

**GENERAL DATA DISSEMINATION SYSTEM
(GDDS PHASE 2 Project)**

MISSION 3 – SIERRA LEONE - FINAL REPORT

**TECHNICAL ASSISTANCE
Statistics Sierra Leone (SSL) and the Ministry of Agriculture
Freetown, Sierra Leone**

January / February 2009

**By Christopher Hill
Consultant to the World Bank
baobab_lodge@yahoo.co.uk
February 2009**

Provision of technical assistance

Topic: Agriculture Statistics For: Sierra Leone, Statistics Sierra Leone (SSL) and the Ministry of Agriculture 3rd Mission

Introduction

1. Summary and main findings.

The purpose of this mission was to provide training to a group of employees on the analysis of Agriculture Statistics.

The mission had the following objectives:

- The consultant will provide further training sessions on data analysis and guidance to a selected number of staff from SSL and MAFFS on the creation of outputs from survey files for a public database.
- Support the National Accounts in the identification of data needs
- Facilitate the standardization and definitions of agricultural concepts
- The consultant will help the staffs evaluate the most appropriate tool for building the database.
- A database should be created and be placed on the SSL web site.

10 persons participated in the workshop of which one was from the Ministry of Agriculture. All the others were from Statistics Sierra Leone.

The aim of this training was to provide more advanced practical training using data from Sierra Leone. The original plan was to use the data from “The Farm Production, Food Security and Nutrition Surveys” (VAM) undertaken by the World Food Programme for both 2005 and 2007. The survey report for the 2007 data has not yet been approved and the 2007 data has still not been made available to SSL. During the previous mission the consultant was provided with the Agricultural data from the Sierra Leone Integrated Household Survey (SLIHS) November 2002 to October 2003. It was therefore decided to work with the 2005 survey and the 2002 SLIHS.

It was noted in the report on the earlier mission with respect to SLIHS that “This data source had some serious problems. There were weaknesses in its design and various problems including confusion in the coding, coding errors and duplications. The consultant had to take some time cleaning the data.” As the VAM 2005 survey had been approved and published it was expected that this source would be clean. In practice the VAM proved to also have serious problems. There was no documentation made available. In a number of variables there were clearly erroneous values. The weighting field appeared to be seriously in error, provided values that were far too high and it was not possible to reproduce estimates in the report.

During the first part of the training a limited set of standard measures were made from both surveys for basic crops and key livestock items with the objective of producing the first steps in a time series. The consultant stressed that different concepts and ways of collecting the data should be seen as limitations on the comparability of the two sources. Nevertheless they should produce results that allowed a reasonable comparison. After the

third day it became apparent that estimates were very different. It was concluded that the weights on the VAM files were totally erroneous. They gave estimates of rural household counts between 4 and 5 times the number of rural households estimated in the Census of Population. It was decided to re-estimate the weights. A simple method was used. From the survey it was possible to identify the districts for each household and hence the sample size per district. From the Census of Population it was possible to get 100% estimates of household counts. Combining this information it was possible to re-estimate weights by district. A revised database for analysis was created using these weights. Now the estimates from the two sources showed results that were sufficiently comparable to be believable.

2. Observations about the situation in the Country and Office

Given that this country has recently come out of a period of conflict, the participants demonstrated a very high level commitment and participation in the training.

Statistics Sierra Leone has overall responsibility for the management and collection of Government Statistics. There are plans in 2009 to repeat the **Sierra Leone Farm Production, Food Security and Nutrition Survey** which was undertaken within the Ministry of Agriculture with support from the World Food Programme in 2005 and 2007. It is essential the SSL are actively involved in this survey and the questionnaire includes items needed to provide basic statistical information on agriculture and analysis can be undertaken to provide data from this survey to be used in National Accounts. SSL have recently appointed a statistician to work in the Ministry of Agriculture and a Statistician in National Accounts responsible for Agricultural data. These two persons have an important role to play in ensuring this takes place. It is essential that the authority of SSL is clearly established and strong and effect links are established between SSL and the relevant Directorate in the Ministry of Agriculture.

There are also plans to do another SLIHS. This survey which is under the direct management of SSL provides very useful data on Agriculture. The limitation of this survey is that it is taken over 12 months different household being covered each month, therefore there is a moving reference period with respect to agriculture. The 12 month moving reference period is entirely suitable for a survey of living standards, but not deal for an agricultural survey. The SLIHS should provide good data on own consumption.

Ideally the content of VAM and SLIHS should related so that VAM provides good basic agricultural information and SLIHS provided information on own consumption of agricultural production.

3 Remarks about the Agenda

The consultant proposed meeting with data users early in the mission. The management of SSL preferred that he concentrate on revised training and left these meetings to the end. The first part of the mission therefore consisted of training, revising the skills developed using Power Point with practical demonstrations during the first training.

During the first week some of the presentations used in the first training were reused to remind the SSL staff of the various techniques. Estimates were made for the key data from VAM 2005 and SLIHS 2002. In the second week the VAM weights were revised and all the estimates recalculated using the revised rates.

4. Priorities and Objectives

The objective of this mission was to provide more advanced training in SPSS and provide estimates for a user database.

The participants were 10 government employees. Most of the participants were more experienced Statisticians and most of them had participated in the previous training. One weakness was that none of the persons from the Ministry of Agriculture who took part in the first training were in the second training.

A key element of this training was that it should be very practical and use data from Sierra Leone. It was proposed to use data from both years of the Sierra Leone **Farm Production, Food Security and Nutrition Survey** for this purpose. Unfortunately the 2007 data was not available. Instead consultant used the Agriculture data from the Sierra Leone Integrated Household Survey (SLIHS) November 2002 to October 2003 and the 2005 **Farm Production, Food Security and Nutrition Survey**.

The SLIHS data were already in SPSS format. The data to be used were all the files of agriculture data plus the household roster. The Farm Production, Food Security and Nutrition Survey was in excel format. It was in a series of separate files. The data made available was limited to the Farm Production component. As mentioned above, the weights of this survey appeared to in error. There were however many other problems in the data. The key identification information for about 20 records was erroneous. This meant that the record could not be allocated to a district or region. The section of fruit crops proved to be very seriously mixed up. Many records had the coded data displaced with respect to variable names, due to confusion in the design of the questionnaire. A result of this was that many of the response for the year in which trees were planted did not have a year, it had another code and then a year occurred in different variable. Unfortunately there were a number of such errors making it very difficult to clean at this late stage. This data was not used.

The second level training reinforced the lessons learned in the first course including:

- SSL and Agriculture need to establish standard codes and classifications for key concepts such as crops and land area and use these throughout any statistical work.
- A standard set of transformations from traditional units such as “three penny pan” or “bushel” into a common unit such as “kilogram” or “metric ton” need to be adopted. The possibility that crops may be measured in different states such of “grain” or “in the pod” also need to be taken into account
- The total possible range of traditional units needs to be limited. Only the major ones should be used
- Units of measurement cannot have an open ended category. It is not possible to recode either land area “other” or unit of weight “other”.

- The number of observations classified as “other” should be minimised. For example the category “other poultry” is more common than most of the specific livestock types.

Some data cleaning at a late stage in the survey may be necessary, but the earlier the data are cleaned the better. Some of the problems identified included:

- Invalid codes that should have been eliminated during data entry. In fact the data entry system should not allow these codes to be entered.
- Extreme outliers, such as “land area over 4000 thousand acres” should also be detected in the field or during data entry.
- Duplicated, incomplete or miscoded records. The livestock section should have been a consistent set of 12 records per holding covering all 12 livestock items. The number per holding varied from 7 to 13 and included obvious duplication or miscodes.

Every software package has its strengths and its limitations. SPSS is a very powerful tool for statistical analysis. It also handles the weighting of data effectively. Most activities that are undertaken by a Government Statistical Office use only a small sub-set of the statistical capabilities of the software. On the other hand SPSS includes many functions that allow the user to organise data including:

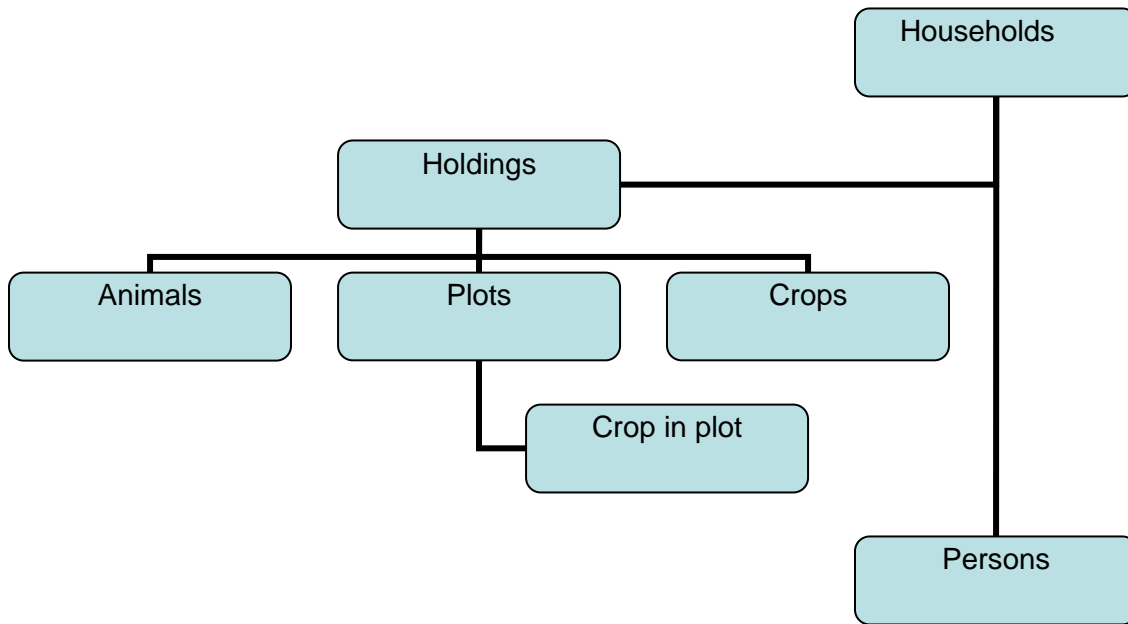
- Identifying duplicates
- Identifying and recoding missing data
- Calculating new variables
- Recoding variables
- Counting values
- Merging files
- Restructuring files

All these capabilities were taught during the first training course and reinforced during the second. At the end of the second course a number of the participants felt much more confident about using these capabilities.

SPSS is not a database system. Each data file is independent. This is a limitation. Most government statistical surveys consist of a set of related entities. The data files for these entities should ideally be linked into a database. By using the functions listed above, this limitation can be overcome.

But care needs to be taken in making these links and documenting the actions taken. The diagram below shows a simplified version of the linkages between files, that can be overcome, but not in a way. Often during survey processing and analysis one or more independent files need to be linked

Simplified example of Example Data Structure



A database system such as Microsoft “Access” allows the user to link and process provides a simpler and more transparent method of linking files through “Queries” and makes it easy to keep track of all actions taken. However if all actions that are made in SPSS are saved as “Syntax” files using the option “paste” it is possible to effectively track all the actions taken. Keeping track of processing actions requires a discipline on the part of the analyst. Emphasis was made during the training on the need to use this capability in order to document and keep track of actions. This facilitates communication among analysts. It also allows the user to design systems in advance with test data or during pilot phases of a survey that can then be implemented on the final data.

5. Activities

The mission undertook the set of activities proposed for the mission in the terms of reference. Noted below in italics qualifies or amplifies on the TOR activities.

1. SSL and MAFFS are expected to make the survey files from the 2005 and 2007 Sierra Leone Farm Production, Food Security and Nutrition Survey available to the consultant for this mission. These surveys could be supplemented by data from Sierra Leone Integrated Household Survey (SLIHS) and the Census of Population and Housing. *The 2007 data was no made available.*
2. The consultant will provide a training session and subsequently provide guidance to a selected number of staff from SSL and MAFFS on the creation of outputs from survey files for a public database. These should be either more experienced staff or junior staff with strong capacity. *There was only one participant from MFF and he had not been in the earlier training.*
3. The outputs created should be aimed at supporting the needs of key stakeholders. Outputs to support the National Accounts unit at SSL should also be produced even if they are not part of the public database. *A meeting was held with MAFF and WFP at the end of the mission only. Detailed discussions were held with the National Accounts and the results of these discussions are reported.*
4. The consultant should facilitate the standardization and definitions to be used in Agriculture Surveys. The need for this should be evident to the staff as they try to construct consistent time series. Recommendations should be made for future improvements to the surveys with respect to definitions and units of measure. *Some notes are attached on these recommendations.*
5. The consultant will help the staffs evaluate the most appropriate tool for building the database. In all likelihood, either MS ACCESS or EXCEL will be available for the team to work with. *It was agreed to use Excel in conjunction with SPSS.*
6. A database should created and be placed on the SSL web site. *A limited database was created, but needs to be reviewed further before putting on the web*
7. Instructions in terms of procedures to be followed to update and maintain the agriculture database should be provided to the manager of the agriculture statistics unit.
8. Recommendations for improving the existing data sources should be made based on the experiences of the consultant and the team in constructing the output database. *These recommendations are included in this report*

The main product to be delivered was the on going training towards producing estimates to be used in a public use database. This was given to 10 persons from the Statistics Sierra Leone and one from the Ministry of Agriculture. Data problems in the VAMS limited the degree of progress in producing the final output, but help the participants understand better the issues involved

7. Design and content of the module

See discussion elsewhere.

8. Wider issues

The Statistics Act of Sierra Leone gives Statistics Sierra Leone overall responsibility for collection and management of National Statistics. The collection of the **Sierra Leone** Farm Production, Food Security and Nutrition Survey was however undertaken within the Ministry of Agriculture with support from the World Food Programme. Recently SSL has appointed statisticians to work in the Ministry of Agriculture.

The report on the Survey has been written but not yet released. It was proposed that data from this survey should be used in the training. It was not possible to obtain the data in advance of the mission from either the 2005 or 2007 surveys. The data for the 2005 survey were made available in the middle of the second week of the earlier training programme.

It was stated in the TOR that the 2007 data should be made available for this training. However the report has not been released and the data has still not been made available to SSL. These 2007 data remains with the WFP.

The data collected by and for Government Agencies in Sierra Leone should be the property of Sierra Leone and be held by the SSL together with any other implementing agency.

The relationship between National Agencies can be a sensitive issue. Often Ministries of Agriculture have suitable field staff who can undertake data collection although Statistical Agencies have staff specifically trained in this activity. The location of statisticians in line Ministries is one option for the management of statistics. The earlier training course brought together personnel from both SLL and the Ministry of Agriculture in a cooperative activity. This latter training had only one person from the Ministry and he had not participated in the first training. The consultant met with personnel in the ministry and it became very clear that their capacity to collect data is weak.

9. Intended deliverables of the country

The earlier report stated that “The data provided from the SLIHS were an acceptable alternative, but not ideal.” In retrospect it is apparent that the 2005 VAM data was no better. The 2007 VAM is still not available to SSL and the consultant

The SSL were also responsible for providing the facilities to be used during the training and technical support to the training activity. The quality of this support was excellent. Every participant had access to a computer and the informatics personnel of SSL were extremely helpful in copying files to allow the smooth functioning of the training. **The only negative aspect of this work is that care needs to be taken by SSL to update their anti-virus systems.**

10. Agreed timeframe of the actions and the dates for the next missions.

This is the last mission in this programme of activities. The mission coincided with another mission on the development of the National Strategy for the Development of statistics (NSDS) in Sierra Leone. Together with this team the need for further work in strengthening the strategy with respect to agricultural was indentified.

11. Working relations.

The support given to the consultant was at all times excellent. Particular mention should be given to Mr Philip Amara the Director of Economic Statistics, and to Mr J.J from the National Account who organised the training programme.

The Head of Informatics and his staff were very efficient in making data and copies of the presentations available to all the participants. I would also like to express particular thanks to the Deputy Statistician General for the strong support he provided throughout the training. The level of involvement of most of the participants was also excellent.

12. Preparation issues.

The consultant felt that his work was handicapped by the unavailability of the 2007 VAM data. The serious problem in the 2005 VAM also made it necessary to revise plans in the middle of the training course.

13. Recommendations on priorities, objectives, activities and working conditions, the TOR.

Statistics Sierra Leone has overall responsibility for Government Statistics in Sierra Leone. Agricultural Statistics is a key component of this. Strengthen and developing agricultural statistics is a shared responsibility of the two institutions and needs strong cooperation.

Statistical Information on Sierra Leone is the property of the Government of Sierra Leone and must be in the hands of the Statistics Sierra Leone.

Data from the 2005 **Sierra Leone** Farm Production, Food Security and Nutrition Survey were made available to the consultant late in the second week of the earlier course. This was to be used in the final mission. The other data files from this survey were not made available. The data for 2007 was also not made available.

14. Other Donors involved.

No one from WFP participated in this training. The consultant did however meet with the Representative from WFP. She recognised the serious weaknesses in the data and WFP report. She also expressed a very positive attitude towards improving participation with SSL

15 On travel.

Travel from one side of Africa to the other is not easy. A request was made to travel Business Class. This was agreed late, but fortunately was I time for the arranging of business class travel. There were no disruptions to the journey this time.

16. On contracting.

No comments to make

17. On the next mission

This was the last mission.

18. Recommendation for Next Mission

There are plans to do a further “The Farm Production, Food Security and Nutrition Survey” (VAM) and a second Sierra Leone Integrated Household Survey (SLIHS) during 2009. Further support could strengthen the usefulness of these surveys.

19. Final reflections.

The consultant hoped that SSL would have access to 2007 VAM data for the training. He also did not realise how poor the 2005 VAM data was. Despite this there was very positive involvement of the participants greatly facilitated the work.

Appendix 1

Standardization of Concepts and Improving Quality of Questionnaires

Introduction

Two surveys with agricultural data are currently available in Sierra Leone and the VAM and SLIHS. Both these surveys suffer from defects in their design that limits the quality of data. Future surveys must attempt to minimize these problems.

Two key issues are

- the standardization of statistical concepts
- and improving the linkage between the questionnaires and the normal practices in the country

Standardization of Concepts

Crop Codes

The SLIHS used different crop codes in different sections and also has very confusing lists of crop codes. The VAM survey on the other hand uses crop names with no associated code. It also breaks down rice into five different categories. It is recommended that one fixed list of codes be used as a standard for all surveys and that this list be organized into crop categories such as Basic Food Crops, Annual cash Crops, Vegetables, Fruit Trees. The following table shows the codes used in Mozambique with some additions to take account of the crops grown in Sierra Leone.

As far as rice is concerned it does not seem necessary to list 5 different types. A distinction could be made between Upland Rice and Swamp rice. The an additional question could be asked concerning the type of environment in which swamp rice is grown

Food crops
1 Maize
2 Rice
3 Sorghum
4 Millet
5 Groundnuts (Large)
6 Groundnuts (Small)
7 Common Beans
8 Cowpea
10 Pigeon Pea
11 Irish Potatoes

12 Cassava
13 Orange fleshed sweet. Potato
14 Non-orange fleshed sweet. Potato
15 Green Beans
16 Mung bean

Cash
20 Cotton
21 Tobacco
22 Sisal
23 Tea
24 Sugar Cane
25 Sun flower
26 Sesame
27 Soy beans
28 Paprika
29 Ginger

Vegetables			
311 Pumpkins	351 Onions	401 Water melon	441 Okra
321 Lettuce	361 Carrots	411 Cucumber	451 Cabbage
331 Garlic	371 Kale	421 Pepper	461 Tomatoes
341 Egg-plant	381 Peas	430 Chillies	490 Other

Cacao	501
Coffee	502
Cashew	503
Coconut	504
Oil Palm	505
Avocado	511
Pineapple	521
Banana	540
Guava	551
Orange	561
Lemon	571

Mangoes	621
Papaya	631
Tangerine	661

OTHER FRUIT TREES

531 Jack fruit	671	Grape fruit
581 Litch	681	Grapes
591 Apple	691	Granadilla
601 Chinese Apple	700	Other fruit tree, spec.
641 Pears		
651 Peach		

Units of Measurement of Crops

Another problem is the measures being used. The SLIHS has a long list of measure for crops and this list is open not linked to any particular crop. Also this list includes traditional measures such as bushels and three penny pan, English measures and metric measures. In practice only a few of the measures are actual applicable to any specific crop.

On the other hand the VAM has one and only one unit for measuring production. It is not clear whether this is too limited and some respondent had difficulty fitting this single unit of measurement.

It is recommended that SSL adopt a short fixed list of measurements and confirm the conversion factors between these measures and a standard unit of kilogrammes.

Unit of area

Unit of area is also confusing. SLIHS includes this question

“What Is the Unit In Which Plot Areas Are Recorded?”

- 1= Cropping Rate e.g. Bushels
- 2= Acre
- 3= Rope
- 4= Other.....
(Specify)”

It is not desirable to use an open ended value “other” in a key variable such as area.

On the other hand VAM records area only in Acres. It was noted however that the yields reported in the 2007 Report included a number of values far higher than are realistic.

The consultant observes two problems.

- Farmers often do not have a good idea of the area of their fields
- Apart from rice it is common to grow crops in association or intercropping

It is recommended that Agricultural Surveys in Sierra Leone adopt a practice of doing some area measurements and also consider ways of estimating the distribution of crops under intercropping

The consultant is attaching information in English from Mozambique that covers these issues.

Improving the linkage between the questionnaires and the normal practices in the country

In Sierra Leone as in many countries in Africa the system of traditional agriculture raises number of measurement difficulties. There are two growing seasons, intercropping and many irregular fields of unknown size. Some crops, notably cassava and sweet potatoes are harvested over an extended period. Certain crops are processed on the farm. This means that the quantities may depend upon the specific state of the crop. For example a large part of the cassava crop is converted into gari and generally sold in this form.

Detailed discussion of how to deal with these issues will not be discussed here, but consideration needs to be given to the following:

- Annual production needs to take account of the separate growing seasons
- There is a need to objective field measurement
- There is a need to estimate crop distributions with intercropping
- Methods must be adopted to measure continuous cropping
- When estimate production, sales etc the state of the crop must be specified or in some cases it may be necessary to measure the crop quantity using different states

Appendix 2

List of Items Needed for National Accounts

Crop Information Basic Crops

<i>Item</i>	<i>Needed for NA</i>
<i>Area Cultivated</i>	Yes
<i>Total Production(- seeds)</i>	Yes
<i>Yield</i>	Yes
<i>Payment in kind to land owner to labour</i>	?
<i>Own Consumption</i>	Yes
<i>Sales</i>	Yes
<i>Price per unit sold</i>	Yes
<i>Total exports</i>	Yes
<i>Input costs</i>	Yes
<i>Post Harvest Losses</i>	Yes
<i>Labour cost</i>	Yes

Crop Information Cash Crops

Item	Needed for NA
Area Cultivated	Yes
Total Production	Yes
Yield	Yes
Sales / Income	Yes
Price per unit sold	Yes
Total Exports	Yes

Crop Information Fruit Crops

Item	Needed for NA
Number of HH owning	
Number or area	
Sales or Production	

Price?	
--------	--

Crop Information Vegetables

- *Number of HH producing*
- *Area Cultivated?*
- *Sales or Production*
 - Suggest Sales better measure than Production
- *Price?*

Livestock

Item	Needed for NA
Number of HH owning	
Stock	
Births	
Purchase or Given	
Total lost	
Total sold	
Total consumed	
Price of sales	
Vaccines	
Imports of livestock	

Tracking livestock numbers

Household	National
Stocks t0	Stocks t0
+Births	+Births
+Purchases and gifts	+Imports
-Deaths (losses)	-Deaths
-Own consumption	-Consumption (own)
-Live sales	-Consumption (commercial)
(sale to slaughter house)	-Exports
Sale as meat	
Stocks t1	Stocks t1

Farming Population and Land

- *Number of Holdings*
- *Total Area Cultivated*
- *Family members working on Holdings by gender*
- *Employees by gender*

Inputs

- *Tools/Equipment*
 - Capital
 - Capital formation
- *Investments in fertilisers, pesticides, vaccines, seeds etc*
- *Credit (cost of credit)*
- *Rental of Machinery*

Additional Information Provided

In additional 5 files are attached. There are two power point presentations. The first one of these sets out the work programme and the second was a presentation given at te end of the mission.

There are four SPSS databases covering Livestock and Basic Food Crops for VAM and SLIHS

Finally there is an excel file showing the results obtained.