REPUBLIC OF RWANDA

NATIONAL INSTITUTE OF STATISTICS OF RWANDA
P.O. Box 6139 KIGALI - RWANDA

THE SYSTEM OF AGRICULTURAL STATISTICS IN RWANDA

(Improvement action plan)

March 2012

Table of Contents

A. GENERAL INTRODUCTION

1. Background

Most estimates in many countries are still based on subjective agricultural systems. In fact, if all else fails, and it sometimes does, the opinions of experts (rural agricultural agents), the use of windshield surveys, administrative records, rapid rural appraisals and subjective samples of farmers or farmers' addresses are widely used to provide estimates. The accuracy of such estimates depends on the knowledge and ability of the person that puts them together and there is no way to estimate the possible error that may be contained therein.

Subjective estimates for some items in many countries are derived from summing area and production of crops, livestock and other items that are reported to a government agency by each farmer in order to obtain permission for marketing, or to receive a subsidy, or simply because the report is required by law. This source of information should be examined closely. First of all, it may be significantly incomplete because a great part of the production may be marketed informally as contraband with only a minimum reported to maintain eligibility. Or, on the contrary, exaggerated production may be reported to obtain a larger subsidy. *There are many reasons why farmers may not be inclined to report accurately on their farms.*

In most countries that formerly had a centralized economic system, national agricultural statistics were based on complete censuses or complete registers, and probability sample surveys were seldom utilized or even learned. But in practice, in many cases, such supposedly complete censuses and registers are more properly described as being subjective sample surveys.

Crop production in Rwanda is characterized by large numbers of farmers working on relatively small plots of land. The main crops include standing crops such as bananas and cassava; root crops such as Irish potatoes and sweet potatoes; cereals such as maize and sorghum; pulses, especially beans; as well as other vegetables and fruit.

Currently, the only regular statistics available every season are the so-called “crop assessments” undertaken jointly by MINAGRI with the participation of NISR. These are forecasts of potential production produced before the end of every season and carried out principally for food security purposes.

Post harvest surveys have been conducted from time to time, for example by MINAGRI in the period 1999 to 2001 and by NISR in 2006 to 2008. The results of the 2008 large scale National Agricultural Survey (NAS) were published in February 2010.

A further source of information, albeit indirect, is the EICV, an integrated survey of living conditions, in which consumption of food (both purchased and own-produced) is recorded. This survey (lasting 12 months) was conducted in 2000/1 and again in 2005/6. The last one (2010/11) has been ended in October 2011.

2. Challenges

Agricultural sector is facing important mutations. The transformation of Agriculture that is taking place in rural area is making an impact on food production and thus statistical measurement.

The crop assessment (CA) estimates are forecasts and the survey focuses on food availability in a period of 6 months. As the total expected availability of food crops is measured in Crop Assessment Survey (CA), regardless of whether or not the entire crop (e.g., cassava) is harvested, the estimated volume levels are assumed to be generally higher than actual harvest.

Therefore, the estimation of agriculture production for the Gross Domestic Product (GDP) purpose is still challenging and actual production needs to be captured to provide a basis for revising the preliminary estimates from the CA.

As in many countries, the crop forecasting data are available before any others, and thus are timelier. The problems come in when they are used as the final actual crop production while compiling the GDP (the reliability is challenged).

3. The project

Given the importance of this sector, the National Institute of Statistics of Rwanda is committed to build a new and improved System of Agricultural Statistics. This by emphasizing on the following:

- Put in place a strong and modern system of agricultural statistics system (an integrated system of agricultural censuses and surveys);
- Produce high quality, sustainable, harmonised and cost-effective crop production statistics for food security monitoring, programme evaluation, national accounting and other purposes;
- Review the crop assessment estimates; and
- Build capacity for the responsible staff for the sustainability of the new system.

In order to achieve this, the NISR has successfully recruited key consultants from Agricultural Assessment International Corporation (AAIC). They have undertaken their first mission to Rwanda from 20th February to 3rd March 2011, and had the following main tasks:

- Review the methods used in Crop Assessment Survey;
- Prepare a document proposal for the agriculture statistics system
- Hold a stakeholder meeting to present the proposed new methodology and to obtain feedback;

Much progress has already been made in this first mission in transferring modern agriculture statistics designs to the Rwandan institutions that will be responsible for maintaining the multiple frame statistics design. However, we are just beginning a long and important work in order for the advanced statistical design methodology to be transferred. Moreover, a major survey is scheduled to start in October 2012.

This document presents the technical report on outcomes and achievements of the first AAIC mission in appendix I, some basic definitions for probability sample surveys in appendix II, and the proposed work plan in appendix III.

B. ACTIVITIES SO FAR DONE:

<u>Appendix 1: THE AAIC FIRST PERIOD ACHIEVEMENTS</u>

1. Introduction

The Agricultural Assessments International Corporation (AAIC) sent two senior consultants to Rwanda from February 19, 2012 to March 3, 2012. This team consisted of William H Wigton, President of AAIC and Dr. Alvaro Gonzalez Villalobos, Team Leader. While in Rwanda, Dr. Gonzalez wrote a technical paper that will serve as the basis for building an agriculture statistics system for the Institutions of Rwanda.

With the global shift towards market economies, the need for timely and reliable agricultural information has increased importance in the decision making process. This statement is true at both national and international levels. Much of the required information for the agricultural sector, such as crop production estimates, livestock inventories and basic social and economic data is obtained through *Current Agricultural Surveys*. These surveys are periodic (annual or seasonal), national, multiple-purpose, agricultural data collection programmes. The establishment and development of such survey programmes is, therefore, a fundamental component of an agricultural information system.

The paper delineates achievements made by Agricultural Assessments International Corporation (AAIC) during its first mission in Rwanda.

2. Analysis of the current methods used to make crop assessments and forecast

Most estimates in many countries are still based on subjective agricultural systems. In fact, if all else fails, and it sometimes does, the opinions of experts (rural agricultural agents), the use of windshield surveys, administrative records, rapid rural appraisals and subjective samples of farmers or farmers' addresses are widely used to provide estimates. The accuracy of such estimates depends on the knowledge and ability of the person that puts them together and there is no way to estimate the possible error that may be contained therein.

During this two week period Dr. Gonzalez and Mr. Wigton reviewed the current situation with regard to agricultural statistics that are being generated by the Rwanda institutions.

Crop Assessments Survey (CAS) that the MINAGRI currently implements is comprehensive. The questionnaire (presented by Mr. Mugabe Stephane, Principal Agricultural Statistician at NISR) is extensive and the summary is impressive. It was decided that this questionnaire will be the basis of any new questionnaires that will be designed for a new system to be implemented in the future. In fact, the entire CAS system has a comprehensive questionnaire, a publication that is result from data collected. Data users need and depend on this publication. Questionnaires, publications and data users are the main pillars of any agriculture statistics system and without any one of the pillars; a statistics system will not be sustainable.

AAIC has criteria that help to evaluate agricultural statistics designs. These criteria include: 1) standard methodology, 2) accurate, 3) timely, 4) objective, 5) comprehensive, 6) sustainable, 7) able to detect change, and 8) flexible. It is the opinion of the AAIC consultants that the current CAS has a standard methodology, and it is comprehensive and being sustained, there are other criteria that can be improved. The accuracy cannot be determined because it is more subjective and it seems to the AAIC consultants that there are both extensive sampling errors as well as nonsampling errors that need to be controlled. Timeliness has to do with time between when data are collected and when data are available to data users. AAIC consultants think that improvements can be made on this criterion. Objective data make it credible and consultants again find data users who think credibility needs improvement. The system isn’t as good at detecting change. So AAIC consultants have some excellent points they want to keep in the new system and some improvements that need to be made when NISR moves forward with a new system design.

After careful discussion, it was decided to build on this system and improve the periodic survey and data collection procedures.

It turns out that the current CAS implemented by MINAGRI provides an excellent guide to the required data needed by the stakeholders in Rwanda. This CAS system provides an excellent target system in many respects for the future system.

CAS questionnaires are comprehensive in the listing of crops and in addition, have space for other vegetables. The first page of the questionnaire has enough identification information so that, if follow up time is required, it is efficient to find the farmer for re-interview or supervisor checks. The AAIC consultant, Seghir

Bouzaffour is currently adapting the questionnaire so that its content can be used in an area frame application.

This adaption will be continued in April when Mr. Bouzaffour travels to Rwanda. For the first survey, AAIC consultants have suggested using an extremely simple questionnaire and they intend to add complexity as supervisors and enumerators become more experienced with the concepts of AFS.

MINAGRI staff is important partners because in the end they have to agree to the changes if there are any in the estimates. Moreover, the MINAGRI can help NISR and AAIC consultants move forward to improve efficiency because they are familiar with all the data requirements and needs assessments of data users. This aspect will need to be clear with the data users that AAIC consultants are helping to design an improved data collection system but in the meantime, there will be two data summary systems producing two crop estimates. The traditional MINAGRI system and the new system will provide different estimates of the same crops. Rwanda will need to keep these two systems running in order to adjust estimates backward in time. Eventually, one system will be dropped and there will only be one system.

Unfortunately, it is not clear whether the data being generated in the Crop Assessments Report each year by the MINAGRI is accurate or not. The CAS estimates do not have credibility with all data users and stakeholders and credibility is extremely important. According to AAIC consultants, there is no way to assess the accuracy of the current estimates because they are not probability based and therefore the coefficient of variation is not available. In addition, there is no way to assess nonsampling errors.

In the new program, NISR and MINAGRI will be able to measure the sampling errors and assess the nonsampling errors. AAIC consultants hope that the MINAGRI will continue the Crop Assessment Surveys until the MINAGRI can fold their program into the new National Program being designed by NISR, MINAGRI and the AAIC.

The proposed new program will be as comprehensive as the current MINAGRI program and have the advantage of higher credibility, timeliness and measuring change than the current system. In the future, all data being collected from the rural sector (livestock, fruits, expenditure and poverty data) will eventually be collected from the new system.

3. Proposal for the agriculture statistics system

3.1. Introduction

After careful analysis and informative discussions, the AAIC proposal for a basis of an agriculture statistics system in Rwanda is to establish an appropriate Programme of Current Seasonal Probability Surveys based on Multiple Frame (Area and List) Sampling Methods. In addition, a monthly Agricultural Reporting System based on a national network of agricultural experts and a ten year full enumeration census when possible is recommended.

During this two week consultancy, Dr. Gonzalez and Mr. Wigton accompanied the NISR agriculture statistics section to Musanze to hold a workshop set up by NISR for major stakeholders of agriculture statistics in Rwanda. From 27 to 29 February, NISR and AAIC presented options and discussed stakeholder preferences to design an agricultural statistics system in Rwanda.

The workshop attendees included staff from NISR, MINAGRI (including RAB and NAEB), Central Bank and other main institutions. Each attendee presented data requirements and data surveys from their institution. AAIC consultant, Dr. Alvaro Gonzalez presented his professional opinion of the best agriculture statistics design and basis for a statistical design in Rwanda called Multiple Frame Sampling (MFS). Mr. Wigton presented area frame sampling methods, the back bone of the MFS system.

On the last day of the workshop in Musanze (February 29, 2012), Mr. Sebastien Manzi, asked each participant of the workshop to provide feedback regarding the recommended agriculture statistics programs presented as options. Each attendee indicated that he approved the agriculture statistics design that was presented by Dr. Gonzalez.

3.2. Seasonal Probability Surveys based on Multiple Frame Sampling Methods

Timely and reliable national statistics of a country's agricultural sector can only be provided by the establishment of an adequate, periodic, national agricultural survey based on probability sampling methods. AAIC plans to provide technical assistance so that the NISR, MINAGRI and other stakeholders in agricultural statistics have credible, reliable, and timely statistics.

The First Pillar:

Seasonal Agricultural Surveys based on probability sampling and estimation methods are the backbone of the proposal to improve agriculture statistics in Rwanda. Moreover, the probability agricultural surveys recommended by AAIC will use M*ultiple Frame Surveys* (MFS). MFS consist of an area frame used conjointly with list frames. The MFS recommended consider list and area frame construction and sample selection methods that take advantage of the excellent cartography available and that require a minimum of resources and specialized staff. The MFS technology takes advantage of the strengths of the area frame, covering all farms in Rwanda, and the list frames, covering the most important farms in Rwanda.

AAIC also proposes statistical designs that take advantage of the availability of informatics methods and instruments, for instance Geographic Information Systems-GIS-, Personal Digital Assistants (PDA) and Geographic Positioning Systems–GPS-, and satellite imagery (when the orthophotos need to be updated when land use has changed extensively. These powerful systems including software and hardware improve data entry, processing, analysis and dissemination.

The Programme of seasonal MFS recommended seems to be the most practical way for the country to produce the required timely, accurate and seasonal data for the agricultural sector. These MFS methods will be fine tuned as the system is adjusted to conditions in Rwanda. Fine tune adjustments can occur when the Rwandan professionals have fully grasped the MFS technology and the consultants have understood the conditions in Rwandan agriculture.

The Second Pillar:

The second pillar AAIC consultants recommend is an agricultural statistical system in Rwanda to establish an appropriate Monthly Agricultural Reporting System based on a national network of agricultural experts. Such a System will be useful to provide expert subjective estimates of the main agricultural commodities. However, these subjective estimates need to be calibrated each season by unbiased estimates provided by probability surveys.

The Third Pillar:

The third pillar on which the agricultural statistical system is based is a full enumeration agricultural census to be conducted every ten years. The census will obtain information from all the farms in the country. This will provide a number of list frames to be used by the MFS conducted on a seasonal basis as well as information for small areas. AAIC realizes that the full enumeration census is very expensive and countries in Africa do not usually undertake this full enumeration census. However, AAIC consultants do want Rwanda to figure a way to build list frames for the multiple frame surveys because a full census provides historic small area (district) estimates which provide a way to subdivide the national level and provincial level estimates coming from the probability surveys.

Resources Required for Implementing Data Collection (First pillar)

The NISR Economic Statistics Unit, Agriculture Statistics Division, MINAGRI staff and the AAIC consultants have recommended 400 (*20-hectare*) segments for the entire country for the first survey rounds. It is suggested to have 30 teams (one for each district) of 4 members each. Each team will consist of a car and driver, a supervisor agronomist (trainer in cartography), and three enumerator/ interviewers. These 30 teams will be required for one month in order to complete the 400 segments each survey round. It is recommended to have two survey rounds each year. In addition, there must be one small survey in the marshland stratum in order to estimate the expanded area in the marshlands.

Appendix II: SOME BASIC DEFINITIONS FOR PROBABILITY SAMPLE SURVEYS

Agricultural sample surveys which use probability sampling methods that allow calculation of the statistical precision of the estimates are called probability sample surveys. In such surveys the statistical precision of the estimates have a precise mathematical meaning if nonsampling errors are controlled.

Planning an agricultural probability sample survey requires the following additional specifications. In order to define the probability model on which the estimates for the area of interest are based one must address the following:

1) The Survey population; Sampling units; Reporting units; and the Probability sampling procedure.
2) Rules of association between the sampling units and the reporting units (a sampling unit must either be associated with one and only one reporting unit or there must be a known rule to associate sampling units with a group of reporting units).
3) The survey variables should be defined in each sampling unit as a function of its values in the group of associated reporting units.

Then, the estimator for each survey variable is a statistic or random variable - A numerical function defined for each possible sample - the value of the estimator for the *selected sample* provides the *estimate* of the survey variable. The estimate of each variable for the survey area is based on the values of the variable in a sample of *sampling units* associated with the *reporting units* of the variables (the farms or the tracts of the survey area), the variance of the above estimators, which are also random variables, provide the precision of the survey estimates. For a given variable, the value of the variance of the corresponding random variable for the *selected sample* provides the sampling error of the estimate. The level of accuracy or desired degree of precision of the estimates should be established; the target survey population is the population to be surveyed. Because of practical constraints, the survey population, which is the population actually sampled (the set of sampling units), and for which inferences are valid for obtaining the survey estimates, may be different from the target population.

The **sample design** of a sample survey refers to the techniques for selecting a probability sample and the methods to obtain the estimates of the survey variables from the selected sample.

Sample designs involve sampling units, rules to assign probabilities of selection to sampling units, sampling fractions, possible stratification and clustering procedures and different types of estimation methods. Some designs involve several sampling selection stages in which case, for each sampling stage, the sampling units, probabilities of selection and sampling method have to be established to obtain the final survey estimates.

For a given sampling stage of a probability sampling design, a sampling frame is the *total set of sampling units with their probabilities of selection*, that is, the list of sampling units from which the sample is selected together with their probabilities of selection. In addition, each sampling unit in the sampling frame must have an address that shows how to get to it in Case it is selected. A frame is needed and has to be constructed for each sampling selection stage and a non-zero probability of selection has to be assigned to each sampling unit of the frame.

Sampling Frames for Agricultural Probability Sample Surveys

There are two basic types of sampling frames used for the *last-stage* of selection of an agricultural probability sample survey: area frames and list frames.

The final sampling units of an area frame are *land areas* called *segments,* and the selection probability of each segment is *proportional to its area measure*. Usually in the design of an area sample frame we have a target size of the final sampling unit and therefore the probabilities are equal. The sampling units of a list frame are usually the farms households or farmers addresses.

Therefore, there are two basic types of sample designs in terms of the *last-stage sampling unit* and the *rules to assign their probabilities of selection*, namely area sample designs and list sample designs, also known as area frame sample surveys and list frame sample surveys.

Multiple frame surveys are those agricultural probability sample designs that combine more than one sample frame to obtain the survey estimates, combining area frame designs with list frame designs. A multiple frame survey usually consists of an area frame sample and a list frame sample.

For the current agricultural probability sample surveys considered, it will be assumed that most data will be obtained by enumerators through personal interviews in the field. The aim is also to introduce PDA and GPS instruments for data collection, whenever possible. For the first rounds, paper questionnaires will be used.

The field data collection of an agricultural survey will include data collection from the sampling unit and may also include expert observation or identification and objective measurement of crop areas or other agricultural characteristics of special interest.

List Frame Sample Surveys

List sample designs are the most commonly used sampling procedures for agricultural surveys.

As defined above, a probability sample that is not an area sample is called a list sample. The usual reporting unit is the *farm* and the commonly used last-stage sampling unit is the *farm or the farmer address*.

A list frame is a frame used for the last selection stage of a list sample design. Therefore, list frames are often formed by farmers or farmers addresses.

Appendix III: NISR-AAIC WORK PLAN FOR IMPROVING AGRICULTURE STATISTICS IN RWANDA (2012)

		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec
1	**Sampling frame Design**												
1.1	Construct sampling frame for 3 districts												
1.2	Construct sampling frame for 10 districts												
1.3	Construct sampling frame in Kigali for 10 Districts												
1.4	Construct sampling frame for 30 districts												
2	**General Administration**												
2.1	Establish resources for "Census" and training												
2.2	Develop agenda for the stakeholders meeting												
2.3	Revise resources for sample census (3 rounds)												
2.4	Assist NISR develop a plan for agriculture survey												
2.5	Produce an options paper												
3	**Develop Data User Requirements**												
3.1	Stakeholder workshop												
3.2	Develop data requirements												

3.3	Develop data user needs (after every survey)												
3.4	Undertake a desk top analysis of available data												
3.5	Study current and post harvests methods												
3.6	Prepare a proposal for the Oct agriculture survey												
3.7	Assist in Dev of National Action Plan												
4	**Develop survey questionnaires and manuals**												
4.1	Design questionnaire based on data user needs												
4.2	Design survey manual based on questionnaire												
4.3	Undertake a desk top analysis of available data												
4.4	Prepare documents for the agriculture census and surveys												
4.5	Prepare a proposal for agriculture production surveys												
4.6	Produce final proposals												
5	**Training for NISR and MINAGRI staff**												
5.1	Training for NISR and MINAGRI with AAIC												
5.2	Training agenda for enumerators (area and yield)												
5.3	Training for NISR supervisors												
5.4	Training for enumerators												
6	**Develop & test data entry systems**												

6.1	Design and train data entry edit and summary programs												
6.2	Summary tables to support data user needs												
6.3	Design publication of data												
6.4	Go back to stakeholders and share report												
6.5	Workshop to share results												
6.6	Develop data user needs (after every survey)												
7	**Conducting Agriculture Survey (Multiple Frame Survey)**												
7.1	Conducting a pilot survey (30 segments across the country)												
7.2	Analysing Pilot survey results												
7.3	Conducting the 2013 Season A survey (400 segments)												
7.4	Analysing the 2013 Season A results												