Global Strategy
IMPROVING AG-STATISTICS

Implementation of the Research Plan
July 2013

The programme of activities for 2013 was discussed at the high level meeting held in FAO Headquarters in December 2012 and was approved by the Global Steering Committee in February 2013.

For all research topics, the plan of action has been defined and relevant information concerning ongoing or already completed research activities has been collected.

For 10 research topics, significant progress has been made:

  - information on the ongoing or already completed research activities has been collected;
  - the partnerships have been established and the activities of the partners have been coordinated;
  - the relevant literature has been identified and reviewed and the reports has been prepared.
- The next steps till the end of 2013 are:
  - identify and analyse gaps and remaining methodological issues and propose possible solutions;
  - organize a workshop on the results of these activities;
  - test the proposed methodological improvements (where appropriate, field tests will be designed and conducted).
- Main activities foreseen for 2014 for these research topics:
  - Prepare technical reports on findings and recommendations for possible solutions to methodological issues;
  - submit the technical reports to peer review and validation;
  - disseminate the results;
  - finalize a methodological publication.

For 8 research topics, the activities have started, but still require efforts for reaching the objectives; and for the research topic "Improving methods for estimating livestock and livestock products" the activities are behind of expectations in terms of delivery.

Identifying the most qualified partners, possibly from developing countries, which can contribute to the implementation of the research topics foreseen in the work plan for 2013 has presented some difficulties, since some of them are not available at the moment.

The administrative work for establishing the partnerships has been long and complex. Contractual arrangements are very time consuming, particularly at the initial stages of the partnership. This has delayed the implementation for several research topics.

The research activities contribute to the achievement of the output 3 of the Global Strategy to Improve Agricultural and Rural Statistics "New cost effective methods for data collection, analysis and dissemination developed and disseminated".

Some research topics are grouped in thematic domains; e.g., the research topics 3.1.1 Conceptual framework (SEEA – Agri) and 3.1.2 Integrated survey framework are grouped under the thematic domain 3.1 Framework for agricultural statistics.

Output 3 - New cost effective methods for data collection, analysis and dissemination developed and disseminated	
3.1	**Framework for agricultural statistics** This thematic domain is constituted by two research topics: • Conceptual framework (SEEA – Agri) • Integrated survey framework

3.1.1 The Conceptual framework (SEEA – Agri)

The conceptual framework (SEEA-AGRI) can be defined as a comprehensive and standard satellite account for the integration of agricultural and environmental data based upon internationally agreed concepts, definitions, classifications and inter-related tables and accounts universally valid, regardless of the stage of economic development reached by the country.

The SEEA-AGRI aims to translate policy issues into data needs and requirements in a standard and coherent manner by:

- Enhancing the use of existing agricultural statistics and related common frameworks (supply and utilization tables and food balances, etc.) through the integration of basic statistics consistent with the System of National Accounts (SNA);
- Providing a consistent, comprehensive, and coordinating framework to link data collected by different surveys and censuses together to build up an integrated database;
- Providing a sound basis for the measurement of a set of economic, social, and environmental indicators for agriculture and rural development aligned with FAO's narrow and broad definitions of agriculture, respectively;
- Providing a framework to expand the analytical capabilities of the original FAO SEAFA and related past FAO initiatives (Fishery and Forestry Accounts);
- Providing a framework that links to other SEEA subsystems being articulated by other agencies (Ecosystems, Energy, etc.).

The collaboration has been agreed with Carl Obst, Australia and other FAO Divisions on the following topics: forestry, fishery, land, water, energy and agriculture.

The activities for this thematic domain were foreseen to initiate in 2012 and to be completed in 2013. In fact, the implementation is initiated in 2013 and will be finalized in 2014.

3.1.2 The Integrated survey framework

This research topic focuses on the strategic objective of producing integrated and consistent agricultural statistics on phenomena related to three different target populations, the units of which are respectively:

- agricultural plots (for what concerns the environmental aspects);
- households (for the social aspect)
- farms (for the economic dimension).

The stress on integration allows achieving a better coverage of specific statistics for which a suitable solution cannot be found by surveying only a specific units. The research focuses mainly on:

- sampling design methods (how to collect the data of interest);
- estimation methods (how to produce estimators of the target parameters using existing data)
- quality framework.

The research topic is being implemented with the support of researchers from the Mozambique National Institute of Statistics, the Italian National Institute of Statistics and the University la Sapienza of Rome.

A first research report on a conceptual note for integrated survey framework has been released

	and discussed in a internal seminar at FAO. It is expected to release the first version of guidelines at the end of January 2014. Some presentations describing the first research results have been submitted and accepted in some scientific meetings: • ISI meeting (end of august) • ICAS IV (end of October) • Symposium of Statistics Canada on imperfect frames (December). The activities for this thematic domain were foreseen to initiate in 2012 and to be completed in 2013. In fact, the implementation is initiated in 2013 and will be finalized in February 2014.
3.2	**Improved methodologies for master sampling frames** The master sample frame is the sampling frame to be used for selecting probability based samples which allow collecting data concerning the land, the farms and, in case the households, for producing some social statistics. According to the characteristics of the country (kind of agriculture, statistical system, landscape, size of farms and fields etc.) different approaches are the most reliable for producing agricultural and rural statistics. The main ones are: sample surveys based on the list of farms (in case a complete and updated list of farms is available) and sample surveys based on area frames. The optimal approach is combining the area frame with the list frame while building the frame (georeferencing the parcels of land of the farms). However, in countries where the list of the farms is out of date and the parcels are not georeferenced, this approach cannot be followed and a feasible alternative is combining estimates obtained from an area frame with the ones produced through a short list of large, commercial farms. The kind of area frame, the way of linking the list and the area frame, the way of using the geographic information varies according to the specific situations and the aim of the research is improving these methods. Thus the following research topics have been identified: • Identifying the most appropriate area frame for specific landscape types • Improving methods for linking area frames with list frames • Improving the use of GPS, GIS and RS for setting up a master sampling frame **3.2.1 Identifying the most appropriate area frame for specific landscape types** The research identifies the most appropriate geographic sampling frames, called area frame, for the different categories of countries, according to the kind of agriculture, statistical system, landscape, size of farms and fields, kind of physical boundaries, etc. Several kinds of area frames, have been developed in the last decades: parcels of land, with physical or regular, theoretical boundaries (generally generated by a regular grids), rectangular parcels called transects, very small circles called points, clusters of points, and so on. A collaboration has been set up with Javier Gallego, Joint Research Centre and Luis Ambrosio, University of Madrid, Spain. The report on the review of related projects and literature has been completed. **3.2.2 Improving methods for linking area frames with list frames** The research identifies the most appropriate list frames for the different categories of countries

	and improves the methods available in the literature for combining area frames with list frames at the frame level, as well as at the estimator level; taking into consideration the different kinds of area and list frame which can be the most appropriate for the different kinds of countries. A collaboration has been set up with Cristiano Ferraz, University of Pernambuco, Brazil. A draft report on the literature review has been prepared. **3.2.3 Improving the use of GPS, GIS and RS for setting up a master sampling frame** The traditional approach to set up an area sampling frame was based on collection of printed maps and aerial photographs (not always ortho-rectified) and involved a large amount of manual work. Current technologies, in particular the ability of GIS to efficiently handle different layers of geographic information, in particular RS-based thematic maps, have made this task much lighter. Stratification for example can be performed in a more efficient way. The evolution of Global Positioning System (GPS), with sufficiently accurate devices at affordable prices, has substantially changed the field work more than the definition of a sampling frame, but both aspects cannot be separated, because the choices in the definition of the sampling frame needs to take into account the field survey aspects. Research is being conducted for improving the use of GPS, GIS and remote sensing for setting up a master sampling frame for integrated survey for the various categories of countries, according to the landscape, the economic structure, the size of farms, the spatial distribution of important crops and livestock species, and the kind of data sources available in the country. A collaboration has been set up with Luis Iglesias, University of Madrid, Spain. The report on the review of related projects and literature has been prepared. The implementation of this research topic will be completed at the end of 2013.
3.3	**Improving methods for estimating cost of production in developing countries** The importance of rural economy is very high in several developing countries and the economical sustainability of the farms, particularly in a period of price volatility, is strictly linked to their cost structure. However, although most countries produce estimates on cost of production, the quality of these estimates is low. The research focuses on nomenclature, questionnaires, sample designs, sample size, taking into account the different typologies of countries (e.g. availability of an updated list frame of farms in the country or not, level of informal economy in the country etc.), evaluation of the level of bias in estimating the cost of production. The activities for this research topic initiated in 2012. A draft of the handbook on best practices has been produced and a meeting of several countries has been held in Vietnam, in order to facilitate the collection of best practices applicable in developing countries. This handbook will be revised by a small group of experts. The literature review has been carried out by the Joint Research Centre as in-kind contribution. The final handbook will be delivered by the end of 2014.
3.4	**Improving methods for estimating Livestock and livestock products** The aim of this research topic is improving the methods for collecting data on livestock, including cattle, sheep, pigs, goats, and poultry, because livestock production is a major contributor to food supply and income. Consumption increases as countries develop, therefore

	resulting in more livestock consuming grain and adding to methane emissions. Accurate estimation of livestock numbers and production is a challenge in many countries, particularly in Africa because of the nomadic and semi-nomadic livestock systems. Social constraints also create difficulties in obtaining accurate numbers on livestock in pastoral societies and estimation of livestock products, especially with regards to small animals. After a careful analysis, till now we have not found a partner from a developing country who can give an important contribution to the implementation of this research topic and is available, in these months. Due to the delay in the identification of the partner, likely most of the activities foreseen in 2013 will be postponed to 2014.
3.5	**Improving methods for estimating Post harvest losses** Quantitative food losses refer to the decrease in edible food mass available for human consumption throughout the different segments of the supply chain. In addition to quantitative losses, food products can also face a deterioration of quality, leading to a loss of economic and nutritional value. Post harvest losses can have a strong impact on food security in several developing countries. The research will analyze the main factors which influence post harvest losses in order to identify the most appropriate methodology/methodologies for estimating them. A collaboration has been established with Robert VanOtterdijk, FAO – AGS, ERS - USDA and Joint Research Center.
3.6	**Improving methods for crops estimates** The problem of estimating the area, the yield and thus the production of crops still has not a satisfactory solution. The problem becomes even more difficult to solve in case of mixed crops, repeated and continuous cropping and when the yield of root crops has to be estimated, like for the very important crop cassava. The result of this research will be improved methodologies for estimating area, yield and production of important crops in developing countries, at different stages of the growing season, taking into consideration the different kinds of crops and typologies of countries. At early stages of the growing season, the acreage of crops is often estimated through a mixture of data sources, including interviews on sowing intention and multi-temporal classification of remote sensing data. The limitations of each of these methods for given conditions have not been sufficiently analysed, in particular the potential bias and subjectivity. Often, yield forecasting and early warning are based on agro-meteorological models, which generally involve regression analysis of time series of ground data and official statistics and sometimes various kinds of remote sending data. At later stages, ground observations, through area or list frames or farmers interviews are used for estimating the area and the yield of crops. **3.6.1 improvement of estimation of crop area, yield and production** The aim of this research topic is improving the methods for: • Estimating the area of crops at the different stages of the growing season. These methods are based mainly on ground observations, sometimes combined with remote sensing data or farmers interviews;

- Estimating the yield of crops at the different stages of the growing season, taking into consideration the specific problems posed by the different kinds of crops. These methods are based mainly on ground observations and time series regression models (e.g. regression models based on the number of ears per plot etc.), sometimes combined with remote sensing data or farmers interviews;
- Yield forecasting (mainly agro-meteorological models) and early warning (mainly early warning systems). These methods are generally based on interpolation and integration of meteorological data, data on soil, sowing date, e.g. models which combine evapo-transpiration models, like Penman's model, yield time series, vegetation indexes, like NDVI. Generally, these models are not based on extensive collection of yield data on the ground and generally involve remote sensing data.

For the sub-topic 1, a collaboration with Mike Craig, USDA, USA has been established, for the sub-topic 2, with George Hanuschak, USDA, USA and for sub-topic 3 with Bruno Basso, Michigan State University, USA.

The report on the review of projects and literature has been completed for the 3 sub-topics.

3.6.2 Improving methods for estimating crop area, yield, production of mixed crops, repeated and continuous cropping

In many countries, the practice of sowing mixed crops in the same field is very common. Mixed cropping provide protection to farmers against weather uncertainties. But the method of sowing mixed crops is not uniform. The crops in mixture are sown either row-wise separately or mixed altogether. Also the technique for apportioning the area for different inter-crops differs from country to country and sometimes, like in India, from state to state. Therefore, a standard statistical methodology for estimating the area and the yield under mixed crops, repeated and continuous crops needs to be developed.

After a careful analysis, the partner institution which seems to be particularly appropriate for improving the methods for estimating crop area, yield and production of mixed crops, repeated and continuous cropping is the Indian Agricultural Statistics Research Institute (IASRI), given its experience in this field. The content of the plan of action was agreed in April; however the IASRI can cooperate only in case the Indian Ministry of Agriculture (ICAR) approves the plan of action and the cooperation. Unfortunately, the authorization is still pending.

3.6.3 Developing methods for estimating yields of root crops

The difficulties in properly estimating the yield of crops is more challenging for root crops, particularly when the crop is harvested in small quantities, over extended periods of time because of better in-ground storability, often even spanning across agricultural seasons.

This is the case, for example, for crops like cassava, that is a particularly important staple crop in many Sub-Saharan African countries and discrepancies in estimation from different sources points to substantial problems. As a result, traditional recall methods commonly used in household surveys may be highly inaccurate and thus questionable.

While different methods have been proposed and applied in the field, the lack of best practices remains a constraining factor in providing opportune technical advice to countries on the correct method. Although crop cutting remains the gold standard, it is impractical in most large scale surveys, particularly if multi-purpose in nature. At the other end of the spectrum, it is increasingly accepted that using recall methods spanning over several months, yields inaccurate estimates of continuous crops like cassava.

	In order to create synergies, an agreement has been reached with the World Bank for collaborating in this field. After a quick literature review, the field tests have been designed. In order to compare the different methods for measuring cassava production quantities and cultivated area, different methods will be compared: crop cutting, diaries and recall. For area measurement, traversing, Global Positioning System (GPS) and farmer's self-reporting will be compared. Other potential issues in properly quantifying cassava production relate to (1) the state of the crop (fresh vs. dry) and (2) the almost universal use of non-standard units e.g. units, heaps, etc.. To solve the first problem, all quantities for each method will be for fresh cassava. For the crop cutting exercise, both fresh and dry weights will be taken and conversion factors estimated. Regarding the second issue, visual aid in printed and electronic format will be used to facilitate the quantification of units of different sizes to be later converted into standard units using uniform conversion factors. The study will take crop-cutting as the gold standard for the quantification of production while compass and rope will be used to benchmark the GPS and self-reported measures of land area.
3.7	**Improving the methodology for using Remote Sensing** A document on best practices for crop area estimation with remote sensing has been prepared by GEOSS (GEOSS, 2009), focusing on the use of remote sensing data as an auxiliary variable for improving the precision of estimates for specific crops. However, several methodological aspects of the use of remote sensing data for producing agricultural and rural statistics still have to be improved. **3.7.1 Developing more efficient and accurate methods for using remote sensing** The following research lines have been identified: • Analysis of the new technologies of remote sensing and their influence on possible improvements in the design and/or estimation methods, with particular attention to the application of very high resolution remote sensing data and to the quality of data available free of charge (such as Google Earth); • Development of more efficient and robust statistical methods for using remote sensing data at the design level (mainly for stratification or selection criteria of the first phase sample, such as probability proportional to size and multiple probability propositional to size designs); • Development of some extensions of the regression and calibration estimators traditionally adopted, in order to improve the efficiency of the estimators and to solve some problems posed by the use of some typologies of sampling units, namely points, households etc.; • Analysis of the robustness of the estimators currently adopted for producing agricultural and rural statistics and development of more robust ones; • Comparison of regression and calibration estimators with small area estimators, for the different kinds of remote sensing data and landscape, from the efficiency and robustness viewpoints; • Comparison of design based and model based sampling for collecting ground data information for training remote sensing data and for estimating agricultural and rural parameters; • Improvement of methods for integrating remote sensing data and ground surveys, including households surveys; • Improvement of statistical methods for quality assessment of land use/land cover

databases and of methods for change detection of land cover;
- Review and improvement of methods adopted for including remote sensing data in yield forecasting models (including agro-meteorological models) and in early warning systems.

For each research line, particular attention is being devoted to the analysis of the peculiarities of each type of statistical units. The robustness of the proposed methods are being verified in presence of eventual violations of the underlying hypothesis regarding the model that links remotely sensed and ground data. All mentioned research lines are being carried out focusing on main characteristics of agriculture in developing countries, e.g. main crops, small and very small fields, mixed crops etc.

A collaboration with Roberto Benedetti, University of Chieti-Pescara, Italy and Javier Gallego, Joint Research Center has been established.

The report on the review of projects and literature has been completed.

3.7.2 Evaluating the cost-efficiency of remote sensing in developing countries

A wide scientific literature has been developed on the application of remote sensing to agricultural statistics (mainly crop area estimation and yield forecasting) and some papers and reports have been published on the cost-efficiency of remote sensing for agricultural statistics in developed countries, but very little can be found focused on developing countries. This research topic aims at assessing the cost efficiency of the use of remote sensing in developing countries, and improving the methods for assessing the cost-efficiency of remote sensing, in order to take into consideration the specificities of developing countries and the different approaches which can be followed for using remote sensing data for producing agricultural statistics.

After a careful analysis, the partner which seemed to be particularly appropriate for evaluating the cost-efficiency of remote sensing in developing countries and for improving available methodology is Redouane Arrach, Director of the Division of Statistics and Information of the Ministry of Agriculture and Fisheries, Morocco and his group.

Unfortunately, the research of the partner took much time, also because it was focused on experts and institutions from developing countries; thus the contracts still have to be prepared.

3.7.3 Improving methods for using existing land cover – land use data bases

During the last decades, there has been a growing interest on application of the satellite remote sensing technology to model the status of land cover and monitoring its changes. In addition, the application of the satellite remote sensing technology is advocated and promoted as a powerful tool for agricultural statistics. Despite the recognition of such importance, current users of land cover information have to cope with three main inadequacies:

- lack access to sufficient reliable or comparable baseline data;
- lack of consistent and validated guidelines for land cover database classification;
- lack of consistent and validated guidelines for land cover data management for agricultural statistics.

The evolution of consistent and comparable land cover products inherently depends on the development of a consensus based approach to land cover classification, one that is ratified through an internationally recognised process.

A collaboration with John Latham and Ilaria Rosati, FAO-NRL geospatial unit, within the Land

	and Water Division, FAO, has been established. The report on the review of projects and literature has been prepared.
3.8	**Adoption of new technology for field data capture, compilation, transfer and dissemination** New technologies such as GPS, PDA, remotely sensed data from satellite and aircraft as well as geographic information systems (GIS) will play an important role in the development of cost effective data collection methods and the improvement of data quality. In 2013, a prototype for Computer Assisted Personal Interview (CAPI) is being developed in partnership with the World Bank.
3.9	**Improving quality and use of administrative data to produce agricultural statistics** In developed countries, governmental interventions such as subsidies, regulation and legislation often require agricultural holders to report information on acreage. Land ownership and cadastral surveys provide useful information for constructing registers. However, more research is needed for identifying where, how and under which conditions, administrative data can be used for producing agricultural, rural and agri-environmental statistics, with particular reference to developing countries. Moreover, the quality of this kind of registers is lower in many developing countries. Other kinds of administrative data are the reports of agricultural extension development officers and experts guesses. Given resource and capacity limitations, administrative data will continue to play an important role in any agricultural statistics system in the majority of developing countries where sample surveys will continue to have multi-annual frequency with estimates not available at very low geographical level in a context of growing demand for such data. Also, there will always be a need of early warning and pre-harvest crop forecasting data to inform Governments on the prospective food situation which is mainly derived from crop monitoring and other auxiliary information produced by field staff. Methods for assessing and improving the quality of these kinds of data will be developed, as well as methods for using these kinds of data for producing cost-efficient agricultural statistics. A partner has been identified.
3.10	**Improving methodology of food balance sheets** The accuracy of food balance sheets, typically compiled at the national level, is dependent on the reliability of the underlying basic statistics of supply and utilization of foods, their nutritive value and conversion factors that govern the transformation of primary commodities into processed products and vice versa. As the major dietary commodities are rarely made available for consumption in their primary or originating form, wheat grain viz a viz wheat flour is an example, the importance of accurately capturing the transformation is paramount for monitoring food security. Compounding this task is the growing complexity and diversity of products that now enter food baskets, as well as new and enhanced technologies that are being introduced to process food. Through obtaining empirical evidence supported by model-based estimates, this research topic seeks to establish and verify appropriate technical conversion factors, i.e. milling rates, crushing rates, extraction rates, conversion and processing factors, carcass weights, milk yields, egg weights, etc., that indicate the average national rate at which commodities are converted. The benefits of the research are manifold. Not only will a more accurate depiction of food

	availabilities be obtained thereby enhancing food security metrics, underperforming food processing chains will also be identified.
3.11	**Improving methodology for small scale fishery** Accurate estimations of small scale fisheries and aquaculture are a challenge in many countries, which are an important source of food security, nutrition (especially protein and trace nutrients), and livelihoods in many countries. The collection of data on inland fisheries and aquaculture (both commercial and subsistence) will depend on developing the appropriate methods for data collection and estimation. A module to be included in the questionnaires for agricultural censuses which will collect information, particularly on the socio-economic characteristics of small scale fishery and aquaculture is being developed. A partnership with Sachiko Tsuji, Senior Fishery Statistician FIPS – FAO and Jennifer GEE has been established. A report on the literature review is being prepared.
3.12	**Better integration of geographic information and statistics** Better methods are needed for taking advantage of geographic information for producing agricultural statistics in the field of data collection, analysis and dissemination and for connecting statistical data collected through sample surveys (which can be represented by points in the space) and other kinds of geo-referenced data (polygons and lines). Research is needed for improving statistically based methods for spatial disaggregation and for integrating various kinds of geographical information and geo-referenced survey data, which is essential for crop forecasting and early warning. Under this thematic domain, the research topic "Developing robust and statistically based methods for spatial disaggregation and for integration of various kinds of geographical information and georeferenced survey data" is being implemented. **3.12.1 Developing robust and statistically based methods for spatial disaggregation and for integration of various kinds of geographical information and geo-referenced survey data** Geographical and statistical information are often available for large spatial (administrative) units. However, for policy making, it may be important to have more detailed information, whether small administrative units or geometric units, such as grid cells of a given resolution. With limited sample size, it is generally unfeasible to produce accurate estimates with the most traditional statistical methods that separately consider the information inside each small unit. The research compares several approaches to disaggregate data with the help of external co-variables. In addition, more efficient methods are studied for small area estimation using geographic information as auxiliary variables. In some cases, spatial data have to be aggregated, in order to have the different information layers with the same spatial resolution, for developing spatial models. Different aggregation methods have been developed for different purposes and have different impacts on the models. This research topic addresses the need for better integration of geographic information and statistics. Particularly, the result of the research activities will be new, more effective and robust methods for the use of geographic information to produce more accurate agricultural and rural statistics; including for connecting economic and social indicators to land use. In addition, more efficient methods will be developed for small area estimation using geographic information as auxiliary variables. The research addresses also the problem of developing

	robust and statistically based methods for spatial disaggregation and for integrating various kinds of geographical information and geo-referenced survey data, which is essential for crop forecasting and early warning. Finally, the problem of assessing the impact of aggregation, disaggregation, interpolation and integration of geographic data on the models adopted for crops forecasting and early warning is addressed. The main research lines are: • Data disaggregation; • Data interpolation; • Data aggregation; • Data integration; • Assessment of the impact of aggregation, disaggregation, interpolation and integration of different data layers on the models adopted for crops forecasting and early warning. A collaboration with Monica Pratesi, University of Pisa, (data disaggregation and interpolation), Alessandra Petrucci, University of Firenze, (data aggregation and integration) and Simone Maffei, Italy (assessment of the impact of these operations on the models) has been established. The report on the review of projects and literature has been prepared.