Food and Agriculture Organization
of the United Nations

Information on Land in the Context of Agricultural Statistics

Publication prepared in the framework of the
Global Strategy to improve Agricultural and Rural Statistics

August 2016

Information on Land in the Context of Agricultural Statistics

Table of Contents

List of Figures

List of Tables

Acknowledgements

The report was developed by John S. Latham and Ilaria Rosati, FAO Geospatial Unit Deputy Director Office for Natural Resources, DDNS, DDN Division. The draft Guidelines were further revised and expanded by Chris Hill.

The authors would like to thank all the contributors and technical experts who have provided useful suggestions and comments for the compilation of this report. The authors particularly wish to thank the following professionals for their several important contributions: Elisabetta Carfagna for her continuous support and valuable suggestions, Mario Bloise (for the data sets and support kindly provided), Antonio Di Gregorio (for refining the LCCS3 concept table and classification systems), Banchi Girma, Luisa Basile, Nancy Gallagher, Angela Piersante and Dramane Bako.

Norah de Falco coordinated the design and communication aspects. The document was edited by Sarah Pasetto.

This publication was prepared with the support of the Trust Fund of the Global Strategy, funded by the UK's Department for International Development (DfID) and the Bill & Melinda Gates Foundation.

Acronyms and Abbreviations

ADG	Advanced Database Gateway (software)
AEZ	Agro-Ecological Zone
ALOS	Advanced Land Observing Satellite
AMIS	Agricultural Market Information System
AVHRR	Advanced Very High Resolution Radiometer
BAU	Business As Usual
DEM	Digital Elevation Model
EA	Enumeration Area
EO	Earth Observation
ESA	European Space Agency
ETM	Enhanced Thematic Mapper
FAO	Food and Agriculture Organization of the United Nations
FS	Farming System
GAEZ	Global Agro-Ecological Zone
GEO	Group on Earth Observation
GEOGLA	GEO Global Agricultural Monitoring
GIS	Geographic Information System
GLC2000	Global Land Cover 2000
GLCN	Global Land Cover Network
HH	Household
IGBP	International Geosphere-Biosphere Programme
IRS	Indian Remote Sensing
ISO	International Organization for Standardization
LAI	Leaf Area Index
LCCS	Land Cover Classification System
LCML	Land Cover Meta- Language
MADCAT	Mapping Device Change Analysis Tools
MERIS	Medium Resolution Imaging Spectrometer Instrument
MMA	Minimum Mapping Area
MSCD	Minimum Set of Core Data
NASA	National Aeronautics and Space Administration
MSF	Master Sampling Frame

NRL	Land and Water Division (of FAO)
Radar	Radio Detection and Ranging
RS	Remote Sensing (of the Earth's Surface by Earth Observation Satellites)
SDTS	Spatial Data Transfer Standard
SIGMA	Stimulating Innovation for Global Monitoring of Agriculture
SPOT	*Satellite Pour l'Observation de la Terre*
SUPARC	Space and Upper Atmosphere Research COmmission (Pakistan)
TM	Thematic Mapper
UML	Unified Modelling Language
UN	United Nations
UNEP	United Nations Environment Programme
USGS	United States Geological Survey
UTM	Universal Transverse Mercator
WGS84	World Geodetic System 1984

Preface

The United Nations Statistics Division convened the first meeting of the United Nations Expert Group on the Integration of Statistical and Geospatial Information on 30 October – 1 November 2013 in New York, USA[1]. The meeting sought to debate the development of a statistical spatial framework as a global standard for the integration of statistical and geospatial information. This objective recognizes the importance of the location and integration of geospatial, socioeconomic and statistical information in supporting social, economic and environmental policy and decision-making. In particular, the meeting brought together geospatial and statistical communities to discuss the development of an international statistical geospatial framework. This work included the review of geospatial frameworks for statistical systems, geospatial coding and statistical geographies, and standardizations of classifications, boundaries and metadata.

In October 2014, the Independent Expert Advisory Group on Data Revolution for Sustainable Development (IEAG)[2] emphasized the opportunities and challenges posed by the efforts to improve data for sustainable development. In its report, the IAEG identified the level of investment required to improve countries' statistical systems. In July 2015, the Third International Conference on Financing for Development, held in Addis Ababa, suggested the main global- and national-level actions that the community should consider and recommend to improve core data systems, which are crucial to monitoring and implementation of the Sustainable Development Goals (SDGs). In particular, national governments should take action to "*incorporate new data collection tools and technologies into SDG monitoring frameworks, such as earth observations, geospatial mapping, new sensors, and cell phone-based data*".

Existing data collection and information systems have been developed for different survey purposes and, notably, different geographies. The list below briefly describes each of these systems:

- **Censuses** – the systematic recording of information from all members of a given population

[1] http://ggim.un.org/UN_Statistical_Geospatial_EGM.html.

[2] http://unsdsn.org/wp-content/uploads/2015/07/Data-For-Development-An-Action-Plan-July-2015.pdf.

- **Household surveys –** national or subnational sample of randomly selected households (HHs) that provides data on demographic and socioeconomic characteristics
- **Agricultural surveys –** survey of agricultural holdings, ranches and individuals who operate related enterprises, and that includes data on crop yields, economic variables and environmental data
- **Geospatial data/infrastructure and facility inventories –** data with location-specific information (including the other data inputs mentioned above) and spatial visualization, including facility inventories and core geographic data layers
- **Civil Registration and Vital Statistics (CRVS) –** administrative data that records vital events in a person's life, including birth, marriage, divorce, adoption and death
- **Administrative data –** data holdings containing information collected primarily for administrative (not statistical) purposes by government departments and other organizations, usually during the delivery of a service or for the purposes of registration, record-keeping, or documentation of a transaction
- **Economic statistics –** financial and economic performance measurements, including labour force and establishments surveys, economic performance, employment, taxation, import and export and other industrial activities
- **Environmental data –** real-time monitoring, ground stations, and satellite imagery for a range of environmental variables, including biodiversity, air quality, water resources and forest and land use change

National agencies have developed and maintained many of these information systems; however, a significant overhaul should be considered, to ascertain their appropriate role in supporting the development of agricultural statistics, and the integration of the latter with other national statistical systems to generate comparable data for areas. Area sampling frames foster the production of accurate agricultural statistics that enable the integration of agricultural, rural and agro environmental surveys with Earth observation (EO) and remote sensing (RS) data, thus giving rise to effective systems for collecting, analysing and monitoring the status of the Earth and developing territorial and environmental databases.

Despite the increasing availability of high-resolution and high-frequency EO data sets, for certain agricultural and rural monitoring purposes (such as yield estimation and forecasting), the satellite data remain limited. However, the processed image data can be used effectively to discriminate broad classes of agroecological zones and land cover, cropped areas, irrigated areas etc., and can provide a foundation for designing and developing the Master Sample Frame (MSF) database, which combines non-spatial and spatial information across the whole survey area or country. The value of the MSF lies essentially in reducing the costs of survey planning and implementation and sample selection by using the same sampling frame for a number of surveys; this can justify the additional effort required to develop the MSF. Georeferenced spatial information and technology assume an important role in the preliminary classification and characterization of the territory; this information enables analysis of land use and landscape, such that a land cover baseline (which can also be linked to other georeferenced data, such as vector, raster and tabular data) can be created to facilitate identification of the different characters and elements of the territory and data integration (Carfagna, 2014) [41]. The MSF's maintenance and durability are important to preserve its value. An MSF is likely to have a long lifespan; however, its value declines over time due to factors such as population growth and changes in land use. The use of satellite image data in developing the MSF also enables detection of substantial land cover changes or changes deriving from vegetative indices analysed over time, which may promote action to update the MSF.

The Global Strategy to Improve Agricultural and Rural Statistics (hereinafter, Global Strategy; see World Bank et al., 2011) was designed to engage in research and formulate recommendations to improve the quality of agricultural statistics. The programme has enjoyed the support of the Geospatial Unit (DDNS) and of the Office of the Deputy Director General of the Food and Agriculture Organization of the United Nations (FAO; see FAO et al., 2012) in exploring how EO systems, as well as land monitoring and surveys, can feed and support statistical systems in developing a land cover/land use database with current geospatial technologies.

The objective is to develop a methodology to generate a reference baseline that can support the construction of an MSF and improve stratification and sample design. The sampling frame covers all units in the population that comply with the definition of the population (land area) and the population's physical arrangement. This combines the digital geospatial information available from land cover mapping, land use, administrative areas and enumeration areas

(EAs), and potential listings and areas of larger holdings or commercial farms (multiple frames). Therefore, the combination of these data sets is capable of supporting several related applications (e.g. crop condition assessment, crop estimation and crop type discrimination). Furthermore, the implementation of the land cover database can also be considered an opportunity to develop, improve and sustain the geomatic cartographic Remote Sensing/Geographic Information System (RS/GIS) capacities of local institutions and agencies, thereby facilitating their involvement in analysis and cartographic presentation. Linking the land cover parcels with the EAs and village and HH information enables the identification of strata, which provide homogeneous areas and thus improve the area sampling frame's efficiency. For this purpose, it is necessary to adopt a robust land cover classification that can assist in delineating the strata, as well as a consistent set of classes that are commensurate with the resolution of the satellite data and/or aerial interpretation. The imagery can also support field-based actions – such as the identification of parcels and the measurement of plot boundaries and physical information captured in the fields – by integrating field georeferenced capture tools and Global Positioning System (GPS) instruments.

This report presents the possible uses of geospatial data in developing the MSF, the basic concepts of land and landscape, and the land cover classification schema to support the stratification of the sampling unit of the area frame (AF). The report also illustrates the benefits of creating a GIS database with satellite imagery are also illustrated, and outlines the preliminary phases of tests regarding the use of RS/GIS techniques.

The second part of the report describes the implementation phases and results achieved in the pilot testing. The aim is to provide guidelines for the application of different RS products in various agricultural landscapes, and for the preparation of an integrated database to be used as a baseline in constructing the MSF.

1

Introduction

This report provides an overview of the role and applications of land geospatial databases in agricultural statistics, with a focus on environmental and technical aspects. Geospatial technologies are excellent tools for observing the Earth and detecting information on land, its cover and related land cover changes, and have strong potential to provide accurate and timely data sets to support applications in the agricultural sector.

RS technology has already been used in agricultural statistics for several important activities, such as:

- Monitoring land cover/land use;
- AF construction;
- Support to census and survey fieldwork;
- Land parcel mapping and holding size estimation;
- Crop acreage estimation; and
- Crop yield forecasting.

The frequency of land cover mapping requirements varies with the country's level of development and environmental change; however, a five- or ten-year interval is typically adopted. The cost of geospatial technologies currently varies according to the RS products (image resolution) and management systems applied. However, the opportunity to leverage a variety of images that are available free of charge, on a daily basis (e.g. Sentinel 2 10m resolution multi-spectral data), may reduce at least data acquisition costs – which would enhance the viability of national geospatial land cover/land use databases.

The Global Strategy has identified a Minimum Set of Core Data (MSCD) and aims to enable countries to produce and disseminate a reliable MSCD[3] on a regular basis. These data should constitute inputs for national accounts and global balances of supply and demand for food and other agricultural products, as well as for several indicators required to monitor and evaluate development policies, food security and the progress made towards meeting the Millennium

---

[3] World Bank, FAO, UN. 2010. Global Strategy to improve Agricultural and Rural Statistics, Report No. 56719-GLB, pp. 16-17. Available at http://www.fao.org/docrep/015/am082e/am082e00.pdf. Accessed on 12 August 2016.

Development Goals (MDGs)[4] and the Sustainable Development Targets (SDTs)[5]. Many of these data are directly linked to land/land cover/land use parameters, and are also used to sustain their social and economic dimensions. Indeed, certain items (e.g. SDGs 2.3, 2.4, 6.6, 15.1, 15.2, 15.3) contain information that could be included in a land cover database.

The efforts of the Global Strategy's research program to foster cost-effective methods for the improvement of data quality have been supported by FAO DDN's initiative to develop a land cover/land use database and land cover classification standards.

1.1. Objectives

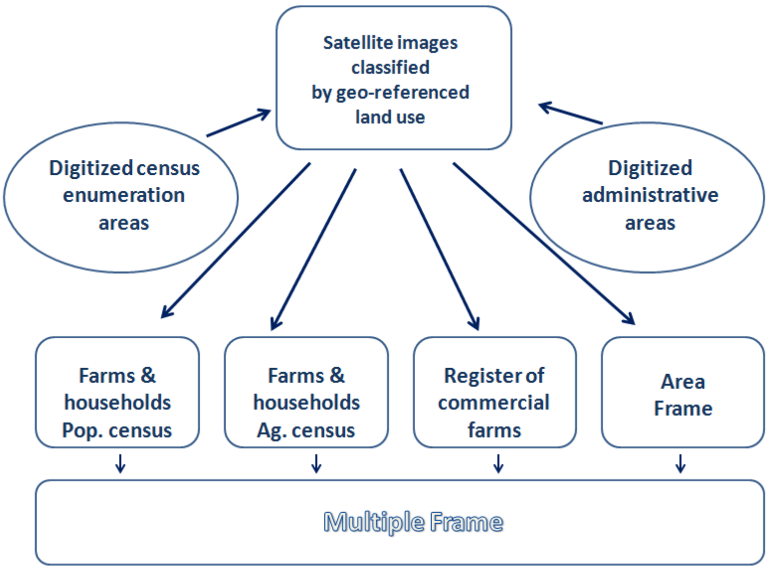
The main objective of this study is to propose a geospatial database structure, using a model adopted by FAO, to support the development of an MSF and stratification for sampling efficiency, enabling cross-sectoral analysis to link agricultural, socioeconomic and population data.

The development of an MSF is one element of the integration of agricultural statistics into national statistical systems. Indeed, the MSF can be used as the basis of all data collection efforts performed through sample surveys and censuses in a given sector; it enables the selection of samples for several different surveys or different rounds of the same survey, as opposed to building an ad hoc sampling frame for each survey (Carfagna, 2013). Setting up an MSF requires a large amount of information from a variety of sources. This information should be sufficiently comprehensive to build survey sampling frames that integrate the agricultural sector's economic, social and environmental dimensions.

Figure 1 is a stylized illustration of the process of constructing an MSF and a multiple sample frame, based on geospatially-collated data layers from different areal units.

[4] The official list of MDG indicators is available on the website of the United Nations Statistics Division: http://mdgs.un.org/unsd/mdg/Host.aspx?Content=Indicators/OfficialList.htm.
[5] Sustainable Development Goals Knowledge Platform: https://sustainabledevelopment.un.org/?menu=1300.

Figure 1: The construction of an MSF[6].



The process for building an MSF is as follows:

1) Frame design and identification of statistical units (small rectangular boxes in Figure 1):

- The minimum statistical unit of population censuses is the HH. However, in developing countries, the traditional sector or subsistence agriculture represents a very high proportion of the HH's total agricultural activity, such that it is possible to assume that a "one-to-one" correspondence exists between HHs and agricultural holdings. This, in turn, means that collecting data on traditional agriculture is equivalent to collecting data on HHs.
- In agricultural censuses, the statistical units are the holdings of the HH sector and the farmers of the non-HH sector.

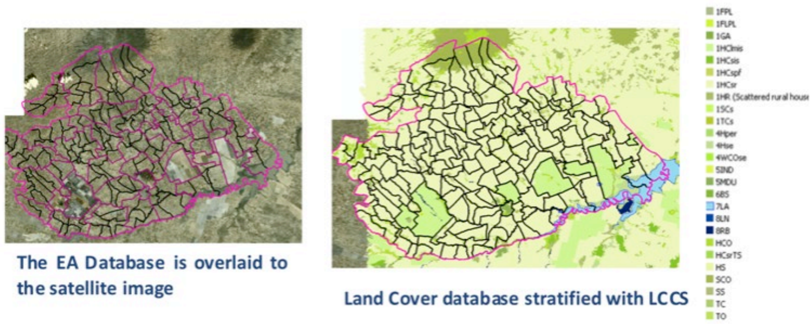
[6] Taken from Global Strategy. 2015. Technical Report on Improving the Use of GPS, GIS and Remote Sensing in Setting up Master Sampling Frames. Figure reproduced from Vogel, F. & Carletto, G. 2012. Master Sampling Frames for Agricultural and Rural Statistics. Paper prepared for the Global strategy to improve agricultural and rural statistics. High Level Stakeholders Meeting on the Global Strategy - From Plan to Action. Master Sampling Frames (MSF) for Agricultural and Rural Statistics, 3-5 December 2012. Rome, FAO, WFP. Available at http://gsars.org/en/technical-report-on-improving-the-use-of-gps-gis-and-remote-sensing-in-setting-up-master-sampling-frames/. Accessed on 12 August 2016.

- The statistical units of the register of commercial holdings are the big, commercial farms.
- In the AF, the statistical units are portions of territory, points or polygons.

2) Satellite images (highest rectangle in Figure 1) can be used as source of georeferenced data (e.g. land cover, derived land use and other land features).

Other baselines relating to the territory (e.g. EAs and administrative areas: the circle in Figure 1) can be digitized or collated from existing sources or maps. Often, the relevant authority will already have a digitized version of these.

Figure 2: Visualization of EA and LCCS stratification.



Handling RS products and georeferenced data in a GIS environment enhances the capacity for analysis and sampling. It is possible to simultaneously display geospatial data, visualize associated tabular information in the geographical space, correlate the different data sets revealing statistical and geostatistical relationships, patterns and trends, use the data for modelling and as a basis for serpentine numbering, etc.

Figure 2 illustrates a SPOT5 (5m resolution) image, overlaid with the land cover vector data sets obtained by visual interpretation and a digitized EA vector file. All data are georeferenced in Universal Transverse Mercator (UTM) projection. The data was extracted from the Ethiopia case study (see Chapter 7) to report the contribution of RS and GIS technologies to constructing and analysing AFs.

3) The MSF is developed as a multiple sample frame in which more than one sampling frame (which may be an AF and or a list frame) are combined with others to ensure that the sample is more representative and efficient. In this case, the MSF combines AFs from land cover, administrative and population

EAs with list frames (e.g. commercial farm listings, HH listings and livestock inventories).

The more methodologically robust the baseline (which guides the structure of the land cover database), the more accurate and reliable the information on the georeferenced land cover/land use will be. For these reasons, international standard classification, classes and parameters are crucial in providing harmonized information and improving data quality. The development of a robust land cover/land use database has multiple uses beyond stratification and statistical sampling; these can help to justify committing expenditure and efforts to the development and updating of the land cover database.

This work presents the theoretical and practical uses of RS and GIS technologies in developing and maintaining the MSF's geospatial database. The work also focuses on the generation of land cover/land use inputs for the stratification of the area sampling frame, and on the use of the geo-spatial data for data collection to reduce survey errors.

1.2. Report structure

This report is part of a series of outputs from the study. These outputs are divided into two Parts, and include a series of country-level case studies and processing outcomes.

Part I (this Report) covers the following topics:

- The basic concepts of land cover, land use and land cover change, as these are useful in the agricultural context, including the definition of "landscape" as a socioeconomic concept;
- Options for managing and developing a land cover geospatial database;
- FAO's standardized land cover methodology (LCCS classification)
- Evaluation of the effectiveness of Google Earth[7] as a baseline for mapping;
- Evaluation of the accuracy of existing maps and databases, and ways to improve their quality;

[7] This report makes reference mainly to Google Earth, considering that it may be used free of charge and that it may be easily connected with the main GIS software (ArcGIS/FAO toolbox). It does not preclude use of other virtual global maps or public domain images resources (e.g. Bing); however, the use of Bing in an ArcGIS environment is subject to a fee payable to the Environmental Systems Research Institute (ESRI) based on the usage level and individual cases, in accordance with the users' specific applications and purposes. In addition, BING focuses more specifically upon the United States and Europe.

- Case studies on land cover production, in support of AF production and statistical sample design.

Part II examines the different stages of image processing and interpretation for data extraction and manipulation. The following topics are examined:

- Georeferenced databases of multi-resolution images, analysis of satellite imagery products and extraction of relevant information and features on the land (agriculture, forest, etc.), land cover change and monitoring;
- Satellite imagery as a source of independent georeferenced information on crop type and acreage;
- The production chain to extract data from RS images concerning the territory, that can be applied at local, regional and country scales;
- LCCS3 as an approach for systematic categorization of land classes and related attributes through an object-oriented approach;
- The potential and costs of Earth observation products, technologies and derived products in support of agricultural monitoring and agricultural statistics.

The country-level reports cover the processing steps operated in Indonesia, Pakistan, Nepal, Cambodia, Darfur (Sudan and South Sudan), the Democratic Republic of Congo and Kenya; the countries were selected to ensure that various rain-fed and irrigated agricultural systems are represented.

2

Basic concepts of land in agriculture

2.1. Introduction

Land cover assessment and monitoring are essential to the sustainable management of agricultural and natural resources, environmental protection, and for food security and humanitarian programs more generally. Data on land also constitute core information that can be considered as a fundamental feature of any environmental study, including those upon agriculture, and an explicit geographical feature that other disciplines (e.g. land use, climatic or ecological studies) may use as a geographical reference. These data are fundamental in fulfilling the mandates of several United Nations (UN) agencies and programs, and international and national institutions whose mandates relate to the topics of food security, environmental degradation and climate change.

Land cover may be classified in accordance with systematic frameworks or classification systems that define their objects or classes and the criteria used to distinguish land cover, in accordance with a legend that is the expression of the classification system applied at a specific location, on a defined scale. However, the users of land cover data still lack access to sufficient reliable, consistent or comparable baselines, because many land cover mapping exercises produce legends that are a hybrid between land cover and land use classes. The classification may also be subject to the characteristics of the data sources from which they are derived; therefore, the class structure, class definitions and mapping procedures may reflect these limitations.

Land cover may be observed directly in the field or by means of RS. However, the observations of land use and its changes must generally be integrated with natural and social scientific methods (e.g. expert knowledge or interviews with land managers) to determine the human activities occurring in the different parts of the landscape, even when the land cover appears to be the same.

Examples of such land cover are:

- areas covered by woody vegetation may represent undisturbed natural shrubland (e.g. cover = shrub, use = natural land);
- a forest preserve recovering from a fire (e.g. cover = shrub, use = conservation);
- growth following tree harvest (e.g. cover = shrub, use = natural land/forestry);
- a plantation of immature rubber trees (e.g. cover = shrub, use = plantation/agriculture), swidden (cut and burned) agricultural plots that are in between periods of clearing for annual crop production (e.g. cover = shrub, use = agriculture food production/shifting cultivation), or an irrigated tea plantation.

As a result, the scientific investigation of the causes and consequences of land use and land cover changes requires an interdisciplinary approach that integrates both natural and social scientific survey methods.

This document proposes a common system of concepts, definitions and classification; it also presents the two methods recommended by FAO to analyse the landscape from an agricultural perspective, as well as relevant accuracy and data quality considerations.

2.2. Concepts and definitions of land cover, land use and land change

Land cover and land use are often confused with one another and combined within mapping and map attributes. This makes it impossible to separate the two components. However, the distinction may be important in the context of developing and stratifying agricultural sampling frames, depending on the survey purpose. Even where the land cover is the result of a particular land use (e.g. plantation), the class structures could be separately distinguished and classified as land cover to enable consistent classes. FAO (1997a) and FAO, together with the United Nations Environmental Programme (UNEP; 1999), define the separate elements of land cover and land use as follows:

Land cover	Land use
Land cover is the observed (bio) physical cover on the Earth’s surface. When considering land cover in a strict sense, it should be confined to the description of vegetation and man-made features (Di Gregorio & Jansen, 2000) [4]. In a holistic sense, it includes areas devoid of vegetation, the surface of which consists of bare rock or bare soil or water; these too may be described as land cover.	“Land use” is the use made of land by humans. It involves the management and modification of the natural environment or wilderness into built environment, such as fields, pastures and settlements. It also has been defined as the set of “arrangements, activities and inputs to a certain land cover type to produce, change or maintain it"[5]. It establishes a direct link between land cover on one hand, and socioeconomic actions and community livelihoods on the other.

Figure 3 illustrates some examples based on the above definitions.

The two images emphasize the differences between land cover considered as “grassland”, separating the classes of use as rangeland area (3a) and as dedicated to recreational purposes as a golf course (3b).

Figure 3: Distinctions in land cover and land use classification.

<table>
<tr><td></td><td></td></tr>
<tr><td>Figure 3a. “Grassland” land cover as rangeland land use.</td><td>Figure 3b. “Grassland” land cover may also refer to a “recreation area” such as a “golf course”.</td></tr>
</table>

Land Cover Change (LCC) is a general term indicating the changes occurring to the land cover over time. These may be natural successional changes, natural events, or changes due to climate change or human interventions. Many land cover types are essentially anthropogenic, created and maintained by human intervention and management; thus, many land cover classes may change as land management and use evolve. Land cover changes tend to be slow and successional in nature, and do not generally require assessment on an annual basis. Similarly, Land Use Change (LUC) may be detected over time through human modification for use, as well as through abandonment and degradation.

Monitoring changes may be especially important when determining whether the AF should be updated. As with the mapping of land cover and land use, it may be simpler to detect LCC than LUC, since the latter may require ground observations and interviews for validation.

Despite the importance of LCC and LUC, the methods adopted to study the subjects are somewhat vague, ambiguous and non-standardized. To date, several different techniques have been applied to study LCC, with a consequent variability in the type and quality of results and thus difficulties in comparing them. As a result, it is difficult to correctly quantify the global and regional changes that have occurred over the last three to four decades. According to the United States Geological Survey (USGS), land change studies attempt to describe and explain:

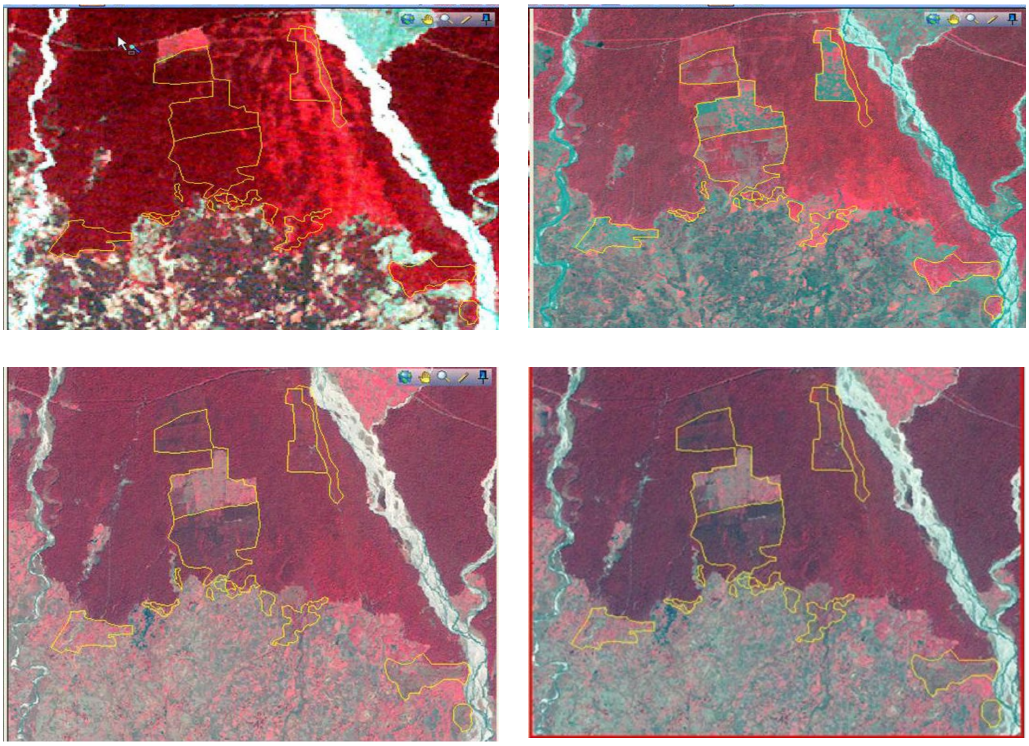
1. Where change is occurring;
2. What land cover types are changing (according to a "from-to" schema);
3. The types of transformation taking place;
4. The rates or amounts of land change;
5. The time over which such changes occur; and
6. The driving forces and proximate causes of change.

Figure 4 illustrates the study carried out by FAO's Land and Water Division (FAO-NRL) in Kenya [35]. The changes occurring in the country over the past three decades were analysed and statistics showing the direction of the changes (a "from-to" table) calculated. Figure 4 shows the four Landsat images in false colour composite (432); the red area represents Mount Kenya's forest tree cover.

Figure 4: LCC analysis in Mont Kenya over four decades (1970 – 2000).

1970 - The coverage is mostly uniform; a small pink patch in the upper part shows a deforested area where agricultural activity is occurring.

1980 - In the central area, the uniform reddish forested coverage was converted to agriculture.



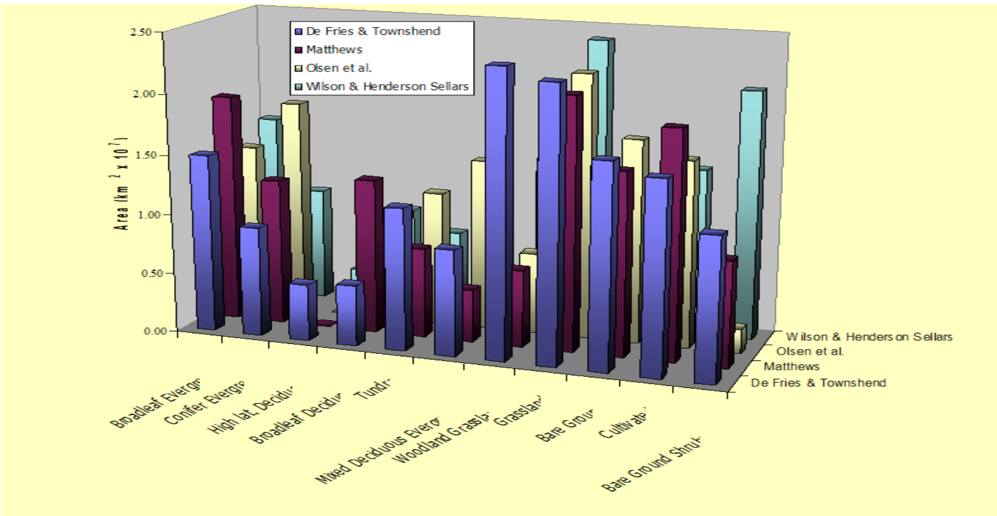
1990 - Reforestation has taken place in two patches in the central area, and in the upper right spot.

2000 - The situation shows no significant land cover change compared to the previous decade.

2.3. FAO LCCS: an international standard tool for mapping

Land cover class definitions may cause confusion not only in terms of the distinction between land cover and land use, but also in the use of conventional land cover names; such confusion may arise even when the names are followed by complete and exhaustive definitions, and when variations in land cover area classes may be observed. An example is given in Figure 5, which compares the areas of 11 cover classes (Broadleaf, Evergreen, Conifer Evergreen, High Lat. Deciduous, etc.) derived from four different global databases (developed respectively by Wilson et al., Olsen et al., Matthews, De Fries and Townshend [42]). The different definitions adopted for the classes affected the final assessment of the areas of the land cover types. This does not mean that the figures are inaccurate (as they relate to different definitions of class membership); rather, they are inconsistent, especially where few classes are used.

Figure 5: Comparing global land cover databases.



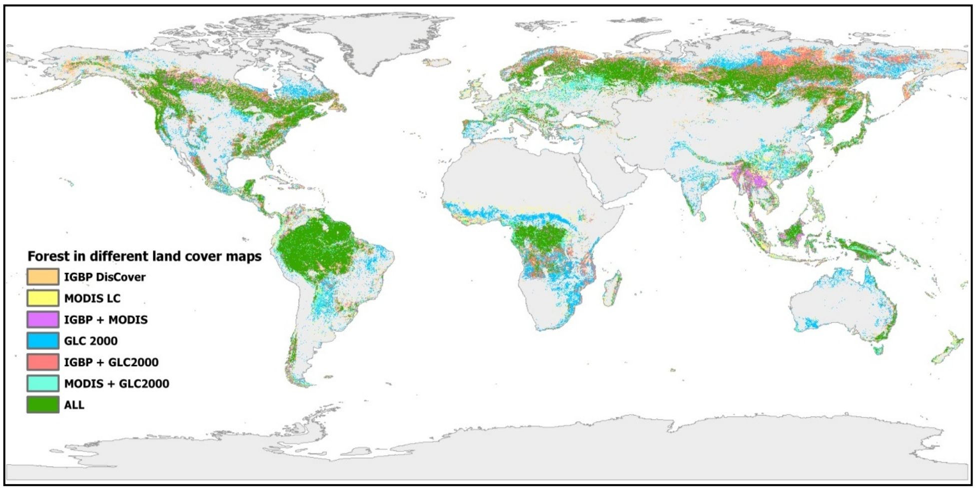
Consequently, the extents of the global classes vary from one database to another, in accordance with the definition of the class. Figure 6 below shows the extents of the "Forest" class reported by six different global classification systems (see the legend in Figure 6).

For example, with reference to the International Geosphere-Biosphere Programme (IGBP) and the Global Land Cover 2000 (GLC2000) project, the following definitions were adopted:

- IGBP legend: tree cover per cent > 60 per cent ; tree height > 2m
- GLC2000 legend**:** tree cover per cent > 15 per cent ; tree height > 3m

As a result, the global extension of the tree coverage is very different (the same coverage areas are shown in green). These differences clearly affect how the data may be used when stratifying an AF – and, consequently, the sampling errors – and may also affect comparisons over time, should the classifications in later surveys change.

Figure 6: Forest extents in different global databases.



A common classification schema would prevent this type of mismatching, as it would depend upon the semantic definition of the land cover class. This is particularly important for projects operating at the global scale, in which data comparability is crucial. Although it may not be feasible to develop a single standard legend, as it may be necessary to take into account the specificities of local cover, use and spatial scales, it is possible to adopt standard principles for the classification system.

The FAO and UNEP have collaborated to develop a land cover mapping schema that provides a common reference system and is also recognized as an ISO standard (ISO 19144-1 and ISO 19144-2): the Land Cover Classification System (LCCS)[8]. Development of the standard began in 1995 with the Land Cover Classification System (LCCS2); this evolved into the Land Cover Meta-Language (LCML), which became an ISO standard on Geographic Information ISO TC-211 (ISO 19144-2:2012) and was finally approved in July 2012. The LCML is a Unified Modelling Language (UML)[9] metamodel that provides a

[8] ISO 19144-2, available at http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=44342.

[9] "Metalanguage: a specialized form of language or set of symbols used when discussing or describing the structure of a language". "Metalanguage". *The Cambridge Dictionary* [online]. Available at:[http://dictionary.cambridge.org/dictionary/english/metalanguage. Accessed on 12 August 2016.

basis for any land cover classification system; it is capable of adopting different classes, although within a common structure. The UML diagram (see Figure 10 below) provides a visual representation of the classification model.

The LCCS3 introduces important innovations with respect to previous classification systems (e.g. the CORINE land cover and the USGS Anderson Classification). The application of the LCML helps users to describe the basic elements of the land cover, rather than compelling them to adopt strict definitions that may be difficult to fit to specific situations on the ground; the LCML also addresses repetitions, overlaps, and even gaps in classification.

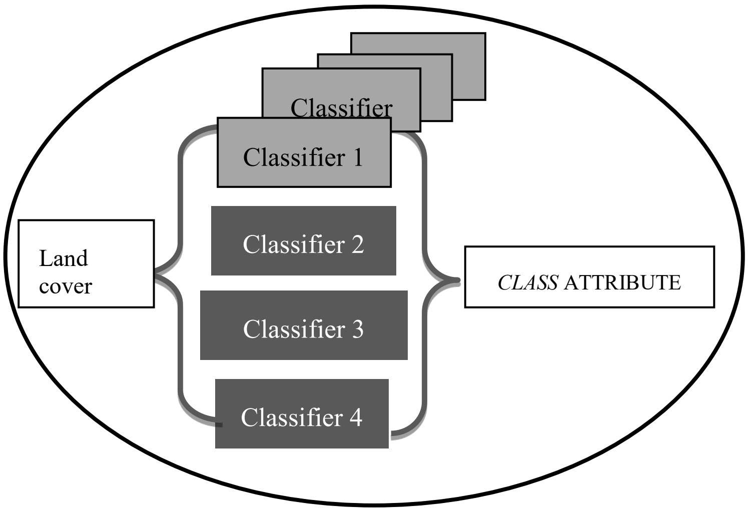
The LCCS has been comprehensively tested in various geographic locations. All of its versions (the most recent LCCS3, but also the previous LCCS2 and LCCS) have been widely used in national mapping programmes in Less Developed Countries (LDCs). Currently, the LCCS is the only universally applicable classification system in operational use.

The classification system is composed of classes, which are based on a combination of land cover diagnostic attributes; these are called "classifiers" in the LCCS2 and "basic elements" (or "basic objects") in LCCS3. The basic objects can be enhanced with further attributes to provide additional strata and describe detailed and distinctive land cover situations (e.g. down to individual crop types).

Figure 7 represents the land cover class as a combination of attributes or classifiers. These classifiers help to describe more complex situations (including different semantic meanings) arising in each land cover class, in any separate application ontology (classification or legend).

The Unified Modeling Language® (**UML®**) of the Object Management Group (OMG) assists in the specification, visualization and documentation of models of software systems, including their structure and design, such that all of these requirements are met. UML can also be for business modeling and modeling other non-software systems. See http://www.uml.org/#home; http://www.omg.org/gettingstarted/gettingstartedindex.htm; and http://www.omg.org/spec/UML/ 12/05/2016.UML is the OMG's most used specification, and the method adopted worldwide to model not only application structure, behavior, and architecture, but also business process and data structure.

Figure 7: LCCS Class and Classifiers.



As is the case in several classifications (e.g. CORINE), there is no predefined list of names. Rather, the classification is developed by the user, based on his or her specific mapping requirements. Therefore, heavy reliance is placed upon the user's idea of the class, and the capacity to materialize this idea using a meaningful and representative sequence of classifiers or elements. This means that classes derived by LCML can be customized to user requirements while retaining a common identity among users. The details within the land cover classes relate to the number of attributes (both biotic and abiotic) used to develop the classification; the more classifiers or elements are incorporated, the more detailed the class will be.

Class boundaries are defined either by the number of classifiers or by the presence of one or more different types of classifiers. Thus, the emphasis is no longer placed upon the class name, but rather on the set of indicators used to define the class. The LCCS3's innovative value lies in its open approach to creating the class and its boundaries; the limitations imposed in LCCS2 are eliminated. A Concept Note [11] was developed to present several aspects and benefits of applying the LCCS legend style in preparing the database. All subsequent mapping should adopt the latest version, LCCS3; however, existing national databases using LCCS2 and other land cover classification systems can be translated into LCCS3 standards.

Table 1 summarizes the main features of the LCCS2 and the LCCS3.

Table 1 – Comparison of LCCS2 and LCCS3 approaches.

LCCS2	LCCS3
Classification system	Tool to apply the rules of the Land Cover Meta-Language (LCML)
Scale-independent	Scale-independent
Source-independent	Source-independent
A priori, hierarchical system	Object-oriented system
Mathematical language used to formalize the meaning	Meaning is formalized on the basis of an UML schema
Predefined database used to generate user-defined land cover classes	LCML-derived classes described by an .xsd[10] file
Use of tree structure with a dichotomous phase, to define the main land cover group or class	Use of LCML basic objects, enhanced by property and characteristics
The element of "artificiality" is used at a higher level of the classification	Artificiality is to be linked to the LCML basic objects at any level
Many complex definitions Still in use	Complex definitions mainly avoided by layering biotic and abiotic LCML elements

The most distinctive theoretical characteristics of the LCML are:

1. The language's essential elements. Identifying the language's essential elements allows the classification to deliver a global standardization of land cover terms which is sufficiently detailed to ensure its practical applicability at national and subnational levels. In other words, the "elements" must be selected such that the number of classes is as low, while still being capable of representing distinctive land cover

[10] XSD (XML Schema Definition), a recommendation of the World Wide Web Consortium (W3C), specifies how to formally describe the elements in an Extensible Markup Language (XML) document. It can be used by programmers to verify each piece of item content in a document and to check whether it adheres to the description of the element in which it is placed.

Like all XML schema languages, XSD can be used to express a set of rules to which an XML document must conform if it is to be considered "valid" according to that schema. However, unlike most other schema languages, XSD was also designed with the intent that determination of a documents validity would produce a collection of information adhering to specific data types. Such a post-validation set of information can be useful in the development of XML-processing software.

situations. The classes derived by LCML can be customized to user requirements while retaining common identities between users.

2. Reducing, as much as possible, the complexity of descriptions and definitions. The object-oriented approach adopted in the LCCS3 enables reclassification to higher class levels, while maintaining detailed classifications to be applied as required.

The complexity of the classification has significant implications for the data's validation and utility. Within the agricultural statistics context, a simple, high-level classification may suffice, while other areas of application may require finer discrimination. The LCML and the LCCS3 can accommodate both requirements; however, the more complex and detailed the classification, the greater the costs and time necessary for production and validation.

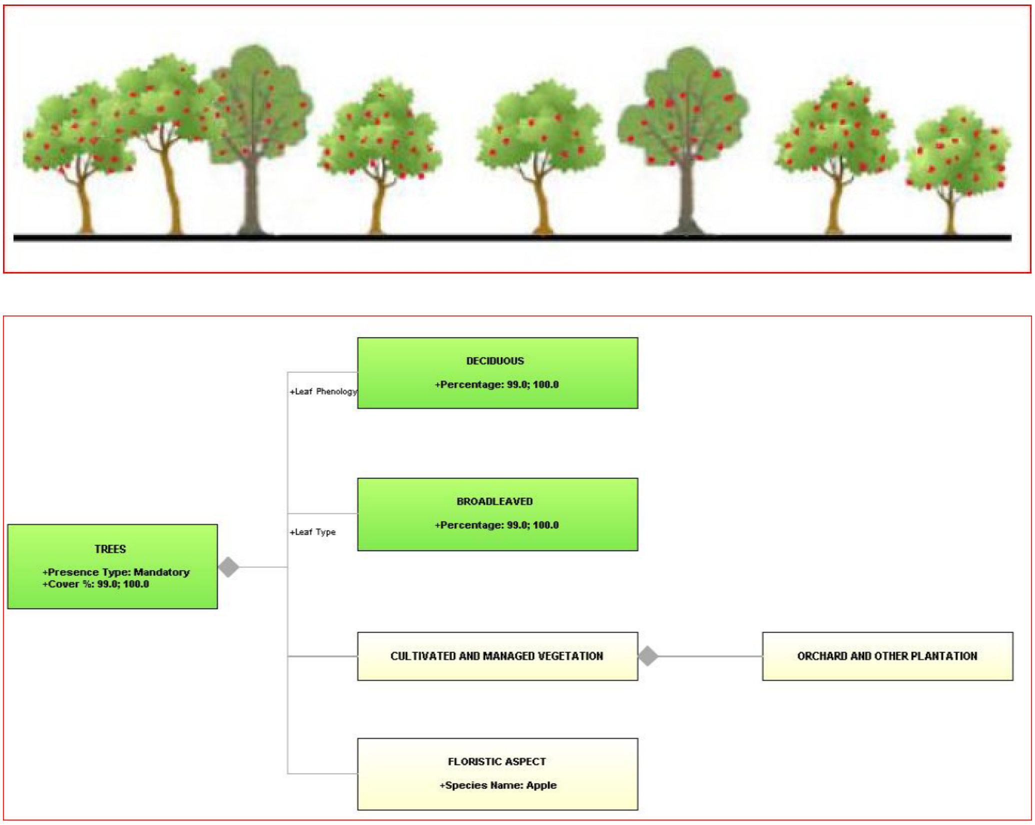
The purpose of ISO Section 19144 is to define a common reference structure for the comparison and integration of data for any generic land cover classification system. The approach has been used to define an LCML expressed as a UML model, which enables description of different land cover classification systems, and provides a rigorous logical framework for the description of any land cover classification system. The approach can also improve the harmonization and integration of spatial data sets, which were defined using different land cover classifications and the legends or nomenclatures developed from these systems, and enable their comparison and integration. ISO 19144-2:2012 specifies an LCML expressed as a UML meta-model that allows description of different land cover classification systems based on their physiognomic aspects. ISO 19144-2:2012 also specifies the detailed structure of a register to extend the LCML; however, it omits specifying how the register should be maintained. It provides a common reference structure for the comparison and integration of data for any generic land cover classification system, but does not seek to replace those classification systems (ISO 19144-2 2012).

The land cover classes adopted in the national classification system are important common reference structures for the country. Once a classification is established, subsequent modifications tend to be rare. Adopting the LCCS3 makes it possible to bear national peculiarities into consideration, while also embracing ISO standards and harmonization at all scales. This approach supports the Global Strategy's needs as well as those pertaining to the aggregation of national classifications into global products, such as GLC-

SHARE [29] and the Global Land Cover Network (GLCN), which have adopted the LCML.

Figure 8 presents the structure of the LCML for the “Orchard” class and illustrates the object-oriented approach to classification. The individual elements making up the land cover class are added to a systematic structure, thus providing an unambiguous description. There is no limitation to the number of blocks that may be added if more information is available, although it is generally sought to limit class structures.

Figure 8 - Example of “Orchard of Apple Trees” land cover class, expressed in LCCS3.



2.4. The “landscape” analysis

Agricultural activity occurs within a landscape.

Beddow, Wood et al. [13] define agriculture within a landscape context as follows: “agriculture is a complex spatial process with yields being greatly influenced by local factors such as weather and climate, soils, and pest pressures. Consequently, agricultural production and productivity are especially sensitive to spatial and inter-temporal variations in natural factors of production.”

While agricultural production is typically surveyed on a parcel basis and reported at the administrative level, agricultural production occurs within a landscape context of natural and semi-natural land covers, physical and biophysical variation, climate and weather variation, and human interventions. These factors affect the types of crop cultivated, the farming system adopted (subsistence, intensive etc.), the yields, etc. This georeferenced information can be used to develop and stratify AFs, increasing their efficiency.

The above considerations suggest the importance of adopting a comprehensive approach when studying the land, as well as a definition of the landscape that is based on appropriate parameters. The latter step is important for especially two reasons:

1. The interrelationships between the different components may be captured, as the definitions of appropriate indicators enable cross-cutting analysis of these components. The analysis of the agricultural landscape and its components support the preliminary characterization of the area of interest.
2. Providing appropriate definitions of the landscape elements contributes to the identification of homogeneous environmental subgroups. This facilitates the identification of clusters that are based on common elements.

The landscape can be characterized and defined using social, environmental and economic parameters such as field size, irrigation types, types of farm, the Digital Elevation Model (DEM), soil classes and land suitability; these factors are highly significant in agriculture.

FAO adopts two different approaches: (i) the Global Agro-Ecological Zoning System (GAEZ); and (ii) Farming Systems (FS). Although these systems were not specifically created for agricultural statistics, they identify different parameters that define the landscape. The two models are still not appropriate for application at the global or regional levels; however, both subdivide the land into homogenous strata, using elements such as climatic, soil, environmental and social parameters.

With regard to country-level agricultural statistics, it must be recalled that zoning with a broad scale provides a limited contribution to AF construction and stratification. Furthermore, some of the elements (especially the anthropic ones) that characterize a specific landscape may change over time, which reduces the efficiency of the stratification.

1) The GAEZ approach

To model the spatial dynamic of agriculture, FAO has proposed the Agro-Ecological Zoning System (AEZ) approach, which identifies areas having similar climatic, land form/soil and land cover features (further details are available at http://www.fao.org/nr/gaez/en/).

The GAEZ provides a framework for establishing a spatial inventory of land resources, compiled from global environmental data sets and assembled to quantify the many spatial characteristics required to assess land productivity under location-specific agroecological conditions. The GAEZ land resources inventory includes multiple spatial environmental and socioeconomic data layers: climate, soil, terrain, land use/land cover, protected areas, population density, livestock density, accessibility and administrative boundaries. It can generate modelled outputs, such as land suitability. The GAEZ is specifically created to define agroecological areas at a global scale; however, it can also be downscaled to provide output at national and regional levels.

The National Agro-Economic Zoning (NAEZ) for major crops establishes a procedure for national agroeconomic zoning that strengthens the capacity to manage issues of land use and land use planning, as well as the sustainable management of natural resources in the country.

A number of AEZ national models have already been completed, in collaboration with the beneficiary countries (e.g China, Libya, Kenya, Mauritius, Thailand and the Syrian Arab Republic).

2) The Farming Systems approach

The structural complexity and the interrelationships between the various components of farms – whether smallholdings or large corporations – may be summarized with the concept of Farming System (FS) zones.

The literature proposes a wide variety of definitions for the terms "farm system" and "farming system", each of which emphasizes different aspects. Examples are the systems components and systems interrelationships approach proposed by Dillon et al. (1992) and Shaner et al. (1982); or the complementary biophysical and socioeconomic processes presented by Norman et al. (1980) [12 and 25].

FAO [14 and 21] defines the FS as a "population of individual farm systems that have broadly similar resource bases, enterprise patterns, household livelihoods and constraints, and for which similar development strategies and interventions would be appropriate. Depending on the scale of the analysis, a farming system can encompass a few dozen or many millions of households".

Further, "[t]he variety of natural resources available to farm families normally includes different types of land, various water sources and access to common property resources – including ponds, grazing areas and forest. To these basic natural resources, one can add climate and biodiversity, as well as human, social and financial capital. The different livelihood is characterized by diversity. Each individual farm has its own specific characteristics arising from variations in resource endowments and family circumstances. The household, its resources, and the resource flow and interactions at this individual farm level are together referred to as a farm system. The biophysical, socio-economic and human elements of a farm are interdependent, and thus farms can be analysed as systems from various points of view. The resource endowment of any particular farm depends, inter alia, on population density, the distribution of resources among households and the effectiveness of institutions in determining access to resources. Regardless of their size, individual farm systems are organized to produce food and to meet other household goals through the management of available resources - whether owned, rented or jointly managed - within the existing social, economic and institutional environment. They often consist of a range of interdependent gathering, production and post-harvest processes, so that besides cropping and livestock keeping, household livelihoods can encompass fishing, agro-forestry, as well as hunting and gathering activities. Off-farm incomes, which make a significant contribution to the livelihoods of many poor rural households, are also included. Farm systems are not found only in rural areas; significant levels of urban agriculture exist in many cities and towns in a wide range of developing countries".

These concepts are used to delineate the major FSs. The various FSs existing in developing regions, as specified in [14 and 21], were classified on the basis of the following criteria:

- The natural resource base available, including water, land, grazing areas and forest; climate, of which altitude is an important determinant; landscape, including slope; farm size, tenure and organization; and
- The dominant pattern of farm activities and HH livelihoods, including field crops, livestock, trees, aquaculture, hunting and gathering, processing and off-farm activities; the main technologies used were also taken into consideration, as these determine the intensity of production and integration of crops, livestock and other activities.

Based on these criteria, broad categories were defined and are set out in detail in Annex III to this document. The criteria and broad grouping of farming systems mentioned above were applied to the six main regions of the developing world: Africa (AFR), Middle East and North Africa (MNA), Eastern Europe and Central Asia (ECA), South Asia (SAS), East Asia and the Pacific (EAP) and Latin America and the Caribbean (LAC).

Currently, the parameterization of the landscape and the definitions of farming systems are not standardized, unlike the case of land cover through LCCS3. The Menu of Indicators set out in Annex I presents some options for these contributory attributes. It is recommended to evaluate the role of delineating landscape and farming systems in the creation and maintenance of MSFs, with a view to identifying appropriate approaches to developing and maintaining these additional classification systems. It is also recommended to propose implementation of an AEZ at national scale, to enhance the definitions of the country's agroecological zones.

3

Options for handling a land cover database

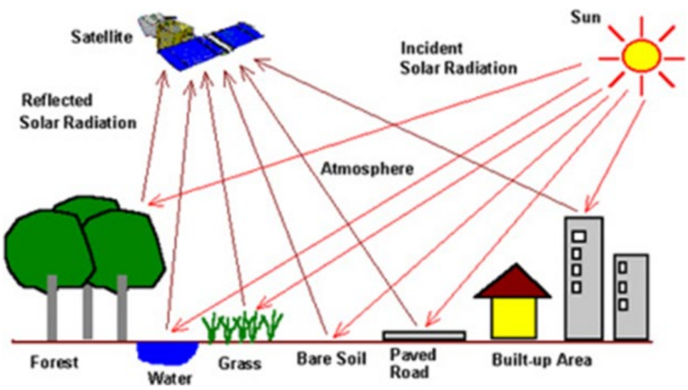
3.1. The contribution of RS and GIS technologies to the development of the land cover database

A key element in the development of the MSF is the creation of and provision of support to the data management systems required to operate the survey sampling. These approaches will depend, to a certain degree, on the technical capacity, IT infrastructure and budgets available for the establishment and maintenance of the systems. The requirements include the core data sets for the land cover and land use databases, the related environmental and socioeconomic variables, the geospatial files from other AFs (administration, population, enumeration etc.) and – potentially – the list frames. The requirements may also depend on the availability of existing national land cover or land use databases and their management, and may refer to the approaches adopted for global land cover products (e.g. GLOBCOVER and GLC-SHARE).

All of these components can also be maintained within a geospatial database environment, although currently, this is not typically the case. The MSF database may also include distributed architecture, according to which the system's components are maintained by relevant authorities having particular mandates and are combined at the point of use.

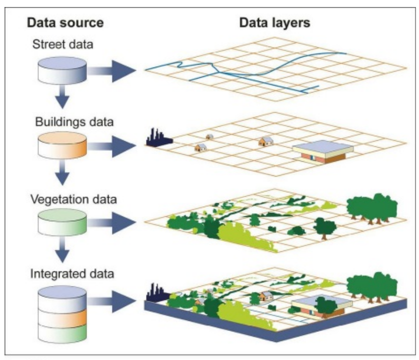
The term "remote sensing" (or RS) covers all techniques relating to the capture and analysis and use of data from non-proximate surveys; typically, this includes environmental and Earth resource satellites (e.g. Meteosat, NOAA-AVHRR, PROBA, Landsat, SPOT, Sentinel and ERS-SAR) and material from aerial photographs, UAVs, drones, etc. The main function of RS is to map and monitor the Earth's resources. The sources of EO and aerial data are typically supplied as grid-based or raster data sets at various temporal and spatial resolutions, and from which vector data sets (of classified features) are often derived by digitization or object-based methods.

Figure 9 – RS components.



The capability to combine different maps into a single operation, known as "overlaying", is one of the most important functions of GISs, together with modelling and site selection. The overlay functions enable coincident views of data at the same geographic location, thus allowing enquiry upon multiple factors or layers at a single location. Combining raster or grid-based data is also possible, although this process engages in analysis and reclassification through image processing or geospatial analysis to derive index products (e.g. the Normalized Difference Vegetation index, or NDVI). In terms of the land cover/land use database, these products may be grid-or vector-based. However, the approach used in the FAO LCCS3 workflow provides a vector-based classified output.

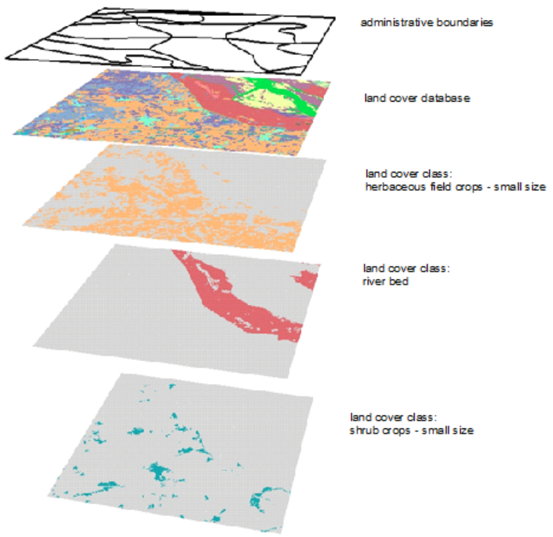
Figure 10 – GIS overlaying.



Source: GAO.

Figure 11 below illustrates a simplified land cover database as part of a combined geo-database of data layers contributing to the MSF; in particular, the administrative and land cover classification core layers are combined.

Figure 11 – Example of stratification based on a land cover database.



The contribution to agriculture statistics and monitoring of this overlay and analytical intersection is illustrated in Table 2 below (for further details, see also Concept Note 11). The result often consists in a combination of different data sources – mainly remote and proximate sensing and surveys – as both techniques are subject to error and require calibration and validation to support accuracy assessments.

Table 2: Value and uses of linked geospatial databases.

Main agricultural activities	Supporting sub-activities
Monitoring land changes and modifications (e.g. land cover change, soil loss)	- Description of land cover and its elements - Support to characterization of landscape or farm systems - Estimation of parcel or farm size - Classification of land cover and land use
Crop acreage estimation	- Calculation of extent of agricultural features (crop acreage, cultivated land, irrigated/rain-fed land) and non-agricultural features (forest land, bare ground, water); and of seasonal and phonological changes based on multi-date imagery - Support crop mask development
AF construction	- Provision of georeferenced information for designing the area sampling frame - Support to stratification
Support to census or survey fieldwork	- Linking relevant tabular data and attributes (e.g. number of HH members employed by farm or non farm) to the cover feature located in the geographical space - Guiding sample allocation and ground survey components - Enabling field boundary digitization against up-to-date imagery, or as linked with GPS technologies
Crop yield monitoring and forecasting	- Locating field areas on which the RS vegetation index time series can generate proxy of yields - Monitoring variance and decadal change for early warning, on basis of historic vegetation index and change data

3.2. Different approaches to managing the land cover database

Depending on the quality of existing national land cover maps, resource availability and country capacity, countries may choose different strategies to build a baseline land cover data. The chosen strategy must optimize the use of available resources while meeting the required criteria for the data outputs.

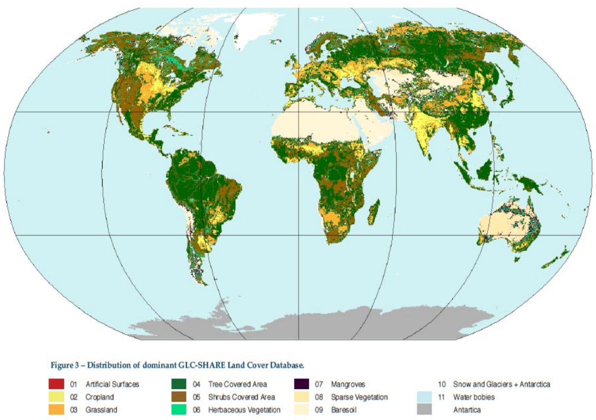
The three main options are the following:

Option 1: Using the existing land cover data

This choice sustains the *status quo* (Business As Usual - BAU) with regard to the creation of a new land cover database. The project does not include the use of resources and time for mapping production or improvement. It continues with the current arrangements and adopts the existing material, even if it is available at the global level. Figures 7 and 8 above present the most accurate global databases available free of charge, and schematically set out their technical details. Nevertheless, the quality and the accuracy of the other sources available at national and subnational levels should still be evaluated.

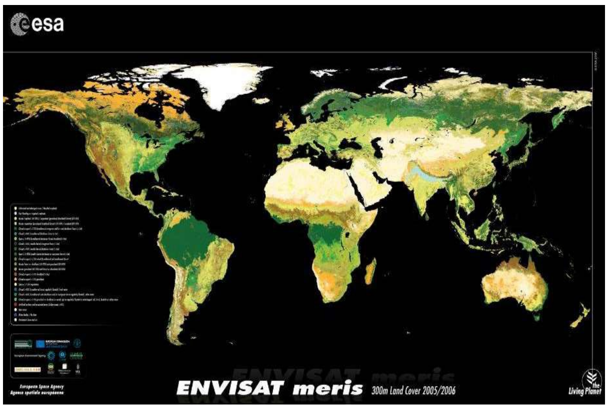
Figures 12 and 13 briefly illustrate the technical characteristics of the global land cover databases.

Figure 12*:* FAO GLC-Share global land cover database (2014).



Producer: FAO
Date of production: 2014
Coverage: Global
Projection: Geographic decimal degree
Resolution: 30 arc sec.
Legend adopted: Land cover legend proposed by the System of Environmental-Economic Accounting (SEEA) – 11 classes
Legend style: LCCS3
Source: Various, at different resolutions assembled according with data fusion criteria (see [29])

Figure 13: ESA GLOBCOVER global land cover database (2006-2009).



Producer: European Space Agency (ESA)
Date of production: 2009
Coverage: Global
Projection: Plate-Carrée with a Geographic Lat/Lon representation (GCTP_GEO)
Resolution: 300m pixel size
Legend adopted at global scale: 22 classes
Legend style: LCCS2
Source: MERIS (remote sensed-based image processing)

Several other products may be available at regional and national levels. The accuracy of the available data sets and maps should be evaluated with a view to ascertaining their agri-statistical functionality. Chapter 4 sets out a number of recommendations in this respect.

Option 2: Improving available land cover data

Improving the existing land cover data is a cost-effective option. These actions may seek to improve and standardize the legends – e.g. by translation to a standard format – although the varying class definitions between different surveys may still affect the outcome. The methods and software to be used in these translation efforts are discussed in Chapters 5 and 6. Possible scenarios to improve the baseline information are:

- Translating the existing land cover classes of the database or map to the standard LCCS3 structures;
- Adopting automatic image processing (pixel-counting) programs to improve cartographic information (this may be faster, albeit less accurate);
- Producing simplified land cover maps using free-of-charge global imagery (e.g. Google Earth) and various mapping tools.

To improve and correct the database while keeping costs under control, a viable option would be to use high-resolution free-of-charge RS products, such as the images and derived products at 10-m resolution made available by ESA (Sentinel-2).

The products' effectiveness for improving existing data may require validation (see Chapter 5) for accuracy and currency; indeed, some data may have been derived from older image resources and may thus not support long-term use within an MSF.

Option 3: Developing a new baseline

If the capacity and resources are available, the best option may be to develop a new classification. Indeed, this solution would provide up-to-date products (subject to the source data) and a classification wholly tailored to the MSF's requirements, thereby prolonging the MSF's viability. It would also provide coincident imagery and a classification that enables optimal use and support to field enumeration. Furthermore, it would enable an effective validation and accuracy assessment that the translation of past land cover cannot achieve.

Developing a new land cover database through a georeferenced digital baseline enhances the capabilities of data analysis and integration with RS data and other geospatial information, to facilitate management of the following statistical activities:

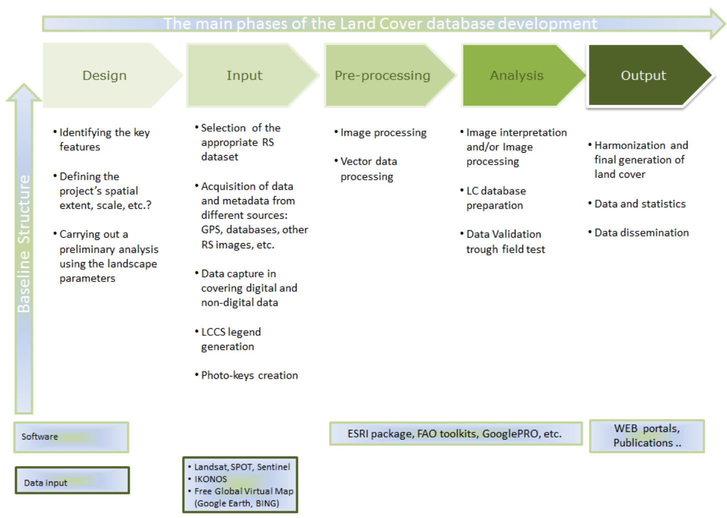
- Integration of information relevant to agricultural statistics, such as field size, crop types, multiple crop production and rotation practices;
- Reduction of the sampling variance within each stratum through photo-interpretation or automatic classification of imagery into land cover classes, using a standardized classification schema to systematically store the data;
- Supporting AF construction by improving the definition of the physical boundaries of the Primary and Secondary Statistical Units – respectively, PSUs and SSUs – (by means of photo-interpretation of satellite imagery) if an area sampling frame with physical boundaries is adopted, as well as a distinction between agricultural and non-agricultural land (with accurate stratification), to avoid allocating sample units where the probability of finding a crop is close to zero; this would considerably increase the stratification's relative efficiency at almost zero cost.

Moreover, developing a new land cover database provides an opportunity to obtain a more comprehensive country snapshot on food security, rural poverty reduction, climate change, water scarcity, ecosystem degradation and biodiversity loss.

The possibility of generating a new land cover classification should be considered at an early stage in the development of an MSF project. The opportunity to integrate the development of the land cover database with other initiatives (e.g. forest assessments) should be taken. FAO provides open-source tools to support these efforts (see Chapter 4 and Annex IV).

Figure 14 synthesizes the main steps in generating a new land cover database:

Figure 14 – Developing a land cover database.



The five main phases of developing a land cover database are:

1. The Design phase, which describes the project's goal, the database's features, and the reference features of the base map and the coordinate system. These features may be identified by means of various analyses, such as those occurring during image selection, or through landscape analysis:

    - Preliminary analysis of the project parameters and the area of interest;
    - Definition of the interpretation strategy, in accordance with the production time and costs;
    - Landscape analysis to contextualize the agricultural sector within the study area (average field size, prevailing farming system, typical agricultural practices, etc.);
    - Selection of the software for image processing, interpretation and GIS analysis;
    - Involvement of national counterparts and other parties, and development of ad hoc training courses and workshops to determine methodology and processes.

2. The Input stage includes all phases relating to the collection of the necessary and ancillary data sets. Data analysis is essential to create the

preliminary legend and to define the photo-keys. The main data sources are the following:

- *Data selection and acquisition.* Appropriate remotely sensed satellite imagery must be chosen and complemented with ancillary resources. Currently, several different sets of satellite imagery are available from international data suppliers, for a wide range of ground resolutions, spectral characteristics and costs (see Annex VI). In addition, the question of what would be the best window for image acquisition (assuming that the analysis is undertaken using single- rather than multi-date imagery) should also be considered;
- *Legend generation*. The legend is the key to the database. It requires a preliminary detailed study of land cover classes in the country, to ensure the database's correct development and accurate communication of the information held therein to data users;
- *Creation of the photo-key table, in the preliminary phase of the interpretation.* The photo-keys orient and guide the interpretation process. Their characteristics are identified with reference to the RS image, visits on the ground (with photos, if available), Google Earth, and the land cover legend. This ensures the common visualization and comprehension of the land cover classes, especially important when a large group of experts are involved.

3. Pre-processing prepares the images for analysis, as follows:

    - Use of ad hoc tools and techniques for the images;
    - Produce a vector file output through segmentation. This object-based segmentation, performed using products such as eCognition, generates a vector-polygon file for image interpretation. Object-based approaches can integrate additional spatial data sets that characterize elements of the landscape, to assist the image segmentation.

4. The Analysis phase interprets the images and produces geospatial information and/or extracts features of interest. Image classification may be partially automated; however, these procedures use a

combination of automated and manual interpretation. The design phase establishes the interpretation strategy.

- Image analysis with two main options
- Automatic classification: Fast, low-cost, with varying levels of accuracy that depend on the validation points that can be selected;
- Visual interpretation: this requires a medium-to-high level of commitment in terms of cost, time and human resources. However, the result is certainly more reliable and has a greater probability of longevity.

The preliminary interpretation requires adequate knowledge of the terrain.
The preliminary photointerpretation must be assessed in terms of its consistency and quality, through an independent process. The accuracy of the land cover database, through ground data collection on an independent field sample basis should be assessed.
A validation campaign in the field is necessary to confirm the preliminary interpretation especially in areas of uncertainty. If it is not possible to collect data on the ground, a possible alternative would be to use recent aerial photographs and/or very high-resolution satellite imagery (also free available in Google Earth) as Pseudo ground truth point source.

5. In the Output phase, the main processes are:

   - Harmonization of a final generation of the land cover database, to ensure topological consistency;
   - Provision of indicators and statistics from the summarization of the data sets classified;
   - Provision, where possible and with the agreement of the data owners, of full access to the land cover data sets, by generating dedicated Internet portals or their integration into existing data portals.

FAO has implemented the latter option. Chapter 4 describes the main steps of the implementation, while Part II describes the test performed in the pilot countries.

4

The proposal for a new land cover database

The method adopted by FAO to construct a land cover database begins with a multifaceted mapping approach, derived from several years of experience with developing detailed national land cover databases. This process is dynamic, and has evolved in several theoretical and technical aspects involving GIS and RS technology.

FAO's approach has been comprehensively tested and validated, as it has been applied at country and regional levels in many areas of the world, to different land cover types and in collaboration with local experts and counterparts. In addition, it was developed with specific and standardized mapping procedures, to obtain a reliable national land cover vector database that could be useful to achieve a range of purposes and meet the needs of several different types of users.

The process is the following:

1) DESIGN

a. Selection of the appropriate software and tools designed and developed by FAO for mapping; these are freely available and can be downloaded from the FAO-GLCN portal (http://www.glcn.org/index_en.jsp; see Annex V for further details). Updated image processing techniques should be applied, using ad hoc tools and techniques.
b. Training and capacity development activities should be carried out on site. The involvement of national staff is strongly supported in all phases: collection of ancillary information, interpretation, validation in the field, as well as data maintenance and distribution.

2) INPUT

a. Selection and acquisition of appropriate satellite image products using crop calendars and the Leaf Area Index (LAI), which are based on image processing products and vegetation indices.
b. Legend generation. The FAO-designed LCCS (LCCS2 or LCCS 3) should be applied. The LCCS is an international standard and comprehensive classification system that is capable of meeting the needs of all potential users. Indeed, it is able to describe all types of coverage, and possesses sufficient flexibility to enable mapping at different scales and levels, thereby improving the land cover description; in addition, the FAO LCCS permits cross-referencing among the various components.
c. Preparation of photo-keys on the basis of the legend, to support photo-interpretation.

3) PRE-PROCESSING

The specific image pre-processing procedures implemented depend on the nature of the images acquired. They may include processes such as georectification and mosaicking. Image segmentation uses object-based processing to generate vector boundaries of consistent parcels. This process may combine supplementary data sets to improve parcel definition. Pre-processing may also be iterative, to develop a vector-based output at the appropriate resolution that can support interpretation; this would be conform the classification legend proposed at the Input stage.

4) ANALYSIS

The interpretation relies upon "object-based" segmentation data layers, which may apply a mix of visual and semi-automatic methods to enhance the possibility to discriminate between land cover features.

The accuracy of the land cover database should be assessed on a sample basis, through ground data collection or the photointerpretation of very high-resolution EO data.

5) OUTPUT

a. Preparation of a variety of outputs, derived products and statistics. These should include a standardized metadata entry that describes the land cover database.
b. Dissemination of data. The FAO Atlas, publications and specific webpages are prepared. The data are also disseminated using GeoNetwork, FAO's data repository for standardized metadata.

Each phase includes a number of specific steps, which are detailed in Part II of this study.

5

Evaluation of open-access remotely sensed image products

5.1. Introduction

Option 2 ("Improving available land cover data") is a viable possibility for managing the land cover data set, through the use of accessible, free-of-charge remotely sensed image archives. The specific roles that such data can play depend on their resolution and currency, and on the license terms of the source data.

The most commonly utilized high-resolution products for land monitoring that are accessible at no charge are:

- Landsat coverage (with WELD 30-m composited mosaic); and
- ESA's Sentinel-2, the new generation of EO satellite images to be made available by ESA and the European Commission (see https://sentinel.esa.int/web/sentinel/missions).

The main characteristics of these products are listed in Annex V to this document, which also provides a brief overview of the satellite-based sensors available on the market, particularly as to their spatial and temporal resolution and costs[11].

In addition, it must be noted that the interactive Virtual Global Maps, such as Google Earth, Marble and Bing are the most commonly used global maps. Google also provides Google Earth Engine (GEE), a cloud-based geospatial platform for environmental data analysis. GEE offers large amounts of historical satellite data free of charge, particularly of the Landsat and MODIS satellite series, as well as a web-based interface for analysing them.

[11] A complete list of the available products may be found at http://earthexplorer.usgs.gov/ and http://glovis.usgs.gov/AboutBrowse.shtml.

This analysis focuses on Google Earth, as it is considered a valuable source of information and a powerful tool for mapping, given its freeware connection with the main GIS software (e.g. the ArcGIS/FAO toolbox).

5.2. Google earth products

Google Earth provides access and source catalogues of free images that can be of great support to the geospatial community. Useful information may be extrapolated from the satellite images stored in the global database, which may also be used for mapping exercises.

Google Earth is free to download and install, and provides viewers with a realistic full view of the Earth. Its internal coordinate system is latitude/longitude on the World Geodetic System of 1984 (WGS84) datum. The Projection is in World Mercator.

Google Earth is a software tool, a standard for exchanging spatial data (via KML, see footnote 13[12]), a catalogue of archived satellite and aerial imagery, and a community of people interested in developing new applications and services. Google Earth has limited processing and analysis functions compared to those commonly found in GISs. However, its core functionality endows it with high accessibility and its 3-D viewing capabilities make it highly immersive.

Google Earth hosts high-resolution imagery that spans over than twenty per cent of the Earth's land surface. This high-resolution archive is rapidly expanding, and has become an important resource for scientific research. However, the baseline resolution is difficult to assess because different product sets are used around the world. Globally, this resolution is generally of approximately 10-15 m, but varies subject to the quality of the satellite or aerial photographs uploaded. Typical high-resolution images are found in Europe and USA. In Africa and Asia, several Digital Globe images have already been uploaded; these, however, do not fully cover the entire world. The data currency also varies widely from tile to tile, and the online platform does not

[12] KML (Keyhole Markup Language) is an XML-based file format used to display geographic data in an Earth browser such as Google Earth, Google Maps, and Google Maps for mobile. KML enables display of virtually every feature on a map.

contain precise metadata for every tile. This makes it difficult to match images to the mapping seasons selected.

Several other sources of images, tools for mapping and for obtaining small tile map images from Google Maps are available online for free or at a reasonable cost (Google Maps, Google Maps Downloader[13], Google Engine for the Landsat data archive, Google Earth Pro, etc.). FAO has developed the MADCAT software, which exploits the imagery data coverage stored in Google Earth. The mapping activity is carried out online, and the result is a vector or raster database classified according to an LCCS legend, if available; an example will be described in this chapter.

There are open source alternatives to Google Earth API (e.g. Cesium[14]). These may be used free of charge, even for commercial purposes; however, none of them have the imagery resolution and aerial photo coverage offered by Google Earth and Google Maps.

5.3. Potential and limitations of Google Earth's high-resolution images

Data distribution and availability are central to several different applications. The shortage of quality information on land and its spatial dynamics has been largely attributed to the cost, effort and expertise required to process and to manipulate these data over large areas and for different periods, and to make these data available by means of suitable portals. Innovative platforms such as Google Earth provide new potential for large data ingestion and manipulation, and visualization of multi-temporal imagery.

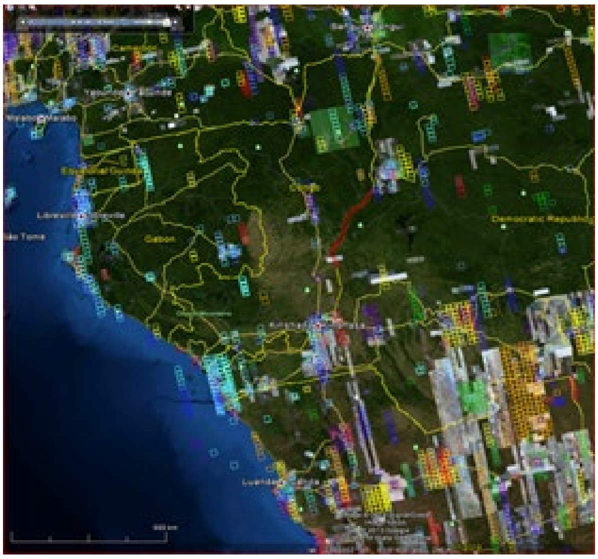
- **Timeliness and currency.** Google acquires imagery from a variety of providers. Some images are provided by local or state governments. The age of the imagery varies greatly, but most of the high-resolution imagery is between six months and five years of age. Due to the varied origin of the imagery, the process to insert it into Google Earth is complex and time- and effort-intensive. Once the imagery is captured, it is processed by commercial providers and then made available to customers, one of which is Google. Google then determines whether the new imagery is better than that currently in its possession. Once an image is selected, it must be processed in accordance with the format

[13] http://google-maps-downloader.software.informer.com/6.6/.
[14] https://cesiumjs.org/index.html.

and coordinate system applied within Google Earth's databases. Then, it must be subjected to a quality control process and fed into a processing system before being distributed to Google Earth's live database servers. The lengthiness of this process is one reason why Google Earth generally does not contain imagery less than six months old, and why updates only occur once every 60 days. Not all of the imagery in Google Earth originates from satellites; a large proportion is taken by aerial photographers, mostly from airplanes using special high-resolution cameras.

- Time consistency. The dates of the coverage imagery vary considerably and over short spatial extents. The dates of the copyright information do not necessarily provide specific dates of the image tiles. This uncertainty may make the tiles harder to use for validation purposes. A rapid visual assessment may be made using the "Layers" option, selecting "More" and thus displaying the index of the images (SPOT and Digital Globes by year). The high-resolution images available per year may be highlighted using the "Chronological Bar". An example is shown in Figure 15: the frames of the data sets available are superimposed over the images.

Figure 15 - Google Earth analysis of available images



Google Earth uses the DEM data collected by NASA's Shuttle Radar Topography Mission (SRTM). Thus, the entire Earth may be viewed in three dimensions. Since November 2006, the 3-D views of several mountain ranges have been improved by the use of supplementary DEM data to fill the gaps in SRTM coverage.

Considering the enormous potential of Google Earth, it can be an important source for several purposes of agricultural statistics, especially if the financial and technical resources available are scarce. However, its limitations concern the limited specification of the image periods and image resolution. In addition, license constraints limit the subsequent use of data derived from GE.

- **Accuracy**. It is difficult to ascertain the accuracy and resolution of the data, as this information is not published through the Google Earth portal. The values of horizontal and vertical accuracy vary from place to place and may change depending upon the time period in question. Most land areas are covered by satellite imagery having a resolution of approximately 15 m per pixel. This base imagery is derived from a 30-m multispectral Landsat, pan-sharpened with the 15-m (panchromatic) Landsat imagery. However, Google actively replaces this base imagery with 2.5-m SPOT-Image imagery and several higher-resolution data sets, mentioned above. Some highly populated areas are also covered by aircraft imagery (orthophotography) at sub-metre resolutions. Figures 16a and 16b provide examples of images at different resolutions and indicate the feasibility of discriminating different features at different resolutions.

Figure 16 (a and b) – Comparing Google image resolutions.

Potere [22] assessed image accuracy (georegistration) by comparing Google Earth with Landsat GeoCover scenes over a global sample of 436 control points located in 109 cities around the world. Landsat GeoCover is an ortho-rectified product having a positional accuracy lower than 50 m, as estimated by the root mean square error (RMSE), a measurement of the distance between recorded points and their true ground location. The positional accuracy was estimated to be 39.7 m RMSE. In this regard, it is important to note that the magnitude of errors ranges from 0.4 to 171.6 m. In 2012, a study [23] compared point coordinates measured by Google Earth with GPS receiver coordinates over 16 checkpoints located in Khartoum (Sudan). The RMSE computed for the horizontal coordinates was of 1.59 m. In October 2012, a slight improvement was achieved compared to the results calculated the previous month. A study conducted by a Japanese group of experts [24] repeated the same exercise in 2013 for ten cities in different region of the world. The study compared Google Earth, Bing Maps and Open Street Map with imagery from the Advanced Land Observation Satellite (ALOS) and Panchromatic Remote-sensing Instrument for Stereo Mapping (PRISM), illustrating that the values yielded by these systems were, respectively, 8.2 m, 7.9 m and 11.1 m.

Figure 17 – High-resolution image from Google Earth.

- **Vertical accuracy.** The precision of elevation data for flat terrain tends to be high. In mountainous terrain, the elevation data is less accurate because the points are not measured directly but are, rather, obtained by an interpolation procedure. The values are extrapolated from the SRTM, which implies that the extrapolated values are not related to the real *terrain* elevation but to the *surface* elevation – in other words, the point elevation is not measured on the terrain (ground) but on a higher, surface point (e.g. vegetation, buildings or other structures, which may be natural or non-natural).

Google Earth is an important source of information. However, the availability of data and its positional accuracy must be accurately evaluated for each area of interest. The ultimate reliability of Google Earth's accuracy for the scope of the project must be assessed in each individual case, as well as the degree to which it is capable of supporting the consistent collection of field points within the given area for validation activities.

5.4. Other sources

The tables in Annex VI briefly illustrate the other RS products available on the market, noting their different costs and resolutions. The official data providers for FAO projects under Long-Term Agreements (LTAs) are:

- MDA Information Systems Inc. (Canada) and
- SPOT Image – AIRBUS-Astrium (France)

These providers offer support for research and image selection in specific areas of study, processing and preparation of specific outputs, and suggestions for technical solutions. Special discounts are available, and the Agreement stipulated with the United Nations includes the grant of a multi-license (up to five users).

Sentinel imagery became available from ESA in 2015, together with derived data products (see Tables 3a and 3b). This data is particularly relevant to land cover and agricultural applications (crop monitoring, natural resource disaster mapping, etc.) as it provides medium-resolution data on a frequent return interval and free of charge. Sentinel 2 provides optical images at a 10-m resolution, with an acquisition revisit time of five days.

Another supplier of satellite data is SKYbox Imaging (a Google subsidiary), which provides not only satellite coverage from commercial EO satellites, but more generally designs and builds solutions for any need relating to geospatial analysis and EO monitoring and analysis[15].

[15] http://www.skyboximaging.com/.

5.5. Conclusion

Google Earth offers a vast catalogue of images, free of charge. It contains several functionalities that may be valuable in many applications. It provides:

- A synoptic view of the study area, enabling visualization and identification of its main features (location of agricultural areas, types of fields, road networks, etc.);
- A detailed view of the study area using the high-resolution images available;
- Using geographical information programs, it may be useful in the preparation of simplified land maps and/or improvement of existing data sets.

However, there are limitations upon the use of Google Earth data: licenses must be obtained to use the information outside Google platforms.

Google Earth Pro and Google Earth Enterprise Client became free tools in 2015; however, restrictions remain upon the use that can be made of derived data. While the imagery can be used effectively for validation purposes, often replacing or together with ground surveys, there are restrictions upon the direct acquisition of point-based or vector data for use in external packages. In addition, the currency of data is often not clear to Google Earth users and may not match the ground survey dates; in areas where significant changes may have occurred to land cover and land use, this may pose a severe limitation. Google Earth's compilation of the best imagery available also means that some mosaics may not be contemporary, or may not have been optimized in relation to seasonal crop calendars. This may limit its use in practical agricultural applications and in stratification.

Using Google Earth (which is free of charge) or other virtual global maps (with various conditions of usage) with dedicated software (e.g. ARCMap or FAO's MADCAT), it is possible to exploit the vast imagery data sets available online and directly prepare or improve a more elaborate georeferenced land cover database.

6

Quality framework

6.1. Introduction

Generating land cover geospatial data requires a robust sampling strategy for validation and accuracy assessment, as well as the establishment of quality measures, which should be reported alongside the data and documentation through metadata and data set descriptions. Accuracy assessments and data quality evaluations may be considered against a series of parameters, and are important considerations within an MSF as they affect its usability, suitability, data volumes and costs.

6.2. The parameters to define data quality

Quality can be defined as the fitness for use (or usability) within the context of a specific data set. Indeed, data suitable for one application may not be fit for use for another. The overall quality of the output data may also depend on the range quality parameters of the primary and supplementary data used in its generation.

The United States' Spatial Data Transfer Standard (SDTS, 1997) has proposed a series of instructions for defining and documenting the quality of the GIS database, as a basis for the data quality report that can be drafted within the standards-compliant metadata and data description. In particular, five components have been identified:

1. Lineage
2. Positional accuracy
3. Attribute accuracy
4. Logical consistency
5. Completeness

1. Lineage

The lineage of data is metadata concerned with historical and compilation aspects, such as:

- Data source;
- Data content;
- Data capture specifications (including data currency);
- Geographic coverage of the data;
- Compilation method of the data, e.g. digitized or scanned;
- Transformation/pre-processing methods applied to the data; and
- Use of algorithms during compilation (e.g. resampling, linear simplification, feature generalization).

For satellite-based data, further parameters may be cloud cover, geometric corrections, and atmospheric corrections relating to the image used in processing, as well as issues concerning the instrument calibration of the radiometric values.

2. Positional Accuracy

Positional accuracy is the degree to which information on a map or in a digital database matches true or accepted location values within horizontal and vertical planes. These indicators may consist of scale, resolution and precision attributes (see Annex XII). The scale represents the ratio between distance on the image and that on the ground; the resolution is the ability to discriminate features at a particular scale (resolution in terms of pixel size); precision is the measurements' degree of exactness. When developing an MSF, vertical accuracy is less important. As explained above, positional accuracy is usually measured by the RSME, a measurement of how far recorded points are from their true ground location.

Two basic considerations should be made upon the accuracy of maps or databases:

- The level of accuracy and precision required for particular applications varies greatly; and
- Highly precise data can be very difficult and costly to produce and/or collect.

High precision does not necessarily indicate a high accuracy; nor does a high accuracy imply high precision. However, both high accuracy and high precision are expensive to achieve.

3. Attribute accuracy

Attribute consistency (thematic or semantic consistency) relates to whether a feature on the ground is correctly classified in the digital data. Attribute accuracy considers the quality of the interpretation provided of the data classes – the degree of correspondence between the data and the real world, recognizing the classification applied.

The ability to achieve this correspondence also depends on the development of a realistic legend and class structure. The development of the legend and the photo-keys helps to ensure that the features are correctly interpreted; in addition, the field validation can resolve other problems due to vagueness or ambiguity. . The ability to achieve the correspondence may also be affected by the size of the polygons determined to attain the classification, and the minimum mappable units selected to reduce the mixed classes observed on the ground and in the image interpretation.

In RS studies, this is typically described by a matrix of omission and commission (confusion matrix) within the classification interpretation (for manual or automated classification systems). Such matrices are typically generated against the remote classification and ground survey validation samples. Validation may also be undertaken against classifications made from higher-resolution imagery or aerial surveys (e.g. Google Earth), rather than all through field-based methods; this yields corresponding cost savings.

4. Logical Consistency

Logical consistency may be non-spatial or spatial. This component concerns the authenticity and coherence of the data structure of a data set. Non-spatial consistency may arise from the different dates of the imagery, the different seasons and timescale of image acquisition (i.e. different lineage parameters) or different pre-processing and processing procedures.

Spatial consistency describes the quality of the data representation, such as incorrect line intersections, duplicate lines or boundaries, or gaps in data sets; these may be referred to as spatial or topological errors. Topological errors should be limited by the adoption of the object-based segmentation and automated tools for topological validation.

5. Completeness

The final component of quality concerns the data set's completeness. This requires consideration of any holes in the data, unclassified areas, and any compilation procedures that are intrinsic to the database and that are difficult to remove.

Additional parameters may be useful to the quality reports, such as the accuracy and completeness of the metadata.

6.3. Evaluating the accuracy of land maps and databases

Validating the land cover classifications requires a number of stages:

- Development of a sampling scheme that provides a representative sample of points within each class and their spatial extents;
- Collect the reference data (either through field surveys at sampling locations or through use of high-resolution images); it must be possible to accurately identify the classes against the definitions for these reference data sets, as they are intended to represent the "ground truth";
- Compare the reference data to the classified map;
- Compute accuracy metrics for individual classes and overall accuracy assessment metrics.

Evaluating the accuracy of a map or database requires a certain degree of quantification or representation of the parameters measuring quality, and their evaluation against the data set's purpose.

Lineage.

All the information on lineage should be always linked to the data set. Indeed, it is a component of the ISO Standard 19115 Metadata – Geographic Information and reported in the section titled "Data Quality".

Geometric or positional accuracy.

To evaluate the accuracy of a map, the positions of points whose locations are shown in the digital data are compared with the corresponding positions on the ground as determined by surveys of a higher accurate source. Basically, there are two types of data source that can be used in support of the accuracy assessments: other SR data (independent sources with higher accuracy, e.g. high-resolution satellite images or aerial photographs) and ground-based data.

As explained above, geometric accuracy can be assessed using the RMSE, which is the average of the distances (also known as the *residuals*) between each pair of data and ground control points. The RMSE is the square root of the average of the set of squared differences between the data set coordinate values and those from an independent source of higher accuracy, for identical points. The relevant references are fully documented in [20]. Whether an RMSE value is acceptably low depends on the nature of the project and the scale of analysis [18].

Semantic and content accuracy. Qualitative or thematic accuracy refers to the correct labelling to be assigned to the specific features or polygons. For example, a pine forest may have been incorrectly labelled as a spruce forest, thereby introducing an error that may not be known or noticeable to the map or data user. The level of error may depend upon the nature of the misclassifications, for example, if a pine forest has been classified as agricultural land (this is referred to as fuzzy accuracy, in which the classes are not merely right or wrong). Certain features may have been omitted from the map or spatial database due to oversight or by design (e.g. the spatial resolution and minimum mapping units). The inaccuracy of the thematic content affects the quality and precision of the final applications, because it provides incorrect information.

Qualitative data (e.g. concerning land use/land cover) is usually assessed using a cross-tabulation of encoded and “actual” classes at sample locations. This produces a classification error matrix (contingency matrix or confusion matrix) that compares, on a class-by-class basis, the relationship between known reference data (ground truth) and the corresponding results of the interpretation of the land cover classification. It is calculated by summarizing the misclassification rates for each land cover class.

These classification errors of omission and commission can be calculated by assessing the classified image against a reference. Typically, this reference is a set of ground control points where the classification is determined, or where higher-resolution imagery is used to classify “control points” which are then compared to image classification values.

For each land cover class, it is possible to calculate measures of user and producer accuracy, as well as omission and commission errors, expressed as percentage values. A series of summary measures are available to assess overall classification accuracy, such as Overall Accuracy and the Kappa Index. Errors of commission occur when classes on the classified image are found to be of a different type from those on the ground; conversely, omission errors occur when the class on the ground is not of the same type as the classified image. Producer accuracy focuses on errors of omission and is given by the number of correct elements (points or polygons) in a category over the total number of elements in that class, on the basis of reference data. User accuracy focuses on commission errors (seeking to ascertain the number of correct elements in a category over the total number of elements classified in that same category). The overall accuracy is a measure of the total correct elements compared to the total sampled elements. It is calculated by dividing the total number of correctly classified elements (sum of elements along the major diagonal) by the total number of elements:

$$\mathbf{TA = \Sigma\ correct\ X / \Sigma X * 100\%.}$$

A useful reference for calculating the error matrix is available in [39].

The Kappa Index is a measure of the difference between the observed agreement between two maps and that allows for statistical comparison. The Kappa Index is useful for comparisons between different classification systems.

The FAO DDN is a geospatial unit that uses a point-based approach to assess the accuracy of the land cover database, creating a grid point shape file in which the distance between points is congruent with the scale and resolution of the land cover map. The grid is overlaid upon an up-to-date high-resolution image data set (Google Earth images, IKONOS satellite imagery, aerial photographs, etc.) and coded with LCCS legend classes. This is the TRUTH classification. The TRUTH points are then intersected or linked with the existing (INTERPRETED) land cover map, thereby creating a matrix. An evaluation point is considered to be correctly classified only if the dominant vegetation type assigned on the existing (INTERPRETED) map is wholly identical to the observed value on the TRUTH map.

6.4. Logical consistency and completeness

Data producers report on components in the various phases of the production chain, in terms of both spatial and non-spatial parameters (e.g. measures of topological correctness or the consistency of data sources across the mapping programme). Guidelines to analyse the accuracy of the existing maps or database may be tested using an available map.

This test was conducted with an existing land cover map of the Mekong river basins available from FAO's Library. It seeks to simulate a typical situation in which a hardcopy map is found and must be adopted as the baseline.

The test attempts to illustrate the operational steps required to assess the value of the reference information available.

6.4.1. Qualitative assessment

The information reported in Table 3 provides a qualitative evaluation of the map.

Table 3 - Qualitative assessment of data quality

Query	Example of Information
Title of the map	Land cover map of Cambodia
What is the age of the data?	1992-1993
Where did it come from?	Mekong Secretariat, Land Use Mapping, Office of Ministry of Agriculture of Cambodia
In what medium was it originally produced?	Map printed on paper to be used directly by the human user
What is the areal coverage of the data?	Country Map 
To what map scale was the data digitized?	250,00
What projection, coordinate system, and datum were used in maps?	WGS84
What was the density of observations used for its compilation?	Unknown
How accurate are positional and attribute features?	A displacement of an xxx circa was calculated (RMSE test)
Does the data seem logical and consistent?	Yes but not many details are provided
Do cartographic representations look "clean ?"	Clean but not accurate
Is the data relevant to the project at hand?	Yes the map provides agriculture features of the country
The classes of interest are included and correctly represented?	The class of interest are included but not very well detailed
In what format is the data kept?	Paper – no digital format
How was the data checked?	
Why was the data compiled?	
What is the reliability of the provider?	

7

Case Studies

Table 4 provides a schematic illustration of two case studies in which land cover mapping has been implemented for the specific purpose of supporting agricultural statistical analysis and stratification.

The mapping activities raise the following important issues:

- the approach must be considered as a medium- or long-term activity;
- the product must be considered as a multipurpose database;
- the involvement of local institutions is crucial to develop appropriate legends and for validation activities.

Table 4: Project components of the case studies analysed.

	Ethiopia	**Pakistan**
Technical support agency	FAO	FAO
Implementing agency	CSA Central Statistical Authority EMA Ethiopia Mapping Agency	SUPARCO National Space Agency of Pakistan
Data used	SPOT 5 5m resolution 2085 scenes Dated 2006 - 2007	SPOT 5 5m resolution 310 scenes Dated 2012
Team	12-15 persons with no background in RS	10 persons With background in RS
Methods	Segmentation Object-based	Segmentation Object - based
LC Classification System Adopted	FAO – LCCS2	FAO – LCCS3
Mapping Period	From 2008 up to date intermittent	From 2012 up to date continuous
Training	Initial course of 2 weeks	Initial Course of 2 weeks
Supervision	2 land cover experts constantly present for the first years	After the first 6 months 2 weeks land cover expert supervision
To date % of country mapped	50% circa Where the main agricultural activities of the country are located	30% circa Where the main agricultural activities of the country are located
Completion of the mapping activities	2016	2014

The case studies have been extracted from FAO's final mapping reports.

7.1. Case study 1 – Ethiopia land cover mapping

For the past six years, Ethiopia has made considerable efforts to use satellite imagery to construct a land cover database in the LCCS style, with the main objective of supporting AF construction. The government institutions involved in the project are Ethiopia's Central Statistical Agency (CSA) and the Ethiopia Mapping Agency (EMA), in direct collaboration with FAO's Land and Water Department (NRL) [32].

The satellite coverage utilized was SPOT 4-5 m resolution of 2005, as this was available in the country; a team of 12-15 experts was involved in the interpretation exercise and in database generation and processing. The work was initially carried out with the support of international experts; at a later stage, the team was able to continue the mapping activity without assistance. Initially, the land cover database was generated as an extra effort to create a multipurpose baseline to support agricultural statistics; it subsequently served to provide the national institutions with a reliable, up-to-date database.

A brief history of agricultural statistical programs in Ethiopia illustrates how the national institutions have continually improved the methods, on the basis of technological advancements. Four distinct periods of agricultural statistical development may be distinguished, as follows:

- Prior to 1974: Ad-hoc surveys
- 1974 – 1979: Annual Agricultural Sample Surveys
- 1980 – 1992: National Integrated Household Survey Program
- 1992 – 2008: Rural Integrated Household Survey Program
- 2008 to date: in a part of the country, AF sampling with advanced technologies

Prior to 1974: Agricultural statistics were just beginning to be incorporated into programs. Statistical services in general, and agricultural statistics in particular, were inadequate in terms of data coverage, timeliness and reliability. FAO provided technical assistance to both the Ministry of Agriculture and Rural Development (MoARD) and the CSA to develop the methodology.

Between 1974 to 1979: The MoARD continued to receive assistance from FAO. Six Annual Agricultural Sample Surveys were conducted, including a small-scale agricultural sample census in 1976/77. Regional and field supervisors of the MoARD and temporary enumerators undertook fieldwork during each survey. The CSA cooperated closely with the MoARD, supplying supervisors, vehicles and equipment. MoARD staff members manually processed the collected data.

Between 1980 to 1992: Before 1980, the socioeconomic and demographic data available in Ethiopia were seriously deficient and obsolete. There were no national statistical programs that could ensure a continuous flow of the socioeconomic and demographic data required for planning, monitoring and evaluating ongoing development programs. As a result, a National Integrated Household Survey Program (NIHSP) was established and ran from 1980 to 1992. The NIHSP enabled the CSA to run a number of annual national socioeconomic and demographic surveys. The CSA's available infrastructure, field staff (enumerators, supervisors, and drivers), logistical support, and data processing facilities were fully utilized. Since the CSA possessed the necessary resources and had been entrusted to produce these important agricultural statistics, the CSA carried out the Integrated System of Food and Agricultural Statistics Program and then implemented 13 Annual Agricultural Sample Surveys between 1980 and 1992. In addition, a Rural Integrated Household Survey Program (RIHSP) was established in 1980 as a component of the NIHSP.

1992 to date: The latest developments in the Ethiopian agricultural statistical program include the migration from the traditional HH survey (list frame) to the AF. The CSA has implemented pilot studies in AF sampling for estimating crop areas in the Oromiya Region, and has constructed AFs in three additional regions including Tigray, Amhara and the SNNP. The EAs constructed for the population census were overlaid with the LCCS database, and stratified on the basis of the EAs' percentage of cropland derived from the LCCS land cover database. The satellite imagery provides a stable foundation and enables certain controls over field data collection that traditional HH lists systems fail to provide.

With the preliminary results of this test, the CSA concluded that the estimates can be improved by using the LCCS database in selecting the stratification sample, and in regression and calibration for estimation, although some constraints were noted:

1. Deep stratification may present the following advantages: (i) the total number of segments in the population N is reduced; and (ii) the segments are more homogeneous with respect to the crops targeted for deep stratification. The frame was produced in the 2013 survey and led to excellent results. The stratification could be refined, completely eliminating Stratum 5 (Cropland 0-15%) from the sampling frame. This choice introduced a downward bias; however, it improved the precision of estimates.

2. The efficiency of the stratification based on the LCCS database is not the same for all crops, because it is not crop-specific. Therefore, it does not enable optimization of the sample allocation for each item. In Ethiopia, the AF is used to estimate over 100 different land covers, with vastly different distributions and optimum sample allocations; indeed, the optimum allocation for teff is not optimal for sorghum and maize. The EAs constructed for a population census are overlaid onto the land cover database. The EAs include city, farm and non-farm parts. The only characteristic shared by all EAs is that they include between 150 and 300 HHs. Their sizes vary: in cities, they tend to be small, but in rural and suburban areas, they may be vast if housing density is low: indeed, they may even reach 4,000 hectares (ha), or 100 forty-ha segments. Each EA is assigned to a stratum on the basis of the percentage of cropland within it. The stratification was performed manually; the EAs were not constructed from a single type of stratum. This procedure was convenient; however, it was not efficient. Indeed, the stratification turned out to be a weak stratification; in any case, the use of the LCCS database for stratification made it possible to select a smaller sample [32]. Land cover data can be used more effectively in stratification and sample selection. When field data collection is properly managed, including crops and farm inputs, the AF is a more suitable estimator than the traditional HH list survey for all variables associated with the land. With AF design, the CSA has better control over the data collection component of the survey.

3. The LCCS properly identifies non-agricultural land. The LCCS can be used in regression analysis to calibrate estimates and improve individual crop estimates. As long as the LCCS data are correlated to the AF segment data, the estimates will be improved using that which has been identified as calibration. Calibration simply adjusts the sample for non-representativeness by using some auxiliary information relating to the variables of interest.

4. The CSA maximized the use of the LCCS by using data from the LCCS land cover in a regression program that calibrates the estimates. This reduced the variances (Coefficient of Variation, or CV) and improved accuracy (mean squared error (MSE). However, this experience raised many questions concerning the use of the detailed land cover database in AF stratification. The first question related to the extensions of the cropland. In Oromiya, the LCCS showed over nine million ha of cropland, while the most reliable source evaluated the cropland as being closer to six million ha. The accuracy of the LCCS classification was estimated at approximately 95 per cent.

5. Another issue arose directly from the interpreted database. In Ethiopia, forest areas are important for certain crops grown in the shade of trees (namely, coffee and other permanent shade crops). The land cover database does correctly identify forest areas, which can be extracted and placed in a separate stratum from coffee and other shade crops.

The interpretation of the imagery enables assessment of field sizes. In Ethiopia, the 40-ha segment was adopted. This size appears to be correct, considering that the optimum segment size is sufficiently large to include a great number of fields, and yet small enough for enumerators to manage in the field. It was also proposed to use 20-ha segments; however, the field supervisors in the Adama Agriculture Zone indicate that the 40-ha segments are manageable. For crops, 40 ha is slightly larger than necessary; for HHs, this figure does not contain more than 15 HHs on average. The team of consultants and CSA professionals had discussed changing the system to have three segments of 20 ha each. Since the socioeconomic data have not been analysed, this question remains open.

It was suggested that the open segment data collection for the AF include the information obtained through the interviews. In this connection, it should be noted that the HHs identified within the segment are interviewed and total farm-level data are collected.

The case study has illustrated some of the important implications of sampling and possible areas for refinement. It was observed that there is a large difference between the list frame (underestimation) and the AF (overestimation); these errors can be eliminated by better statistical programs that highlight any mistakes in data entry. Ethiopia will have to implement multiple frame (MF) sampling. MF sampling involves the use of two or more sampling frames which are sampled conjointly (i.e. area and list). Some large farmers are too important to be left out (and have therefore been included on the list frame) and several small farmers who cannot be listed due to their sheer quantity (and have thus been added to the AF).

7.2. Case study 2 – Pakistan

The land cover mapping of Pakistan was prepared by Pakistan's Space and Upper Atmosphere Research Commission (SUPARCO) in collaboration with FAO, using the LCCS, initially in the Sindh & Punjab provinces of Pakistan. The land cover mapping is gradually covering the entire country. The project was jointly formulated by SUPARCO and FAO, as part of the ongoing project titled "Building Provincial Capacity in Pakistan for Crop Forecasting and Estimation" [31].

The objective of the land cover database is to establish a robust statistical AF methodology and assist the development of an improved capacity for the monitoring and management of natural resources.

The database is categorized using 13 main land cover classes, which are further subdivided into 36 sub-classes. The land cover database of Punjab and Sindh Provinces includes several categories, containing rocky areas, natural vegetation, etc. According to the FAO-UN approach, land cover also includes those lands affected by human intervention, such as built-up areas, agriculture etc.

To ensure standardization across the country, the national legend was devised using the LCCS. As noted above, 36 classes have been created, considering the country's geophysical aspects. The structure of the database developed using the LCCS legend has been harmonized with FAO's GLCN. SPOT-5 5m resolution satellite images were segmented into homogeneous polygons and interpreted according to the FAO land cover mapping methodology to produce a seamless and detailed land cover database.

In Pakistan, a traditional AF was adopted. The land was stratified using SPOT-derived land cover; the strata were then subdivided into PSUs. The final segment size was of 30 ha, which was highly efficient. The stratification was performed manually. Currently, the technical reports do not contain any information on man-hours and the type of statistical or machine algorithms used for classification and stratification.

To estimate the crop area, an AF was developed. The first step was the stratification, using ten land cover classes, a probability proportion, a raising factor (RF) and seasonal ground truthing. AF segments were identified (e.g. rice zones), the Frame Sample (PSU) was designed and the samples were selected. A further division was performed in consideration of the ecological zones of the two provinces examined [34].

The statistical analysis is still ongoing, and the final result of the use of the land cover database in LCCS will be fully documented. However, at this stage it can be stated that Pakistan is using the LCCS effectively. For example, all orchards have been discriminated and the list thereof may be sampled in a double sampling frame. In addition, the stratification will be refined and the substrata will be further homogenized.

8

Conclusion and recommendations

This research was conducted to analyse the development of land cover databases in support of sampling and surveys in design and operation activities in the agricultural sector.

The proposal suggests the adoption of a robust baseline land cover with the following components:

- Land cover maps based on a comprehensive, standardized a priori classification system, created for mapping exercises and independent of the scale or mapping method applied. The system could therefore serve as an internationally agreed basis for land cover in the development and maintenance of MSFs. The methodology is applicable at any scale and is comprehensive, in the sense that any land cover identified anywhere in the world can be readily accommodated using the same semantic description. The suggested classification is FAO LCCS3 or LCML.
- Geospatial database development is capable of providing useful information to agriculture, forestry and urban development planning activities, environmental protection efforts, and many other applications. The data collected in the database would enable spatial analyses and integration of other relevant AFs and list frames with the standardized metadata. FAO's toolkit is capable of carrying out all necessary steps, from pre-processing to interpretation.
- Large-scale mapping programmes enable the preparation of data in consistent formats and to known accuracy levels across countries and regions. The availability of free medium-resolution satellite imagery and processing tools allows for a more cost-effective operation. The online accessibility of higher-resolution images through products such as Google Earth provide opportunities for cost-effective approaches to validation.

Part II of this project reports on the tests of the procedures for land cover classification, and GAEZ and Farming Systems support to the development of AFs and MSFs.

The tests were conducted with multi-resolution images, for the purpose of illustrating the capability and efficiency to discriminate and manage land feature data in different agricultural systems and with different image products. The crucial steps in the selection of representative areas with intensive or rain-fed agricultural activities and in the research of a complete RS coverage have been carried out. Finally, the image provider and the specific RS products have been identified, and the acquisition campaign has already started. The selection criteria for the countries of origin and areas has been studied, as well as the criteria to identify the best acquisition period for RS products, which may vary with the crop calendar. The technical approaches have been presented in this Part I. In addition, the following actions have been taken or are ongoing:

- download of the free-of-charge Landsat imagery data sets covering the areas of interest;
- collection of ancillary information;
- contacting and sending requests of acquisition of RS products to FAO's main data providers and discussing the best solutions with them;
- sending requests of support and collaboration to the relevant national counterparts.

The specific actions to be taken in the selected validation areas with the acquired images include the following:

- The accuracy and statistics of the land cover classes will be analysed using the available global land cover databases.
- The calculated statistics will be assessed to quantify the bias.
- The geometric accuracy will be also checked, to the extent that appropriate data from the field will be available. A part of this test has already been conducted and reported upon ([11]); this work will be integrated, comparing the results obtainable with the new RS images.

In the selected image tiles, the following studies were conducted using the new RS products:

- Analysis of the databases available for a preliminary characterization of the country/area of interest;
- Analysis of the area characteristics;
- Creation of an integrated database with other available data;
- Application of a global agriculture legend in LCCS3, maintaining the national classification system;
- Segmentation and interpretation of the image tiles;
- Geographic and statistical analysis of the vector data outputs;
- Application of the visual and semi-automatic interpretation method using FAO/GLCN tools;
- Automatic interpretation using the LCCS legend for raster data output;
- Geographic and statistical analysis of the raster data outputs.

The outcomes of the test offer a complete overview of the RS applications in an agricultural environment.

Geospatial technologies are valuable in supporting MSF development and database management, through their support for the creation of consistent land cover classifications and ancillary data (AEZ, FSs, etc.) and the collation of other multiple AFs and list frames. This process aims to yield greater efficiency in supporting agricultural statistics and, in practical terms, reduces the number of ground validation visits; moreover, it can contribute to reduce non-sampling errors, particularly in field enumeration activities. Three possible approaches have been observed, based on international practices and the availability of resources: i) using the existing land cover databases; ii) improving the available land cover data; and iii) developing a new baseline. Subject to the resources available, the development of a new baseline is preferable in terms of the currency and consistency of information and the longevity of an MSF. This process can be supported by FAO-NRL, in collaboration with country mapping and statistical agencies, to improve agricultural statistics.

The positive impact on the production of useful data and the integration of economic, social and environmental sectors support the Global Strategy's objectives and the needs of decision- and policymaking, across the sectors of renewable natural resources in agriculture, forestry and fisheries.

Recommendations

Although RS imagery is increasingly available, the application of the environmental analysis is often constrained by costs and the absence of systematic approaches to generate products to support agricultural statistics. Developing and implementing standardized and consistent procedures for land cover database generation, tailored to agricultural statistical needs, should be considered to be a core requirement.

It is necessary to evaluate the delineation of landscape and farming systems within the development and maintenance of MSFs, such that other appropriate standardized approaches may be identified. It is also recommended to propose an implementation of an AEZ at national scale for the better definition of such zones within countries.

A comprehensive technical assistance program on the use of RS for agricultural statistics and sampling must be held to bridge this knowledge gap. FAO can support capacity development to develop in-country, highly trained, multi-disciplinary, motivated and sustainable teams with the skills necessary to develop and maintain the MSF and engage in other multi-disciplinary uses of the land cover data.

Free satellite image collections provided by new-generation products may provide sustainable solutions to reduce the costs of land cover database management. FAO can further support cost-effective image selection and acquisition.

The adoption of available software to facilitate image processing, photo-interpretation capacity and GISs would significantly improve the analysis process.

References

World Bank, FAO & UN. 2010. *The Global Strategy to Improve Agricultural and Rural Statistics*. Report No. 56719-GLB. World Bank Publication: Washington, D.C. Available at: http://www.fao.org/docrep/015/am082e/am082e00.pdf. Accessed 12 August 2016.

World Bank, FAO & UN. 2012. *Action Plan of the Global Strategy to improve agricultural and rural statistics: For food security, sustainable agriculture, and rural development*. World Bank Publication: Washington, D.C. Available at: http://www.fao.org/docrep/016/i3082e/i3082e.pdf. Accessed 12 August 2016.

FAO & UNEP. 1997. *The Future of Our Land: Facing the Challenge. Guidelines for Integrated Planning for Sustainable Management of Land Resources.* FAO Publication: Rome. Available at: http://www.fao.org/docrep/004/x3810e/x3810e04.htm. Accessed 12 August 2016.

Di Gregorio, A. & Jansen, L.J.M. 2005. *Land Cover Classification System. Classification concepts and user manual. Software version 2*. Environmental and Natural Resources series, n. 8. Geo-Spatial data and Information. FAO Publication: Rome.

FAO. 1997a. *Terminology for integrated resources planning and management*. FAO Publication: Rome. Available at https://www.mpl.ird.fr/crea/taller-colombia/FAO/AGLL/pdfdocs/landglos.pdf. Accessed on 12 August 2016.

FAO & UNEP. 1999. *Land Resources*. FAO Publication: Rome. Available at: http://www.fao.org/nr/land/use/en/. Accessed on 12 August 2016.

Asner, G.P., Scurlock, J.M.O. & Hiche, J.A. 2003. Global synthesis of leaf area index observations: implications for ecological and remote sensing studies. *Global Ecology & Biogeography* 12: 191–205.

FAO, European Commission, OECD, UN, World Bank. 2013. *The System of Environmental-Economic Accounting Central Framework (SEEA)*. Available at: http://unstats.un.org/unsd/envaccounting/seeaRev/SEEA_CF_Final_en.pdf. Accessed on 12 August 2016. .

Foote, K.E. & Huebner, D.J. 1995. *Error, Accuracy, and Precision*. University of Colorado Publication: Boulder, CO, USA. Available at http://www.colorado.edu/geography/gcraft/notes/error/error.html. Accessed on 12 August 2016.

DiBiase, D., Sloan, J., Baxter, R., Dutton, J.A. & King, B. 2014. "3. Geospatial Data Quality", in *GEOG 482: The Nature of Geographic Information.* Pennsylvania State University Publication: University Park, PA, USA. Available at: https://www.e-education.psu.edu/natureofgeoinfo/c5_p3.html. Accessed on 12 August 2016.

Latham, J.S. & Rosati, I. 2013. *The land cover database in LCCS style: a dynamic tool for agricultural statistics.* FAO Concept Note. Unpublished internal document.

Sayer J., Sunderland, T. Ghazoul, J., Pfund, J.L., Meijaard, E., Venter, M., Boedhihartono, A.K., Day, M., Garcia, C., van Oosten, C. & Buck, L.E. 2013. *Ten principles for a landscape approach to reconciling agriculture, conservation, and other competing land uses*. Proceedings of the National Academy of Sciences, 110(21): 8345-8348.

Beddow, J. M., Pardey, P. G., Koo, J. & Wood, S. 2010. "The Changing Landscape of Global Agriculture". In Alston, J.M., Babcock, B.A. & Pardey, P.G. (eds). *The Shifting Patterns of Agricultural Production and Productivity Worldwide*. Midwest Agribusiness Trade Research and Information Center: Ames, IA, USA.

Dixon, J., Gulliver, A. & Gibbon, D. 2001. *Farming Systems and Poverty: Improving Farmers' Livelihoods in a Changing World*. FAO & World Bank Publication: Rome & Washington, D.C.

Burrough, P.A. 1986. *Principles of Geographical Information Systems for Land Resource Assessment*. Oxford University Press: Oxford, UK.

Carfagna, E. & Marzialetti, J. 2009. *Continuous Innovation of the Quality Control of Remote Sensing Data for Territory Management*. In Erto, P. (ed.). *Statistics for Innovation: Statistical Design of "Continuous" Product Innovation*, pp. 145-160. Springer-Verlag Italia: Milan, Italy.

Carfagna, E. & Marzialetti, J. 2009. Sequential design in quality control and validation of land cover databases. *Journal of Applied Stochastic Models in Business and Industry (ASMBI), 25(2): 195-205.*

ERDAS, Inc. 1999. *ERDAS Field Guide*. 5th ed.B. ERDAS Publication: Atlanta, GA, USA.

Taylor, J., Sannier, C. Delincé, J. Gallego, F.J. 1997. *Regional crop inventory in Europe assisted by remote sensing: 1988 -1993. Synthesis Report of the MARS Project-Action* 1. European Communities Publication: Luxembourg.

US Geological Survey. 1998. *Geospatial Positioning Accuracy Standards Part 3: National Standard for Spatial Data Accuracy*. National Data Infrastructure (NSDI) FGDC-STD-007.3-1998. Federal Geographic Data Committee Publication: Reston, VA, USA.

Garrity, D., Dixon, J. & Boffa, J.M. 2012. *Understanding African Farming Systems. Science and Policy Implications*. Paper prepared for the Conference on *Food Security in Africa: Bridging Research and Practice*, 29-30 November 2012. Sydney, Australia.
Available at:
http://aciar.gov.au/aifsc/sites/default/files/images/understanding_african_farming_systems_report_for_aifsc_conference.pdf 12/08/2016. Accessed on 12 August 2016.

Potere, D. 2008. Horizontal Positional Accuracy of Google Earth's High-Resolution Imagery Archive. *Sensors*, 8: 7973-7981. Available at: http://www.mdpi.com/1424-8220/8/12/7973 12/08/2016. Accessed on 12 August 2016.

Nagi, Z.M., Ghazi, A. & Hussam, E.M. 2013. Positional Accuracy Testing Google Earth. *International Journal of multidisciplinary Sciences and Engineering*, 4(6): [6 http://www.ijmse.org/Volume4/Issue6/paper2.pdf. Accessed on 12 August 2016.

Ubukawa, T. 2013. *An Evaluation of the horizontal positional accuracy of Google and Bing satellite imagery and three road data sets based on high resolution satellite imagery*.
Available at:
http://www.ciesin.columbia.edu/confluence/download/attachments/19726351/Ubukawa_PositionalAccuracy_march2013.pdf?version=1&modificationDate=1363795265000. Accessed on 12 August 2016.Various references on FS approaches in http://www.fao.org/docrep/w7365e/w7365e04.htm.

ISO. 2012. *ISO 19144-2 – International Standard – Geographic Information – Classification system Part 2 - Land Cover Meta Language (LCML).* First edition 2012-07-01.
Available at:
http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=44342. Accessed on 12 August 2016.ISO Publication: Geneva, Switzerland.

Sabins, F.F. 1986. *Remote Sensing: Principles and Interpretation*. 2nd ed. W.H. Freeman & Co.: New York, USA and Oxford, UK.

Lillesand, T.N., Kiefer, R.W., Chipman, E. 2004. *Remote Sensing and Image Interpretation*. John Wiley & Sons: New York, USA.

Latham, J.S., Cumani, R., Rosati, I. Bloise, M. 2013. *GLC-SHARE Global Land Cover Share*. FAO Publication: Rome.

Latham, J.S. 1983. *Utilization of Remote Sensing and field data for production of agricultural statistics on crop acreage. Implementation Stage Part I. Survey design and field data collection*. FAO Technical Report. Unpublished internal document.

FAO & SUPARCO Crop Reporting Service. 2014. *Land Cover Atlas of Pakistan: The Punjab and Singh Province*. FAO Publication: Rome.

Wigton, W.H., Chandra, H. *Trip Report - Statistics from Space Support to Ethiopia to Improve Agricultural Statistics in Ethiopia Project*. MTF/ETH/078/BMGF (July 26 to August 18, 2013) FAO Final Report. Unpublished internal document.

FAO & SUPARCO Crop Reporting Service. 2012. *PUNJAB CRS: Base Line Survey: Agriculture Information System: Building Provincial Capacity for Crop Forecasting and Estimation*. 978-969-9102-10-3. FAO Publication: Rome.

FAO. 2014. *Handbook and Instructions for Assessing Country Capacity to Produce Agricultural and Rural Statistics*. Development Series 1b. FAO Publication: Rome.

FAO. 2011. *Creation of the Land Cover Change Database for Kenya (Update of the AFRICOVER database).* FAO **[document type?].** Unpublished internal document.

Fischer, G., Granat, G., & Makowski, M. 1998. *AEZWIN. An Interactive Multiple-Criteria Analysis Tool for Land Resources Appraisal*. Interim Report IR-98-051, International Institute for Applied Systems Analysis Publication: Laxenburg, Austria.

Harris Geospatial Solutions. 2016. *Calculate Confusion Matrices.* Available at: http://www.exelisvis.com/docs/CalculatingConfusionMatrices.html. Accessed on 12 August 2016.

Caprioli, M. & Tarantino, E. 2003. *Standards and Quality in GIS Contexts.* Paper prepared for the FIG Working Week 2003, 13-17 April 2003. Paris. Available at: http://www.geo.upm.es/postgrado/CarlosLopez/papers/StandardsAndQualityInGIScontexts.pdf. Accessed on 12 August 2016.

Srivastava, M. 2012. *Global Strategy to Improve Agricultural and Rural Statistics.* Presentation prepared for the Training Workshop on Country Assessment, 27-29 August 2012. Addis Ababa. Available at: http://www.google.it/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&ved=0CCYQFjAB&url=http%3A%2F%2Fwww.fao.org%2Ffileadmin%2Ftemplates%2Fess%2Fess_test_folder%2FWorkshops_Events%2FWorkshop_Addis_2012%2F6.FAO_Global_Strategy_Ethiopia-MKS-_Aug_2012.ppt&ei=W-a_VJaWKMnaat2pgugH&usg=AFQjCNGIUuhrqZLI1b9OH4AVWX4etEqm2A&bvm=bv.83829542,d.bGQ. Accessed on 12 August 2016.

Global Strategy to Improve Agricultural and Rural Statistics (GSARS). 2014. *Handbook on the Master Sampling Frame for Agricultur*al Statistics. Available at: http://gsars.org/wp-content/uploads/2016/02/MSF-010216-web.pdf. Accessed on 12 August 2016.

Carfagna, E. 2013. *Using Satellite Imagery and geo-referencing technology for building a master sampling frame*.
Available at http://www.statistics.gov.hk/wsc/IPS110-P1-S.pdf. Accessed 12 August 2016.

Matthews, J.A., Bartlein, P.J., Briffa, K.R., Dawson, A.G., De Vernal, A., Denham, T., Fritz, S.C. & Oldfield, F. (eds). 2012. *The SAGE Handbook of Environmental Change. Volume 1.* SAGE Publications Ltd.: London.

Di Gregorio, A. 2016. *Land Cover Classification System: Classification concepts*. *Software version 3*. Available at: http://www.fao.org/3/a-i5232e.pdf. Accessed on 12 August 2016. FAO Publication: Rome. (Card URL: http://www.fao.org/documents/card/en/c/93c7a130-2f12-4f7a-a317-011eef91c5af/).

Di Gregorio, A. & Leonardi, U. 2016. *Land Cover Classification System: User manual. Software version 3.* Available at: http://www.fao.org/3/a-i5428e.pdf. Accessed on 12 August 2016. FAO Publication: Rome. (Card URL: http://www.fao.org/documents/card/en/c/c41f08a4-e612-45d8-b569-b751f27a3542/).

Annex

Annex 1: Menu of indicators

	INDICATOR	DATA REQUIREMENTS	DATA SOURCES	TECHNICAL NOTES
	SECTOR WIDE INDICATORS FOR AGRICULTURE AND RURAL DEVELOPMENT			
1	Gross domestic product (GDP).		Censuses and surveys of firms, farms, and households for small holders.	Value added should include unreported activities as well as the value of informal or small-scale operations. Annual estimates between census or surveys based on extrapolations based on other indicators.
2	GDP growth from agriculture value added	Estimates of total production and value for all commodities produced in the country, including that from small holders and household plots minus estimates of the cost of inputs such as seed, feed, energy, fertilizer, labor, etc. Agriculture includes forestry and fisheries.	Censuses and surveys, agricultural enterprises, farm and rural households, and administrative and processor data.	SNA concepts followed. Problems include estimation of output consumed by the household and the annual coverage of all commodities for which only periodic census data are available. Annual estimates made using previous census and other administrative data if available.
3	Amount of public spending on agriculture, subsidies, and infrastructure	Government budget allocations and spending related to agriculture. Agriculture includes forestry and fisheries.	Ministry of finance, national accounts, planning commissions, donor reports.	The definition for public spending on agriculture should follow the U.N. Classification of Functions of Government (COFOG) for agriculture.
4	Amount of public spending on rural infrastructure, health, and education	Government budget allocations and spending related to rural areas.	Ministry of finance, national accounts, planning commissions, and donor reports.	Rural defined using national description.
5	Change in investment in capital stock	Inventories of machinery and equipment owned by agricultural holdings, buildings such as milking purposes, animal breeding stock, area of semi-permanent crops such as trees and vineyards, number of trees and vines.	Agricultural resource surveys of holdings and agricultural enterprises.	Machinery and equipment inventories should be by purpose (tillage, harvesting, etc.) and size.
6	Demographics of agricultural and rural population	Rural population and number of rural households, number of agricultural households and population living in them, age and education levels. Agriculture includes forestry and fisheries	Census of Population, Census of Agriculture, Household surveys, administrative records	Rural defined using national description
7	Rural poor as a percent of total poor population	Household income and consumption estimates for national and rural poverty lines. Purchasing Power Parities (PPPs) for comparisons across countries.	Household surveys. International Comparison Program for comparisons across countries.	Countries should use poverty estimates based on PPPs and extrapolate between ICP benchmarks.
8	Rural hungry as a percent of total poor population	Household income and food consumption estimates for national minimum energy requirements.	Household surveys. International Comparison Program for comparisons across countries.	Countries should use hunger estimates for monitoring food deprivation levels.
9	Food production index	Area, production, and yield for food crops, livestock numbers, and production of meat, milk, eggs, fish captured and cultured, and other food products, nonfood use of food products, and food imports and exports.	Agricultural census, surveys of agricultural enterprises, processors, fish landings, administrative data such as imports and exports. Food balances and household consumption surveys.	Follow FAO guidelines for inclusions and exclusions.

3 Indicators should be disaggregated by gender.

continued

	INDICATOR	DATA REQUIREMENTS	DATA SOURCES	TECHNICAL NOTES
10	Change in value of trade—imports and exports	Imports and exports—quantities and values of agricultural products including fishery and forest products.	Customs inspections—in some countries the customs offices collect the data, which then are turned over to the national statistical office for compilation.	National statistical offices should collaborate with customs officials to ensure coding and classifications follow international guidelines.
	INDICATORS FOR AGRICULTURAL SUBSECTORS AND RURAL AREAS			
11	Productivity of crop production as measured by crop yields	Quantity harvested per unit of area such as hectare and area harvested. Area harvested distinguished between irrigated harvested crops and rainfed harvested crops.	Census of agriculture, crop-cutting surveys. Production sample surveys, processor surveys, such as oil seed crushers and cotton ginners.	Difficult to measure with multi-cropping or with crops that can be harvested more than once a year. Crop cutting can over estimate yields.
12	Change in components of crop balances	Area harvested, quantity harvested, quantities imported or exported, change in stocks, quantities by utilization such as food, biofuels, own consumption for every crop including those produced for fiber and oil.	Surveys of agricultural enterprises, administrative data on trade, processors by utilization, and household surveys for own consumption.	Crop balances should reflect the growing cycle and marketing year, which could be different from the calendar year.
13	Livestock value added	Estimates of quantity and value of production of meat, poultry, milk, eggs, by-products such as hides and skins and wool mohair minus costs of inputs such as feed and replacement stock.	Surveys of agricultural holdings, enterprises such as slaughter plants, dairies, and processors. Household surveys for own consumption	Own consumption should be included, difficult to measure.
14	Change in components of livestock and poultry balances by species	Number of animals born, acquired, slaughtered, and deaths from disease. Number of animals by purpose such as breeding, meat, milk, wool, and by age breakdowns relevant to specie (see FAO 2010 Census).	Surveys of agricultural holdings at least annually but more often for species with more frequent births during a reference period. This ranges from annually for cattle to monthly for egg production.	Data collection intervals should reflect the reproductive cycles. This suggests annual for cattle, semiannual for pork, and quarterly or shorter for poultry and milk.
15	Change in productivity of capture fish production	Quantity of fish taken by unit of fishing effort; scientific estimates of fish stock and exploitation rates.	National fishery surveys, surveys at landing sites, onboard observers, national, regional, and global assessment results.	
16	Change in productivity of aquaculture	Estimates of quantity and value of production of fish by species minus costs and quantity of inputs such as seed, feed, and fertilizers.	Surveys of aquaculture enterprise, and holdings, aquaculture census, market certifications,	
17	Change in components of fish balances	Quantities and value of captures from coastal and offshore waters, rivers, and lakes including nonlanded catch; quantities and value of products from aquaculture; utilizations including own consumption and discards, imports and exports.	National fishery surveys, fishery census, aquaculture census, surveys of fishery and aquaculture enterprises, processors, market information, and administrative and inspection sources.	See *CWP Handbook* and FAO coding and classification.
18	Change in components of forestry balances	Quantity and value of removals of products from forested areas and respective utilizations.	Appropriate ministries, satellite imagery, price surveys, or processor data.	
19	Commodity price indexes	Market reports of prices being offered by commodity and location. Prices received by the enterprise at the first point of sale.	Market observers, surveys of enterprises, agro-enterprises purchasing commodities from agricultural enterprises.	Care needed to ensure units of measure for pricing are comparable.
20	Consumer price indexes	Monthly or seasonal prices paid by the consumer.	Consumer price index.	Care is needed to ensure highly seasonal products do not distort the price series.
21	Early warning of change in food security.	Monthly or seasonal prices paid by the consumer.	Windshield surveys of crop conditions, amount of precipitation, satellite imagery of vegetative indexes, changes in trade data, and animal disease outbreak.	These do not have to be statistically rigorous, mainly to provide an early warning that other interventions are needed.

	INDICATOR	DATA REQUIREMENTS	DATA SOURCES	TECHNICAL NOTES
	CLIMATE CHANGE, LAND, AND THE ENVIRONMENT			
22	Change in land cover and use	Land Cover Classification System (LCCS), area and georeferenced for cultivated land, grass or pasture, inland water, marine water, wetlands, shrubland, woodland, fallow or idle cultivated land, barren land, urban or developed areas, areas equipped for irrigation.	Land use surveys, satellite imagery. Georeferenced data on economic situation of agricultural holdings needed to understand effect of policy decisions on land use.	Ground truth data required to provide more detailed breakdowns of cultivated land, especially for crops in small plots. Difficult to apply in detail where multicropping is used.
23	Change in proportion of land area covered by forests, rate of deforestation	Area georeferenced to map materials.	Ministry responsible for forestry, satellite imagery.	Follow LCCS classification.
24	Percent of land and water area formally established as protected areas	Land and water area and georeferenced to mapping material.	Responsible ministry—satellite imagery.	Follow LCCS coding with expansion covering inland and marine water bodies.
25	Irrigated land as percent of total cropland Productivity of irrigation	Total cropland and area irrigated by source of water for irrigation (surface water, groundwater, treated wastewater, etc.) and by method (surface, sprinkler, localized irrigation). Crop yields from irrigated land compared to yields from nonirrigated areas.	Agricultural census, other crop-related surveys or water-user survey.	Irrigation refers to the artificial application of water to assist in the growing of crops (and pastures). Can be done by letting water flow over the land ("surface irrigation"), by spraying water under pressure over the land concerned ("sprinkler irrigation"), or by bringing it directly to the plant ("localized irrigation").
26	Withdrawal of water for agriculture as a percent of total water withdrawal	Area under irrigation, number of irrigations, irrigation intensity and requirements by crop, water withdrawal and turnover rate for aquaculture consumption, and per capita consumption by people and animals.	Appropriate ministries, special studies or surveys to estimate water use in agriculture and aquaculture, and surveys of aquaculture enterprises and holdings.	Should include both surface and ground water. Coding and classifications should be defined.
27	Change in soil loss from watersheds	Reduction in crop yields, reduction in area of cultivated land.	Appropriate ministries, georeferenced data with satellite imagery.	
28	Change in affect of inputs on the environment	Fertilizer, pesticide, and other chemicals applied to the soil, water bodies, and plants by type of crop and watershed area, stocking.	Agricultural census and or follow-up surveys to measure fertilize and chemical use, tillage methods.	Data should be georeferenced to land cover and use.
	THE AGRICULTURAL AND RURAL ECONOMY			
29	Number of family and hired workers on the holding	Include unpaid labor of the operator of the holding and family members plus number of hired workers.	Labor force surveys of holdings.	Need to establish standards for minimum ages of workers and the number of hours worked per week to be considered a worker. Need to define reference period. Need to ensure female workers are counted.
30	Number of household members employed by farm and nonfarm	The employment status for work off the agricultural holding for each household member.	Labor force surveys—household surveys.	Need to distinguish defined employment from unpaid household service work such as domestic chores.
31	Change in farm and rural nonfarm household income from all sources	Income to the household by sector, crop, livestock, etc. Income from investments or employment outside the agricultural holding.	Rural household survey.	Rural to be defined using national definitions.
32	Percent of rural population using services of formal banking institutions	Total number of rural households, number using credit or savings services.	Central bank or commercial banks, special surveys, agricultural census.	
33	Change in sales of agro-enterprises	Sales, net profits of enterprises providing services to agriculture.	Special surveys.	Use standard accounting principles.

Source: [1]

Annex 2: Minimum Set of Core Data

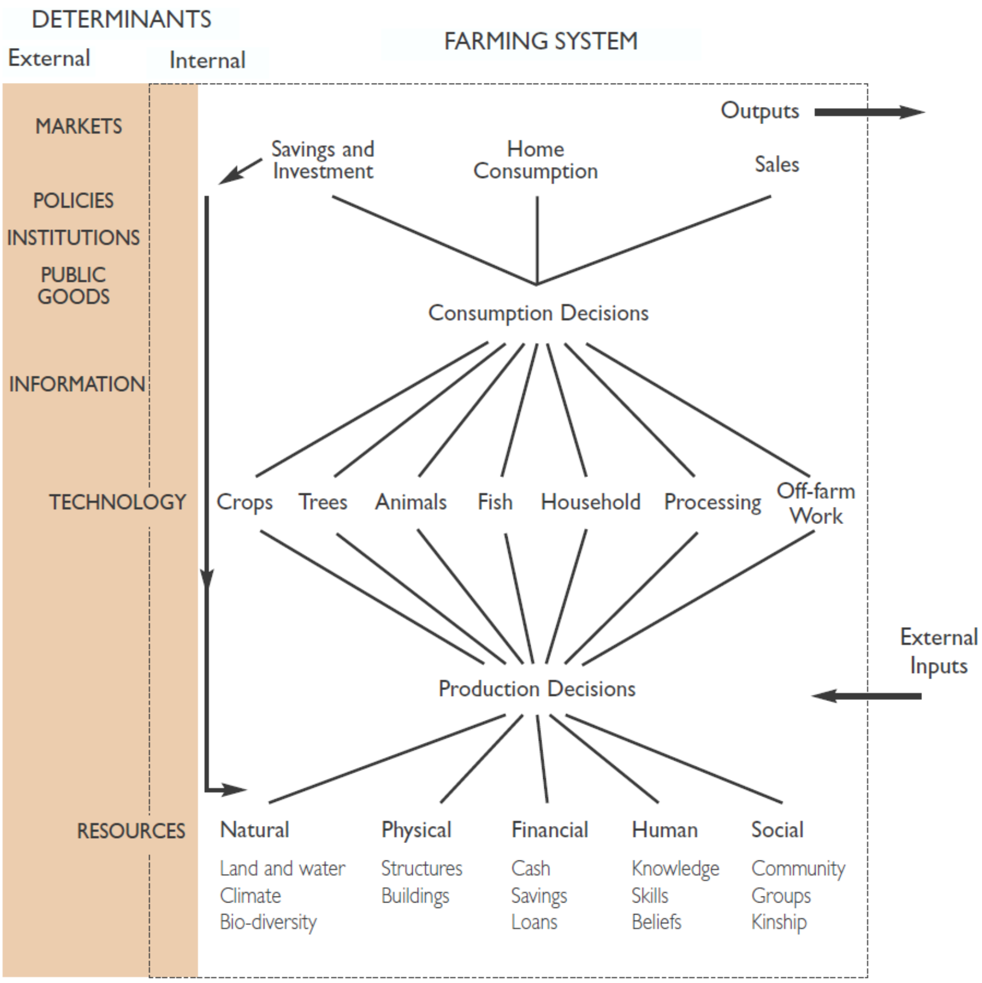
GROUP OF VARIABLES	KEY VARIABLES	CORE DATA ITEMS	FREQUENCY[a]
ECONOMIC			
Output	Production	Core crops (e.g., wheat, rice, etc.) Core livestock (e.g., cattle, sheep, pigs, etc.) Core forestry products Core fishery and aquaculture products	Annual
	Area harvested and planted	Core crops (e.g., wheat, rice, etc.)	Annual
	Yield/births/productivity	Core crops, core livestock, core forestry, core fishery	Annual
Trade	Exports in quantity and value	Core crops, core livestock, core forestry, core fishery	Annual
	Imports in quantity and value	Core crops, core livestock, core forestry, core fishery	Annual
Stocks	Quantities in storage at beginning of harvest	Core crops	Annual
Stock of resources	Land cover and use	Land area	
	Economically active population	Number of people in working age by sex	
	Livestock	Number of live animals	
	Machinery	Number of tractors, harvesters, seeders, etc.	
Inputs	Water	Quantity of water withdrawn for agricultural irrigation	
	Fertilizers in quantity and value	Core fertilizers by core crops	
	Pesticides in quantity and value	Core pesticides (e.g., fungicides herbicides, insecticides, disinfectants) by core crops	
	Seeds in quantity and value	By core crops	
	Feed in quantity and value	By core crops	
Agro processing	Volume of core crops/livestock/fishery used in processing food	By industry	
	Value of output of processed food	By industry	
	Other uses (e.g., biofuels)		
Prices	Producer prices	Core crops, core livestock, core forestry, core fishery	
	Consumer prices	Core crops, core livestock, core forestry, core fishery	
Final expenditure	Government expenditure on agriculture and rural development	Public investments, subsidies, etc.	
	Private investments	Investment in machinery, in research and development, in infrastructure	
	Household consumption	Consumption of core crops/livestock/etc. in quantity and value	
Rural infrastructure (capital stock)	Irrigation/roads/railways/communications	Area equipped for irrigation/roads in km/railways in km/ communications	
International transfer	ODA[b] for agriculture and rural development		
SOCIAL			
Demographics of urban and rural population	Sex		
	Age in completed years	By sex	

a The frequency for the items not specified will be established by the framework provided in the Global Strategy to determine the national priorities for content, scope, and frequency. The frequency requirement will also be considered in the establishment of the integrated survey framework where the data sources will be defined.
b ODA = Official Development Assistance

GROUP OF VARIABLES	KEY VARIABLES	CORE DATA ITEMS	FREQUENCY
	Country of birth	By sex	
	Highest level of education completed	One digit ISCED by sex	
	Labor status	Employed, unemployed, inactive by sex	
	Status in employment	Self employment and employee by sex	
	Economic sector in employment	International standard industrial classification by sex	
	Occupation in employment	International standard classification of occupations by sex	
	Total income of the household		
	Household composition	By sex	
	Number of family/hired workers on the holding	By sex	
	Housing conditions	Type of building, building character, main material, etc.	
ENVIRONMENTAL			
Land	Soil degradation	Variables will be based on above core items on land cover and use, water use, and other inputs to production.	
Water	Pollution due to agriculture		
Air	Emissions due to agriculture		
GEOGRAPHIC LOCATION			
GIS coordinates	Location of the statistical unit	Parcel, province, region, country	
Degree of urbanization	Urban/Rural area		

Annex 3: Farming System

a) Main elements of the Farming System



Source: [14], [21]

b) Classification of Farming System by broad category

CLASSIFICATION OF FARMING SYSTEMS BY BROAD CATEGORY

•••

AFR	MNA	ECA	SAS	EAP	LAC
SMALLHOLDER IRRIGATED					
Irrigated	Irrigated				Irrigated
WETLAND RICE BASED					
			Rice Rice-Wheat	Lowland Rice	
SMALLHOLDER RAINFED HUMID					
Forest Based Rice-Tree Crop Root Crop Cereal-Root Crop Mixed Maize Mixed			Tree Crop	Root-Tuber Temperate Mixed	Forest Based Intensive Mixed Maize-Beans (Mesoamerican)
SMALLHOLDER RAINFED HIGHLAND					
Highland Perennial Highland Temperate Mixed	Highland Mixed		Highland Mixed Sparse (Mountain)	Upland Intensive Mixed Highland Extensive Mixed	Intensive Highland Mixed High Altitude Mixed (Central Andes) Moist Temperate Mixed Forest-Livestock

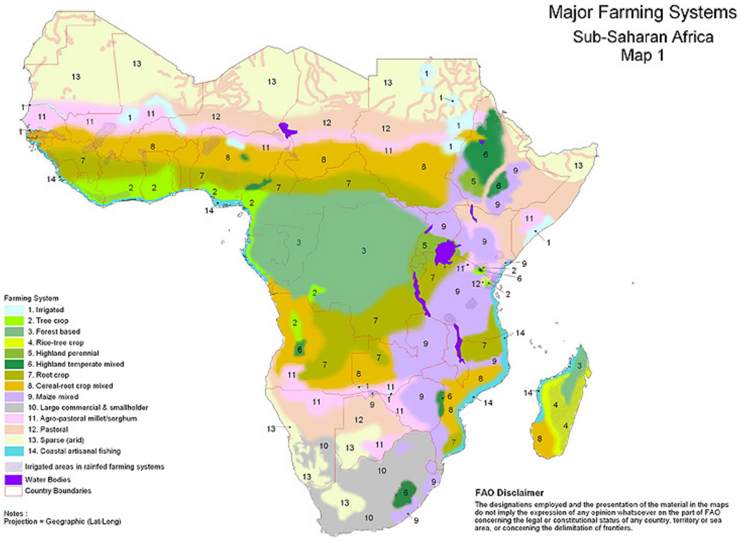
Source: [14], [21]

AFR	MNA	ECA	SAS	EAP	LAC
SMALLHOLDER RAINFED DRY/COLD					
Agropastoral Millet/Sorghum Pastoral Sparse (Arid)	Rainfed Mixed Dryland Mixed Pastoral Sparse (Arid)	Small scale Cereal-Livestock Sparse (Arid)	Rainfed Mixed Dry Rainfed Pastoral Sparse (Arid)	Pastoral Sparse (Arid) Sparse (Forest)	Dryland Mixed Pastoral Sparse (Forest)
DUALISTIC					
Tree Crop Large Commercial & Smallholder		Irrigated Mixed Forest Based Livestock Horticulture Mixed Large scale Cereal-Vegetable Extensive Cereal-Livestock Pastoral Sparse (Cold)		Tree Crop Mixed	Coastal Plantation & Mixed Extensive Mixed (Cerrados & Llanos) Cereal-Livestock (Campos) Temperate Mixed (Pampas) Extensive Dryland Mixed (Gran Chaco)
COASTAL ARTISANAL FISHING					
Coastal Artisanal Fishing	Coastal Artisanal Fishing		Coastal Artisanal Fishing	Coastal Artisanal Fishing	
URBAN BASED					
Urban Based	Urban Based	Urban Based	Urban Based	Urban Based	Urban Based

c) Farming Systems of Africa

Farming Systems	Land Area (% of region)	Agric. Popn. (% of region)	Principal Livelihoods	Prevalence of Poverty
Irrigated	1	2	Rice, cotton, vegetables, rainfed crops, cattle, poultry	Limited
Tree Crop	3	6	Cocoa, coffee, oil palm, rubber, yams, maize, off-farm work	Limited-moderate
Forest Based	11	7	Cassava, maize, beans, cocoyams	Extensive
Rice-Tree Crop	1	2	Rice, banana, coffee, maize, cassava, legumes, livestock, off-farm work	Moderate
Highland Perennial	1	8	Banana, plantain, enset, coffee, cassava, sweet potato, beans, cereals, livestock, poultry, off-farm work	Extensive
Highland Temperate Mixed	2	7	Wheat barley, tef, peas, lentils, broadbeans, rape, potatoes, sheep, goats, livestock, poultry, off-farm work	Moderate-extensive
Root Crop	11	11	Yams, cassava, legumes, off-farm work	Limited-moderate
Cereal-Root Crop Mixed	13	15	Maize, sorghum, millet, cassava, yams, legumes, cattle	Limited
Maize Mixed	10	15	Maize, tobacco, cotton, cattle, goats, poultry, off-farm work	Moderate
Large Commercial and Smallholder	5	4	Maize, pulses, sunflower, cattle, sheep, goats, remittances	Moderate
Agro-Pastoral Millet/Sorghum	8	8	Sorghum, pearl millet, pulses. sesame, cattle, sheep, goats, poultry, off-farm work	Extensive
Pastoral	14	7	Cattle, camels, sheep, goats, remittances	Extensive
Sparse (Arid)	17	1	Irrigated maize, vegetables, date palms, cattle, off-farm work	Extensive
Coastal Artisanal Fishing	2	3	Marine fish, coconuts, cashew, banana, yams, fruit, goats, poultry, off-farm work	Moderate
Urban Based	<1	3	Fruit, vegetables, dairy, cattle, goats, poultry, off-farm work	Moderate

Source: FAO data and expert knowledge.
Note: Prevalence of poverty refers to number in poverty, not depth of poverty, and is a relative assessment for this region.


Major Farming Systems
Sub-Saharan Africa
Map 1
Farming System
1. Irrigated
2. Tree crop
3. Forest based
4. Rice-tree crop
5. Highland perennial
6. Highland temperate mixed
7. Root crop
8. Cereal-root crop mixed
9. Maize mixed
10. Large commercial & smallholder
11. Agro-pastoral millet/sorghum
12. Pastoral
13. Sparse (arid)
14. Coastal artisanal fishing
Irrigated areas in rainfed farming systems
Water Bodies
Country Boundaries
Notes :
Projection = Geographic (Lat/Long)
FAO Disclaimer
The designations employed and the presentation of the material in the maps do not imply the expression of any opinion whatsoever on the part of FAO concerning the legal or constitutional status of any country, territory or sea area, or concerning the delimitation of frontiers.

Annex 4: FAO-NRL software for data interpretation and analysis

The FAO/GLCN provides a number of effective tools to support the development of consistent land cover classifications. The software comprises a number of applications:

- The Geographical Vector Interpretation System (GeoVis) is a vector-based editing system specifically designed for thematic interpretation. Therefore, it facilitates and accelerates all mapping activities based on RS data. It is a user-friendly system that embeds the main tools of vector drawing and editing, including topological functions, with advanced capabilities of raster management and a direct link with the LCCS.

- The MApping Device–Change Analysis Tool (MADCAT) is a stand-alone application that provides a combination of two different, but fully integrated sets of functions: 1) Land cover mapping support: the software allows users to generate vector-based land cover data sets using different interpretation techniques (visual, semi-automatic, fully automatic); 2) Detection and validation, on a statistical basis, of changes in the land cover changes. Changes can be detected in different ways: (a) on an existing vector layer (whole-to-whole, on the basis of polygons selected by a "pattern recognition" filter); and (b) on statistically sound sample areas. The software is directly linked with LCCS and is able to manage LCCS classes and LCCS cartographic standards.

- The Land Cover Classification System (LCCS) is the only universally applicable classification system currently in operational use. It enables a comparison of land cover classes, regardless of data source, economic sector or country. The LCCS method enhances the standardization process and minimizes the problem of dealing with a very large amount of predefined classes. The FAO/GLCN initiative has generated two versions of LCCS: LCCS2 [4] and LCCS3, [43] or LCML [26].

- The Mapping Accuracy Program (MAP) is a statistical program that enables the automatic calculation of the thematic mapping accuracy, using different methods and different expected levels of statistical confidence. The accuracy assessment methods used are both qualitative and quantitative. The assessment can be performed by single point, by multiple points or by polygon area. MAP software is included in the GeoVis package.

- The Advanced Database Gateway (ADG) is cross-cutting interrogation software that enables the straightforward and fast recombination of land cover polygons according to the requirements of individual end users. Aggregated land cover classes can be generated not only by name, but also on the basis of the set of existing classifiers.
  For tutorials and other related documentation, see
  http://www.glcn.org/index_en.jsp.

Annex 5: Other RS products

Sensor	Spectral Resolution: Channel	Spectral Resolution: Wavelength range (µm)	Spatial Resolution (m)	Revisit time (days)	Recommended maximum working scale (approx.)	Cost (approx.)
Low/medium						
NOAA-AVHRR	Channel 1	0.58-0.68 (red)	1,100 x 1,100	1	1:1.5 m	US$ 0.00015/km^2
	Channel 2	0.725-1.1 (near IR)				
	Channel 3	3.55-3.93 (mid. IR)				
	Channel 4	11.3-12.6 (ther. IR)				
	Channel 5	11.4-12.4 (ther. IR)				
MODIS	Channel 36	From 0.620 To 2.155	250-1,000	1	1:1.5 m	Free
MERIS	Channel 15	From 0.39 To 1.04	300-1,200	3	1:1.5 m	Free
PROBA-V	Channel 1	0.438-0.486 (blue)	100-350	1	1 m	Free
	Channel 2	0.615-0.696 (red)				
	Channel 3	0772-0.914 (N. IR)				
	Channel 4	1564-1634 (SWIR)				
SPOT VGT	Channel 1	0.50-0.59 (green)	1,150 x 1,150	1	1:1.5 M	Free
	Channel 2	0.61-0.68 (red)				
	Channel 3	0.79-0.89 (near IR)				
	Channel 4	1.58-1.75 (sh. w. IR)				
High						
Landsat	Channel 1	0.5-0.6 (green)	80 x 60	18	1:200,000/ 1:100,000	Free
MSS	Channel 2	0.6-0.7 (red)				
	Channel 3	0.7-0.8 (near IR)				
	Channel 4	0.8-1.1 (near IR)				
TM/ETM	Channel 1	0.45-0.52 (blue)	30 x 30	16		
	Channel 1	0.52-0.60 (green)				
	Channel 3	0.63-0.69 (red)				
	Channel 4	0.76-0.9 (near IR)				
	Channel 5	1.55-1.75 (sh. w. IR)				
	Channel 6	10.4-12.5 (ther. IR)	120 x 120			
	Channel 7	2.08-2.35 (sh. w. IR)	30 x 30			

ETM	Pancro	0.52-0.9	15 x 15			
Landsat 8	See Table 3					
SPOT-XS	Band 1	0.50-0.59 (green)	20x20 5- 2.5	26	1:100,000 1:5,000	
	Band 2	0.61-0.68 (red)				
	Band 3	0.79-0.89 (near-IR)				
SPOT	Pancro	0.51-0.73	10 x 10		1:50,000	
ASTER	3 Bands	VNIR (near-IR)	15	16	1:100,000	US$ 0.0152/km^2
	6 Bands	SWIR (near-IR)	30			
	5 Bands	TIR (ther. IR)	90			
Indian Remote Sensing (IRS)	Pancro	0.5-0.75	5.8 x 5.8	24	1:15,000	US$ 0.33/km^2
IRS LISS	Band 2	0.52-0.59 (green)	23.5 x 23.5		1:100,000	US$ 0.114 US/km^2
	Band 3	0.62-0.68 (red)				
	Band 4	0.77-0.86 (near-IR)				
	Band 5	1.55-1.7 (near-IR)				
IRS WiFS	Band 3	0.62-0.68 (red)	188 x 188	5	1:500,000	
	Band 4	0.77-0.86 (near-IR)				
DMC	Band 1	0.52-0.60 (green)	22	3/1	1:100,000	Ortho-rectified US$ 592 per image US$ 0.12 per km^2
	Band 2	0.63-0.69 (red)				
	Band 3	0.77-0.90 (near-IR)				
DMC Nigeria Sat-2 2015	Pancro	0.52-0.898	2.5	Now on demand	1:100,000 1:5,000	
	Band 1	0.448-0.517 (blue)	5			
	Band 2	0.527-0.606 (green)				
	Band 3	0.63-0.691 (red)				
	Band 4	0.776-0.898 (near-IR)				
DMC3 2015	Pancro		1			
	R,G,B,NIR		4			
ALOS	Band 1	0.42-0.50 (blue)	10	46	1:100,000	US$ 0.1004/km^2
	Band 2	0.52-0.60 (green)				
	Band 3	0.61-0.69 (red)				
	Band 4	0.76-0.89 (near-IR)				

DEIMOS 1-2 (Urthecast)	Band 1	red	22 m 75 cm	2	1:100,000	EUR 12/km^2
	Band 2	green				
	Band 3	blue				
	Band 4	Near-IR				
Very high						
Pleiades	Pancro	0.47-0.83	0.7	1 day (43° off-nadir) 4 days (30° off-nadir) 5 days (20° off-nadir) 13 days (5° off-nadir)	Very detailed: < 1:5,000	US$ 13/km^2
	Band 1 (B0)	0.43-0.515(blue)	2.8			
	Band 2 (B1)	0.50-0.62 (green)				
	Band 3 (B2)	0.59-0.71 (red)				
	Band 4 (B3)	0.74-0.94 (near-IR)				
RapidEye	Band 1	0.44-0.51 (blue)	5	5	Very detailed < 1:5,000	US$ 1.28/km^2
	Band 2	0.52-0.59 (green)				
	Band 3	0.63-0.685 (red)				
	Band 4	0.76-0.85 (near-IR)				
GeoEye-1	Pancro	0.40-0.80	0.5	21	Very detailed < 1:5,000	25-40 US$/km^2
	Band 1	0.45-0.51 (blue)	2.0			
	Band 2	0.51-0.80 (green)				
	Band 3	0.655-0.69 (red)				
	Band 4	0.78-0.92 (near IR)				
Quickbird	Pancro	0.45-0.90	0.6-2.4	1-3.5 days, depending on latitude (30° off-nadir)	Very detailed < 1:5,000	US$ 28-40/km^2
	Band 1	0.45-0.52 (blue)				
	Band 2	0.50-0.60 (green)				
	Band 3	0.63-0.69 (red)				
	Band 4	0.76-0.90 (near-IR)				
IKONOS	Pancro	0.45-0.90	0.8-3.2	Approx. 3 days at 40° latitude	Very detailed < 1:5,000	US$ 20-35/km^2
	Band 1	0.45-0.52 (blue)				
	Band 2	0.50-0.60 (green)				
	Band 3	0.63-0.69 (red)				
	Band 4	0.76-0.90 (near-IR)				
Sentinel	See table 4					

Annex 6: Technical characteristics of selected satellite sensors

Landsat 8

Date of launch	February 2013
Instruments on board	Operational Land Imager (OLI) Thermal Infrared Sensor (TIRS)
Processing	Level 1T - Terrain Corrected
Pixel size	OLI Multispectral bands: 30 m OLI panchromatic band: 15 m TIRS Thermal bands: 100 m (resampled to 30 m, to match multispectral bands)
Spectral bands	Landsat OLI image data 9 spectral bands (band designations) with a spatial resolution of 30 m for Bands 1 through 7 and Band 9. Band 8 (panchromatic) is 15 m. Landsat TIRS image data will consist of two thermal bands with a spatial resolution of 100 m (resampled to 30) for Bands 10 and 11. For data distribution, the TIRS data will be packaged with the OLI.
Scene size	170 km North-South by 183 km East-West
Data characteristics	GeoTIFF data format Cubic Convolution (CC) resampling North Up (MAP) orientation UTM map projection (Polar Stereographic for Antarctica) World Geodetic System (WGS) 84 datum 12 m circular error, 90% confidence global accuracy for OLI 41 m circular error, 90% confidence global accuracy for TIRS 16-bit pixel values
Data delivery	ftp download within 24 hours of acquisition

Sentinel-1

Date of launch	2013
Instruments on board	Sentinel-1 is a polar-orbiting, all-weather, day-and-night radar imaging C-band imaging radar mission, comprising a constellation of 2 satellites
Processing	Not available
Pixel size	5 × 20
Spectral bands	A single C-band synthetic aperture radar instrument operating in 4 modes (varied swath widths and coverage)
Scene size	Swath width of 250 km
Data characteristics	Includes C-band imaging operating in four exclusive imaging modes with different resolution (down to 5 m) and coverage (up to 400 km). It provides delivery within 24 hrs and 2 acquisitions every 12 days.
Data delivery	Radar data is delivered within an hour of acquisition.

Sentinel-2

<table>
<tr><td>Date of launch</td><td colspan="2">2014</td></tr>
<tr><td>Instruments on board</td><td colspan="2">Sentinel-2 is a polar-orbiting, multispectral high-resolution imaging mission</td></tr>
<tr><td rowspan="3">Processing</td><td>Level 1 image processing includes:</td><td>a) Radiometric corrections: straylight/crosstalk correction and defective pixels exclusion, de-noising, de-convolution, relative and absolute calibration;
b) Geometric corrections: co-registration inter-bands and inter-detectors, ortho-rectification.</td></tr>
<tr><td>Level 2 image processing includes:</td><td>a) Cloud screening
b) Atmospheric corrections including thin cirrus, slope and adjacency effect correction
c) Geophysical variables retrieval algorithms e.g. fAPAR, leaf chlorophyll content, leaf area index, land cover classification.</td></tr>
<tr><td>Level 3 provides spatio-temporal synthesis.</td><td>Simulation of cloud corrections within a Level 2 image</td></tr>
<tr><td>Pixel size</td><td colspan="2">< 1 ha MMU (Minimum Mapping Unit) fully achievable with 10 m;
4 bands at 10 m, 6 bands at 20 m and 3 bands at 60 m spatial resolution (the latter is dedicated to atmospheric corrections and cloud screening)</td></tr>
</table>

Spectral bands	Optical payload with visible, near-infrared and short-wave infrared sensors comprising 13 spectral bands). 0.4-2.4 µm (VNIR + SWIR)
Scene size	Swath width of 290 km
Data characteristics	Revisit time of 5 days at the Equator (under cloud-free conditions) and 2–3 days at mid-latitudes
Data delivery	Unknown