yield estimation and shares results from current research on this topic. The 50x203-suppported work by Azzari et al. (2021), which serves as an input into the guidelines, showcases the use of survey data as an input into high-resolution satellite-based crop type maps (seen in **Figure 1**), as well as the importance of the georeferencing protocols deployed.

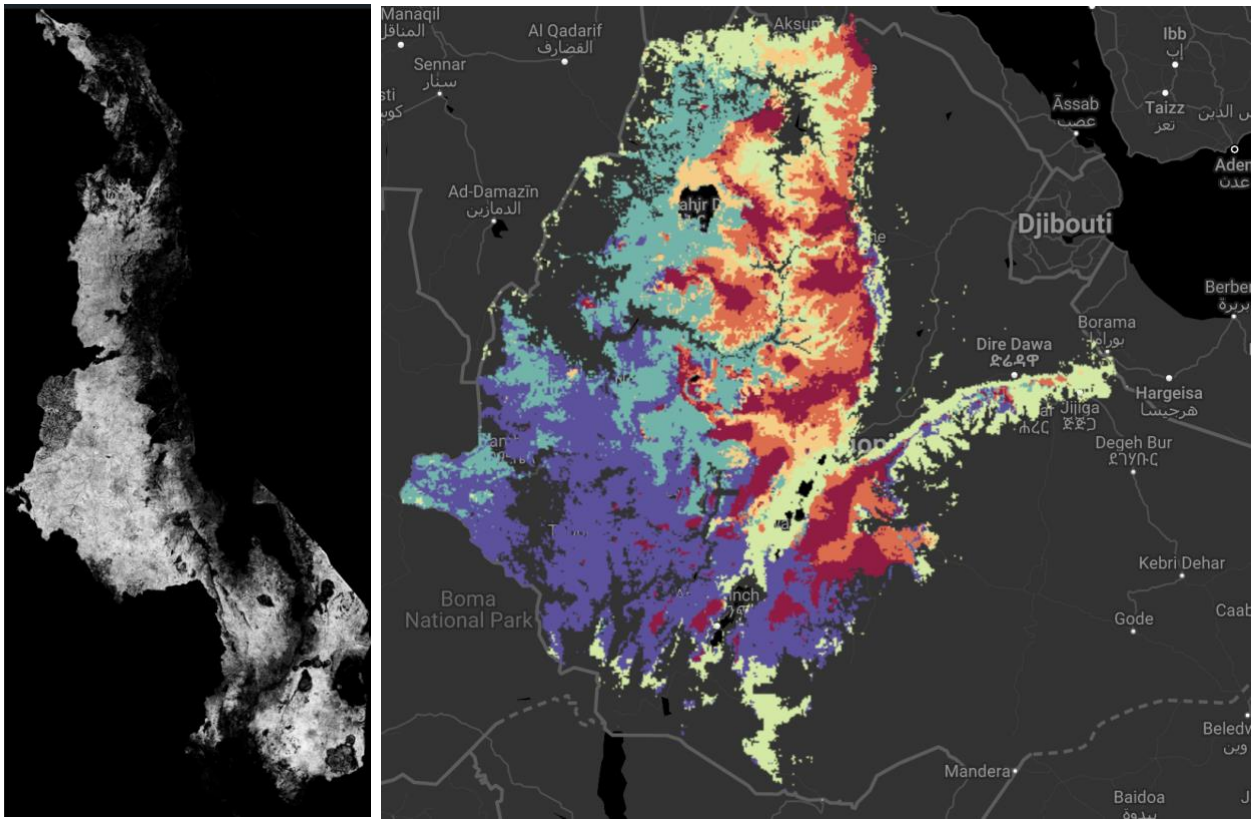
Georeferencing is perhaps the most critical factor in enabling the integration of survey data with machine learning models; location is the unifying attribute linking the survey information with geospatial data from other sources. For example, an important step in this research was to accurately link the georeferenced plot-level survey data to publicly available Sentinel-2 imagery and other ancillary remote sensing data for the reference agricultural season. This pre-processing is necessary to make the data analysis ready for use in the remote sensing based experimental research. The guidebook highlights factors in the survey planning process that could enhance the utility of collected data for informing machine learning models, without diverting efforts from primary survey objectives. The forthcoming recommendations are intended to minimize additional burden on field teams and be mindful of cost implications, while at the same time maintaining data quality and guarding against information loss.

---

[8] Katalin Tóth, 2021. "Georeferenced Agricultural Data for Statistical Reuse," *Geosciences* 2018, *8*(5), 188; https://doi.org/10.3390/geosciences8050188.

[9] Azzari, G., Jain, S., Jeffries, G., Kilic, T., & Murray, S. (2021). Understanding the requirements for surveys to support satellite-based crop type mapping: Evidence from sub-Saharan Africa. *Remote Sensing*, *13*(23), 4749.

Figure 1. Left: Malawi maize area map for 2016. Right: Ethiopia crop area map for 2018 *(deep red: barley; dark orange: teff; light green: sorghum; bluish green: maize; purple: other crops. These are snippets from the 10-meter resolution geotiff files produced by Azzari et al. (2021).*



2. METHODS FOR GEOREFERENCING IN SURVEYS

Protocols for collecting georeferenced data can vary across countries, resulting in non-comparable and/or non-optimized data collection practices. The guidance provided herein will help standardize data collection protocols and procedures, filling the data gap in countries where georeferenced information is missing or insufficient. This guidance is developed based on existing best practices on georeferencing of household and agriculture surveys as well as methodological research conducted by various researchers, including under the 50x2030 Initiative. This section describes different methods and recommendations for collecting georeferenced data in the context of 50x2030-supported surveys.

What to collect?

In general, georeferenced data can be captured for points (i.e., coordinates), vectors, and polygons, which refer respectively to locations, paths/lines, and the boundaries of parcels or plots of land. The most common, and arguably the most critical, georeferenced point is the location of the household's dwelling. This allows for geographic analysis at more granular levels, beyond national or provincial levels, for topics such as climatic zones or ethnic regions.

For land owned and/or used by the households or agricultural holdings, georeferenced data can be used to capture location of parcels and plots, as well as to measure land area.[10] For the collection of GPS coordinates of land owned and/or used by the household, the recommendation for 50x2030 surveys is to collect coordinates from the center of each cultivated plot (see Box 1).[11] Plot center points are estimated by the enumerator, after understanding the boundary of the plot with the help of the respondent. The plot center point should always fall within the plot boundary, even in the case of irregularly shaped plots (see Figure 2 for examples). Additional details on how to implement coordinate collection are provided in the next section.

Figure 2. Examples of plot shape outlines (green) and estimated plot center points (orange)

When to collect?

GPS coordinates of households (that is, their dwelling) can be captured as part of the household listing (for sampling), as part of the household questionnaire, or as part of both for internal data verification. If a listing is conducted as part of the survey, GPS locations can be collected for all residential structures or for all agricultural households. This will help survey teams to track and find the sampled households and can be used for sample analysis and fieldwork verification. In surveys that use pre-existing listing data (panel surveys, surveys shortly after a census, other relevant listing activity, etc.) that do not include coordinates, GPS coordinates must be collected during the questionnaire interview. Similarly, for the non-household sector, GPS coordinates may be available or collected through the listing or otherwise collected at the time of the interview. The 50x2030 reference questionnaires include collection of the farm location coordinates (generally the household dwelling in the case of the household sector or main agricultural building or main parcel in the case of the non-household sector) in the beginning of the interview, when the interviewer arrives at the location and is completing the cover page.[12]

The collection of coordinates for agricultural plots can be collected during the regular course of the interview. For plots which are further away from the household, protocols must be established as to

[10] Using GPS to measure the area of plots provides more complete information on, for example, the total land cultivated by the household or holding. GPS area measurement (together with respondent self-reporting) is the recommended method for capturing land area in 50x2030 surveys. For detailed guidance on measuring land area – methods as well as protocols/guidance for measuring area and storing plot or parcel outlines – refer to the 50x2030 Technical Note on Land Area Measurement.

[11] Azzari, G.; Jain, S.; Jeffries, G.; Kilic, T.; Murray, S. Understanding the Requirements for Surveys to Support Satellite-Based Crop Type Mapping: Evidence from Sub-Saharan Africa. *Remote Sens.* 2021, *13*, 4749. https://doi.org/10.3390/rs13234749

[12] In areas with weak satellite connection, where it may take longer for the tablet to acquire the signal necessary to record the GPS coordinates of the farm location, this may be done at the end of the interview, to limit waiting time of the respondent. Note, however, that the interviewer must return the appropriate location to complete the farm location coordinates if, for example, they finish the interview in the field or other location.

which distant plots should be geolocated as well as how to go about visiting these plots. Interviewers should coordinate with their team supervisors on the best schedule for data collection for these distant plots. For example, a protocol may include coordinating at the team level to schedule visits to plots that are within/near the EA but not within walking distance of the interviewed farm. Note, however, that the survey respondents will need to accompany interviewers to identify the plots. Additionally, survey operations should establish overarching protocols for which land is required to be visited. This could be based on, for example, a maximum walking distance, maximum travel time by any means of transport, or administrative boundaries such as parish. The 50x2030 reference questionnaires include collection of the plot coordinates as part of the plot roster.

How to collect GPS coordinates?

GPS coordinates can be collected within the existing framework of the survey and field teams can be trained relatively quickly on the best methods for capturing this location data, with some additional training time needed to practice capturing georeferenced land area measurements (described in the 50x2030 Technical Note on Land Area Measurement). In good conditions, collecting GPS points takes roughly the same time as asking one survey question, whether they are collected directly through a CAPI application or through a handheld GPS device, though the later approach requires manual entry of the coordinates into the CAPI application or naming of the waypoint and therefore can require more time overall (more on this below). Collection of GPS points may take more time if the satellite signal is weak and/or the accuracy requirements are set in CAPI are too high for the context. The collection is straightforward but requires interviewers to be physically standing in the place from which they will collect coordinates. This implies that additional time will be required for traveling to locations away from the household dwelling that require GPS data. Using GPS technology to measure land area, by walking the perimeter of plots, takes more time as it requires both visiting the plot and walking the full perimeter (more details are found in the 50x2030 Technical Note on Land Area Measurement).[13]

Collection instruments

Georeferenced information can be collected by GPS receivers in external handheld devices, directly in CAPI questionnaires using the GPS receivers embedded in tablets, and even in smartphones. Household and farm surveys have often used external, handheld receivers due to their higher accuracy and the limited availability of other methods. However, as CAPI-embedded technologies have become cheaper and more reliable, especially when connected to WiFi or 5G networks, this is becoming the most widely used receiver for collecting GPS coordinates in household surveys (though separate handheld GPS devices are still recommended for GPS-based land area measurement).[14] CAPI interview software, such as Survey Solutions, typically offer built-in options for easily capturing GPS coordinates, further facilitating the use of CAPI-embedded georeferencing. Using a built-in GPS receiver on the CAPI devices that are already in use by interviewers streamlines procurement requirements and makes data collection procedures less cumbersome. More importantly, it reduces data transfers and minimizes human error in transcribing information, as coordinates are automatically recorded in the CAPI questionnaires. Additionally, in cases where area measurement is implemented using an external handheld GPS device (which is currently the recommendation for 50x2030 surveys), the coordinates collected through the CAPI tablet can facilitate cross-checking and cleaning of plot identifiers entered transcribed from the handheld GPS devices, which are often prone to data-entry errors.

[13] For complete guidance on using GPS to collect land area measurement, see the 50x2030 Technical Note on Land Area Measurement.

[14] The 50x2030 Initiative is currently undertaking work to validate the use of tablet-based land area measurement. Due to concerns about the accuracy of the GPS sensors embedded in tablets, the current recommendation is to use handheld GPS devices, separate from the CAPI tablet, for area measurement. This recommendation may be updated pending finalization of the validation exercise.

Collection challenges

The accuracy of the GPS receiver's calculated location depends on the strength and number of satellite signals it receives. Obstacles such as buildings, mountains, and tree canopies can distort the signals and introduce error. Even more serious, user mistakes such as inaccurate or incomplete coordinate identification can be extremely difficult to rectify after the field teams have completed data collection, though this is less of a problem for GPS coordinates collected directly in the CAPI application as no manual recording of the household or plot is necessary. In general, external handheld GPS devices have superior accuracy to the GPS sensors embedded in CAPI tablets. The accuracy of the collection instrument, whether it be a separate handheld GPS device or a CAPI tablet with an embedded GPS sensor, needs to be carefully considered as it could have significant implications for data quality. Additionally, quality controls can be implemented in the CAPI software to ensure that coordinates are recorded from the appropriate areas, such as country or province.

Recommendations for 50x2030 surveys

Taking into consideration the array of desired data uses, the recommendation for 50x2030 surveys is to collect georeferenced coordinates for the farm location, generally the household dwelling in the case of the household sector or the main agricultural buildings/holding for the non-household sector, as well as the center point of all cultivated plots. It is also strongly encouraged to collect and save plot outlines for land area measurement. Collecting additional GPS data is up to the particular needs of a given survey. See Box 1 (above) for additional details.

In the case of multi-visit surveys, coordinates for cultivated plots should be collected during the post-planting visit. During this visit the dimensions of the plots will be visible (with the support of the respondent), since plots are defined as a function of the area cultivated with a specific crop or crop combination and those crops will be in the ground at the time of the post-planting visit. At the time of the post-harvest visit, those crops will generally have been harvested already, making the identification of plot dimensions less reliable. Therefore, for surveys implemented using a single post-harvest visit, coordinates should instead be collected from the center of cultivated (or partially cultivated) *parcels*, as these boundaries do not change over time.[15] Parcel-level coordinates do not provide the same level of detail as plot-level coordinates, but without a post-planting visit this is a second-best solution. Note that the 50x2030 questionnaire instruments do not collect georeferenced plot-level data for the non-household sector, which are typically administered in a single visit, but this can be added based on country context and survey needs.

Regarding the use of different GPS receivers, the recommendation is to use CAPI-embedded GPS for collecting coordinates – dwelling and plot centers – and handheld GPS devices for area measurement and plot outline delineation. In surveys where plot outlines are collected, plot center GPS points become optional but are still recommended as they can facilitate data management and cleaning. For example, because the GPS embedded in CAPI automatically links the GPS point to the farm identification data, it can be overlaid with the handheld GPS data to cross-check and correct any errors in household identification, which are manually assigned to the handheld device data.

[15] A **parcel** is defined as "any piece of land of one land tenure type entirely surrounded by other land, water, road, forest or other features not forming part of the holding, or forming part of the holding under a different land tenure type" (FAO, 2015, p.45). A **plot** is defined as a continuous piece of land on which a specific crop or a mixture of crops is grown or which is fallow is waiting to be planted, under a uniform, consistent crop management system. A parcel consists of one or more plots.

3. PRACTICAL GUIDANCE FOR GPS DATA COLLECTION IN 50X2030 SURVEYS

It is essential to plan for resources, equipment needs, and quality assurance protocols for collecting GPS data before the start of any fieldwork activities. The specifics of the surveys will determine how some decisions will be made; however, there are some details that will be common to all surveys.

This section includes guidance for planning and managing GPS data collection as a component of overall survey data collection for 50x2030 surveys.

Equipment

Equipment should be purchased as soon as possible to ensure sufficient time for procurement processing, equipment testing, and software set up. For large-scale surveys, procuring the necessary number of devices can take many months depending on the available supply and internal procurement processes.

Equipment for CAPI

The following equipment is required for conducting interviews using CAPI:

Android Tablets. One tablet per field worker plus some as reserve/back-up in case of damage or loss. The tablet will be used for survey data collection (interviews), which can include the collection of GPS coordinates. Tablets need to be capable of data transfer via internet or SIM cards. Although WiFi or network connection is not required at the time of the collection of georeferenced data, being connected can improve signal acquisition speed and strength.

Power Banks. Using GPS during data collection consumes additional tablet power. Providing one power bank for each interviewer (or at least a few per team) helps ensure they can work a full day and avoid loss of fieldwork days due to power interruptions.

Car chargers. If the field teams have a vehicle, car chargers can be useful backups in case of power disruptions or in areas without electricity.

No additional or GPS-specific equipment is required for using the embedded GPS feature on CAPI devices.

GPS considerations for CAPI tablets

The tablet specifications for using CAPI tablets for GPS data collection are not limited to the embedded GPS receiver and should factor into the overall device selection process. In addition to the general specifications under consideration for CAPI tablets, also consider the following minimum requirements when using the embedded GPS receiver:

Device battery life. The more you expect to use the embedded GPS, the more important battery life becomes. Consider opting for a tablet with a battery life that lasts at least an entire fieldwork day.

Memory and processing capacity. Tablets differ in terms of their internal memory storage and processing capacity (RAM). Requirements will differ based on expected use the GPS as well as expected frequency of downloads and uploads. For example, infrequent ability for downloading and uploading requires greater processing and memory capacity between downloads.

Global Navigation Satellite Systems (GNSS). The compatibility of tablets with various global navigation satellite system should be considered, keeping in mind the location where the survey is being implemented. There are several global navigation satellite systems, with the most prominent being the Global Positioning System (GPS), operated by the United States, and the Global Navigation Satellite System (GLONASS), operated by Russia. Tablets that can access both systems are generally

best suited for collecting GPS location data, as they have access to a greater number of satellites for measurement.

Equipment for handheld GPS

If utilizing a separate handheld GPS device, the following is needed:

GPS handheld device. Typically, one device is needed per interviewer plus some as reserve/back-up in case of damage or loss.

Power supply. If the GPS receivers do not come with their own charger (or if charging opportunities are limited), survey planners should budget for backup batteries and provide each team with multiple sets of replacement batteries, depending on battery quality and expected intensity of GPS use. The use of rechargeable batteries and a battery charger offers a good alternative to replacement batteries, where available. Nonetheless, interviewers should be sure to power off the devices when they are not in use to avoid excessive battery drainage.

Connection cords for PC/tablet. Most GPS receivers are sold with their own cord which usually has a USB connection. Extra cords may be needed, as they are easily misplaced.

GPS utility software. To transfer the GPS data from the device to a computer, most devices do not require any specialized software; be sure to check the specifications of the devices your team will use. For data analysis there are a wide range of GPS utility software options. If coordinates are being saved in a handheld GPS device to be exported and linked with the farm survey data, predetermined naming conventions will need to be designed and followed to allow for this linkage.

Back-up record. Although the handheld GPS device automatically records location data, consider whether your protocols will require field workers to also record some or all of the data externally, to use as a verification cross-check for the data stored in the handheld device. This can be, for example, manually entering the GPS data in the CAPI tablet as part of the main questionnaire. A template example for this is part of Annex 1.

Guidelines for selecting a GPS device

There are a wide variety of GPS devices available for purchase. When selecting one for data collection, the following minimum specifications are recommended:

- Ability to store at least 500 waypoints.
- Allows for use of various global navigation satellite systems (i.e., GPS and GLONASS).
- In addition to GPS coordinates, also records altitude, if of interest to the study.
- Track saving functionality and capacity for storing plot outlines, if that is to be undertaken as part of the survey.
- Power supply (batteries vs charging cords) that is appropriate for the supply needs of the survey location.

Generally, the number of device features is negatively correlated with battery life, and this should be taken into consideration when determining the optimal suite of features. Additionally, devices with simple and intuitive interfaces can minimize training requirements and errors during data collection.

Two models that meet these requirements and are often used in household and farm surveys are the Garmin eTrex Legend H and the Garmin eTrex 30. A field guide for the eTrex Legend is included in Annex 2.

Questionnaire design & CAPI programing

The decision to use computer-assisted personal interviewing (CAPI) to collect GPS location data requires planning, preparation, and piloting of the CAPI questionnaire. Once the location questions

are incorporated into the questionnaire design and training manual, the questions must be programmed into the CAPI questionnaire and piloted before the start of the fieldworker training.

In most CAPI programs, GPS coordinates can be easily programed as a single GPS question when capturing a single point (for example, the location of the household's dwelling or plot center point). For each specific location point, a GPS question will record a set of geographic coordinates (latitude, longitude, accuracy, altitude), which identify the physical location of the tablet.

Additionally, there are some features in CAPI applications that can support data quality and monitoring. In Survey Solutions, for example, it is possible to preload maps that the interviewers can view offline. These maps, which are uploaded at the Headquarters level and assigned to relevant interviewers, can be set up such that interviewers have a map for each enumeration area or other relevant level, and allow them to view their current location on a map.[16] Additionally, there are functions in Survey Solutions that can support the programming of warning or error messages to be displayed to interviewers in real time if coordinates collected during the interview fall outside of expected geographic boundaries, such as the enumeration area bounds, if known, or other administrative level.[17]

There are also CAPI features that can be enabled for measuring agricultural land areas, though the trade-off between convenience and accuracy will need to be considered. Methodological research by the 50x2030 Initiative is ongoing to validate the use of this approach (see the 50x2030 Technical Note on Land Area Measurement for more details).

Annex 1 shows various examples of how GPS questions can be formatted (and visualized) on both paper and CAPI questionnaires. Guidance on programming GPS interview questions in Survey Solutions is available in Annex 2. For those interested in programming GPS questions in CSPro, refer to the CSPro Users Guide.[18]

Quality assurance

GPS quality assurance monitor

Adequate training of personnel is crucial for proper use of the GPS receivers, regardless of whether those are separate handheld devices or embedded in CAPI tablets, including troubleshooting any problems that may occur in the field. Someone at the project level – NSO technical support, implementation manager, etc. – can be designated as a GPS monitor early in the fieldwork preparation process. This should be someone who has existing knowledge and/or experience with GPS data or the willingness and ability to learn quickly. This can be someone directly involved in the processing and analysis of the GPS data, or it can be someone on the project management team who can coordinate with the GPS data analyst or analysis team as needed for adjustments, updates, and/or data clarifications during training and data collection.

GPS training

The GPS monitor can train the supervisors who should understand the basic operation of the GPS receiver and know how to reset and modify the system settings (e.g. coordinate system, datum, measurement units), if needed, to ensure they are in line with the predetermined settings that are to

[16] For more information on how to upload and assign maps for offline use, refer to the Survey Solutions guidance at: https://docs.mysurvey.solutions/headquarters/mapsmanage/map-files/

[17] In Survey Solutions Designer, you can, for example, use *InRectangle* to confirm that the coordinates collected fall within a predefined rectangular area. Other functions, like GpsDistance allow for real time calculation of the distance between two sets of coordinates collected during the interview (e.g., distance between plot and dwelling). For more information, see: https://docs.mysurvey.solutions/syntax-guide/questions/syntax-guide-gps-questions/

[18] Specifically, refer to the section beginning on page 631 of the user guide (available at: https://www2.census.gov/software/cspro/documentation/cspro80.pdf)

be used in the survey. The GPS monitor or the team supervisors can train the interviewers, in collaboration with the GPS data specialists.

The field team, supervisors, and interviewers must be trained on how to collect data using GPS and the survey's data collection protocols. To avoid delays or inaccurate collection, everyone on the field team should have a basic understanding of how the GPS receivers work, including simple troubleshooting techniques.

It is essential to incorporate GPS training into the survey training schedule. This must include a hands-on session with adequate time allotted for everyone at training to practice collecting GPS data. For GPS training the team will need access to a field or park where there is a clear view of the sky. For surveys collecting only GPS coordinates, training is relatively straightforward and does not require much time, especially if collecting coordinates directly in a CAPI tablet. When also implementing land area measurement via a GPS device, training will need to be longer and allow for practical exercise. This training can last from a few hours or a half-day depending on the number of people and survey specifics.

Supervision during data collection

During data collection, the team supervisors can support interviewers' requests and conduct spot checks to ensure interviewers are following established protocols for properly collecting GPS data, that all necessary GPS data is collected, and that all survey data – including GPS data – are regularly synced or downloaded.

The GPS monitor should regularly review incoming data, verifying that all required data is being collected and synchronized/downloaded/entered regularly and that the format and data are logical (no invalid entries or entries with incorrect household identification). If GPS coordinates are collected through the CAPI tablet, they will be automatically synchronized with/part of the survey data. If GPS data is collected on an external device, the GPS data will need to be downloaded to a computer/laptop and shared with the survey managers. When done on a regular and frequent basis this also allows for revisitation of the farm/plot when needed to correct errors. The GPS monitor can also share incoming GPS data with the GPS data analysis team for further data analysis and validation. This should begin early in the collection timeline, or even during training, to ensure any issues are quickly identified and addressed.

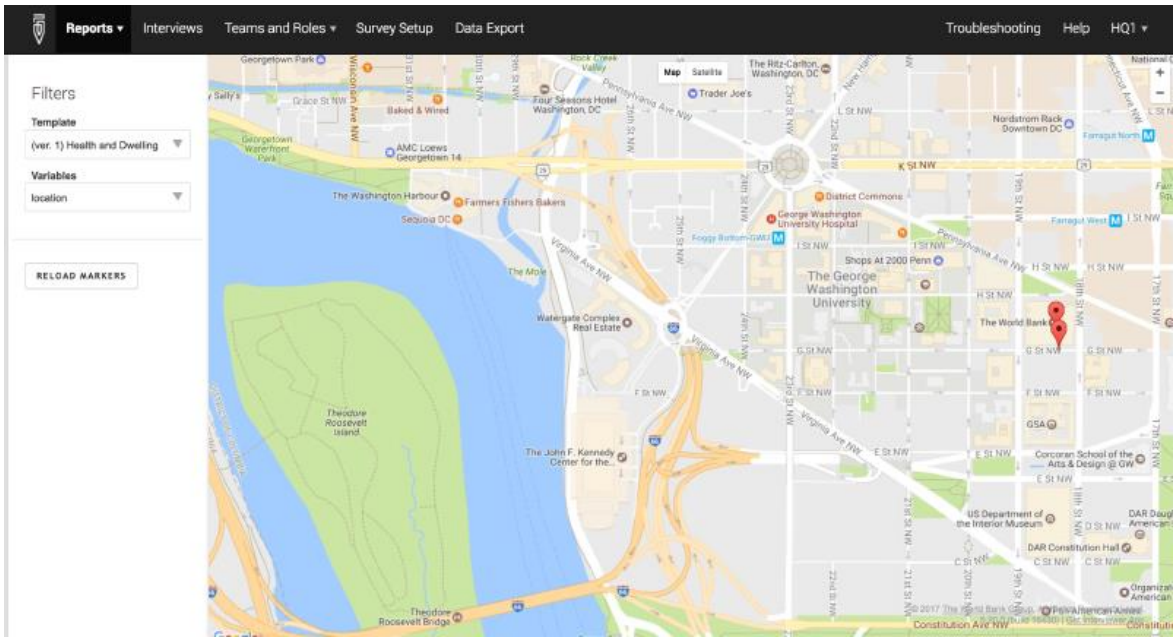
During fieldwork, the GPS coordinates collected through the CAPI questionnaire can also support data fieldwork monitoring. In Survey Solutions, for example, survey managers can visualize the locations of GPS coordinates directly on a map, allowing for verification that the field teams are interviewing in the appropriate areas and monitoring overall completion of assigned areas. An example of a "Map Report" visible in the Survey Solutions software is presented in Figure 3.

Data privacy & anonymization

Because GPS data are inherently data that can identify the respondent, maintaining confidentiality is of utmost importance. Collection of georeferenced data should be mentioned in the informed consent statement at the beginning of the interview. If respondents have any concerns about the use of such data, interviewers should explain that the location data is collected for survey management and data analysis purposes, and not for, for example, government cadastral records. Caution is needed whenever there are plans for disseminating georeferenced data. While specific 50x2030 protocols for anonymization and dissemination of georeferenced data are under development, existing protocols

that can be referenced include those established by the Demographic Health Surveys (DHS) program and utilized by the LSMS-ISA program.[19]

Figure 3. Example of a Survey Solutions Map Report. This report plots the location of selected georeferenced variables on a map, allowing for visual assessment of the survey progress and location of interviews. Image source: Survey Solutions reports documentation.

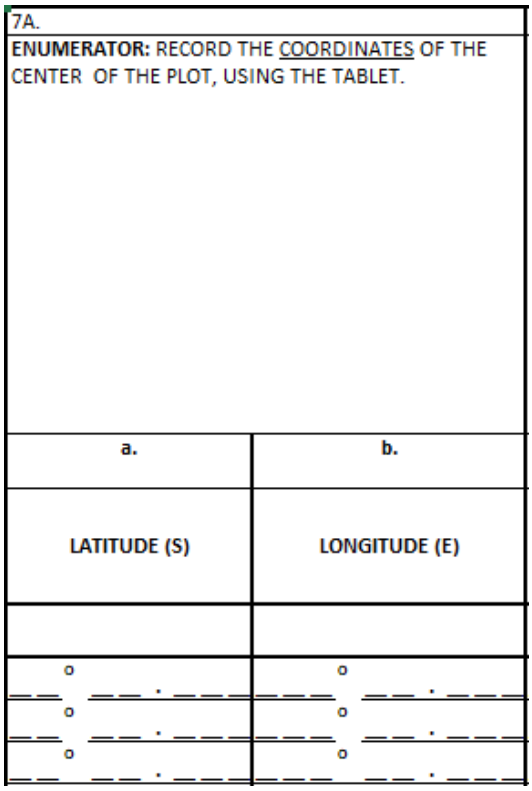


[19] Burgert, C. R., Colston, J., Roy, T., & Zachary, B. 2013. Geographic displacement procedure and georeferenced data release policy for the Demographic and Health Surveys. Available at: https://dhsprogram.com/publications/publication-SAR7-Spatial-Analysis-Reports.cfm

ANNEX 1: EXAMPLES OF GPS QUESTION FORMATTING

Below are examples of how GPS questions can be formatted in both paper and CAPI questionnaires.

Excerpt from paper version of the 50x2030 CORE questionnaire (plot roster)



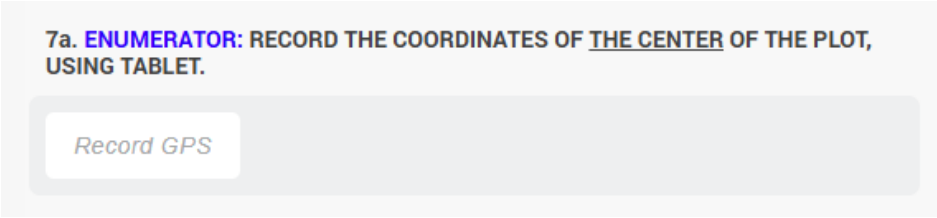
7A.
ENUMERATOR: RECORD THE COORDINATES OF THE CENTER OF THE PLOT, USING THE TABLET.

a.	b.
LATITUDE (S)	LONGITUDE (E)
o __ __ . __ __ __	o __ __ __ . __ __ __
o __ __ . __ __ __	o __ __ __ . __ __ __
o __ __ . __ __ __	o __ __ __ . __ __ __

Note, it is critical that the questionnaire tool accurately reflects whether the country is in the north or south latitudes, and east or west longitudes. For countries that span across both north and south, and/or east and west, a space should be added for the interviewer to record N or S and E or W.

Excerpt from Survey Solutions version of the 50x2030 CORE questionnaire (plot roster)

The GPS coordinates are automatically recorded by the system when the interviewer clicks on the "Record GPS" button. The interviewer must only click that button when they are physically in the location for which they want to capture coordinates.



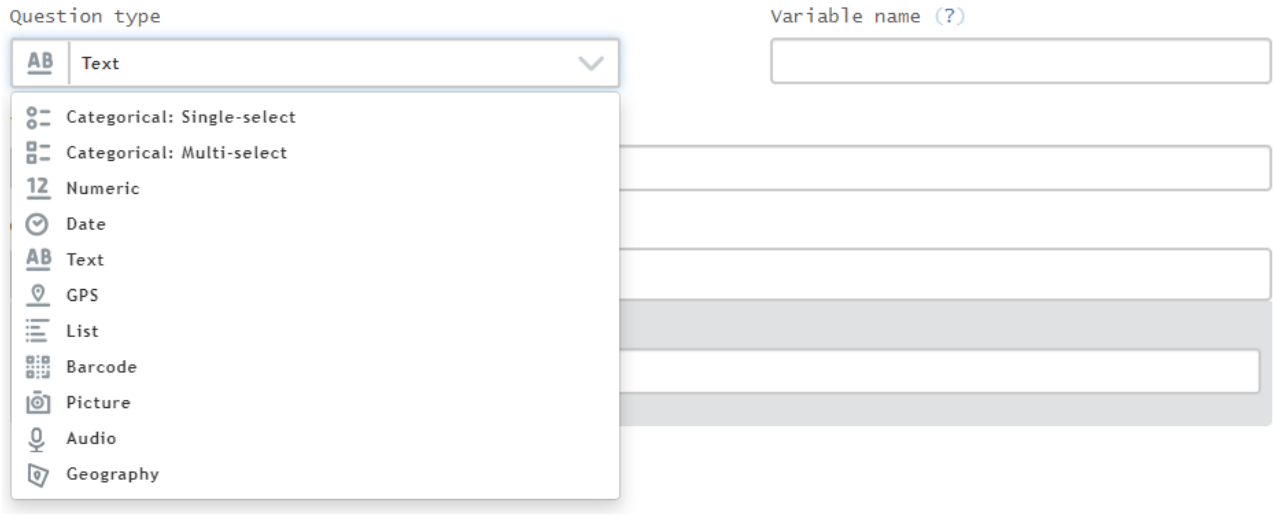
ANNEX 2: SURVEY SOLUTIONS IMPLEMENTATION

Survey Solutions questionnaires are programmed using the Survey Solutions Designer website. Survey Solutions allows for various question types, including questions to capture GPS data. There are two types of questions that can be used to collect georeferenced data in Survey Solutions: "GPS" question type, which collects coordinates (location data), and "Geographic" question type, which allows for various types of geographic information such as area (polygons), lines (polylines), single points (point), and multiple points (multipoint).

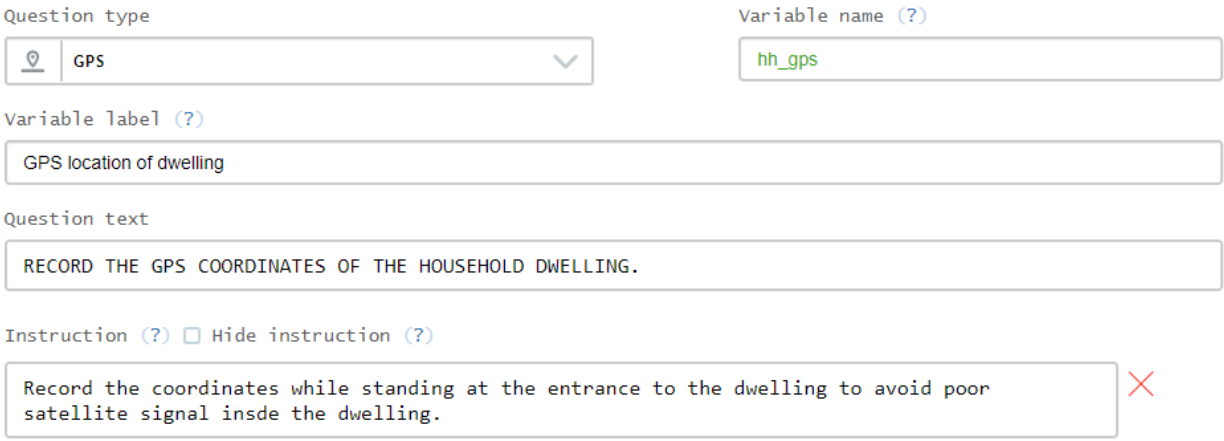
The information below explains the design and use of the **GPS question type**, which is the most straightforward option when aiming to collect a single set of coordinates.[20] The following guidance is based on documentation provided by Survey Solutions.[21]

Survey design: Adding a question to collect GPS coordinates to the questionnaire

Including a question to collect GPS coordinates into a Survey Solutions questionnaire is simple. First, click "Add question" in the section in which you want to include the GPS coordinate collection. This could be in the cover page or the plot roster, for example. Then, to create a GPS question, click on the question type text box and select GPS from the drop-down menu (image below).



You can then add a variable name and variable label, which will be included when exporting the survey data, as well as the question text that the interviewer will read and instructions for the interviewer, as needed. See the example below of how this appears in Designer.

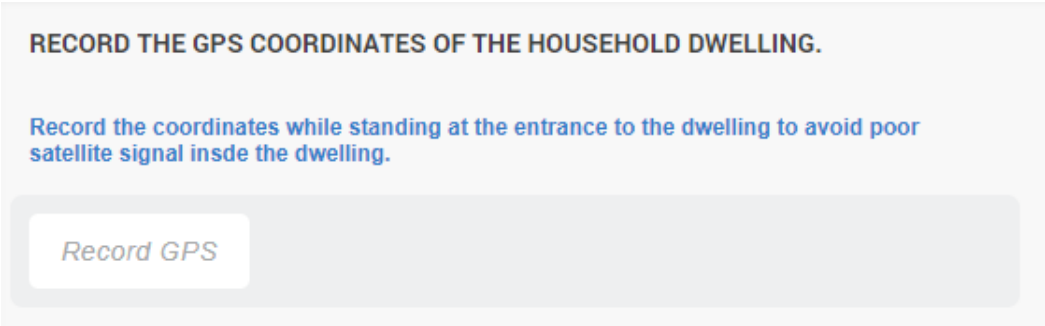


[20] Alternatively, GPS coordinates can be collected using the "geography" question type in Survey Solutions. The geography question type allows for the collection not only of a single point, but also of multiple points, lines, any polygons. Users employing this question type must also be aware of the parameters that need to be set at the workspace level, including the required accuracy (in meters) and the interval of automatic coordinate collection (in seconds). For more information on the geography question type, visit: https://docs.mysurvey.solutions/questionnaire-designer/questions/geography-question/

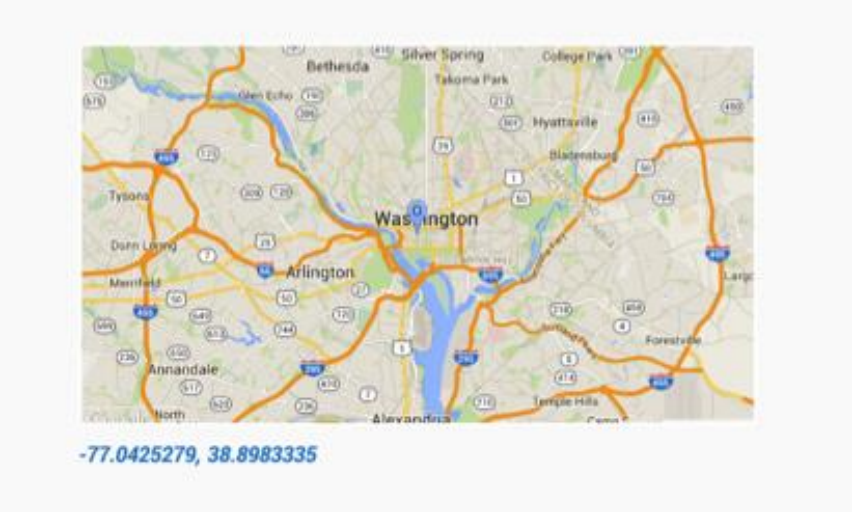
[21] Available at https://docs.mysurvey.solutions/questionnaire-designer/questions/gps-question/

Interviewer perspective: Collecting GPS coordinates in Survey Solutions during an interview

In the survey Solutions Interviewer app, the collection of coordinates is very straightforward. For the example programmed above in Designer, the interviewer will see, as one of the questions in their questionnaire:



To record GPS coordinates, **while standing in the exact location** from they wish to collect coordinates, the interviewer simply taps on the *Record GPS* button and the geographical coordinates will be captured and displayed along with a map of the location:



If the interviewer needs to re-capture the coordinates, in the case of poor accuracy for example, they can simply click *Record GPS* again and the revised coordinates will be displayed and saved.

Exporting GPS coordinates from Survey Solutions

Data collected using the GPS question type is exported as five separate variables: latitude, longitude, accuracy, altitude, and time stamp (example below).

Unanswered question exported in tab-delimited file

J	K	L	M	N
gps__Latitude	gps__Longitude	gps__Accuracy	gps__Altitude	gps__Timestamp

Answered question exported in tab-delimited file

J	K	L	M	N
gps__Latitude	gps__Longitude	gps__Accuracy	gps__Altitude	gps__Timestamp
38.8983029	-77.0425346	95.30500031	0	9/4/15 15:07

ANNEX 3: FIELD GUIDE FOR GPS COORDINATE COLLECTION

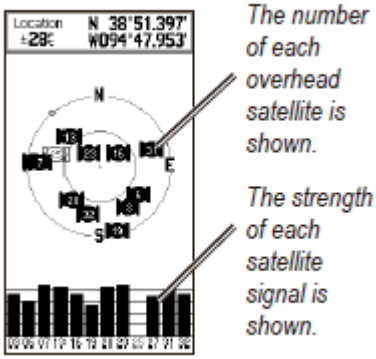
The guidance below, informed by the DHS Field Guide to GPS Data Collection, is based on a Garmin eTrex Legend H model GPS device, but can be adapted to other models of handheld GPS device.[22]

How to Set-up the GPS device:

Insert two AA batteries…

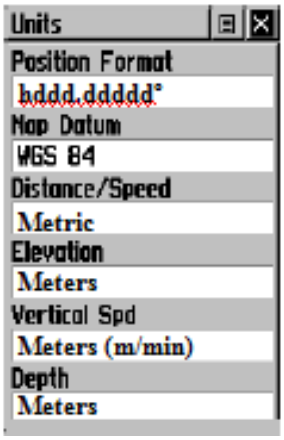
1. Remove the battery panel on the rear of the receiver.
2. Flip the small metal handle and rotate counter-clockwise to open the battery compartment.
3. Insert two AA batteries
4. Replace the battery panel.

Turn GPS Receiver On…



1. Press the POWER button.
2. The first screen says "Wait… Tracking Satellites."
3. After 1-3 minutes, the GPS receiver should acquire a number of satellite signals and display the message "Ready to Navigate. Accuracy: XX m."

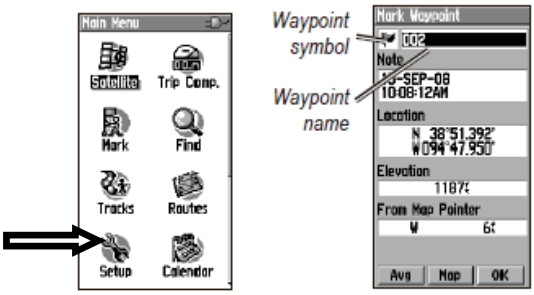
Set position format, datum, and distance unit on <u>all devices</u> to ensure consistency across all interviewers/devices…



1. Press QUIT/PAGE until you arrive at the MENU screen.
2. Using the ROCKER, scroll to SETUP and press the ROCKER.
3. In the SETUP sub-menu, scroll to UNITS. Press the ROCKER.
4. A screen appears called UNITS. Set these units:
   Position Format: hddd.ddddº
   Map Datum: WGS 84
   Distance/Speed: Metric
   Elevation (Vert. Speed): Meters (m/min)
   Depth: Meters
   a. Using the ROCKER, highlight the unit (ex. Position Format) you wish to change. Press the ROCKER.
   b. Scroll to the correct unit using the ROCKER. Press the ROCKER. for each receiver.

[22] USAID. 2013. Incorporating Geographic Information into Demographic and Health Surveys: A Field Guide to GPS Data Collection. Available at https://dhsprogram.com/pubs/pdf/DHSM9/DHS_GPS_Manual_English_A4_24May2013_DHSM9.pdf

Collect a GPS waypoint (set of coordinates)…



1. Turn the GPS receiver on by holding the POWER button.
2. The first screen says "Wait… Locating Satellites."
3. After 1-3 minutes, the GPS receiver should say "Ready to Navigate Accuracy: X meters."
4. Press PAGE until you reach the MENU screen. Press the ROCKER.
5. Highlight MARK on the MENU screen. Press the ROCKER.
6. Using the ROCKER, highlight the GPS point number then press the ROCKER.
7. A screen appears with a keypad. Using the ROCKER, change the waypoint name to reflect the household id (for household location coordinates) or the household ID – Parcel ID – Plot ID (for plot coordinates). Press the OK.
8. Again using the ROCKER, scroll to OK and press the ROCKER.
9. Record the latitude and longitude in the questionnaire.

To review the location of a GPS waypoint…

1. From the MENU page, use the ROCKER to highlight FIND. Then press the ROCKER.
2. Using the ROCKER, highlight Waypoints and press the ROCKER.
3. Navigate to the point you would like to edit using the ROCKER and press the ROCKER.
4. A screen appears called Waypoint. Using the ROCKER, highlight the waypoint name and press the ROCKER.
5. To view the location of a stored GPS point, use the ROCKER to highlight the MAP button and press the ROCKER.
6. This will display a rough map of the points collected.

To delete a GPS waypoint…

If the GPS data collector realizes that waypoints are incorrectly paired, (s)he should delete, then re-collect the incorrect waypoints.

1. From the MENU page, use the ROCKER to highlight FIND. Then press the ROCKER.
2. Using the ROCKER, highlight Waypoints and press the ROCKER.
3. Navigate to the point you would like to edit and press the ROCKER.
4. Use the ROCKER, highlight DELETE and press the ROCKER.
5. The message "Do you really want to delete waypoint XXX?" appears. Press the ROCKER to select Yes.
6. Follow the steps above to "Collect a GPS Waypoint."