European Commission

JRC TECHNICAL REPORTS

Using GIS and Remote Sensing to build Master Sampling Frames for Agricultural Statistics

Javier Gallego, Herve Kerdiles, Olivier Leo

2015

Joint Research Centre

EUR 27597 EN

Using GIS and Remote Sensing to build Master Sampling Frames for Agricultural Statistics

JRC Science Hub
https://ec.europa.eu/jrc

JRC98808

EUR 27597 EN

ISBN 978-92-79-54002-8 (PDF)

ISSN 1831-9424 (online)

doi:10.2788/438056 (online)

How to cite: Gallego F.J., Kerdiles H., Leo O., 2015, Using GIS and Remote Sensing to build Master Sampling Frames for Agricultural Statistics, Report EUR 27597 EN, doi:10.27/438056

Table of contents

CONTENTS

Abstract 3

1. Introduction 4

Background. 4

2. Geographic Information Systems (GIS) 7

Types of layers in a GIS. 7

Projections 7

Georeferencing elements in a list frame 8

Using GIS-based administrative registers as a basis to define an area sampling frame. 8

Administrative units: 9

3. Global Navigation Satellite Systems 11

Using GPS to define a sampling frame 11

Using GPS to run a survey (field work) 11

4. Remote Sensing 13

Main types of satellite images. 13

References 15

List of abbreviations 16

List of definitions 17

Abstract

This report is a JRC contribution to the FAO Global Strategy to improve Agricultural and Rural Statistics (GSARS). Its aim is providing guidelines on the suitable ways to use satellite images and geographic information tools to build master sampling frames that can be used both for agricultural and environmental statistics. The main readers to which the report is addressed are agricultural and environmental statisticians in developing countries.

We consider separately the use of technological tools for area sampling frames and for list sampling frames. The use of Global Navigation Satellite Systems (GNSS), better known as GPS, is also discussed, although its use is more connected to carrying out field surveys rather than to the design of sampling frames.

1. Introduction

The purpose of this report is to provide guidelines on the construction and use of a Master Sampling Frame (MSF) for agricultural statistics. One of the main pillars of The Global Strategy to Improve Agricultural and Rural Statistics is that agriculture be integrated into the national statistical system, and that this integration be achieved by the development of a MSF for agricultural and rural statistics. A successful integration means that the use of the same concepts and sampling frame in multiple surveys and survey rounds offers gains in efficiency and quality over separate survey systems. The Action Plan to Implement the Global Strategy identifies a research program to develop best methods for integrating agriculture into the national statistical system and implementing the MSF. The Action Plan also suggests that guidelines be prepared for the development of the MSF and its use.

Background.

There are three main requirements for agricultural statistics. While most people think the data are mainly for policy purposes, statistics are also needed to ensure markets function efficiently requiring very timely information. A serious problem in many developing countries is that that there is no early warning of food security problems. Census data that are several years old are of little use in understanding the current situation. Policy makers require information about many issues including the well being of the farm population and the impact of previous policy decisions relating to the economy and the environment. Bad data or lack of data lead to bad policy decisions, which lead to reduced support for statistics. A third use of statistics is to support decisions about investment by all levels of government and by farm operators. All uses require timely and accurate data that can be efficiently provided by sample surveys. Cost is a limiting factor; however, the use of a well-designed MSF will reduce data collection costs compared to the use of ad hoc data collection methods.

In many countries, a main source of agricultural statistics is the agricultural census conducted every 10 years or so. Where there is no agricultural census, the population and household censuses may provide information about the agriculture sub population. Both provide data at local administrative areas such as counties. Both also provide listings of farms and/or households that are used as sampling frames. The problem with both is that the data become out of date because of the long span between collection periods. In some cases it may take several years before the census data become available making them out of date from the beginning. These censuses are often conducted by the National Statistical Offices or Ministries of Agriculture.

The dilemma is that information about agricultural production needs to be very timely starting with advance estimates of pending production. A dilemma faced by many countries is that in order to fill this void, agricultural statistics come from Administrative Reporting Systems where the data are based on the subjective assessment of field agents working under the auspices of the Ministries of Agriculture. The accuracy and reliability of such data are dependent on the knowledge of each field agent. The data are usually aggregates for administrative areas, often villages where there are no clear boundaries defining their production area. It is usually not possible to describe characteristics of farms and households needed for food security analysis and policy decisions. It is not possible using these data to make inferences about the population at large nor do they provide measures of precision of the estimates. Nor do they provide a link to the environmental situation.

A well-designed sample survey can be completed quickly and with the capability of making inferences to the population with known probabilities and measures of sampling variability. And a well-designed sample for national estimates will require a surprisingly small number of agricultural holdings or households.

However, Administrative Reporting Systems can be useful as they can provide early warnings of rapid changes in crop conditions and may also provide useful information about important, but relatively rare commodities. The World Bank (2008) describes a Windscreen Survey and other rapid appraisal methods based on expert judgment. While these methods will not meet the requirements of the future system of agricultural statistics, they need to be considered in the development of the master sampling frame. In some situations, subjective reports by local experts can also be useful as auxiliary variables to be combined with more objective surveys. The subjective reports may also be useful for small area estimation.

One of the major findings brought forth by the Global Strategy is that the emerging data requirements go beyond what can be provided by the periodic censuses and administrative reporting systems which mainly focus on agricultural production and the farm or agricultural holding as the unit of interest. Emerging data requirements include the well being of the farm and rural households and agriculture's impact on the environment. The FAO's World Program for the Census of Agriculture (FAO2005b) recommends that the census consider the agricultural holding or farm as the basic unit for production and other economic statistics. However, the same report stresses the use of a population census and the collection of agricultural data for households that are not agricultural producers. On the other hand, the increasing demand for agri-environmental information requires that both the farm and households be linked with the land.

The challenge in developing the MSF for agriculture is that it must satisfy the need for three statistical units: the farm or agricultural holding, the household, and the land. And these three units need to be linked, for example, so that household income, health, etc are compared to the farm's economic situation; and both to their general environmental impact. In many cases there will be a one-to-one relationship between the agricultural holding, the household, and the land parcel. In these cases, it will be feasible to collect economic, social, and environmental information from one unit. If these units are geo referenced, the three units can also be associated with land cover. A challenge facing the development of the MSF is where there is not always a one-to-one link between the agricultural holding and a household.

The Master Sampling Frame

The development of the MSF begins by defining the data items to be measured. An example is the total production of maize in the country, the number of beef animals, average education levels by gender, and changes in land cover. Possible sampling frames would be a listing of maize fields, a listing of animals, a listing of people by gender, and a listing of land parcels. It would not be practical or even possible to develop a sampling frame for each data item needed. Instead, the population for each can be defined indirectly as listings of farms or agricultural holdings, households, and parcels of land. In each case sampling units and respective reporting units associated with each need to be defined.

The MSF for agriculture is a listing of sampling units that when associated with reporting units provide complete coverage of the populations of interest and provides a linking of the agricultural holding to the household and land dimensions. This listing of sampling units can be lists of farms from an agricultural census, a list of households from a population and housing census, a list of commercial agricultural enterprises not associated with a household, and a list of area units defined geographically. Multiple Frame Sampling is the joint use of two or more of these listings, and will be defined in more detail below.

If the final sampling unit for the listing is a farm or agricultural holding, the reporting units are the land it operates and the crops and livestock on that land, households associated with that land, and the land geo referenced. In other words, the reporting unit is the holding and all its activities and people living in it. This listing may include commercial agricultural enterprises not associated with a household and households with livestock, but no land. Households furnishing agricultural labor but no land would be excluded as would households with small plots for food consumption, but with production below a threshold. This frame will provide a linkage between the agricultural holding and the household associated with it, but not with other types of rural households. This listing is a sampling frame, more specifically a list frame. However, it is not a complete MSF because it excludes rural households.

If the final sampling unit is a household, the reporting units are the household, the agricultural holding and land to which it is associated, the crops and livestock on that land, and the land geo referenced. The listing, if from a population census, will include all rural households including those not linked to land, but that may have livestock, contribute to the agricultural labour force, or are simply rural nonfarm households. This listing is a sampling frame, and more specifically, a list frame. This can become a MSF if a listing of commercial agricultural enterprises is added to ensure complete coverage of the items to be measured. These agricultural enterprises can be added to the list of households to become part of the same list, or be used as a separate sampling frame in the multiple frame sampling context. Chapter __ provides additional details about area frame sampling.

If the final sampling unit is a segment or parcel of land, the reporting unit can be the holding associated with the land, the household(s) associated with the land, all crops and livestock on the land, and all holding and household characteristics. All rural non farm households within the land parcel are also reporting units. Commercial agricultural enterprises are also reporting units. This listing is an area sampling frame. It is also a MSF. Chapters _6_ provide additional details about area frame sampling including the use of points rather than land parcels as the final sampling unit.

The listings of households and parcels of land described above meet the requirement to be a MSF. This does not mean they are the most statistically efficient; this Handbook addresses those issues. The handbook will also show that the listing of farms or agricultural holdings can become a MSF by including a sampling frame of land parcels or points, thus becoming a MSF based on multiple frame sampling.

2. Geographic Information Systems (GIS)

Geographic Information Systems (GIS) are tools for collecting, storing, retrieving, transforming and displaying spatial data. These tools are used to manipulate, and operate on geographical elements such as points, lines, polygons and continuously varying surfaces. In a wider sense, the concept of GIS can also include the different sets of information stored, often known as layers.

A GIS provides a framework to store and combine different information layers. This can refer to information required to build the sampling frame, select the sample and compute extrapolation coefficients, as well as information generated while carrying out the survey.

There is a wide range of GIS software tools. Many packages are free and most of them are open-source while other tools are commercial, with very heterogeneous prices. The choice of a software environment needs to be carefully studied in function of the price, flexibility, training that can be found and support provided (including the community of users and developers). Among the commercial GIS tools, Arc-GIS is probably the most widely used. It is quite complete but expensive. The most popular free GIS systems might be GRASS and QGIS. Image analysis systems, such as ERDAS have a certain number of GIS analysis capacities.

Types of layers in a GIS.

The elements to be stacked in a GIS may be points, lines, polygons or nearly-continuous surfaces. A polygon can be seen as a polyline that finishes in the starting point, but if we do so in a GIS, the system will not understand the topology, i.e. concepts such as inside or outside. Nearly-continuous surfaces are represented as "rasters", i.e. bidimensional arrays of square cells (pixels) of constant size. Many spatial analysis operations, such as an intersection of maps, may be carried out both as polygon layers and as raster layers, but operations in raster mode are often faster and require less computational power.

Projections

We shall not discuss here the general problem of cartographic projections. We only give a few recommendations on issues that can lead to anomalies if managed without sufficient care when building a sampling frame in a GIS environment.

- All combined layers for a given analysis should be in the same projection.
- Some GIS tools make an automatic "on the fly" re-projection of each layer added to a specific map. For example an analyst may be working in UTM (Universal Transversal Mercator) projection and add a layer in geographic coordinates. The system will display it in UTM even if the corresponding files have the information in geographic coordinates. However, the accuracy of such "on the fly" projections is not always optimal (at least at the time being). Therefore, it is recommended to use an explicit tool for the re-projection before combining different layers.
- Among the properties that may be wished for a projection, the most important one for applications dealing with area estimation is preserving the proportion between areas (equal-area projections). A given projection cannot be conformal (preserving angles) and equal-area at the same time. Some usual projections, such as UTM, are not equal-area, but the area distortion of an UTM projection is extremely small, and in practice, there is no major objection to use an UTM projection. The most important rule is avoiding the direct use of latitude and longitude, in particular for large regions if they are not close to the equator.

Other projection systems have distortions if used far from the area for which they have been conceived. In general the best thing to do is respecting the choice made by the National Geographical Institutions.

Georeferencing elements in a list frame

List frames are often built on the basis of population or agricultural censuses, or on administrative records. Traditional list frames do not contain precise location information (geo-referenced) of the composing elements. The information on the location of a household or a farm does not go beyond the belonging to a given small administrative unit.

List frames may be enriched by adding a precise location information of their elements in a digital format. This may be existing or easy to produce if the elements are administrative units, but this is not always true for small administrative units that are often used in the first sampling stage. They can be communes, villages or Enumeration Areas (EA) specifically defined for a census. Here we call them EAs. Building or updating a digital layer of EA boundaries may be an important investment, but it is usually a cost-efficient one at medium term because it may be used for multiple purposes. In the case of agricultural surveys it allows characterizing EAs on the basis of land cover maps as an intermediate step to stratification or for an ex-post correction of estimates. It also facilitates the integration of list frames and area frames to build a multiple frame.

Georeferencing plots: Some administrative record systems foresee the obligation to georeference their elements, such as parcels or buildings. This is the case of many cadasters for taxing purposes. However using a cadaster as a basis for a sampling frame is not always a good solution because the owner is often not the cultivator and the physical cultivated plots may be very different from the cadastral plots. The potential usefulness of a digital cadaster for a sampling frame should be carefully analyzed case by case.

Geo-referencing a household by the coordinates of the dwelling is relatively easy in principle, but is not always a good geographic reference of its agricultural activity, since the fields managed by the household may be located far from the dwelling. The question becomes even more complicated if we consider **geo-referencing farms** that do not have a 1-to-1 correspondence with households. In this case a good criterion may be the location of the farm headquarters (where most of the machinery or stocking facilities are).

When building a two-stage sampling frame with EAs as Primary Sampling Units (PSUs) and households or farms as SSUs, the list of households or farms in the selected EAs is built or updated. Recording dwelling or headquarters coordinates provides a useful quality assurance tool, in particular in landscapes dominated by scattered houses or huts. In exchange, the usefulness of dwelling coordinates is more debatable when the livelihood is concentrated in villages or towns.

Using GIS-based administrative registers as a basis to define an area sampling frame.

Some countries have GIS layers of single agricultural plots coming from administrative registers. The modality of data collection may be photo-interpretation of ortho-photographic products combined with field observations or farmers declarations. GPS-based coordinate measurement on the ground usually generates a large amount of boundary anomalies that requires a heavy GIS editing work. Therefore it is generally not recommended to undertake this task if the result is used only to build a sampling frame. However if such a layer has been elaborated for other purposes and is available, it provides an excellent information source for an area frame, even if they are not very recent. The usefulness needs to be assessed in the specific case of GIS layers provided by cadastral databases for taxing purposes. In this case experience suggests that the difference between property and management is often so strong that the usefulness to build a sampling frame is very doubtful and needs to be assessed case by case.

Administrative units:

Most countries have good GIS layers with the boundaries of large or intermediate administrative units, but there may be changes in administrative boundaries. The suitable approach to update the GIS boundaries depends on the way the units are legally defined. If physical boundaries (a river, a sequence of mountain peaks) are used for the legal definition, the best way to update boundaries will be photo-interpretation of ortho-photographic products. If they have been graphically defined on hardcopy topographic maps, a suitable procedure may involve digitizing such maps, geometric correction with the help of reference points (consider at least 10-20 points per map sheet) and editing boundaries with the digitized map on the screen background. If the background map is good quality, intersection points of x-y (long-lat) lines will be good reference points and possibly no GPS coordinates capture will be necessary.

The situation is more complex for the small units (not necessarily with a clear administrative meaning): Enumeration Areas (EA), communes, villages, etc. In some countries the legal geographical definition may be unclear. This can be a source of bias difficult to address if such units are used as PSU, whatever the selected approach to build the sampling frame.

Some of the reasons why a GIS layer of small units can be useful are:

- A clear delimitation of small units is available for the implementation and assessment of any type of rural policy.
- Data will be easier to structure for these units. These data can come from different sources such as the census, classified satellite images or subjective expert estimates. If these data are exhaustive, or available for a very large sample, they can be used as covariate in a regression estimator. If there are extension workers associated to EAs on which they regularly provide expert estimates, the units attributed to extension workers are likely to be the most pertinent ones.

For the usefulness of these units both in the sample design phase and in the estimation procedure, their size should be suitable to be used as PSUs in the context of a two-stage sample.

Notice that the term PSU is sometimes used for units in which only one Secondary Sampling Unit (SSU) is selected (Cotter and Tomczac, 1994). In this case exhaustive covariates over the PSUs are not very useful as covariates in the estimation process, because the single SSU observed in the PSU does not provide a "decent" estimate for the PSU that can correlate well with the covariate.

If such units are too large, they may have limitations in the efficiency. For example if a country have 400 large communes, using them as PSUs may lead to a sample of say 40 communes, that might be too small for a good sampling efficiency.

A too heterogeneous size of the PSUs (geographic area, number of farms or households, agricultural area) may also have a negative effect on the efficiency, even if heterogeneity can be managed to some extent with PPS sampling or Horwitz-Thomson estimators. A strong heterogeneity can appear for communes and for census sections or enumeration areas. Size heterogeneity for communes may be due to geographical or to historical reasons. If it is due to geographical reasons (larger units in mountainous or very arid areas), it may be manageable with stratification. Heterogeneous size in census sections or enumeration areas is generally linked to population density and therefore is significantly reduced if stratification separates urban and semi-urban areas. The remaining heterogeneity does not necessarily need to be removed.

3. Global Navigation Satellite Systems

A Global Navigation Satellite Systems (GNSS) is a system based on a network of navigation satellites controlled by ground stations on the earth, which continuously transmit radio signals that are captured by receivers to determine the geolocation of the receiver (longitude, latitude, and elevation) on the Earth's surface. The oldest and most popular GNSS system is the US system GPS (Global Positioning System). For this reason GNSS systems are generally known as GPS. Another GNSS system is the Russian GLONNAS. Two other GNSS systems are being developed: the Chinese system BDS (BeiDou Navigation System), already operational in China and neighbouring areas, and the European Union system GALILEO. Although it is technically more correct to speak about GNSS, we shall use in this text the acronym GPS, far more frequently used.

In general GPS provide a support to field activities: geo-referencing plots, household or farm headquarters, locating sample units whose coordinates are known or measuring the area of a plot or landscape patch. Their use is rather limited at the stage of definition of a master frame.

Using GPS to define a sampling frame

Building a layer of agricultural plots is a costly operation. It is often carried out for taxing or subsidy purposes (cadasters or administrative registers) by computer assisted photo-interpretation of ortho-images, but it may require additional field visits with GPS observations. If field visits are numerous the cost may soar very quickly. In general the cost is difficult to justify for the specific purpose of building a sampling frame.

Another (more reasonable) use of GPS to build an area frame is collecting coordinates of reference points for the geometric correction of images or even maps that are not properly georeferenced.

Using GPS to run a survey (field work)

This is the most important application of GPS for agricultural statistics. When the sampling frame and sample selection has been defined on the basis of coordinates (possibly in a GIS environment), we need to locate in the field points with given coordinates. The two main tools to locate a point are ortho-photographic documents and GPS. If the ortho-photographic documents are good quality, they should have the priority in case of inconsistent indications within the range of accuracy of the GPS device ant the ortho-photographic documents. Relatively often field boundaries have changed after the date of the ortho-image. In this case GPS usually gives a warning, but an additional check of distances is wise.

Measuring plots with GPS:

GPS is very useful to measure the area of single plots on the field. For small fields, GPS area measurement with GPS has been considered insufficiently accurate for small plots at the end of the XX century and beginning of the XXI century. GPS measurements used to be less precise, but faster, than the traditional tape and compass measurement. A small negative bias of GPS measurement has been observed in an FAO pilot project (Keita and Carfagna, 2009). However more recent studies (Carletto et al., 2015) strongly suggest a significant improvement that the technological evolution in more recent years even with moderate-price GPS, in particular when they are able to combine signals of more than one satellite constellation (GPS and GLONASS at the moment, Galileo and BeiDou should follow)

GPS as a tool for quality control.

In some area frame surveys it has happened that surveyors made observations from unsuitable points. Sometimes because of location errors, but also due to a certain level of negligence, in particular under bad weather conditions. This type of weakness can be controlled if surveyors are due to record with a protected GPS device the point from which the observation is made.

GPS for point location control in objective measurement yield surveys.

A possible source of bias in yield surveys with objective measurements on a sample of points is linked to the determination of the point in which the crop sample will be collected. The traditional approach is providing rules on the number of steps the surveyor should move in certain directions. The movements are determined with the help of a random number table. In some pilot projects (Taylor et al, 1997) it has been observed that surveyors are not very rigorous on the application of the rules when supervisors are not present. The sampling process turns out to be more rigorous if coordinates are sampled before the field work and their coordinates are recorded in the field with a GPS device, taking a picture of the location.

4. Remote Sensing

When we talk about Remote Sensing (RS) in this text, we refer to images acquired with a conventional camera or electronic sensors on board aircrafts or satellites. The scenes record radiation in several ranges of the electromagnetic spectrum, including the normal visual range microwave radar, infrared and ultraviolet. Techniques applied to process and interpret remote sensing imagery include visual photo-interpretation and a wide range of numeric algorithms. This section provides guidelines on the use and choice of technology to develop a Sampling Frame and later in the estimation process.

Main types of satellite images.

Most satellite images are produced by optical sensors that measure earth reflectance. Other types of images, such as **SAR (Synthetic Aperture Radar)** or **LIDAR (Laser Imaging, Detection And Ranging)** have had little impact on agricultural statistics so far. SAR images are linked with the roughness of the land cover and LIDAR provides very accurate distance measures. SAR images have the advantage of measuring through clouds. This should be a major advantage in areas with persistent cloud cover, but the high noise/signal ratio has limited their usability so far, except for the delineation of areas cultivated with paddy rice. It seems the recently available SAR images from the Sentinel 1 satellite represent a major improvement and pre-operational applications may be soon available.

The most popular optical images in our everyday life are the **very high resolution (VHR)** images, with a pixel size between 0.5 and 2.5 m. However these images often have strong limitations for agricultural statistics: a full coverage of a country is too expensive and complicated to manage if it is built specifically for the definition of a sampling frame or for the production of estimates in a given year. However if a full ortho-photographic coverage has been produced for other purposes and is available, it provides an excellent basis for the stratification of an area sampling frame. An alternative that may be theoretically considered is building an area sampling frame on the basis of a sample of VHR images. This option would oblige to use PSUs of a size and shape similar to the VHR images and would lead to inefficient sampling schemes (Gallego, 2012, Gallego and Stibig, 2013).

Some leading companies in the information technology sector produce public access image mosaics with global coverage (**Google Earth, Bing**). The potential use of such images for agricultural statistics comes naturally to the floor. A major advantage is that they are available and easily accessible with an efficient interface. Most agricultural areas of the world are covered with VHR images and this is a significant asset, in particular for countries that seldom have recent homogeneous ortho-image coverages. However these public image layers have some limitations that we should keep in mind:

- Image geometry. For example the so-called Google projection is not an equal-area projection. It has been conceived to optimize the display speed at variable scales, but it is not optimal to provide comparable area measurements in different locations. Nevertheless the distortion introduced by the projection is likely to have a minor impact for agricultural statistics.
- Image overlay: when an older image is substituted by a recent one, the overlay between both shows a shift that may be up to 20m. This may introduce inconsistencies if an older image has been used to define a master frame and a different image is used to produce support documents to locate a given point or plot in a specific survey.

- Heterogeneous image dates. Two neighbouring areas are often covered by images that may differ 5-8 years. The impact of using images with such heterogeneous dates to define a master frame may be moderate if the field boundaries are relatively stable, but may be difficult to assess otherwise. Image viewing tools (Google Earth, Bing) report a date for the image on the screen, but it does not seem to coincide always with the date of acquisition of the image.

Images with a resolution of the order of 10-50 meters used to be called "high resolution images" until the 90's. At the time of writing this text they are rather referred to as **medium resolution images**. The most popular satellite series providing medium resolution images is Landsat. The Landsat-TM images have a resolution of 30 m, complemented since Landsat-7 with a panchromatic (black and white) band with a resolution of 15 m. They have been traditionally the most widely used for land cover mapping at national or subnational level and in particular for cropland identification. An important characteristic of satellite images is the swath, i.e. the width of the area covered in each pass of the satellite. TM images have a swath of 180 km, compared to the 60 km of SPOT, another widely used earth observation series of satellites. A non-negligible difference between Landsat and SPOT is the distribution policy since it has been decided to distribute Landsat images free of charge. There is a large number of satellites and sensors of the medium resolution type, whose availability is generally more heterogeneous. A hopefully significant improvement in the panorama should be the forthcoming launch of Sentinel 2, scheduled for June 2015) with 10 m resolution and a 290 km swath.

Coarse resolution images with a pixel size between 250 m and several km generally do not allow to distinguish single agricultural fields, except in countries with very large plots. They have the advantage of high frequency (daily images in general) and are a major tool to monitor the status of vegetation, yield forecasting and early warning, but they have limited interest to define a master sampling frame.

In the last years, the potential use of **drones** (unmanned aerial vehicles) is being widely discussed. The ortho-rectification and mosaicking technique is sufficiently developed to produce documents with an acceptable accuracy and the dates for image acquisition can be chosen with a flexibility similar to a field survey. A limitation of drones in many countries comes from flight regulations, that often forbid the flight beyond the sight of the operator. This limits the size and the shape of the area to be covered by a single flight, preventing in particular long and thin stripes of images that would be much more efficient than compact areas. Thus the limitation of drone's efficiency may be similar to the efficiency of VHR images. The area covered by a drone in a flight may be seen as a surrogate of a segment. Drones can provide images with a spatial resolution around 5 cm. If the great majority of the crops cultivated in a region can be identified with 5 cm resolution images drones may substitute field surveys. In areas with a complex crop pattern, in particular with a significant amount of mixed crops, this is unlikely to happen.

Small low-altitude aircrafts with pilot provide images with a similar resolution (~ 5 cm). They have the advantage of being more frequently authorized to fly long stripes (around 100 km by 100 m), much more efficient in terms of variance than approximately square units with the same area. Efficient stripe-based sampling plans can be defined for small aircrafts, in particular for the estimation of nomadic livestock.

References

Gallego F.J., 2012, The efficiency of sampling very high resolution images for area estimation in the European Union. International Journal or Remote Sensing 33 (6): 1868-1880.

Gallego F.J., Stibig, H.J., 2013, Area estimation from a sample of satellite images: the impact of stratification on the clustering efficiency. Journal of Applied Earth Observation and Geoinformation 22 pp. 139-146

Taylor J., Sannier C., Delincé J, Gallego F.J., 1997, Regional Crop Inventories in Europe Assisted by Remote Sensing: 1988-1993. Synthesis Report . EUR 17319 EN, Office for Pubblications of the EC. Luxembourg. 71pp. http://mars.jrc.ec.europa.eu/mars/Bulletins-Publications/Regional-Crop-Inventories-in-Europe-Assisted-by-Remote-Sensing-1988-1993

Carletto C., Gourlay S., Murray S., Zezza A., 2015, Welcome to Fantasyland: Comparing approaches to land area measurement in household surveys. World Bank Conference on Land and Poverty. Washington DC, March 23-27, 2015. https://www.conftool.com/landandpoverty2015/index.php/Carletto-523-523_paper.pdf?page=downloadPaper&filename=Carletto-523-523_paper.pdf&form_id=523

Keita, N. & Carfagna, E., 2009. Use of Modern Geo-Positioning Devices in Agricultural Censuses and. Surveys: Use of GPS for Crop Area Measurement. *57th ISI session* Durban, South Africa, https://www.statssa.gov.za/isi2009/ScientificProgramme/IPMS/0617.pdf

List of abbreviations

AFS:	Area Frame Sampling
EA	Enumeration Area
EU	European Union
FAO	Food and Agriculture Organization of the United Nations
GIS	Geographical Information Systems
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
LUCAS	Land Use/Cover Area-frame Survey
NDVI	Normalized Difference Vegetation Index
GSARS	Global Strategy to improve Agricultural and Rural Statistics

List of definitions

Agricultural land: total of cropland and permanent meadows and pastures (FAO WCA2020) . Land for aquaculture production is also included in the scope of the Global Strategy to Improve Agriculture and Rural Statistics.

Area Sampling Frame An area frame is a listing of land areas. The listing can be done in single or multiple stages. An area sampling frame where the ultimate sampling unit is either a parcel (segment of land) or a point to which a reporting unit of land associated with a holding is defined**.**

Crop area: "the horizontal projection of a particular extent of earth's surface" which corresponds to the area shown on cadastral maps (FAO 1982). This definition takes care of crop area in plains as well as hilly regions. It also ensures that the total area is equal to the sum of component area which is not the case when an area is measured on slopes..

Cropland: total of arable land and land under permanent crops (FAO WCA2020).

Enumeration Area: small geographical units defined for census enumeration purposes (FAO WCA2020).

Field: piece of land in a parcel separated from the rest of the parcel by easily recognizable demarcation lines, such as paths, cadastral boundaries, fences, waterways or hedges (FAO WCA2020). A field may consist of one or more plots, where a plot is a part or whole of a field on which a specific crop or crop mixture is cultivated.

Frame: A frame is the set of source materials from which the sample is selected (UN, 2005). It is the basis used for identifying all the statistical units to be enumerated in a statistical collection

Global Positioning Systems (GPS): makes it possible to find the geographic position of a point on the earth's surface by longitude and latitude. It can allow geo-referencing of the holding, the household and the land to the appropriate administrative areas.

Multiple Frame Surveys A sample survey based on multiple sampling frames. For agriculture this includes the joint use of area and list sampling frames. The frames are usually not independent; some of the frame units in one of the frames are present in the other frame.

Parcel: 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. A parcel may consist of one or more fields or plots adjacent to each other. The concept of a parcel used in the agricultural census may not be consistent with that used in cadastral work. The reference period is a point of time, usually the day of enumeration.

Plot: part or whole of a field on which a specific crop or crop mixture is cultivated (FAO WCA 2020).

LB-NA-27597-EN-N

JRC Mission

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new methods, tools and standards, and sharing its know-how with the Member States, the scientific community and international partners.

Serving society
Stimulating innovation
Supporting legislation

Publications Office

doi:10.2788/438056

ISBN 978-92-79-54002-8