

ISSUES RELATING TO THE ASSESSMENT OF THE QUALITY OF STATISTICAL DATA IN DEVELOPING COUNTRIES

Report by FAO¹

1. At the 35th session of the Administrative Committee on Coordination (ACC) Subcommittee on Statistical Activities held in Vienna on 18-20 September 2001 the Committee decided that issues relating to assessment of the quality of statistical data and recent experiences in developing countries needed to be summarized in order improve the coordination of statistical activities.

2. This report highlights issues faced by international statistics offices with examples being taken from recent FAO Statistics Division experiences in developing countries in all regions. Whilst this report focuses on FAO Statistics Division's experiences of data quality in developing countries, many of the issues and experiences of the FAO Statistics Division are common with those of other international statistics offices in collection and processing of statistics in their particular field. Assessment of data quality involves many very subjective judgements which can be argued from various methodological perspectives. As statisticians we are quite familiar with hypothesis testing and in the case of data quality it is not as straight forward as accepting or rejecting any null hypothesis on changes in data quality. The situation is complicated by the possibility of evaluating data quality on either, cells, quantities or values. It is natural for data to come to international organizations with some delay. Prima facie evidence of less official data in recent years cannot always be attributed to deterioration in quality but can be the time component/lag.

A. Data Quality and flow

3. The general quality and flow of agricultural statistical data from developing countries leaves much room for improvement. Data quality problems are not new, but have not often been the focus of initiatives in developing statistical capacity. Improvement in data quality has often been seen as a by-product of other programmes. This report focuses on some of the concepts identified by Eurostat (2000) in their definition of quality in statistics as framework for discussion.

Reporting

4. There has been little improvement in the number of developing countries reporting and timeliness over the past decade. Generally, about 50% of developing countries in all regions are regular and timely reporters of agricultural production and trade statistics. About 30% of developing countries are sporadic reporters with the remainder (about 20%) reporting once in the past decade or not at all. This is not to say that FAO does not have up-to-date agricultural production and trade statistics: the sporadic and non-reporting countries (such as small island states) mostly have little or no agricultural production or trade.

¹ Paper prepared by Robert Mayo, Statistics Division, FAO (Robert.Mayo@fao.org)

5. The African region has been a major concern for many years with the level of country reporting deteriorating. In 1999, agricultural production statistics were requested from all African countries, but only about 50% of replies for agricultural production statistics and about 30% for agricultural trade statistics were obtained on time. Ten countries in this region consistently failed to report agricultural production statistics over the past five years and a few for the past ten years. However in the 2000 reporting year, a dramatic reporting turnaround was achieved. For 2000, 75% of the African countries reported agricultural production and 67% reported agricultural trade statistics. This dramatic improvement was primarily due to a particularly enthusiastic FAO statistician visiting and contacting countries directly during the year. The non-reporting countries in the African region also generally have little or no agricultural production or trade as well as poor infrastructure.

Completeness

6. There is considerable variability in the completeness of statistics reported from developing countries. This applies to both the completeness of time-series reporting as well as item levels and variables on questionnaires. The on-going problem of sporadic reporting, results in very incomplete time-series information being available. Problems of completeness are particularly evident for FAO's data collection on livestock production, agro-industrial products, production of some staple food commodities, agricultural inputs and prices. This situation is common among developing countries in all regions.

7. The percentage of official primary production data as a proportion of all primary production data for developing countries varies considerably from country to country and from region to region. During 1999, in the African region about 30% of the primary production data (crops and livestock) was official with the remainder being estimated or coming from semi or unofficial sources (this has improved to about 38% in 2000). It must be remembered that subsistence agriculture in some African countries can be as high as 90% of production of certain crops and livestock. For the Asia and Pacific region about 50% of the primary production data was official with the remainder being estimated or coming from semi or unofficial sources. Over 90% of the value of production and trade can be classified as official for major crop groupings. For the Latin American/Caribbean region about 33% of the data of the primary production data was official. There has been little change in the level of official data during the past decade. For other regions the proportion of official data available remains relatively stable, but still at low levels compared to those of developed countries. The only encouraging sign was in the African region. It however remains to be seen if this recent improvement can be sustained in the African region. The importance of the human element in rigorously pursuing data and data sources cannot be overstressed.

8. FAO Statistics Division has started regular monitoring of country data in the FAOSTAT database for completeness. The Agricultural Bulletin Board on Data Collection, Dissemination and Quality of Statistics project (ABCDQ)² is a significant development in assessing data quality of FAOSTAT. Country data completeness in the FAOSTAT database has been assessed to provide the user with indicators of the availability of official data for the years 1996 to 2000. The indicator (presented as barometer) is calculated for crops and livestock production as well as agricultural trade (export and import) statistics. This monitoring device has proved useful in summarizing data completeness and highlighting countries and regions and that require further attention.

² *The Agricultural Bulletin Board on Data Collection, Dissemination and Quality of Statistics project (ABCDQ)*, <http://faostat.fao.org/abcdq/abcdq.jsp>, FAO Statistics Division.

9. Evaluation of data quality is often complicated or not possible due to the lack of adequate metadata on such items as reliability (survey methodology, sample design and size, sample error estimates, sample bias etc.). Problems relating to the units of measurement, or classification are common with agricultural production and trade statistics.

10. Many developing countries have a large informal economy and usually missed in the data collection exercise. A particular problem faced in agricultural and other trade statistics is that of active tariff avoidance which results substantial undercounting and misclassification and in some countries grossly distorting their agricultural commodity trade statistics. Commonly we have to adjust official figures to reflect the informal sector, which we can gauge for example by exports of cash crops. The resulting adjusted data moves us from using an official figure to an unofficial estimate, but improves data quality from our perspective. Data users (both novice and experienced) however see adjusted data (unofficial estimates) as reflecting lesser quality. This highlights one of the important roles international statistics offices can make in improving data quality.

Coherence

11. There are on-going problems related to the coherence of statistics from developing countries. On a day-to-day basis problems are encountered in the use of different definitions, classifications and related methodological issues. A recent example encountered by FAO involved four different statistics reporters sending data for the same series for one country on the same topic for the same reference period. There was little coherence in the data. The resources required to solve these problems are considerable. Many inherent problems related to the coherence of statistics are becoming apparent as we obtain more detailed metadata from developing countries and this situation is likely to continue for some time.

Accessibility

12. Another major challenge facing international statistics offices is improving the flow of statistics from national statistics offices to the international offices. Recent experiences of FAO Statistics Division highlight the need to improve the flow of statistics. Over the past few years FAO Statistics has been systematically sending statisticians to the national statistics offices to assist in training and other issues. It was surprising to find that many of these offices had data that we were unaware of and had not been sent when requested by the standard questionnaires/requests. Subsequently, this data was available for inclusion in our databases. This highlighted the problem of data being available at the country level in various formats, but the traditional methods of obtaining it not being successful.

13. New technologies for transferring data such as ftp sites and virtual questionnaires are improving the flow of statistics from national statistics offices to the international offices, but more work and support is needed. The flow of data will be a major focus of the new version of FAOSTAT (FAOSTAT2)³. CountryStat, a scaled-down version of the FAOSTAT2 application, which would provide countries with functionality to compile, validate, analyze and disseminate their national data. Outputs from CountryStat could then be easily loaded into FAOSTAT2 for further dissemination through FAO's web site, publications and CD ROMs. The development and publication of an XML-based standard for the exchange of food and agricultural statistics would also facilitate improvements in the quality of FAOSTAT2 data. This standard would be

³ *Modernization of FAOSTAT*, <http://www.fao.org/es/ESS/faostat1.htm>, FAO Statistics Division.

distributed to member countries and other external partners who would be requested to provide their data in files formatted according to the published standard.

B. Contributing causes

14. There are many possible reasons for the problems of data quality highlighted in the previous section. The following section summarizes some of the likely contributing causes:

- For some of the countries where no response is obtained, this may be so because of internal/external conflicts or natural disasters – known to be the main reason currently in five African countries.
- No clearly defined national/country focal points responsible for the reporting of statistics to international organizations and the lack of a transparent division of responsibilities between the various national ministries and national statistical agencies.
- Complicated administrative channels for contacting the national statistical agencies (frequently through the Ministry of Foreign Affairs) has a negative impact on the speed of communication with these institutions and the quality and completeness of the returns.
- The relationship between international and national offices is often one-sided and involves data requests.
- Underlying weaknesses in the technical capacities of national statistics institutions in many developing countries to meet data requirements of users.
- Poor co-ordination of scarce resources at the national level, especially of donor assistance.
- Programmes reflecting the interests of donors rather than those of the country.
- Non-sustainability of externally funded programmes.
- Lack of co-ordination: lack of understanding and co-ordination between statistical agencies producing the data (data producers) and offices undertaking economic analysis, planning and decision-making (data users).
- Lack of government commitment: a key constraint in many countries.
- Lack of resources: government priorities often exclude support to statistical systems.
- Technical and methodological constraints: the complex environment of the agriculture sector in Africa (the bulk of African agricultural production comes from small traditional farmers using a wide variety of agricultural practices), coupled with a lack of documented and factual information on the farming practices used, presents a particular challenge for data collection. It also results in the high cost of data collection in Africa.
- Role of technical cooperation; too much emphasis has been placed in the past on 'ad hoc' and uncoordinated interventions over a short time period aimed at addressing a particular data gap rather than long term capacity building and the establishment of sustainable systems.
- High turnover of statistical staff in the African and Latin American/Caribbean regions.
- A cost and little benefit to developing countries in filing out questionnaires for international statistics offices. The lower ranking official responding to international data requests (with a high risk of errors being made) gets very little “pay-back” for their considerable effort.
- Developing countries are “bombarded” with data requests. Often the data

requests by international agencies are not well thought out, and they want everything. The international agencies rather than asking for sorted, targeted data, ask for everything and want to sort it out back in their offices.

- Data requests by international agencies are not relevant for the current policy problems of the developing countries and hence are not useful to them.

C. Responses to address these data quality issues

15. Many of these problems related to data quality are not new. Considerable resources have been used to address these problems over the past fifty years, but they still exist. The current activities/responses being undertaken by FAO Statistics and other international organisations are a response to problems related to data quality in developing countries:

Technical Documentation Support:

- Handbooks on methods;
- Compilers manuals;
- Concepts and definitions;
- Data quality assessment frameworks;
- Classification frameworks and systems.

Technical Support:

- Training activities on various conceptual and methodological aspects of statistics in an effort to strengthen national capacities – FAO’s regularly hold training sessions include National Statistics Systems, Censuses of agriculture, Supply Utilisation Accounts and Food Balance Sheets, Economic Accounts for agriculture, Area Frame Sampling, etc.;
- Field projects;
- Joint projects such as FAO/World Bank/USDA Initiative for strengthening agricultural statistics systems in Africa;
- PARIS21 initiatives;
- Technology transfer.

D. Concluding remarks

16. As a result of many of the activities mentioned in the previous section, many developing countries (especially in the African region) have been able to produce some baseline structural data on agriculture, and annual agricultural production data following standard international recommendations. Despite these positive results, the overall picture is still far from optimum. The demands for statistical data at the international level far out-reach the data that is available.

17. In the quest to be balanced, a mention should also be made about data from developed countries. It is not uncommon, to find that developed countries consider themselves above international statistics agencies and do not send data. They refer us to their website (which we are required to pay to access) or their publication. They want international statistics agencies to adopt their classifications and standards rather than adjust to the international standards as the developing countries do. Many times, data that was available on individual developed countries previously, is now not available.

18. It is clear that many of these issues relating to statistical data quality are going to continue in the near and not so near future. There however are many success stories and we should not lose sight of the possibility of repeating these successes in the future. While the data produced has proved useful, today's demands for data by users, including FAO, go far beyond what is available in most countries.

Points for discussion

- Do we need to target specific issues related to data quality for developing countries?
- What short-term, medium-term and long-term plans are needed to achieve and sustain sufficient data quality in developing countries?
- What role should the IAMCSA take?
- Do we need to focus on the relationships building between international and national offices?
- Adoption by national stakeholders of a statistical development framework
Government awareness of the implications of the partnership approach and formal commitment and support to the implementation.
- Mobilisation of donors.
- Current data gaps are often met by funds from donor countries. What should be done for statistical topic areas or developing countries/regions that donors are not so willing to fund?
- Should countries be paid for data (processing etc.)?
- Should incentives for providing data be given to the respondents to international questionnaires?
- Is it worth looking at data collection as the priority?
- Is there anything that can be done regarding the problems of high turnover of statistical staff?
- Do we need to look at alternative/new methods of data collection and transfer?
- The role technology transfer in improving data quality.
- Data quality and metadata availability.
- International agencies should review their data requests carefully. Reducing the requests to a few which are relevant and concentrate on the quality of those. leave the rest to specialized and interested private sector or national governments.

References

Eurostat, *Assessment of Quality in Statistics*, paper presented at the UN/ECE Work Session on Statistical Metadata, Washington DC, United States of America, 28-30 November 2000.

Food and Agriculture Organization of the United Nations, Statistics Division, *Data quality assessment and status of agricultural statistics in Africa*, International Round Table Discussions on Current Agricultural Production and Trade Statistics, Meeting organised by FAO/University of Pretoria, Pretoria, South Africa, 23-27 April 2001.

Food and Agriculture Organization of the United Nations, Statistics Division, *Data quality assessment and status of agricultural statistics in Latin American and the Caribbean*, International Round Table Discussions on Current Agricultural Production and External Trade Statistics, Meeting organised by FAO/Oficina Nacional de Estadísticas, Havana, Cuba, 3 – 7 December 2001.

Food and Agriculture Organization of the United Nations, Statistics Division, *Agricultural data quality in Africa as limiting factor in the analysis of food supplies*, International Round Table Discussions on Current Agricultural Production and External Trade Statistics, Tunis, Tunisia, 24 – 27 June 2002.

Food and Agriculture Organization of the United Nations, Statistics Division, *An Overview of the quality of FAO's agricultural production, and trade statistics for North Africa and Western Asian countries*, Paper presented at the Workshop on Data for Food Security in the framework of an Integrated System of Agricultural Statistics, Cairo, Egypt, 10 - 20 February 2002.

Food and Agriculture Organization of the United Nations, Statistics Division, *Data quality assessment and status of agricultural statistics in Africa*, International Round Table Discussions on Current Agricultural Production and Trade Statistics, Meeting organised by FAO/ ENSEA, Abidjan, Côte d'Ivoire, 29 October – 2 November 2001.

Food and Agriculture Organization of the United Nations, Statistics Division, *An Overview of FAO's Statistics on Agricultural Production, and Trade for Asia and Pacific Countries*, Paper presented at the Asia and Pacific Commission on Agricultural Statistics (APCAS) meeting, Bali, Indonesia, 6 - 10 November 2000.