Jour. Ind. Soc. Ag. Statistics
51(2&3), 1998 : 303-314

Africa—A Fertile Ground for Agricultural Statistics Research

O.P. Kathuria[1]
Institute of Statistics and Applied Economics,
Makerere University, Kampala, Uganda

SUMMARY

Area estimation by measurement of fields with prismatic compasses and yield estimation through crop-cutting method are increasingly becoming unaffordable by African countries as these methods are time consuming and expensive and the means of transport and communication in rural areas are poor. A household being the ultimate unit of sampling in Africa, the preparation of sampling frame which could be used in agricultural sample surveys over a period of time is the single most difficult problem in Africa. A national master sample has been prepared in most countries which serves as the sampling frame but its preparation is expensive and it needs revision every 4-5 years. Continuous, mixed and relay cropping are other problems which need to be tackled in area and yield estimation surveys.

Keywords: Area and yield estimation, Sampling frame, National master sample, Continuous, Mixed, Relay cropping.

1. Introduction

The importance of comprehensive statistics of a reasonable quality in formulating and evaluating development plans needs no emphasis. The shortage of reliable statistics and of other basic information is one main obstacle to effective agricultural development planning in African countries. As agriculture is the predominant activity in most of these countries, high priority has to be given to the development of agricultural statistics which constitutes the basis of all agricultural development plans.

Agricultural statistics cover a wide range of topics related to agriculture. These may be divided into two broad categories, namely, (i) basic agricultural statistics and (ii) current agricultural statistics (FAO [2]).

Basic agricultural statistics refer to information on structural characteristics which do not change much from year to year e.g., land use patterns, distribution of holdings, farm practices, use of agricultural resources e.g., manpower, land

1 Commonwealth Expert in Sampling and Agricultural Statistics.

and water, agricultural implements and machinery etc. Such data are usually obtained from a census of agriculture of from administrative records. Since the changes in structure are not generally very rapid, the information need not necessarily be collected on a yearly basis.

Current agricultural statistics are data on agricultural activities which continue from year to year and reflect current changes in agricultural characteristics e.g., planted area, production, utilization, price, inputs used etc. Current agricultural statistics are collected through surveys on a continuous or seasonal basis, possibly several times during the agricultural year. Since agriculture in broad sense also includes livestock and fishery, statistics on livestock numbers, livestock products and fisheries may also be considered to come under the purview of agricultural statistics. While data on total count of livestock population and its composition according to different species are generally collected through a livestock census and would thus form part of basic agricultural statistics the inter-census estimates of livestock numbers, estimates of production of various livestock products etc. form part of current agricultural statistics and may be required for each year or for each season within a year. Such data can be collected only through properly planned and executed sample surveys.

2. Household Surveys as the Major Source of Agricultural Statistics

In Africa, household surveys are the major source of comprehensive quantitative and reliable information on agriculture and indeed on several other aspects such as income, consumption and expenditure, labour force, demographic aspects, health and nutrition and on other household economic activities. Census of agriculture and agro-economic/farm management surveys are the other major sources of agricultural statistics but these are *ad hoc* sources not conducted with any regularity. In recent years crop forecasting and early warning units have been set up in several countries which collect agricultural statistics on sample survey basis and correlate it with meteorological data. A household is a unit of enumeration of enquiry in all these surveys. Thus, for example, in a sample survey for estimation of production of crops, estimation of both area and average yield begins at the household level.

In many African countries, household surveys were initiated in the 1950's. The main subjects covered in these surveys were income, expenditure, consumption and demographic characteristics of households. The frequency of these surveys increased significantly during the 1960's following attainment of independence and in response to increased demand for data for planning and policy formulation in the immediate post-independence period. In the last two

decades, household surveys have come to play a central and strategic role as the major source of National Statistics in Africa. The household surveys in Africa are being implemented in the form of an Integrated Household Survey Program (IHSP) under the auspices of National Household Survey Capability Program (NHSCP) which was established in 1979 in pursuance of a resolution of the Economic and Social Council of the United Nation to be implemented as an inter-regional program in developing countries of the world. As of 1991, the NHSCP was operational in 23 countries of Africa namely, Algeria, Angola, Benin, Botswana, Cameroon, Ethiopia, Ghana, Kenya, Lesotho, Malawi, Mali, Mauritania, Morocco, Mozambique, Niger, Nigeria, Rwanda, Sierra Leone, Sudan, Swaziland, Tanzania, Zambia and Zimbabwe (United Nations [10]). Agriculture constitutes a core component of IHSP in several countries to be included in the survey every year. For example, the Rural Agricultural Survey in Nigeria is one of the core surveys. In Lesotho, Zimbabwe, Zambia, Malawi, Ethiopia etc. the Annual Household Survey of Agriculture is the core survey of IHSP.

3. Types of Data Needed

Food security at household, regional and national levels has been and continues to be the major concern in all African countries. Three main components of food security are production, distribution and consumption. Of these, food production constitutes the core since distribution and consumption depend mainly on what is produced in the country. Therefore, the first and foremost requirement in developing national statistical systems is the collection with adequate degree of timeliness and accuracy data of main crop areas, average yield and their production and to maintain the data series over years.

For development planning purposes, the production data including data on crop areas and yields are not sufficient if only given at the national level. The statistics have to be disaggregated in various ways to satisfy the needs of the different users. They have to be disaggregated (i) by region (physical, administrative, ecological etc.) to permit regional planning and monitor distribution and movement of food stocks in different regions of the country; (ii) in time: for each agricultural cycle separately showing the succession pattern within the year as well as the crop rotation from year to year; and (iii) according to its utilisation, either as self consumed and marketed or as seed, feed, stocks and industrial uses etc. Moreover other production related data on the volume of existing stocks and the rates of wastes (pre-harvest, harvest, transport, storage etc.) are also needed. Data on prices, input use, consumption pattern at household level are also required but their collection is beset with less

difficulties. Governments in Africa also need crop forecasts for improving national capability with regard to Early Warning Systems.

In the livestock sector, statistics are needed on livestock numbers, products, feeds and fodder required and available etc. Data on costs of production and marketing of livestock products are important for development of animal industry.

Fish is a major item of food and a rich source of protein. Africa is abundantly endowed with rivers, and fresh water lakes - many of them among the largest in the world. A large segment of fishermen population earn their livelihood by catching fish from rivers, swamps, lakes and from marine waters in the coastline countries. Fishery resources in Africa have still not been fully harnested. Estimation of fishery potential, areas exploited and available for exploitation, fish yields obtained are important for development of fishery economy and fish industry.

4. Problems of Data Collection in Africa

The various issues related to agricultural statistics enumerated in the preceding section are no different from those encountered in other countries of the world. For example India has been grappling with these problems for several decades with varying degrees of success. What characterises them as distinct in Africa is the historical, geographical and cultural background of this continent which pervades the way of life of the people here. Some of the major problems in collecting reliable agricultural statistics in Africa are enumerated below:

4.1 Preparation of Sampling Frame

Preparation of sampling frame which could be used in agricultural sample surveys over a period of time is the single most difficult problem in Africa. Over large parts of Africa, agriculture is characterised by small-holders for production. In large parts of several countries, ownership of land lies with the village or tribal chief and the farmers till the land as tenant farmers with land-holding changing from year to year. In this scenario, the only thing that is relatively stable over time is the household which was been adopted as the ultimate unit of sampling in all household surveys including agricultural surveys. Therefore, a major problem in determining the sample structure for household surveys in Africa is to create a multiple purpose sampling frame capable of responding to various data needs, which should be cost effective in terms of funds and manpower used and should be manageable and controllable and useable for a period of time. Many countries in the region have constructed

a National Master Sample (NMS) to provide a unified framework for carrying out household surveys including agricultural surveys. A NMS is a sample consisting of units representing the population from which further samples are selected for different surveys or survey rounds (United Nations, 1986). A NMS is usually a two or three stage sampling design. In a two stage NMS, primary sampling units (PSU's) are Enumeration Areas (EA's) based on population census. The sampling design of an agricultural sample survey is a stratified two stage design with administrative/geographical zones as strata, Enumeration Areas or groups thereof as PSUs and households as second stage units. A sample

Table 1. Sample structure for agricultural surveys in selected African countries

Country	Stratification Nature	No. of strata	Sampling stages	Primary sampling units (PSUs)	No. of sample PSUs	Overall sample of agricult. house-holds
Ethiopia	Administrative units	102	2	Farmers' Associations	870	12,500
Kenya	Administrative districts	32	2	Enumeration areas	768	9,300
Sierra Leone	Administrative districts	12	2	Enumeration areas	120	1,731
Lesotho	Ecological zones	24	2	Enumeration areas	80	–
Botswana	Agricultural districts	22	2	Survey blocks	650	4,114
Tanzania	Agro-ecological zones	150	3	Villages	50	3,106
Malawi	Administrative units	110	2	Enumeration Areas (EAs)	768	9,300
Zaire	Rural Communities	–	2	Villages	50	–
Cape Verde	Agricultural practice	–	2	Enumeration districts	108	3,377
Zimbabwe	Administrative province	196	3	Divisions	150	2,400

- Information not available

Sources: Household surveys of agriculture in Africa, United Nations, 1991.

of Enumeration Areas is selected with probability proportional to rural (or small holder farming) population or number of rural households. A NMS is used for a certain period of time (usually 4 or 5 years) after which it is revised or updated. Major problems in use of NMS for agricultural surveys are (i) Enumeration Areas based on population census are not selected on agricultural criteria. (ii) construction and maintenance costs being enormous the NMS is not updated periodically resulting in non-response, (iii) a NMS based on population census may not provide representative sample for crops which may be concentrated or thinly distributed in certain areas resulting in large sampling errors and (iv) in many places, people move from area to area, and in some countries census estimates of proportions of farmers in EA's proved almost useless a few years later. Table 1 gives the sample design and sample structure of agricultural surveys in selected African countries.

4.2 Area and Yield Measurement Surveys

The production of a crop is generally estimated as the product of area under it and its average yield. For accurate estimation of production, it is essential that both area and the average yield should be known accurately. Based on past experience in India (Mahalanobis [5], [6]; Panse [7]; Sukhatme [8]) and elsewhere (Zarkovich [12]) it was observed that accuracy of area and yield estimates could be ensured only by following objective methods of estimation, viz, (i) measurement of fields for the estimation of area and (ii) crop cutting from small areas of specified shape and size and the subsequent threshing, drying and weighing for the purpose of the estimation of yield. Since cadastral maps were available for more than 85 percent of land area in India and a revenue agency existed in the villages who maintained the records of land utilisation, estimates of area under various crops could be obtained fairly accurately in these areas. In the remaining 13 percent area in India, estimates of area under the various crops were obtained through sample surveys by actually measuring the sample of selected fields. Similarly, for estimation of average yields of major crops an elaborate administrative system and field agency was evolved for undertaking work of crop cutting experiments on randomly located plots in sample of fields selected by adopting a stratified multi-stage sampling design (ICAR, [3]). The above system with minor modifications exists today for estimation of area and average yield of all important crops in India. In this respect India can rightly be called a developed country with a long history of statistical work.

For a long time in many African countries objective method of area and yield estimation continued to be followed. Prismatic compasses, measuring tapes and programmable calculators are used for measurement of area of all the fields

owned by the selected households growing the crop for which estimate of area is required. Often the fields are small in size, irregular shaped with no proper boundaries involving considerable time to take the measurements.

The method (FAO, [1]) involves the following steps:

(a) determining plot boundaries;

(b) going round each plot in a clock-wise direction to determine how many corners there are (starting from the point designated by letter A, other corners are designated by letters B, C etc.);

(c) using measuring tapes, poles and compasses, lengths between different corners as well as forward and backward bearings are measured and recorded;

(d) on the basis of the measured lengths of all sides of the field and the bearings, a map of the field is drawn to scale on a piece of paper; if the field does not 'close', that is, if the last side measured does not automatically end at A (this is due to unavoidable errors in taking measurements), the distance between the 'original A' and the 'new A' where the last line in the map will end is called error of closure or closing error;

(e) if the error of closure turns out of be less than 5 percent, then the field measurement results are accepted. If it is 5 percent or more, the field will be re-measured.

(f) for each field whose measurements are accepted, the calculation of area is done by using programmable claculator.

It is obvious from the above that the method of area measurement is complex, expensive and time consuming and requires a skilled enumerator to record measurements. So long as production estimation surveys involving area and yield measurements were of *ad hoc* nature and were done only periodically, the above method of area measurement could be found workable. In recent years with food security becoming the national objective in many countries requiring estimates of production at national, regional and subregional levels, it has become necessary to conduct production estimation surveys annually on a regular basis. Many countries now find it practically impossible to do area measurement surveys using compasses and programmable calculators.

The estimation of average yield is done by crop cutting method by locating one or two random plots of a given shape and size in the field(s) of selected households. The usual procedure of driage experiments based on a subsample

of harvested yield is adopted. Table 2 gives the shape and size of crop cutting plots in selected African countries.

Table 2. Shape and size of crop-cutting plots for crops measured in selected African countries

Country	Shape of YMP	Size of YMP (Sq. m)	Crops measured
Tanzania	Square	25	maize, paddy, sorghum and bulrush millet
Malawi	Square	50	maize, groundnut, millet, wheat, pulses, sorghum and sunflower
Sierra Leone	Square	49 12.25	upland rice swamp rice
Zambia	Circle	2.5	maize
Lesotho	Square	25	–
Zimbabwe	Row	4	maize
Nigeria	Square	100	sorghum, yam, millet
Kenya*	Square	–	maize

YMP: Yield measurement plot

* No regular crop-cutting measurement takes place. In 1987, an experimental crop cutting survey was crried out on maize crop.

Source: Household surveys of agriculture in Africa-A methodological study, United Nations, 1991

Data collection through objective method (physical observation) requires honesty and integrity of a very high order on the part of field investigators and supervisors. In surveys of repetitive nature involving several rounds of field work, which all agricultural sample surveys are, after one or two rounds of field work of a part of the selected sample it is not too difficult for an intelligent field investigator to anticipate the likely responses during subsequent rounds or of the similarly placed remaining sampling households. The stakes for "cooking up" the data become high particularly under adverse logistic and climatic conditions. In India, the success of objective method of area and yield estimation was primarily due to the fact that the country had a long established land record system in 87 percent of the area with local revenue official virtually residing in each village who knew everything about the village and maintained records of areas under various crops in each season. India could therefore establish a permanent field agency in each state for crop-cutting work. The situation is no longer the same now in India. The local revenue official in the village has acquired considerable political clout and can therefore afford to ignore area enumeration with impunity. In fact, the scale of area enumeration

has gone down considerably in recent years. With high travel costs and budgetary constraints restricting travel by supervisory staff, conditions are favourable now for taking crop cuts. The production estimates may not be as objective as they appear to be.

Measurement of crop areas and crop cutting experiments are season bound activities confined to a few weeks in each crop season. Even in areas growing two crops in a year, area enumeration and crop cutting work involve intensive work during short periods close to when the crop is ready for harvest requiring a large number of field investigators to be deployed and also intensively supervised. African countries can not afford to have a permanent field agency for data collection. Means of transport and communication are extremely poor. The level of statistical training is generally low. Inflation rate being high the general economy in most of the African countries does not remain stable beyond 4 - 5 years. Therefore even with *ad hoc* field staff cost of data collection through objective method is extremely high in Africa. That is why in the past, many agricultural surveys were abandoned on withdrawal of donor or international support. Thus African countries find it difficult to afford objective method for data collection.

This has led many countries and international agencies to explore farmer's interview to get accurate estimates of areas planted and average yields. In this context the most significant study done so far is by Verma, Merchant and Scott [11] on maize crop in five African countries, namely, Benin, Central African Republic, Kenya, Niger and Zimbabwe in which crop cutting and farmers interview methods were compared with complete harvest from the selected fields. The study revealed that:

(i) the farmers generally over-estimated the planted area,

(ii) farmers' post-harvest estimates were close to actual production and superior to objective method of crop cutting,

(iii) farmers' preharvest estimates were also good for predicting production levels but had high variance and

(iv) crop cutting method over-estimated average yields by about 30 percent.

The study, however had several weaknesses, namely,

(i) the same enumerators interviewed the farmers, carried out crop cutting work and measured the total harvest of the sample fields,

(ii) the sample size in each country was small leading to high within-country variability and

(iii) the selected farmers were fully aware and willing participants at each stage of the study and thus independence of the methods was lost.

Unpublished results of a study done by India Agricultural Statistics Research Institute, New Delhi on wheat in Lucknow and Aligarh districts of Uttar Pradesh and Ludhiana district of Punjab State during 1988-89 and 1989-90 revealed close agreement between farmers' post-harvest estimate of average yield, yield obtained through crop cutting and the whole field harvest. A study done for FAO on maize crop in Zambia (Kathuria [4]) revealed that the farmers generally tended to over-estimate the areas planted when compared with areas measured with prismatic compass. It is necessary to do more experimentation on farmers' interview method keeping in view the fact that intelligent farmers may tend to under or over estimate the areas and average yields suited to their interests.

4.3 Problems of Continuous Harvesting, Mixed Cropping and Relay Cropping

Cassava is a staple food crop grown in many African countries. It is not completely harvested at any particular time of the year but may be taken out from the ground when required or used as a "reserve crop" in emergencies. In some countries it may be infeasible at any moment to identify the area planted with cassava for sampling, since at any moment some has just been harvested in a continuous cropping sequence.

Mixed cropping as subsistence need is a widely used practice by small farmers in many African countries. In a 'mixed cropping' situation several crops are planted at the same time in the same field with seeds mixed or planted in rows in some given ratios. Harvesting of the mixed crops may not always be at the same time requiring repeated visits by the enumerators to estimate yields.

'Relay cropping' is a situation where one crop is put into the ground before another has been harvested. Thus in some countries sorghum is interplanted while early millet is growing and late millet replaces early millet before sorghum is harvested. In such situations, estimates of cropped area would depend on the time at which measurement is made.

Preparation of sampling frame, sample selection and measurement of area are problems which need to be addressed to in the above situations. Each of the above situations needs different approach for average yield estimation. When and how often the data on yield should be collected in case of continuous and relay cropping are important points for consideration.

5. Limitations of Sampling Designs Used and Some Suggested Remedies

It is obviously clear from the preceding section that the IHSP is a multi-subject survey with the same sampling design used for studying various characteristics in different modules of the survey which has a bearing on the efficiency of the design. Since the surveys were primarily planned to study socio-economic characteristics of households such as income, consumption and expenditure there generally is oversampling of urban stratum. Over the time, many subjects were included in the surveys such as demography, housing, health, nutrition, literacy, labour force etc. Production was estimated as what was sold and consumed. Later on, production estimation through area and yield estimation was added as a separate module taken up each year but the basic design of the survey remained unchanged. Thus, the precision of the agricultural estimates is often affected by inadequate sampling of rural and semi-urban strata.

Households being the units of sampling in the IHSP in Africa, it is necessary that their numbers should be accurately known which are used as weights for building up estimates or where population count is taken as measure of size for pps selection. Often these counts are out of date by several years as no revision of NMS takes place leading to inaccurate use of weights and probability measures. The problem becomes all the more serious in area and yield estimation surveys when many sample households do not happen to be growing any crops or when inaccurate weights are used for building up estimates of area and production of crops.

Area-based frames using remotely sensed data have been suggested and used on experimental basis in some countries as an alternative to household based frame for agricultural surveys. The approach was abandoned as it was not only extremely costly to apply in view of small sizes of fields in Africa but also it required skills in aerial photo-interpretation and other expertise which is not available in these countries.

As stated earlier, Enumeration Areas (EAs) or groups thereof are taken as the PSUs in most of the surveys in Africa. Generally speaking, there is no permanent field agency and the data of a survey are collected by short term appointment of field staff many of whom may not be familiar with region in which they are posted for data collection. While no country in Africa can afford a large field organisation on permanent basis it may be possible to create a permanent cadre of core field staff in each EA preferably from within the region who besides being responsible for survey data collection should periodically update the frame of households and maintain and update the records of some key characteristics in respect of households in the EA under their jurisdiction. For agricultural households, the characteristics on which records could be maintained and periodically updated for each household are size of holding, crops grown, measured area under each crop, livestock owned etc. The sampling

frame can thus be always maintained up-to-date at reasonable costs. Such a step may ultimately lead to preparation of a separate sampling frame for agricultural surveys. Thus an optimum sampling design and an optimum sample size can be selected for estimation of area and production of crops with appropriate weights for grossing up of estimates. This may also obviate the need of taking measurements of all fields every time in crop area estimation surveys.

Mixed cropping and intercropping are common phenomenae in many countries of the world. Several approaches have been tried for apportioning area under each crop and for estimating average yield (FAO [1]). Methods are also available for estimating production of crops involving multiple harvests.

ACKNOWLEDGEMENT

The author wishes to thank the Commonwealth Secretariat, London for the opportunity provided to him to work at Institute of Statistics and Applied Economics, Kampala which enabled him to write the present article.

REFERENCES

[1] FAO, Rome, 1982. Estimation of crop areas and yields in agricultural statistics. *Economic and Social Development Paper No. 22.*

[2] FAO, Rome, 1983. Use of household surveys for collection of food and agricultural statistics.

[3] ICAR, New Delhi, 1951. Sample surveys for the estimation of yield of food crops.

[4] Kathuria, O.P., 1995. A Report of area and yield estimation surveys in Zambia.

[5] Mahalanobis, P.C., 1944. On large-scale sample surveys. *Phil. Trans. Roy. Soc.*, **231B**, 329-451.

[6] Mahalanobis, P.C., 1946. Recent experiments in statistical sampling in the Indian Statistical Institute. *J. Roy. Statist. Soc.*, **109**, 329-378.

[7] Panse, V.G., 1954. Estimation of crop yields. FAO, Rome.

[8] Sukhatme, P.V., 1947. The problem of plot size in large scale yield surveys. *J. Amer. Statist. Assoc.*, **42**, 297-310.

[9] United Nations, New York, 1986. National household survey capability programme.

[10] United Nations, New York, 1991. Household surveys of agriculture in Africa - a methodological study.

[11] Verma, V., Merchant, T. and Scott, C., 1988. Evaluation of crop-cut methods and farmers reports for estimating crop production- results of methodological study in five African countries. Longacre, London.

[12] Zarcovich, S.S., 1966. Quality of statistical data. FAO, Rome.