African Statistical Journal
Journal statistique africain

Analyzing Ranking and Rating Data from Participatory On-Farm Trials

Estimation of the Magnitude and Frequency of Floods in Uganda: A Regional Frequency Analysis Approach

Hypothèse de constance des coefficients techniques : Mythe ou réalité ?

Using Statistics to Assist Decentralization and Local Economic Development in Some West African Countries

Challenges of Statistical Infrastructure in Africa: A Systemic Approach to Statistical Capacity Development

Peer Reviews of African National Statistical Systems

Toward a Vibrant and Effectual Statistics Training Approach

CAST for Africa

Volume 10 – May / mai 2010
African Development Bank Group
Groupe de la Banque africaine de développement

Design, layout and production by Phoenix Design Aid A/S, Denmark. ISO 14001/ISO 9000 certified and approved CO_2 neutral company – www.phoenixdesignaid.dk. Printed on environmentally friendly paper (without chlorine) with vegetable-based inks. The printed matter is recyclable.

African Statistical Journal
Journal statistique africain

Volume 10
May / mai 2010

Table des matières

Points saillants du quatrième trimestre 2009 et du premier trimestre 2010 :

Contents

Highlights for the Fourth Quarter 2009 and First Quarter 2010:

Editorial

Nous souhaitons la bienvenue à nos lecteurs au volume 10 du *Journal statistique africain* (JSA). Ce volume commémore le 5ème anniversaire du journal. Il confirme la mission du journal pour servir de point de diffusion de la recherche et de partage d'information entre statisticiens et utilisateurs d'information statistique principalement dans la région Africaine. Nous voudrions remercier tous ceux qui ont soutenu ce journal depuis son lancement il y a cinq ans. Dans ce volume, huit articles sont publiés. Ils couvrent divers thèmes.

Le premier article analyse le classement statistique et l'estimation des données des essais participatifs dans les fermes agricoles. Le second applique l'approche régionale de fréquences pour estimer l'importance et la fréquence d'apparition des inondations en Ouganda. Le troisième passe en revue l'hypothèse de constance des coefficients techniques qui sont utilisés pour les comptes des branches lorsqu'on élabore les comptes nationaux. Le quatrième article présente une méthodologie statistique pour élaborer des comptes économiques de régions en appliquant au niveau local les principes de la comptabilité nationale. L'article montre également comment cet approche peut contribuer à l'amélioration des politiques de décentralisation et de développement local dans certains pays d'Afrique de l'Ouest. Les quatre derniers articles se focalisent sur le renfoncement des capacités statistiques dans les pays africains, notamment : l'infrastructure statistique, la revue par les pairs des systèmes statistiques nationaux et la formation statistique. Nous tenons à remercier tous les auteurs qui ont soumis leurs articles pour publication dans ce volume du journal et tout ceux qui ont passé en revue les articles. Nous invitons tous nos lecteurs et leurs collègues à continuer de soutenir ce journal en le choisissant comme outil de diffusion préféré pour publier leur travail de recherche et/ou pour partager des expériences.

Depuis notre dernière édition en novembre 2009, plusieurs événements ont été organisés sur le renforcement des capacités statistiques dans les pays africains. Nous mettons en exergue les principaux d'entre eux. D'abord, lors des réunions du consortium PARIS21, tenues à Dakar au Sénégal du 16 au 18 novembre 2009, les participants ont unanimement affirmé qu'un système statistique effectif et efficace est un élément essentiel de la bonne gouvernance, de la politique et de la prise de décision, et qu'il y a besoin d'accroître l'investissement dans le renforcement de capacités statistiques si les données requises pour faire le suivi des objectifs de développement du millénaire doivent être fournies d'ici 2015. En second lieu, le 5ème symposium sur le développement statistique en Afrique (ASSD) a été organisé à Dakar du 19 au 21 novembre 2009 à la suite des réunions du consortium PARIS21. Le thème choisi pour le symposium était : *« La technologie de*

l'information et de communication dans la diffusion de données : Rapprocher davantage fournisseurs et utilisateurs dans le cycle 2010 des recensements de la population et de l'habitat ». Troisièmement, le Comité de coordination statistique africain (ASCC) a tenu sa quatrième réunion à Dakar, Sénégal le 22 novembre 2009. L'objectif principal de la réunion était de discuter des mécanismes pour réaliser une plus grande synergie entre les institutions régionales/continentales, et coordonner des programmes et des activités statistiques en vue d'accroître et de soutenir l'efficacité de l'appui statistique aux pays africains.

Quatrièmement, une réunion des chefs des instituts nationaux de statistique (INS) africains sur l'harmonisation statistique et le renforcement des capacités statistiques pour l'intégration africaine ont eu lieu du 10 au 12 décembre 2009, à East-London en Afrique du Sud. La réunion, qui était co-organisée par la BAD, la CUA et la CEA, a passé en revue la stratégie africaine pour l'harmonisation statistique en vue de l'intégration africaine et la stratégie africaine pour la mise en œuvre du système de comptabilité nationale 2008 (SCN 2008). Cinquièmement, la Commission économique des Nations Unies pour l'Afrique (CEA) a organisé la deuxième réunion de la Commission statistique pour l'Afrique (StatCom-Africa) du 18 au 21 janvier 2010 à Addis Ababa en Ethiopie. Le thème de cette deuxième réunion de StatCom-Africa était *« Mobiliser les ressources financières et techniques pour appuyer le suivi des objectifs de développement du millénaire dans les pays africains. »*

Sixièmement, une conférence des partenaires au développement sur *« La stratégie mondiale pour améliorer les statistiques agricoles : Plan de mise en œuvre pour l'Afrique »* a eu lieu à Hammamet en Tunisie, du 3 au 4 février 2010. La réunion était co-organisée par le Groupe de la Banque africaine de développement (BAD), l'Organisation des Nations Unies pour l'alimentation et l'agriculture (la FAO) et la Fondation Bill et Melinda Gates (BMGF). Les principaux objectifs de la réunion étaient de:(i) finaliser les directives pour préparer un projet de renforcement de capacités pour améliorer les statistiques agricoles dans les pays africains ; (ii) établir une coalition de donateurs pour mobiliser le financement conséquent pour le plan d'action ; (iii) se mettre d'accord sur les perspectives, les prochaines étapes et le calendrier. Le plan de mise en œuvre pour l'Afrique de la stratégie mondiale comportera trois composantes techniques (assistance technique, formation, recherche), un cadre fédérateur et un mécanisme de gouvernance. Chaque composante technique sera développée comme composante autonome mais entièrement intégré et complémentaire aux autres composantes techniques deans le cadre fédérateur. Le plan de mise en œuvre pour l'Afrique aura

une perspective de long terme (10 à 15 ans) mais suivra une approche par étapes avec la première phase couvrant les 5 années à venir (2011-2015).

Nous espérons que vous trouverez ce volume du journal intéressant et instructif.

Dr. Charles Leyeka Lufumpa
co-président, comité de rédaction
Directeur, Département des statistiques
Groupe de la Banque africaine de développement
Tunis, Tunisie
Email : c.lufumpa@afdb.org

Professeur Ben Kiregyera
co-président, comité de rédaction
International Statistical Consultant
DevInfCo
Kampala, Ouganda
Email : bkiregyera@yahoo.com

Editorial

We welcome our readers to volume 10 of the ASJ, which commemorates the fifth anniversary of the Journal. It confirms the mission of the Journal to serve as a research outlet and information platform for statisticians and users of statistical information mainly in the Africa region. We would like to thank all those who have supported this journal since its inception five years ago. In this volume, eight articles have been published covering a variety of topics.

The first article analyzes statistical ranking and rating data from participatory on-farm trials. The second applies the regional frequency approach to estimate the magnitude and frequency of the occurrence of floods in Uganda. The third one reviews the constancy of the technical coefficient hypothesis, which is used for industry accounts while compiling national accounts. The fourth article presents a statistical methodology to compile district economic accounts by applying at the local level the principles of national accounting. The paper also shows how this approach can contribute to the improvement of decentralization and local development policies in some West African countries.

The last four papers focus on statistical capacity building in African countries within a range of areas: statistical infrastructure, peer reviews of national statistical systems, and statistical training. We thank all the authors who submitted their articles for publication in this volume of the journal and all those who reviewed the articles. We invite all our readers and their colleagues to continue to support this journal by selecting it as a preferred outlet of choice for publishing their research work and/or for sharing their experiences.

Since our last edition in November 2009, several events have been organized in support of statistical capacity building in African countries. We highlight the main ones. First, at the PARIS21 Consortium meetings, held in Dakar, Senegal from November 16-18, 2009, participants unanimously affirmed that an effective and efficient statistical system is an essential element of good governance, policy- and decision-making. Furthermore this meeting highlighted the need to scale up investment in statistical capacity building if the data needed to monitor the Millennium Development Goals are to be provided by 2015. Second, the 5th Africa Symposium on Statistical Development (ASSD) was organized in Dakar from November 19-21, 2009, back to back with the PARIS21 Consortium meetings. The theme chosen for the Symposium was: *"Information and Communication Technology in Data Dissemination: Bringing Suppliers and Users Closer in the 2010 Round of Population and Housing Censuses."*

Third, another major event took place in November 2009, when the African Statistical Coordination Committee (ASCC) held its fourth meeting in Dakar, Senegal. The meeting discussed mechanisms for achieving greater synergy between regional/continental institutions, and coordinating statistical programs and activities with a view to enhancing and sustaining the effectiveness of statistical support to African countries.

Fourth, a meeting of the Heads of African National Statistical Offices (NSOs) on the theme *"Statistical Harmonization and Statistical Capacity Building for African Integration"* took place from December 10-12, 2009, in East London, South Africa. The meeting, which was jointly organized by the AfDB, AUC, and ECA, reviewed the African Strategy for Statistical Harmonization for African Integration and the *African Strategy for the Implementation of the 2008 System of National Account (SNA 2008).*

Fifth, the United Nations Economic Commission for Africa (UNECA) organized the second meeting of the Statistical Commission for Africa (StatCom-Africa) from January 18-21, 2010, in Addis Ababa, Ethiopia. The theme of this second meeting of StatCom-Africa was *"Harnessing Financial and Technical Resources in Support of the Monitoring of the Millennium Development Goals in African Countries."*

Sixth, a Conference of Development Partners on the *"Global Strategy for Improving Agricultural Statistics: Implementation Plan for Africa"* took place in Hammamet, Tunisia, from February 3-4, 2010. The meeting was co-organized by the African Development Bank Group (AfDB), the Food and Agriculture Organization of the United Nations (FAO), and the Bill and Melinda Gates Foundation (BMGF). The key objectives of the meeting were: (i) to finalize Guidelines for developing a capacity-building proposal for improving agricultural statistics in African countries; (ii) to build a donor coalition for raising subsequent funding for the action plan; and (iii) to agree on the way forward, next steps and calendar. *The Implementation Plan for Africa of the Global Strategy* will comprise three technical components (Technical Assistance, Training, Research), an umbrella framework and a governance mechanism. Each technical component will be developed as a standalone component but fully integrated and complementary to the others under the framework. The Implementation Plan for Africa will have a long-term perspective (10 to 15 years) but will follow a phased approach with the first phase covering the next 5 years (2011-2015).

We trust that you will find this volume of the journal both interesting and informative.

Dr. Charles Leyeka Lufumpa
Co-Chair, Editorial Board
Director, Statistics Department
African Development Bank Group
Tunis, Tunisia
Email: c.lufumpa@afdb.org

Professor Ben Kiregyera
Co-Chair, Editorial Board
International Statistical Consultant
DevInfCo
Kampala, Uganda
Email: bkiregyera@yahoo.com

Collaborateurs

Leonard K. **Atuhaire** : Institute of Statistics & Applied Economics, Makerere University, Kampala, Uganda. Email: latuhaire@isae.mak.ac.ug

Samuel **Blazyk** : Fonctionnaire, Paris21, Paris, France.
Email : samuel.blazyk@oecd.org

Guest **Charumbira** : Consultant, Paris21, Paris France.
Email: contact@paris21.org

Richard **Coe** : Chercheur, ICRAF, Nairobi, Kenya. Email : r.coe@cgiar.org

Lamine **Diop** : Consultant, Paris21, Paris, France. Email : ldiopsen@yahoo.fr

Kennedy **Jumanyol** : Professeur, Institute of Statistics & Applied Economics, Université de Makerere, Kampala, Ouganda.
Email: kjumanyol@isae.mak.ac.ug

Parin **Kurji** : Professeur, Faculté d'Agriculture, Université de Nairobi, Kenya.
Email: parinkurji@yahoo.com

Michel **Mouyelo-Katoula** : Coordinateur Mondial, Programme de Comparaison Internationale, Banque Mondiale, Washington DC, USA.
Email: mmouyelokatoula@worldbank.org

Xavier R. **Mugisha** : Professeur, Economic Policy Research Centre, Makerere University Campus, Kampala, Uganda. Email: xmugisha@eprc.or.ug

Adalbert **Nshimyumuremyi** : Statisticien Principal, Statistical Capacity Building Division, Banque Africaine de Developpement, Tunis, Tunisie.
Email: a.nshimyumuremyi@afdb.org

Achille **Pegoue** : Statisticien, Institut National de la Statistique, Yaoundé, Cameroun. Email : apegoue@yahoo.com

Roger **Stern** : Biometricien en Chef, Statistical Services Centre, Université de Reading, Reading, Royaume Uni. Email: r.d.stern@reading.ac.uk

Doug **Stirling** : Professeur, Institute of Fundamental Sciences, Massey University, Palmerston North, Nouvelle Zelande. Email: d.stirling@massey.ac.nz

Mary **Strode** : Consultant, Paris21, Paris, France.
Email : strodemary@yahoo.co.uk

Tony **Williams** : Consultant, Paris21, Paris, France.
Email : tonywilliamsstat@aol.com

Contributors

Leonard K. **Atuhaire**: Lecturer, Institute of Statistics & Applied Economics, Makerere University, Kampala, Uganda. Email: latuhaire@isae.mak.ac.ug

Samuel **Blazyk**: Officer, Paris21, Paris, France.
Email: samuel.blazyk@oecd.org

Guest **Charumbira**: Consultant, Paris21, Paris France.
Email: contact@paris21.org

Richard **Coe**: Researcher, ICRAF, Nairobi, Kenya. Email: r.coe@cgiar.org

Lamine **Diop**: Consultant, Paris21, Paris, France. Email: ldiopsen@yahoo.fr

Kennedy **Jumanyol**: Lecturer, Institute of Statistics & Applied Economics, Makerere University, Kampala, Uganda. Email: kjumanyol@isae.mak.ac.ug

Parin **Kurji**: Senior Lecturer, Faculty of Agriculture, University of Nairobi, Kenya. Email: parinkurji@yahoo.com

Michel **Mouyelo-Katoula**: Global Manager of the International Comparison Program, World Bank, Washington DC, USA.
Email: mmouyelokatoula@worldbank.org

Xavier R. **Mugisha**: Lecturer, Economic Policy Research Centre, Makerere University Campus, Kampala, Uganda. Email: xmugisha@eprc.or.ug

Adalbert **Nshimyumuremyi**: Principal Statistician, Statistical Capacity Building Division, African Development Bank, Tunis, Tunisia. Email: a.nshimyumuremyi@afdb.org

Achille **Pegoue**: Statistician, Institut National de la Statistique, Yaoundé, Cameroon. Email: apegoue@yahoo.com

Roger **Stern**: Chief Biometrician, Statistical Services Centre, University of Reading, Reading, United Kingdom. Email: r.d.stern@reading.ac.uk

Doug **Stirling**: Senior Lecturer, Institute of Fundamental Sciences, Massey University, Palmerston North, New Zealand. Email: d.stirling@massey.ac.nz

Mary **Strode**: Consultant, Paris21, Paris, France.
Email : strodemary@yahoo.co.uk

Tony **Williams**: Consultant, Paris21, Paris, France.
Email : tonywilliamsstat@aol.com

1. Analyzing Ranking and Rating Data from Participatory On-Farm Trials[1]

Richard Coe[2]

Abstract
Responses in participatory on-farm trials are often measured as ratings (scores on an ordered but arbitrary scale) or rankings (respondents are simply asked to order treatments). Usual analyses of variance and linear model-based analyses are not appropriate for these data. Alternatives, based on generalized linear models, are described. These methods can be successfully used when the designs are irregular, as typically occurs in participatory trials, and when covariates are measured on each plot or farm in order to identify GxE interaction.

Key words: *ranking, on-farm trials, generalized linear models*

Résumé
Des réponses dans les essais participatifs en milieu réel sont souvent présentées sous forme d'indice (scores sur une échelle ordonnée mais arbitraire) ou classement (les répondants sont simplement invités à ordonner les traitements). L'analyse de la variance habituelle et les analyses des modèle-linéaires de base ne sont pas appropriées pour ces données. Des solutions de rechange, basées sur les modèles linéaires généraux sont décrites. Ces méthodes peuvent être employés avec succès quand les conceptions sont irrégulières, comme cela se passe dans les essais participatifs, et quand des covariantes sont mesurés sur chaque parcelle de terrain ou ferme afin d'identifier l'interaction de GxE.

Mots clés : *classement, essais participatifs en milieu réel, modèles linéaire général*

1. INTRODUCTION

Participatory methods have been widely by adopted by researchers working on applied agricultural problems, including crop breeding. This change in paradigm has implications for the methods used, both for design and analysis. Some of these are summarized in a companion paper (Coe 2007). An assumption of this paper is that formal analysis of systematically collected quantitative data collected from trials is still an important part of the process. Without this, it is difficult to see how the research activity can generate information of relevance to anyone other than the small numbers

1 This paper was originally published as R. Coe (2002), "Analysing ranking and rating data from participatory on-farm trials," in M.R. Bellon and J. Reeves (eds), *Quantitative Analysis of Data from Participatory Methods in Plant Breeding*. Mexico, DF: CIMMYT, pp. 44–64. We thank CIMMYT for permission to republish in this journal.

2 ICRAF, Nairobi, Kenya. Email: r.coe@cgiar.org

of farmers directly involved. Breeders adopting participatory methods have generally recognized this, but faced three difficulties:

- The experimental designs used are often irregular in layout due to the input by the participating farmers (e.g. choosing which varieties to test on their farm), or the constraints arising from the trials being conducted in farmers' fields.
- The focus of analysis often shifts from overall selection of varieties to understanding the variation in responses across farms. This is GxE interaction, but the 'E' may include social or economic variables in addition to biophysical environments.
- Much of the quantitative data collected may be ratings and rankings, for which the more usual methods of analysis may not be appropriate.

This paper seeks to demonstrate that analysis methods are available to deal simultaneously with each of these difficulties.

In Section 2 of the paper the nature of ranking and rating data are summarized and approaches to analysis are reviewed. Section 3 introduces the examples used to illustrate methods. Sections 4 and 5 present a detailed discussion of an approach to analysis of rating and ranking data respectively. The discussion in Section 6 highlights some outstanding problems and implications of the methods presented.

2. TYPES OF DATA AND ANALYSIS

This paper is concerned with the analysis of responses measured in experiments in the form of rankings and ratings, therefore I start with a summary of data types. The nature of the response variable is one determinant of the type of analysis that can be undertaken, whether conducting formal or informal analysis. It is therefore important to understand exactly how the data are collected and what the numbers represent.

2.1 Continuous

Quantities such as crop yield can be measured on a continuous scale, for example in kgm^{-2}. The numbers have the property that "2 really is the average of 1 and 3," making many common statistical procedures appropriate. Such quantities may be on a "ratio" or "interval" scale, the difference being whether the scale has a real zero. A yield of 1 tha^{-1} is 50 percent that of a

yield of 2tha^{-1}, but a temperature of 10°C is not 50 percent of a temperature of 20°C as the zero for temperature is arbitrary.

2.2 Scores or ratings

Here I refer to data that are recorded on a scale from "poor" to "excellent," or "less than enough" to "more than enough." The categories used are often given numeric labels, such as 1, 2, 3, 4, 5. These are called "scores" or "ratings" and such a scale is also described as "ordered categorical". The numerical labels are arbitrary. An observation of 3 is higher than that of 2, but we cannot say it is better by the same amount as an observation of 5 is better than that of 4. An analysis of the data would ideally use the ordering without using the actual numerical label, so the results would be the same whatever numerical labels were used as long as they described the same order. The word labels ("poor", "excellent" etc.) are not arbitrary, and will determine how respondents use the scale and award scores. The quality of data can be enhanced by giving careful thought to these word labels, and by providing explanations and examples to respondents regarding their meaning.

2.3 Binary

Data that are recorded with just two categories are common, for example "yes – no", "dead – alive", "acceptable – not acceptable." Analysis is based on the frequency with which the categories occur. Methods for these data are widely described and they are not elaborated further here.

2.4 Rankings

In many investigations of preference, data are collected by asking respondents to rank alternatives. The options available are placed in order without any attempt to describe by how much one differs from another, or whether any of the alternatives are "good" or "acceptable". We might have variety A ranked above B which is ranked above C, yet none of these varieties might be considered as good. The data would look the same in the case of a respondent who placed them in the same order, but where 1, 2 or all three were considered "acceptable."

Other scales may be hybrids between these.

2.5 Analysis

The steps in the analysis of any data set can be summarized as follows:

1. *Define the analysis objectives*. These drive the rest of the analysis. It is impossible to do a good analysis of data without clear objectives. Often the key graphs and tables can be defined at this stage, even though the results to fill them in may not yet be available.

2. *Prepare the data*. Data sets will have to be entered and checked, suitable transformations made (e.g. to dry weight per unit area), relevant information from different sources (e.g. farm household data and plot level yields) extracted to the same file, and so on.

3. *Exploratory and descriptive analysis*. The aim is to summarize the main patterns and notice further patterns that may be relevant. This step is only covered briefly in this paper as the methods used will depend on the context in which the analysis is carried out, and on the audience for the results.

4. *Formal statistical analysis*. The aim is to add measures of precision and provide estimates from complex situations.

5. *Interpretation and presentation*.

Iteration between the steps will be necessary. Training materials by Coe *et al.* (2001) provide much more information on analysis of experiments. Some comments on the roles of these steps in analysis of participatory experiments are given in Coe (2007).

A common objective in analysis of many participatory breeding trials is to understand the nature of variation in the responses given by different farmers. Many researchers report that participatory on-farm trials produce highly variable results, making interpretation difficult. Certainly, if a standard analysis aimed at identifying differences in varietal means is carried out, the result may well be a very high "residual" variation and a correspondingly large standard error of varietal difference, implying only vague knowledge about the relative performance of the entries. However, the variation can often be understood as GxE interaction.

The environment in which a participatory trial takes place is heterogeneous. There will be many sources of variation that are not apparent in trials in which a researcher has full control and performs the assessment, and these will include social or economic factors as well as the more usual biophysical definitions of environment. For example, male and female farmers may assess varieties differently; or ratings may depend on the level of a farmer's market

integration. The analyses carried out must therefore be able to identify and describe these GxE interactions. When this is done, the results are often the most useful output of the trial, as they allow recommendations to be tuned to particular local conditions.

A spreadsheet package such as Excel is useful for much of the descriptive analysis. Flexible facilities for data selection and transformation, tabulation, and graphics are useful. However, dedicated statistical software is needed for the analyses described here – they cannot be done in Excel. There are several packages with roughly equivalent facilities. All the examples cited here use Genstat (2000), as I find it often the easiest to understand, particularly as methods for different problems can be addressed with a similar sets of commands. The key commands used to produce each analysis are included in the text with the output they produce. SPSS is widely used by social scientists but is not particularly useful for the analyses described here. Further comments on software are made in the last section.

3. EXAMPLES

3.1 Agroforestry/ soil fertility in Malawi

Although this is not a breeding trial, it is included here as the design is typical of many participatory on-farm trials. Three soil fertility strategies are compared over a number of years:

g – mixed intercropping of maize and gliricidia
s – relay planting of maize and sesbania
c – the control of continuous maize.

Forty-one (41) farmers each compared the control with one or both of the other treatments. Crop yield is the response of interest. A number of covariates were measured at the plot or farm level to help identify the reasons for variation across farms. In the analyses below, the data structure "name" identifies the farmers, "trt" represents the treatments to compare, and "score98" the response of interest.

3.2 Maize varieties in Zimbabwe

This was a "baby" trial comparing 12 maize varieties. 146 farmers in 25 different sites took part, each one testing 4 of the 12 varieties. The varieties for each farmer to test were chosen by the researcher. Some household

and field covariates were recorded. The actual crop yields obtained were not available for analysis, so the examples here use simulated yield data but the original field design. In the analyses below, the data structure FARM identifies farmers, ENTRY the varieties to compare.

3.3 Maize varieties in Kenya

This was a "baby" trial comparing 18 varieties of maize, two of them being local controls. 29 farmers were involved, each planting two replicates of all 18 entries. Crop performance was rated on a scale of 1, 2, 3, 4, 5. Sex of the respondent and farm size were also recorded. In the analyses, IDNO identifies farmers and REP identifies blocks within farms.

4. ANALYZING RATINGS OR SCORES

Example 1

The crop yields in Example 1 were actually measured in tonnes per hectare. However, to illustrate the method of analyzing scores, I have here converted them. The conversion is "exact" (i.e. the scores farmers would give if asked to assess yield and could do it without error) so that, for illustration purposes, the results are comparable with those that can be obtained from actual yields. Scores were allocated as:

Yield	Score	Label
y < 1	1	poor
1 <= y < 2	2	ok
2 <= y < 3	3	good
3 <= y	4	excellent

Descriptive analyses of these data have been explained elsewhere. For example, we could tabulate frequencies as:

TABULATE [PRINT=nobs; CLASSIFICATION=trt,score98; MARGINS=no] score98

Nobservd				
score98	**poor**	**ok**	**good**	**excellent**
trt				
c	9	13	6	3
g	5	15	10	9
s	3	7	9	5

This is informative. For example, for treatment g the mode of the distribution is "ok." This shifts to "good" for treatment s. For treatment c the mode is also "ok" but the frequencies of other scores suggest that g is better than c.

This type of analysis has obvious drawbacks:

- It is difficult to know how to handle more complex patterns.
- It seems to ignore some of the structure in the data. For example, we have not used the fact that each farmer rates 2 or 3 treatments.
- It is not obvious how it could be extended to deal with more complex problems such as identifying and describing the effects of covariates to describe GxE.
- It is not obvious how to formalize it so we can give measures of uncertainty (standard errors, confidence intervals, or statistical hypothesis tests).

A common approach is to treat the scores as quantities measured on a continuous scale. Then means can be calculated (see below) and all the methods of analysis of variance, regression, and related modeling could be tried.

	Mean	**Variance**
trt		
c	2.097	0.8903
g	2.590	0.9852
s	2.667	0.9275

There are two reasons to be uncomfortable about this approach:

1. Many of the assumptions of analysis of variance or linear regression modeling may be inappropriate, given the limited range of the observations. A critical assumption is that the variance between observations of the same treatment is constant across treatments. This is commonly not the case, with the extreme entries showing less variation in score than those with a mean of 2 or 3.
2. The method makes some assumptions about the meaning of the scores that may not be appropriate. For example, is the average of "poor" and "good" really "ok"? The seriousness of this objection is plain when it is realized that the scores 1, 2, 3, 4 are just labels but the results depend critically on the labels given. If we used, for example, 0, 1, 5, 100 then the results using this method would look very different, yet logically these are equally acceptable labels.

There are situations in which both these objections are unimportant and a useful analysis can progress along these lines. However, we would like to have something that is theoretically more sound and robust, and which is applicable in a wider range of cases.

A second approach is to dichotomize the response – change it from a 4-level to a 2-level scale. For example, we could group "poor" and "ok" together, and "good" and "excellent" together to give a measure with just two possible values. There are well-established methods for analyzing such data, including models (e.g. logistic regression) that allow the effects of complex arrangements of covariates to be disentangled, and even methods (generalized linear mixed models) that allow random effects to be incorporated, as in the REML analysis of continuous data (Coe 2007). However, this approach is also unsatisfactory. If the variable is originally measured on a 4-point scale and we reduce it to a 2-point scale, then we must be losing information.

There have been methods developed that are valid, that use the all information without making unreasonable assumptions, and that can model the effect of covariates. In order to understand the model, we look first at the data for just two treatments, g and c, and forget about the fact that the observations are paired by farmer. The data are thus the frequencies:

treatment	poor	ok	good	excellent
c	9	13	6	3
g	5	15	10	9

If we combine the top three categories, the data reduce to the 2x2 table:

treatment	**poor**	**ok+good+excellent**
c	9	22
g	5	34

It looks as if g is better than c. A higher proportion of the plots are in the 'ok+good+excellent' category. A common measure of this association is the odds ratio, O, or log odds ratio log(O)=L.

$$O = \frac{\text{odds on g high}}{\text{odds on c high}} = \frac{34/5}{22/9} = 2.78$$

L = log(2.78) = 1.02

Now we could "cut" the categories at a different place, combining "poor" and "ok" to give the data:

treatment	**poor+ok**	**good+excellent**
c	22	9
g	20	19

This table has O=2.32, L=0.84.

A third "cut" is possible, combining "poor," "ok" and "good' to give:

treatment	**poor+ok+good**	**excellent**
c	28	3
g	30	9

O=2.80, L=1.02

In this case the values of O are similar for each cut. If we make the assumption of such "proportional odds," with a constant value of O, then its value and standard error can be estimated without choosing any particular cut. In Genstat the calculations are done using the regression modeling commands. Note that the data have to be arranged so that there is a response variable for each possible response category. The variable for each score contains the number of plots which had that score.

print treat,s1,s2,s3,s4

treat	s1	s2	s3	s4
c	9.000	13.00	6.000	3.000
g	5.000	15.00	10.000	9.000

model [dist=multinomial;yrel=cumulative;link=logit] s1,s2,s3,s4
fit [p=e,a] treat

***** Regression Analysis *****

*** Estimates of parameters ***

	estimate	s.e.	t(*)	antilog of estimate
Cut-point 0/1	-0.927	0.367	-2.53	0.3956
Cut-point 1/2	0.948	0.367	2.58	2.581
Cut-point 2/3	2.161	0.438	4.93	8.680
treat g	0.932	0.452	2.06	2.539

* MESSAGE: s.e.s are based on dispersion parameter with value 1

Parameters for factors are differences compared with the reference level:
Factor Reference level
treat c

*** Accumulated analysis of deviance ***

Change	d.f.	deviance	mean deviance	deviance ratio
+ treat	1	4.37545	4.37545	4.38
Residual	2	0.16035	0.08018	
Total	3	4.53580	1.51193	

* MESSAGE: ratios are based on dispersion parameter with value 1

The analysis of deviance is interpreted similarly to an analysis of variance, comparing the deviance with a chi squared distribution to judge the importance of the effect. In this case there seems to be a "significant" treatment difference.

The parameter estimate treat g measures the difference between treatments g and c. The estimate 0.932 is the log odds ratio = log (odds of g being high v low / odds of c being high v low). Here "high" and "low" refer to being above and below some cut point in the ordered set of scores. It doesn't matter which cut point, as the model constrains this odds ratio to be the same for any choice of cut point.

The value of 0.932 for the log odds ratio means the odds ratio is exp(0.932) = 2.539. This is similar to the average of the three odds ratios found directly from the data. The standard error can be used to test the hypothesis of no difference between g and c (log odds ratio of 0) or to give a confidence interval for the log odds ratio or odds ratio. The cut-point parameters listed by Genstat do not have a useful interpretation in this case.

Now we analyze the whole dataset using the same ideas, and including a term for farm to account for the fact that each farmer is evaluating 2 or 3 plots. There is a row of data for each plot and a column for each possible score (poor, ok, good, excellent or 1, 2, 3, 4), here given the names s98[1], s98[2], s98[3] and s98[4]. The data value is again the number of plots that were given that score, but now these are all just 0 or 1, with a single 1 in each row. A small part of the data is shown:

print name,trt,s98[1...4],score98;10;decimals=0

name	**trt**	**s98[1]**	**s98[2]**	**s98[3]**	**s98[4]**	**score98**
Chakame	g	0	0	0	1	excellent
Chakame	s	0	0	1	0	good
Chakame	c	1	0	0	0	poor
Thobola	g	0	1	0	0	ok
Thobola	s	1	0	0	0	poor
Thobola	c	0	1	0	0	ok
Adisani	g	0	1	0	0	ok
Adisani	c	1	0	0	0	poor
Majoni	g	0	0	0	1	excellent
Majoni	s	0	0	0	1	excellent

model [dist=multinomial;yrel=cumulative;link=logit] s98[1...4]

fit [p=*] name
add [p=a] trt

* MESSAGE: Term name cannot be fully included in the model because 2 parameters are aliased with terms already in the model

(name Komwa(died 97)) = 0
(name Lipenga(died 98)) = 0

***** Regression Analysis *****

*** Estimates of parameters ***

	estimate	**s.e.**	**t(*)**	**antilog of estimate**
Cut-point 0/1	1.90	2.00	0.95	6.716
Cut-point 1/2	7.23	2.27	3.18	1382.
Cut-point 2/3	10.72	2.43	4.41	45305.
name Belo	4.04	2.56	1.58	56.75
name Bisiwiki	0.00	2.76	0.00	1.000
name Chakame	5.63	2.51	2.24	278.9
name Chimimba	0.97	3.57	0.27	2.638
.				
.				
.				
name White	5.56	2.51	2.21	259.7
trt g	3.5898	0.770	4.67	36.51
trt s	2.722	0.786	3.47	15.21

* MESSAGE: s.e.s are based on dispersion parameter with value 1

Parameters for factors are differences compared with the reference level:
Factor Reference level
name Adisani
trt c

*** Accumulated analysis of deviance ***

Change	d.f.	deviance	mean deviance	deviance ratio
+ name	38	115.468	3.039	3.04
+ trt	2	30.218	15.109	15.11
Residual	51	105.977	2.078	
Total	91	251.663	2.766	

* MESSAGE: ratios are based on dispersion parameter with value 1

The analysis of deviance is interpreted in the usual way, using a chi squared distribution to assess the size of contributions. A deviance of 30.2 with 2 d.f. confirms that treatment is having a clear effect on the ratings.

The parameter estimates for each farmer are uninteresting – they reflect the fact that farmers can differ in the mean rating given. The estimates for the treatments are important and give the quantitative summary of the ratings. In this example the control treatment c is the baseline from which the others are measured. Hence the important results are in the table below. For comparison, analysis of the actual yields using a similar method (linear model fitting farmer+treatment effects) is also shown (details in Coe 2007). Remember the scales are different. We cannot hope to recover information on actual yield per hectare from data which have been recorded only as "poor," "ok," etc. What is important to note are the differences and similarities between treatments which are revealed by this analysis.

treatment	rating log odds ratio	s.e.*	yields adjusted mean	s.e.	scaled yields** adjusted mean	s.e.
g	3.60	0.77	2.62	0.15	3.60	0.53
s	2.72	0.79	2.37	0.17	2.68	0.61
c	0	-	1.64	-	0	-

* the s.e. is the standard error of the difference from c
** yield means scaled to match the log odds ratio scale

When the scales are aligned, then the results of the analyses are remarkably similar. The s.e. values for the rating data are higher as ratings contain less information than actual yields.

The value of the analysis becomes clear when we start looking at differences between groups of farmers, or trying to understand the effect of covariates. For example, slope2 is a factor classifying farms into flat or sloping. The variate cec is related to soil fertility. There were hypotheses that g would perform relatively better on flat land and that both g and s would be superior to c when cec is low. These are investigated in the following table:

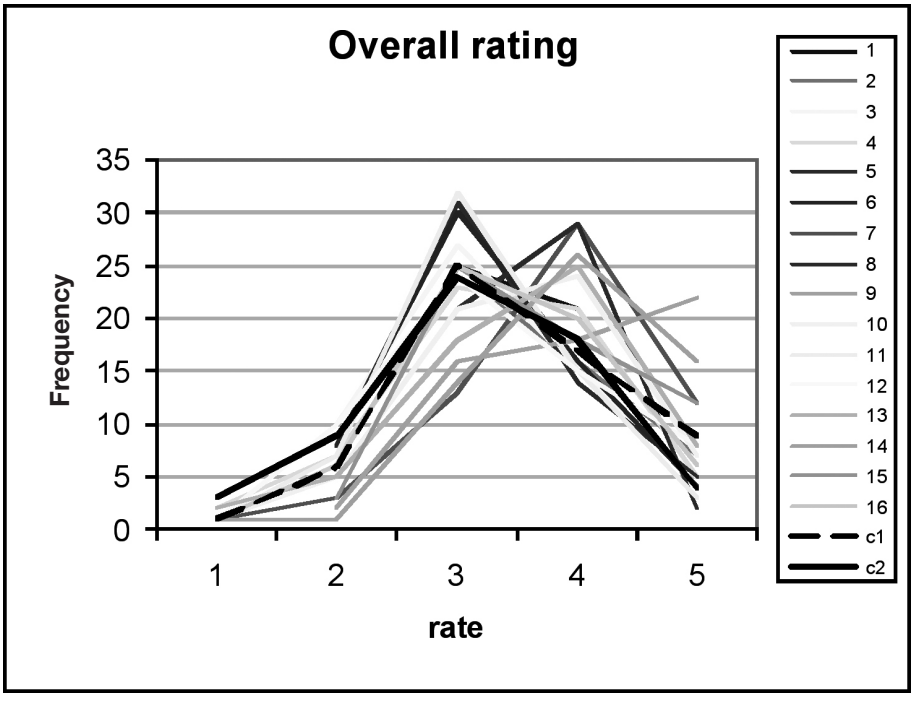
*** Accumulated analysis of deviance ***

Change	**d.f.**	**deviance**	**mean deviance**	**deviance ratio**
+ cec	1	1.945	1.945	1.94
+ slope2	1	8.959	8.959	8.96
+ trt	2	6.259	3.129	3.13
+ name	37	133.823	3.617	3.62
+ cec.trt	2	2.087	1.043	1.04
+ slope2.trt	2	2.543	1.271	1.27
Residual	44	90.606	2.059	

There is no clear evidence for either slope or cec showing an interaction with treatment.

Example 3

In this example, the performances of 18 varieties were rated on a scale from 1 (poor) to 5 (excellent). Criteria were yield, cob size, cob filling, and an overall assessment. The design was straightforward as each of 29 farmers evaluated all 18 varieties, with two plots of each. Simple descriptive statistics can therefore give a useful summary of some characteristics. For example, the graph below shows the frequency of responses for overall rating of each of the 18 varieties. Varieties 17 and 18 are local checks, so have been highlighted.



The varieties seem to fall into two main groups (those with a mode at 3 and those at 4), with entry 9 being rated more highly than all the others.

There are a number of reasons why the modeling analysis is still worthwhile:

- It provides simple, concise summaries with measures of precision.
- It makes inclusion of covariates straightforward. In this case both farm size and sex of the respondent have been recorded.
- It simplifies comparison of the ratings under different criteria.

The analysis follows a similar pattern to the previous example. Note that the layout with two replicates per farm can be explicitly included in the analysis if sensible. Here I have assumed the two replicates correspond to two blocks on each farm. Farms are distinguished by the factor IDNO and blocks within farmers by REP.

```
model [dist=multinomial;yrel=cumulative;link=logit] overall[]
fit [p=*]
add [p=*] IDNO
add [p=*] IDNO.REP
add [p=a,e] ENTRY
```

*** Accumulated analysis of deviance ***

Change	**d.f.**	**deviance**	**mean deviance**	**deviance ratio**
+ IDNO	28	342.309	12.225	12.23
+ IDNO.REP	29	82.354	2.840	2.84
+ ENTRY	17	123.623	7.272	7.27
Residual	966	2189.420	2.266	
Total	040	2737.706	2.632	

*** Estimates of parameters ***

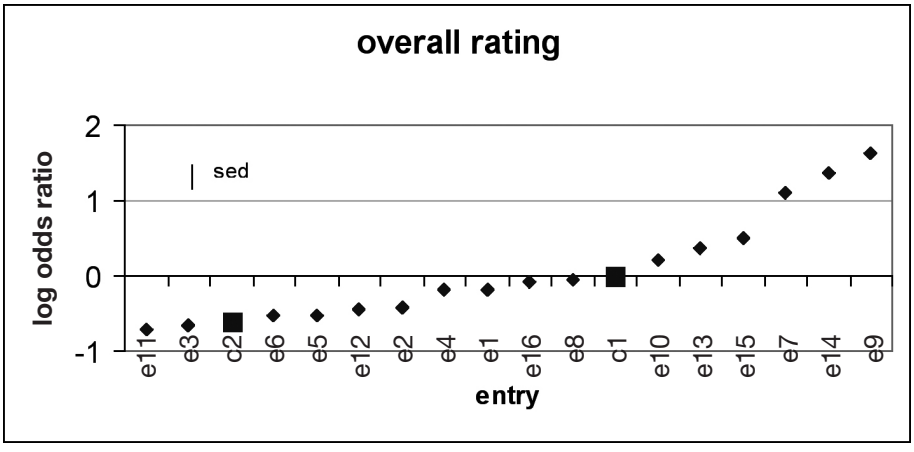
	estimate	**s.e.**	**t(*)**	**antilog of estimate**
Cut-point 0/1	-7.333	0.609	-12.05	0.0006534
Cut-point 1/2	-4.558	0.541	-8.43	0.01048
Cut-point 2/3	-1.560	0.522	-2.99	0.2102
Cut-point 3/4	0.848	0.520	1.63	2.335
IDNO 2	0.170	0.645	0.26	1.185
.				
..				
IDNO 29	-1.993	0.649	-3.07	0.1363
IDNO 1 .REP 2	-1.365	0.641	-2.13	0.2554
.				
..				
IDNO 29 .REP 2	-0.319	0.652	-0.49	0.7271
ENTRY c2	-0.610	0.361	-1.69	0.5433
ENTRY e1	-0.182	0.359	-0.51	0.8338
ENTRY e2	-0.419	0.360	-1.16	0.6575
ENTRY e3	-0.653	0.361	-1.81	0.5206
ENTRY e4	-0.196	0.359	-0.55	0.8219
ENTRY e5	-0.530	0.361	-1.47	0.5883

	estimate	s.e.	t(*)	antilog of estimate
ENTRY e6	-0.539	0.361	-1.49	0.5834
ENTRY e7	1.109	0.360	3.08	3.030
ENTRY e8	-0.049	0.359	-0.14	0.9523
ENTRY e9	1.625	0.365	4.45	5.078
ENTRY e10	0.223	0.358	0.62	1.250
ENTRY e11	-0.701	0.361	-1.94	0.4963
ENTRY e12	-0.438	0.360	-1.22	0.6453
ENTRY e13	0.377	0.358	1.05	1.458
ENTRY e14	1.380	0.362	3.81	3.974
ENTRY e15	0.510	0.358	1.43	1.666
ENTRY e16	-0.078	0.359	-0.22	0.9248

* MESSAGE: s.e.s are based on dispersion parameter with value 1

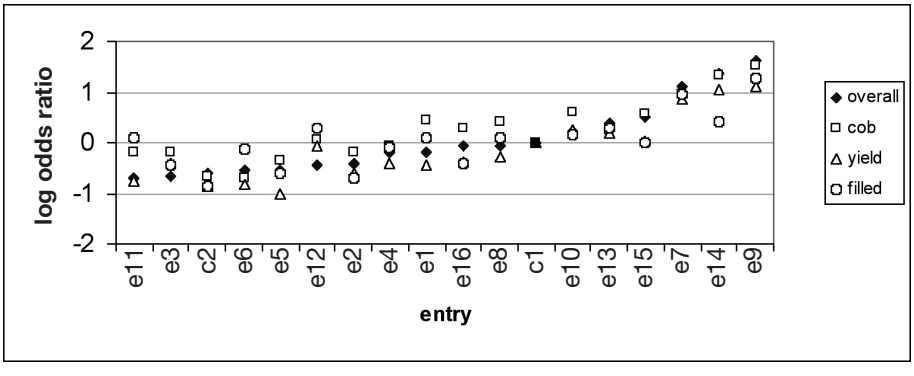
Parameters for factors are differences compared with the reference level:
Factor Reference level
IDNO 1
ENTRY c1

The analysis of deviance suggests that there are large differences between the entries. The parameter estimates summarize these. Remember the estimates are log odds ratios that describe the chance of being in a high response category rather than a low one, for each entry compared with the baseline. The data files have been set up so that the baseline is the first local check, c1. A simple graph reveals the patterns:



Apart from entries 9, 14 and 7, there is a continuous spread of ratings of these varieties, rather than any clear groupings, with one of the local checks towards the lower end of the spread and the other towards the upper end.

The ratings on each criteria can now be compared by repeating the analysis and putting the log odds ratio for each on the same graph. The pattern is much the same for each criterion. Of the three best performing entries, entry 14 does less well on cob filling.



There are two covariates of interest recorded: the sex of the respondent and the farm size. The question of interest is whether males and females tend to rate the entries differently, and whether the relative ratings depend on farm size.

```
model [dist=multinomial;yrel=cumulative;link=logit] overall[]
fit [p=*]
add [p=*] IDNO
add [p=*] IDNO.REP
add [p=*] ENTRY
add [p=*] SEX.ENTRY
add [p=a] SIZE.ENTRY
```

*** Accumulated analysis of deviance ***

Change	**d.f.**	**deviance**	**mean deviance**	**deviance ratio**
+ IDNO	28	342.309	12.225	12.23
+ IDNO.REP	29	82.354	2.840	2.84
+ ENTRY	17	123.623	7.272	7.27
+ ENTRY.SEX	17	21.404	1.259	1.26
+ SIZE.ENTRY	18	21.701	1.206	1.21
Residual	931	2146.315	2.305	
Total	1040	2737.706	2.632	

MESSAGE: ratios are based on dispersion parameter with value 1

Neither of the covariates shows any interaction with entry. Thus we can conclude that the overall rating of varieties is much the same for males and females, and does not have any linear relationship with farm size. The effect of farm size could perhaps be investigated further, for example by putting farms into a few (3 or 4) size categories. This approach removes the assumption of a linear relationship between farm size and the log odds ratios.

5. ANALYZING RANKINGS

At first glance, data from rankings look much the same as rating data. Like ratings, the observations are integers from a limited range, and we want to find out the same sort of information – are there consistencies in the rankings given to different treatments, that can allow us to reach conclusions about which treatments consistently ranked high? However, there are some important differences from rating data that will emerge.

Example 1

Again, to illustrate a method, I have converted yield data from Example 1 to ranks. Each farmer compared two or three treatments. Ranks have been allocated exactly, so that the treatment with the lowest yield on each farm is given rank 1, the next lowest rank 2, and the third (if there is one) rank 3. There were no ties. A small part of the data is shown.

name	**yield98**	**rank**	**trt**
Adisani	1.449	2.000	g
Adisani	0.801	1.000	c
Belo	*	*	c
Belo	2.071	2.000	s
Belo	1.246	1.000	g
Bisiwiki	0.643	1.000	c
Bisiwiki	1.514	2.000	g
Chakame	0.761	1.000	c
Chakame	3.380	3.000	g
Chakame	2.142	2.000	s
Chimimba	1.943	1.000	g
Chimimba	*	*	s
Chimimba	*	*	c
Chinzeka	2.356	3.000	g
Chinzeka	1.477	1.000	s
Chinzeka	1.713	2.000	c

Simple displays of the data can be designed. For example, we can tabulate the number of farmers who rank each treatment as 1, 2, or 3:

TABULATE [PRINT=counts; CLASSIFICATION=trt,rank; MARGINS=no] yield98

Count			
rank **trt**	**1.000**	**2.000**	**3.000**
c	24	6	1
g	9	16	14
s	6	12	6

Unknown Count 14

Treatment g is ranked 3 more often than s, an indication that it is superior. But a difficulty is clear straightaway: it is also ranked 2 and 1 more often than s. The problem arises from the fact that each farmer is only ranking the treatments s/he tests, and these are not the same for each. In the table above, when g=2, we cannot tell whether g was best out of 2 treatments or second out of 3. Changing the ranking method to 1=best does not help. Some authors suggest converting ranks to scores, but of course the problem cannot be fixed by a conversion that simply changes the ranks 1, 2, 3 to another set of numbers.

A more realistic summary comes from studying each treatment pair. If we take, for example, g and c, we can look at all those farmers that compared these two and calculate the proportion that ranked g higher than c.

Pair	**Number of comparisons**	**Number with first of pair ranked higher than second**	**Proportion with first of pair ranked higher than second**
g - c	31	28	0.903
s - c	21	16	0.762
g - s	24	16	0.667

This summary now correctly only relies on the rankings within each farm and is explicit about what is compared with what. Its shortcomings, and the reasons for wanting a formal analysis, are much the same as for the rating data. We need to put measures of precision on results and would like to extend the analysis to look at the effect of covariates or groupings of respondents. The analysis also seems unsatisfactory when we think of Example 2 with its 12 treatments and hence 66 pairs of treatments. A table such as the one

above but with 66 rows to describe performance of 12 varieties would be nothing but opaque!

The modeling approach to this type of data is based on the above table. The idea is to find a score s_i for each treatment such that the probability that treatment i is ranked higher than treatment j when the two are compared depends on the difference between the scores, $s_i - s_j$. If the relationship between scores and probability is a logistic function, then the model can be fitted using standard logistic regression software. Hence we put

p_{ij} = Prob(i ranked above j) and

$\log(p_{ij}/1-p_{ij}) = s_i - s_j$.

Setting up the data to fit the model is slightly messy. There has to be a row for each pair of treatments compared. Thus a farmer with just g and c will contribute one row of data for the pair g-c. A farmer with three treatments g, s and c will contribute three rows of data, g-c, s-c and g-s. Indicator variables are needed for each treatment and the response variable contains 0s and 1s. The first few rows of data are shown:

namel	**firstl**	**secondl**	**c**	**g**	**s**	**compl**
Adisini	g	c	-1	1	0	1
Belo	s	c	-1	0	1	*
Belo	s	g	0	-1	1	1
Belo	g	c	-1	1	0	*
Bisiwiki	g	c	-1	1	0	1
Chakame	g	c	-1	1	0	1
Chakame	s	c	-1	0	1	1
Chakame	s	g	0	-1	1	0
Chiminbo	g	c	-1	1	0	*
Chiminbo	s	g	0	-1	1	*
Chiminbo	s	c	-1	0	1	*
Chinzeka	g	c	-1	1	0	1
Chinzeka	s	g	0	-1	1	0
Chinzeka	s	c	-1	0	1	0

The first row of data shows that Adisini compared g and c. g was ranked higher than c, so when g is the first and c the second, the response is "suc-

cess," indicated by a 1 in the last column. Belo had all three treatments but the observation for c was missing, therefore both the s-c and g-c comparisons are missing.

The modeling now proceeds in a similar way as for other situations.

```
model [dist=b] compl; nbin=1
fit [con=o]g+s+c
```

*** Summary of analysis ***

d.f.	deviance	mean deviance	deviance ratio
Regression	2	*	*
Residual	74	73.49	0.9931
Total	76	*	*

* MESSAGE: ratios are based on dispersion parameter with value 1

*** Estimates of parameters ***

treatment	estimate	s.e.	t(*)	antilog of estimate
g	2.072	0.435	4.76	7.939
s	1.290	0.425	3.04	3.632

* MESSAGE: s.e.s are based on dispersion parameter with value 1

The output looks a little odd because Genstat does not know what to use as a null model when the constant is omitted, so cannot calculate a Total deviance, hence also cannot calculate a Regression deviance. In this case the sensible null model is one of "no preference" between treatments, corresponding to $p_{ij} = 0.5$ for all pairs, or $\log(p_{ij}/1\text{-}p_{ij})=0$. The deviance for this model is given by

```
model [dist=b] compl; nbin=1
fit [con=o]
```

Now the analysis of deviance can be reconstructed:

	d.f.	deviance	mean deviance	deviance ratio
Regression	2	31.91	15.96	15.96
Residual	74	73.49	0.9931	
Total	76	105.4	1.386	

The model appears to explain much of the variation, suggesting real difference between the treatments. When interpreting the parameter estimates, remember that the p_{ij} depend only on the differences s_i - s_j. Hence we only need to estimate two of them and can arbitrarily set the third, in this case c, to zero. Hence the estimates above give an ordering and even magnitude of differences between the treatments. They can be compared with the results from analyzing both actual yields and the scores.

	ranking		rating		yields		scaled yields**	
treatment	s_i	s.e.*	log odds ratio	s.e.	adjusted mean	s.e.	adjusted mean	s.e.
g	2.07	0.44	3.60	0.77	2.62	0.15	2.07	0.32
s	1.29	0.43	2.72	0.79	2.37	0.17	1.54	0.36
c	0	-	0	-	1.64	-	0	-

* the s.e. is the standard error of the difference from c
** yield means scaled to match the s scale of the ranking data

The analysis of ranks has, to within the arbitrary scaling, produced an order and relative difference between treatments which is remarkably similar to that from the actual yield data, yet with larger s.e.d. values: the ranks contain less information than actual yields.

Note the table of pairwise probabilities p_{ij} can be reconstructed from the scores s_i using the relationship

$$p_{ij} = \exp(s_i - s_j)/(1+\exp(s_i - s_j))$$

These are shown in the table below and indicate a reasonable fit of the model.

Pair	Number of comparisons	Number with first of pair ranked higher than second	Proportion with first of pair ranked higher than second	Fitted probabilities p_{ij}
g - c	31	28	0.903	0.888
s - c	21	16	0.762	0.784
g - s	24	16	0.667	0.686

As in other situations, an advantage of using an explicit model to analyze the ranks, rather than relying on more ad hoc methods, is that the effects of covariates can be identified. As an illustration I have looked at slope, classified into 2 levels (0=flat, 1=sloping), as one of the hypotheses was that g would perform relatively less well on sloping land.

add [p=a,e] slopel.(g+s+c)

*** Accumulated analysis of deviance ***

Change	d.f.	deviance	mean deviance	deviance ratio
- Constant				
+ g				
+ s				
+ c	1	*		
+ g.slopel				
+ s.slopel				
+ c.slopel	2	0.778	0.389	0.39
Residual	72	72.715	1.010	
Total	75	*		

* MESSAGE: ratios are based on dispersion parameter with value 1

The analysis of deviance suggests that there is no consistent difference in the way g, s and c are ranked on flat and sloping land. This conclusion is also reflected in the parameter estimates:

*** Estimates of parameters ***

	estimate	**s.e.**	**t(*)**	**antilog of estimate**
g	2.117	0.583	3.63	8.305
s	1.598	0.607	2.63	4.944
g.slopel 0	-0.056	0.901	-0.06	0.9454
g.slopel 1	0	*	*	1.000
s.slopel 0	-0.632	0.858	-0.74	0.5313
s.slopel 1	0	*	*	1.000

* MESSAGE: s.e.s are based on dispersion parameter with value 1

These can be put together into a table of scores, together with standard errors of the difference between treatments within slope categories.

Treatment	**slope=0**	**slope=1**
g	2.117-0.056=2.061	2.117
s	1.598-0.632=0.966	1.598
c	0	0

If the standard errors of the interaction effects were smaller, we would say the results were consistent with the hypothesis – the difference between g and s is greater on flat than on sloping land.

Remember that it is impossible to look at the "main effect" of slope. We cannot determine whether the treatments are generally assessed as better on flat land than sloping. Each participant ranks among the alternatives tested on their farm, and each farm is classed as either sloping or flat. Similarly, we cannot compare the two columns in the table above, comparing g on flat and sloping land. There is no information in the data on this comparison as all rankings are done within farms. The situation would be different if there were farms that had both flat and sloping land.

Example 2

Yields for Example 2 were also converted to ranks for the purpose of illustrating the analysis. Remember, this study has 12 varieties with 146 farmers comparing 4 varieties each. It is difficult to think of a useful simple, descriptive

analysis of this rank data that shows the differences between varieties. The design is very unbalanced, so any simple totaling of ranks will give a biased picture. We could look at all the 66 = 12 x 11 / 2 pairwise comparisons, and find the proportion in which one treatment ranks above another. It is not easy though to view a matrix of 66 values and understand the relative performance of 12 varieties. It has been suggested (Russell 1997) that an overall score be given to each variety by counting the number of times each one ranks above another. However, this requires each to occur equally often. Some sort of average proportion could be devised. However the modeling approach is simple once the data file is set up.

The data file structure and modeling proceeds as in Example 1. Twelve indicator variables, e[1],...e[12] are needed for the 12 varieties. In the statements below, the first FIT gives the correct Total deviance from which the analysis of deviance table is constructed.

```
model [dist=b] compl; nbin=1
fit [con=o]
fit [con=o] e[1...12]
```

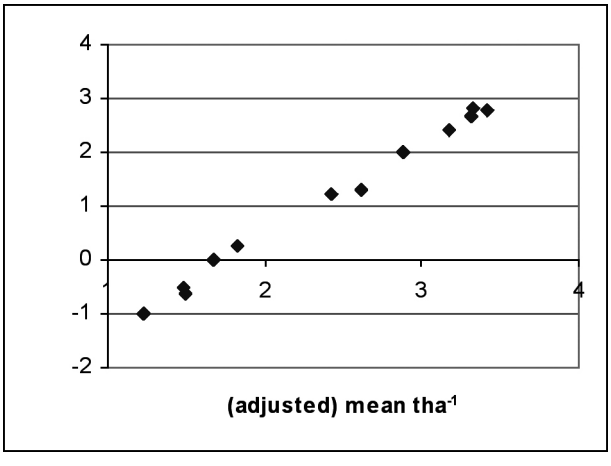
d.f.	**deviance**	**mean deviance**	**deviance ratio**
Regression	11	439.1	39.92
Residual	865	774.9	0.8959
Total	876	1214.	1.386

Genstat has put the score for the last treatment, s_{12}, to zero. The parameter estimates above then give the scores that show the relative performance for each variety. If these are compared with the results based on actual yields, it can be seen (graph below) that the method not only reproduces the ordering of the varieties very closely, but also the relative differences. Of course, in this case the ranks were calculated from the yields without error. Nonetheless, it still seems very surprising that this information about relative performance of the varieties can be recovered from just the four ranks on each farm.

*** Estimates of parameters ***

	estimate	s.e.	t(*)	antilog of estimate
e[1]	-0.996	0.293	-3.40	0.3692
e[2]	1.998	0.300	6.66	7.378
e[3]	1.282	0.284	4.52	3.606
e[4]	2.818	0.317	8.88	16.74
e[5]	-0.523	0.303	-1.73	0.5926
e[6]	2.655	0.312	8.51	14.22
e[7]	0.247	0.277	0.89	1.281
e[8]	1.205	0.281	4.29	3.336
e[9]	-0.637	0.299	-2.13	0.5289
e[10]	2.769	0.318	8.70	15.95
e[11]	2.423	0.303	8.00	11.28

* MESSAGE: s.e.s are based on dispersion parameter with value 1



As each score is relative to the score of zero for variety 12, the s.e.s listed with the estimates are for the comparison of that variety with variety 12. Other s.e. values are most easily found using predict. For example, the difference between scores for variety 1 and 2 is found by:

predict [back=n] e[1...11]; 1,-1,0,0,0,0,0,0,0,0,0

Prediction	s.e.
-2.995	0.335

More complex contrasts between treatments can be calculated in a similar way. For example, varieties 1, 5, 7, 9 and 12 are in one group, namely a. Varieties 4, 10 and 11 form group b. We can calculate the difference between the average scores for groups a and b by taking $(s_1+s_5+s_7+s_9+s_{12})/5 - (s_4+s_{10}+s_{11})/3$. Remembering $s_{12}=0$, predict can be used for this:

predict [back=n] e[1...11]; 0.2,0,0,-0.3333,0.2,0,0.2,0,0.2,-0.3333,-0.3333

Prediction	s.e.
-3.052	0.213

Group a is clearly worse than group b.

As in Example 1, it is simple to turn differences in scores into probabilities of one variety being ranked higher than another. For example, the chance that variety 1 is ranked higher than 2 is given by:

predict e[1...11]; 1,-1,0,0,0,0,0,0,0,0,0

Prediction	s.e.
0.0477	0.0152

Variety 2 is almost certain to be ranked higher than variety 1.

As before, the model can now be extended to look at the extent to which covariates interact with treatment differences. I use two continuous covariates, soil P and sand content. The data file has been set up with a column giving the sand and P for each pairwise comparison.

```
fit [p=*; con=o] e[1...11]
add [p=*; con=o] sandfl.e[1...11]
add [p=a; con=o] Pfl.e[1...11]
```

The analysis of deviance table can be constructed from this output. Note the total degrees of freedom has changed from earlier as there are missing values in the covariates.

	d.f.	**deviance**	**mean deviance**
e[1...11]	11	309.5	28.14
+sandfl.e[1..11]	11	15.64	1.42
+Pfl.e[1...11]	11	15.16	1.38
Residual	561	483.3	0.8614
Total	594	823.5	1.386

The results show that neither P nor sand has a strong interaction with variety. However, in order to show the types of results obtainable, the model with sand is refitted and parameter estimates produced.

fit [con=o;p=e] e[1...11]+sandfl.e[1...11]

*** Estimates of parameters ***

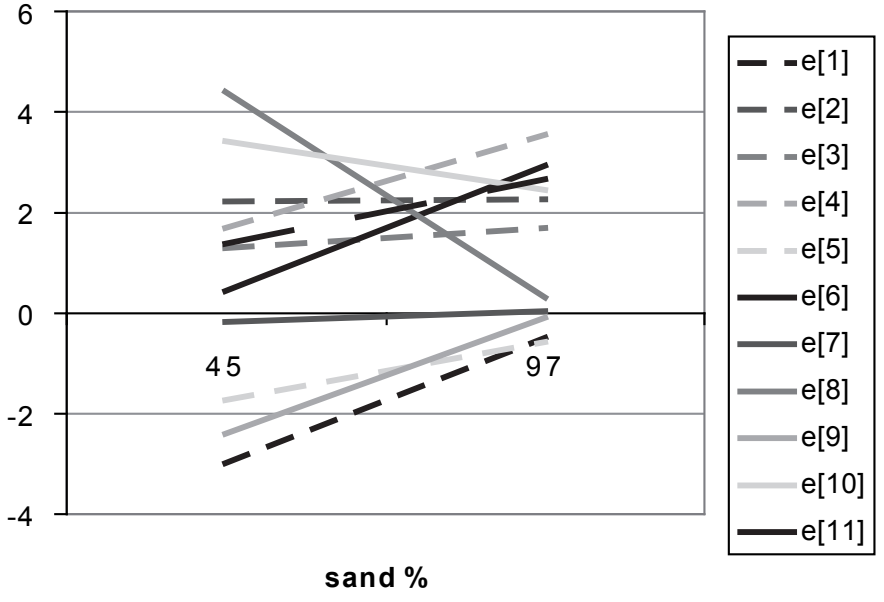
	estimate	**s.e.**	**t(*)**	**antilog of estimate**
e[1]	-5.11	2.83	-1.81	0.006012
e[2]	2.22	3.21	0.69	9.223
e[3]	0.97	2.49	0.39	2.631
e[4]	0.10	3.20	0.03	1.109
e[5]	-2.73	2.61	-1.04	0.06530
e[6]	-1.72	2.75	-0.62	0.1796
e[7]	-0.42	2.38	-0.18	0.6559
e[8]	7.90	3.70	2.14	2699.
e[9]	-4.38	2.64	-1.66	0.01248
e[10]	4.28	3.46	1.24	72.58
e[11]	0.28	4.01	0.07	1.317
e[1].sandfl	0.0475	0.0327	1.45	1.049

	estimate	**s.e.**	**t(*)**	**antilog of estimate**
e[2].sandfl	0.0004	0.0370	0.01	1.000
e[3].sandfl	0.0075	0.0290	0.26	1.008
e[4].sandfl	0.0357	0.0375	0.95	1.036
e[5].sandfl	0.0223	0.0301	0.74	1.023
e[6].sandfl	0.0478	0.0320	1.49	1.049
e[7].sandfl	0.0049	0.0280	0.17	1.005
e[8].sandfl	-0.0781	0.0417	-1.87	0.9249
e[9].sandfl	0.0444	0.0304	1.46	1.045
e[10].sandfl	-0.0187	0.0398	-0.47	0.9814
e[11].sandfl	0.0244	0.0460	0.53	1.025

The scores for each variety now depend on the sand content. For example, the score for variety 1 is

$s_1 = -5.11 + 0.0475$ sand

These are plotted below for the range of sand contents found in the trial, 45 to 97 percent.



Remembering that the scores show the relative performance of varieties, with variety 12 fixed at a score of zero, two main patterns emerge. Several varieties (1, 4, 5, 6, and 9) rank higher than 12 with increasing sand content. Variety 8 ranks distinctly worse with high sand content.

6. DISCUSSION

6.1 General

The methods described above for analysis of ranking and rating data are not new but they are not being routinely used in the analysis of agricultural trials. A discussion of the proportional odds model used for rating data can be found in Agresti (1996). The model for ranks is not so widely used, explaining why common statistical software does not make it immediately available. When the observations are paired comparisons (i.e., each participant is asked to state which of two treatments is superior), then the Bradley-Terry model (Bradley and Terry 1952) has been widely used, particularly in social science applications. Dittrich *et al.* (1998) use the method for paired comparisons when there are categorical covariates and mention that it is possible with continuous covariates. The approach used when more than two treatments are compared is described by Critchlow and Fligner (1991).

Both models involve making assumptions about the nature of the data, however this is true of all statistical analyses. It is a necessary part of attempting to reach conclusions about general patterns. Methods for checking the key assumptions are well developed for established linear model methods (for example, looking for various patterns in residuals) and similar tools need developing for these models. Alternative models may be more appropriate for either ranks or rates. The methods presented here appear to be the simplest that have proved useful in some common situations. Again this is common to all statistical modeling. For example, linear regression analysis is widely used but not because "nature has to be like that", but because the model has proved to be a useful approximation in many problems.

From the examples in this paper, it should be clear that statistical analysis of participatory breeding trials cannot be automatic. When researcher-designed trials were run using a very regular design, it was possible (though probably not wise) to run a standard analysis on each data set. Such an approach will not recover most of the useful information from participatory trials.

6.2 Further discussion on analysis of ranking

The method described above for analyzing data presented in the form of ranks seems to be appealing and powerful. It is able to produce an overall ordering of treatments, and even indicate the relative magnitudes of the differences between the treatments. It can handle awkward incomplete sets of data in which each farmer does not rank all treatments. Most importantly, it can show how treatments interact with covariates. In Example 1, the covariate was a categorical variable, dividing the sample of farmers into groups. In Example 2, continuous covariates were analyzed.

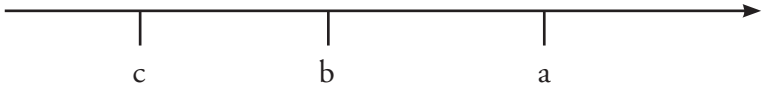
Unlike many other approaches to an analysis of ranks, the method uses estimation and not just testing. The distinction is often made when analyzing continuous variates such as crop yields. It is rarely useful simply to conclude that mean yields "differ significantly" between varieties, or even that variety A yields at a significantly higher level than variety B. We can draw useful conclusions when we can assess by how much A outyields B, and put a confidence interval around this. The same is true when analyzing ranks. It is rarely useful to simply report that treatments differ "significantly" in their ranks, yet this is all that most statistical procedures for an analysis of rankings do. The method presented here shows the relative magnitude of the differences and these can be interpreted. For example, we may show that A and B are ranked "significantly" differently. The scores for the varieties can be converted to a probability p_{AB} that A will be ranked higher than B.

If $p_{AB} = 0.95$ the interpretation is very different than if $p_{AB}=0.55$, yet both could be "significantly different" from the no-preference value of $p_{AB}=0.5$.

There are a number of questions about this analysis, some of which require some theoretical statistical investigation.

1. The model makes assumptions about the nature of the data and the effects of covariates. It is not clear how to check whether they are reasonable or how robust the results are to departures from the assumptions.

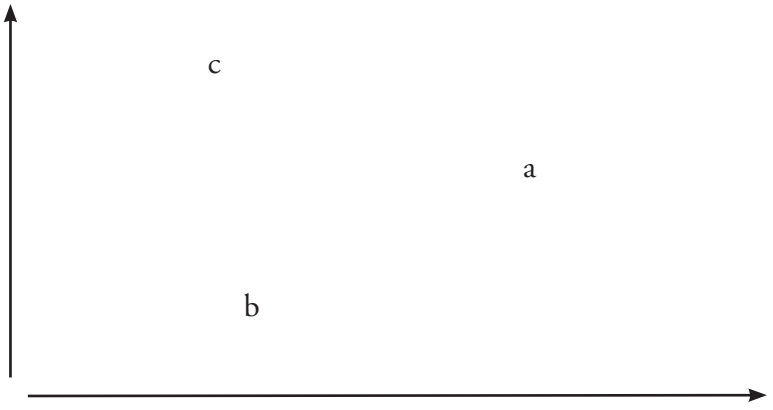
2. The analysis depends on the model, which assumes that the treatments can be allocated scores such that the probability of one ranking higher than another depends on the difference between the scores. This is the "linearity assumption" of Taplin (1997). It is helpful to represent it graphically. If farmers consistently rank a>b>c then we could derive scores that would put the treatments on a line:



It is easy to produce data for which this linearity assumption fails. For example, we can have a>b and a>c equally often, suggesting the ordering should look like the line below.

However, at the same time we can have b>c. This might occur if, for example, different farmers were making the comparisons and using different criteria for each one.

The problem occurs in other examples of "ordination," for example that used by ecologists to describe species occurrence. An answer is to introduce a second dimension. The distances between points a, b, and c can reflect the rankings if they are arranged as:



With just three treatments (as in Example 1), it is clear how an extra term can be introduced in the model to test whether a non-linear arrangement is superior (in Example 1 it is not). However, I do not know how to fit models that follow the usual, useful multivariate approach of gradually increasing the number of dimensions until a suitable fit is obtained.

3. In some ranking procedures, ties are allowed – for example, farmers are allowed to state that they have no preference between two or more of the treatments. Dittrich *et al.* (1998) show how the model can be modified to allow for ties. Extra parameters are included, so that for each pair we estimate the probability that they tie as well as the probability that one is ranked above the other.

4. Coe (2007) illustrates the value of being able to describe variation at different levels in the design with random effects. It is not clear if these ideas are useful or could be used here. In principle we can fit the model with random effects using the GLMM framework. However, this may not be necessary. All the information in the ranks is at the "within farm" level. Hence we can look at treatment differences and interactions of these with farm level (or higher level) covariates. However, we cannot look at any "between farm" effects. It is also not clear if plot-level covariates can be incorporated. For example, suppose farmers ranked treatments but also reported whether each plot was normally fertile or not, so that plots within one farm could have differing values of the fertility covariate. A model that uses these data would have to be built on the probability of a>b, depending on both the difference in treatment scores and the difference in fertility.

5. The analysis described here is suitable for one objective – that of determining treatment effects and their interaction with covariates when

the observations are ranks and the design is incomplete or irregular. If the design is more complete, for example with each respondent comparing all treatments, then other approaches are possible. Different objectives may be of interest, for example, comparison of the rankings under different criteria, or partitioning the sample of respondents into homogeneous groups, when again different methods will be appropriate. Remember also that if the data are rankings produced at, say, a group meeting, so that a consensus is arrived at, then no statistical analysis is necessary. Abeysekera (2001) and Riley and Fielding (2001) describe some of the simple alternatives. Taplin (1997) describes a number of the statistical tests available.

6.3 Ranks vs. rates

Having seen how data from these trials can be analyzed, it is worth looking again at the relative merits of using ranking or rating.

First it should be clear that a response measured on a continuous scale, using an accurate and unbiased instrument, contains more information that the equivalent observation using a rating scale with a few levels, or using ranks. Reasons for not using the continuous variate include:

1. *Time, money and logistics* (e.g. we may not be able to measure crop yield as we are uncertain when farmers will be ready to harvest).

2. *Lack of a suitable instrument*. If we want to assess taste or opinions, there is little alternative to rating or ranking.

3. *Participation*. Collection of ranking and rating data involves participants. Other measurement methods may be alienating.

Methods of collecting rating data have been described (e.g. Ashby 1990) and include tools that can give high-quality, repeatable, and reliable data. It appears that farmers are able to give scores to a large number of alternatives. There are statistical questions regarding the number of levels to use. There is no point in using too many levels, as small differences in rating will probably not reflect real differences in opinion. Note that we do not make a rating scale into a continuous variate simply by using many levels. The fundamental characteristic of a rating scale is that the numbers represent qualitative labels ("very good," "poor," etc.) and the quantitative analogue is missing. This may not be the case if the markers are used to represent the score.

There is a lot of theory and practice from the social science literature that is relevant here. Respondents are often reluctant to use the ends of a scale, particularly the lower end. Hence a 5-point scale may in practice be used as a 3-point scale. Note that some degree of consistency in the use of the scale by different participants, particularly in different locations, can be achieved by explaining what "poor," "excellent," etc. mean. For example, "poor" might correspond to "I would never consider growing this again," while "excellent" might mean, "I would like this to become a main variety for my farm each year."

Ranking is used when it is considered that participants find it easier to order alternatives than to give them a score. One reason for this is clear: participants might have a preference for two alternatives which might score the same (e.g. both "excellent"), and hence be able to give them different ranks. However, a shortcoming is immediately clear. The ranks may be the same if both are also considered "poor." This is an important problem. The information in ranks is all "within respondent," that is, we can identify whether, for example, participants consistently rank A above B, and we can determine whether this is true for both male and female respondents. However, we cannot determine what either group of participants actually thinks of A and B. An important part of any research is to make generalizations and extrapolations and ranked data are often not able to do this. Abeysekera (2001) makes the point that ranked information is considerably enhanced if some sort of baseline is also measured. For example, if a local control variety is included in each participant's set of alternatives, then we could get a rating for the local control, and rank the others relative to this. It is not clear exactly how such data could be analyzed.

A study by ICRAF (1996: 55) assessed the suitability of 12 tree species as firewood. The researcher thought that women could only realistically compare pairs of species. The participants ranked each pair tested, from which it was possible to produce an overall ordination. However, they were also asked the reasons for preferring one species to the other. An alternative design would have used a pilot study of this type to elicit important criteria, then asked for ratings on these for each species tested.

Remember that there is no ranked information on effects of quantities that vary across farms. In Example 1, we were unable to determine whether g was more effective on sloping or flat land.

Overall there seems little reason to collect ranked data unless they are specifically required by the objectives.

6. NOTE ON AVAILABLE SOFTWARE

Since the original version of this paper was published five years ago, there have been important changes in statistical software suitable for this type of analysis.

Genstat, used here to illustrate methods, has been updated to the 12th Edition. The basic commands needed to perform the analyses illustrated have not changed. Most are available to users through simple menus and dialogue boxes. Some details have changed. For example, when the MODEL command is used to fit ordered categorical regression models, the response variable no longer has to be arranged with a separate variate of counts for each category. A single response factor can be given. More importantly, VSNi, the company that produces Genstat, has made the Discovery Edition available free to researchers and educators in the developing world. Details are available from: http://www.vsn-intl.com/products/discovery/ or http://www.worldagroforestrycentre.org/rmg/GDE/index.html.

A second source of high-quality statistical software free to all is R. Development of this open-source software has been by a consortium with many contributors. Details are available and the software can be downloaded from http://cran.r-project.org/

It may take new users a while to learn the basics of R, but the effort is repaid by giving access to a very wide range of statistical tools, often including the very latest developments in statistics methods. As a starter, the following commands will give the analyses of Example 2 from this paper.

```
#Read the data, in this case from the clipboard after copying in Excel
soilfert<-read.table("clipboard", header=TRUE,na.strings="*")
attach(soilfert)

#Change the score column to an ordered factor, make sure name and trt are factors
score98<-ordered(score98, levels=c("poor", "ok","good","excellent"))
name<-as.factor(name)
trt<-as.factor(trt)

#Fit the ordered categorial (proportional odds) model
library(MASS)
ratemod<-polr(score98~name+trt)
summary(ratemod)
```

REFERENCES

Abeysekera, S. (2001). "Analysis Approaches in Participatory Work Involving Ranks or Scores." DfID theme paper (revised). Statistical Services Centre, University of Reading, UK.

Agresti, A. (1996). *An Introduction to Categorical Data Analysis*. New York: Wiley.

Ashby, J.A. (1990). *Evaluating Technology with Farmers: a Handbook*. Cali: CIAT.

Bradley, R.A. and M.E. Terry (1952). "Rank Analysis of Incomplete Block Designs. 1 The Method of Paired Comparisons," *Biometrika*, vol. 63, pp. 255-262.

Coe, R. (2007). "Analysing Data from Participatory On-farm Trials," *African Statistical Journal*, May, pp. 89-112.

Coe, R. and S. Franzel (2000). "Keeping *Research* in Participatory Research." Paper presented at the Third International PRGA Conference, Nairobi, November 6-11.

Coe, R., R. Stern, and E. Allen (2001). "Analysis of Data from Experiments." Training materials, SSC/ICRAF, 250pp.

Critchlow, D.E. and M.A. Fligner (1991). "Paired Comparisons, Triple Comparisons and Ranking Experiments as Generalised Linear Models, and their Implementation in GLIM," *Psychometrika*, vol. 56, pp. 517-533.

Dittrich, R., R. Hatzinger, and W. Katzenbeisser (1998). "Modelling the Effect of Subject-Specific Covariates in Paired Comparison Studies with an Application to University Rankings," *Applied Statistics*, vol. 47, pp. 511-525.

Genstat (2000). *Genstat for Windows* (5th Edition). Oxford: VSN International.

International Centre for Research in Agroforestry (ICRAF) (1996). *Annual Report* 1995. Nairobi: ICRAF.

Riley, J. and W.J. Fielding (2001). "An Illustrated Review of Some Farmer Participatory Research Techniques," *Journal of Agricultural, Biological and Environmental Statistics*, vol. 6, pp. 5-18.

Russell, T. (1997). "Pair Wise Ranking Made Easy," *PLA Notes*, vol. 28, pp. 25-26.

Taplin, R.H. (1997). "The Statistical Analysis of Preference Data," *Applied Statistics*, vol. 46, pp. 493-512.

2. Estimation of the Magnitude and Frequency of Floods in Uganda: A Regional Frequency Analysis Approach

Kennedy Jumanyol,[1] Xavier R. Mugisha,[2] and Leonard K. Atuhaire[3]

Abstract
The estimation of extreme quantiles corresponding to small probabilities of exceedance is commonly required in hydraulic problems. The choice of the underlying flood frequency distribution affects the flood quantile estimates. Statistical distributions cannot be delineated from small sample sizes commonly available in hydraulic work using the at-site approach. Consequently, the Regional Frequency Analysis (RFA) method has been developed to resolve this problem.

This study aimed at conducting an RFA of annual maximum river discharges from selected major rivers in Uganda. Uganda is one of the countries that are frequently affected by floods, yet no flood frequency analysis had been carried out on its rivers. The study recommended that the results presented could be improved in the future when more data on peak discharges become available. Further research is needed to determine how aspects of climatic change could be incorporated into flood frequency analysis for design and planning purposes.

Key words: *Catchment area, delineation of homogeneous regions, discordancy measure test, hierarchical cluster analysis, L –moments, quantiles estimation*

Résumé
L'évaluation des quantiles extrêmes correspondant à de petites probabilités est généralement exigée dans les problèmes hydrauliques. Le choix de la distribution de fréquence d'inondation sous jacente affecte les estimations de quantile d'inondation. Les distributions statistiques ne peuvent pas être décrites à partir d'échantillons petite taille généralement disponibles dans le travail hydraulique en utilisant l'approche terrain. En conséquence, la méthode d'analyse de fréquence régionale (RFA) a été développée pour résoudre ce problème.

Cette étude avait pour objectif de conduire une RFA des décharges maximales annuelles des principaux fleuves choisis en Ouganda. L'Ouganda est l'un des pays qui sont fréquemment affectés par des inondations, pourtant aucune analyse de fréquence d'inondation n'avait été effectuée sur ses fleuves. L'étude a recommandé

1 Institute of Statistics & Applied Economics, Makerere University, Kampala, Uganda. Email: kjumanyol@isae.mak.ac.ug

2 Economic Policy Research Centre, Makerere University Campus, Kampala, Uganda. Email: xmugisha@eprc.or.ug

3 Institute of Statistics & Applied Economics, Makerere University, Kampala, Uganda. Email: latuhaire@isae.mak.ac.ug

que les résultats présentés pourraient être améliorés à l'avenir quand plus de données sur des décharges maximales seront disponibles. Davantage de recherche est nécessaire pour déterminer comment des aspects du changement climatique pourraient être incorporés à l'analyse de fréquence d'inondation pour des buts de conception et de planification.

Mots clés : *Bassin de captation, délimitation des régions homogènes, test de mesure de discordance, analyse hiérarchique en grappes, L-moments, estimation de quantiles*

1. INTRODUCTION

1.1 Background

Flooding is a major natural disaster that poses a significant threat to economic development in many countries. Floods account for about one-third of all natural catastrophes throughout the world, and are responsible for more than half of the fatalities. The high floods that have affected a number of regions in recent years have attracted great attention globally. Trend analysis reveals that major floods and the losses generated by them have increased drastically in recent years. Flood Frequency Analysis is one of the techniques applied by hydrologists to estimate the probabilities associated with design flood[4] events, thus answering questions on the estimation of magnitudes of floods and their probable frequency of occurrence (or return period).

It has been recognized that small sample sizes commonly available in hydrologic work cannot be reliably handled using the at-site approach (Adamowski *et al.*, 1994). Hosking and Wallis (1997) developed the Regional Frequency Analysis (RFA) method to resolve this problem by "trading space for time." That is, data from several sites were used in estimating event frequencies at any one site. They also illustrated that L-moments are efficient in estimating parameters of a wide range of distributions for small samples.

A reliable estimation of magnitude and frequency of occurrence of extreme events such as flooding is of great significance in minimizing damage by facilitating proper planning and design of civil engineering structures such as bridges, barrages and dams (Gelder *et al.* 1997). The choice of the functional form of the underlying flood frequency distributions has a large effect on the flood quantile estimates, especially since the quantiles that are of interest are

4 "Design flood" may be defined as a hypothetical flood representing a specific likelihood of occurrence.

those in the extreme right-hand tail of the distribution. Although Uganda is one of the countries frequently affected by floods in the main rivers and their tributaries, no flood frequency analysis has been carried out on the rivers. Hence, there has been a lack of estimates of the peak discharge for specified return periods of occurrence, which is substantially longer than the available gauged record.

This study aimed at conducting an RFA of annual maximum river discharges from selected major rivers in Uganda. The specific objectives of this study were: to delineate homogeneous regions for an RFA; to establish the underlying statistical distributions; and to estimate the distribution parameters and exceedance quantiles of annual peak discharges.

1.2 Regional Frequency Analysis (RFA)

A "region" in this study means a group of catchments, which is assumed to have data drawn from the same frequency distribution. An RFA is widely used in flood frequency analysis when data are available from more than one site. Regionalization allows for a pooling of data from more than one site in such a way as to produce a single regional flood frequency curve. This curve is applicable, after rescaling using each site's sample mean annual peak discharge, anywhere in that region. The biggest advantage of regional estimation is the increase in record length. Studies by Lettenmaier *et al.* (1987) and Hosking (1990) have shown that flood estimates based on regional information are more accurate (have less absolute error) and are more stable (have less variance) than those based solely on at-site records.

1.3 Estimation of distribution parameters and at-site exceedance flood quantiles

A common problem in statistics is the estimation, from a random sample of size n, of a probability distribution whose specification involves a finite number, p, of unknown parameters. Analogous to the usual method of moments, the "method of L-moments" obtains parameter estimates by equating the first p sample L-moments to the corresponding population quantities.

If X is a (real) random variable with cumulative distribution function $F(x)$ and quantile function, $x(F)$ and if $X_{1:n} \leq X_{2:n} \leq \ldots \leq X_{n:n}$ are the order statistics of a random sample size n drawn from the distribution of X, then the L-moments of X are defined to be the quantiles

$$\lambda_r \equiv r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} E\left(X_{r-k:r}\right) \text{ for } r = 1,2,\ldots \quad (1)$$

To standardize the higher L-moments, λ_r, $r \geq 3$, so that they are independent of the units of measurements of X, the L-moments ratios of are defined as the quantiles

$$\tau_r = \lambda_r / \lambda_2 ; r = 3,4,..... \quad (2)$$

In particular, λ_1 is the mean of the distribution; λ_2 is a measure of the scale or dispersion; τ_3 and τ_4 are measures of skewness and kurtosis, respectively.

For an ordered sample $x_1 \leq x_2 \leq ..... \leq x_n$, estimates of the first few L-moments are:

$$l_1 = \sum_{i=1}^{n} x_i / n ,$$

$$l_2 = \sum_{i>j} (x_i - x_j) / n(n-1) ,$$

$$l_3 = \sum_{i>j>k} 2(x_i - 2x_j + x_k) / n(n-1)(n-2) , \quad (3)$$

General formulae are given by Hosking (1990). l_1 is the sample mean; l_2 is the scale. L-skewness and L-kurtosis are estimated by $t_3 = l_3 / l_2$ and $t_4 = l_4 / l_2$, respectively.

1.3.1 Generalized Extreme Value (GEV) distribution

The Generalized Extreme Value (GEV) distribution combines into a single form the three possible types of limiting distributions for extreme values. The GEV is probably the most widely used distribution when measuring annual peak discharge series of a gauging site. A typical application consists of fitting one type of extreme value limiting distribution to the series of annual maxima.

The GEV distribution function is given by

$$F(x) = \begin{cases} \exp\left[-\left\{ 1 - k\left(\frac{x-\xi}{\alpha} \right) \right\}^{\frac{1}{k}} \right], k \neq 0, \\ \exp\left[-\exp\left\{ -\left(\frac{x-\xi}{\alpha} \right) \right\} \right], k = 0, \end{cases} \quad (4)$$

Where X is bounded by $(\xi + \alpha / k)$ from above if $k > 0$ and from below if $k < 0$, ξ is the location parameter $\alpha(> 0)$ is the scale parameter and k is the shape parameter. The shape parameter determines which type of extreme value distribution is represented.

The type I GEV (Gumbel) distribution corresponds to $k = 0$, the type II GEV (Frechet) distribution corresponds to $k < 0$ and the type III GEV (Gumbel) distribution corresponds to $k > 0$. This distribution is often used in hydrology to analyze extreme low river discharges. Hosking (1990) used L-moments to show that point estimates of the GEV distribution can be obtained by using;

$$z = \frac{2}{(3 + t_3)} - \frac{\log 2}{\log 3},$$

$$\hat{k} \approx 7.8590z + 2.9554z^2,$$

$$\hat{\alpha} = \frac{l_2 \hat{k}}{(1 - 2^{-\hat{k}})\Gamma(1 + \hat{k})},$$

$$\hat{\xi} = l_1 + \frac{\hat{\alpha}\left\{\Gamma\left(1 + \hat{k}\right) - 1\right\}}{\hat{k}} \tag{5}$$

where l_1 is the sample mean of annual peak discharges, l_2 is a measure of scale and t_3 is a measure of skewness.

1.3.2 Generalized logistic distribution

The distribution function for the 3-parameter generalized logistic distribution (GLO) is

$$F(x) = \left[1 + \left(1 - \frac{k}{\alpha}(x - \xi)\right)^{\frac{1}{k}}\right]^{-1} \tag{6}$$

As with the GEV distribution, ξ is the location parameter, $\alpha(> 0)$ is the scale parameter and k is the shape parameter. When $k = 0$, the GLO distribution reduces to the 2-parameter logistic distribution.

Hosking (1990) showed that point estimates of the parameters of the GLO distribution can be obtained via L-moments by using;

$$\hat{k} = -t_3,$$

$$\hat{\alpha} = \frac{l_2}{\Gamma(1+\hat{k})\Gamma(1-\hat{k})},$$

$$\hat{\xi} = l_1 + \frac{l_2 - \hat{\alpha}}{\hat{k}} \tag{7}$$

1.3.3 The index flood approach

The underlying flood frequency distribution at each gauging site is assumed to be identical, except for a scale factor. Consequently we are able to use a straightforward pooling approach. First, the data at each gauging site are normalized by the index flood. The site-specific scaling factor is called the index flood. The index flood was taken to be the sample mean of annual peak discharges at site *i*. Next, the parameters of a dimensionless regional flood frequency curve are estimated. Finally, the parameters are rescaled at the site of interest by a local estimate of the scaling factor, usually the at-site mean annual peak discharge.

The key assumption of an index flood procedure is that the region is homogeneous, that is, the frequency distributions of the N sites in a region are identical, apart from a site-specific scaling factor. The distribution common to all sites in the region is called the regional frequency distribution. It is dimensionless and defined by its (regional) quantiles, $q(F)$, $0 \leq F \leq 1$. It is usually assumed that the form of $q(F)$ is known apart from p undetermined parameters $\theta_1, \theta_2, \ldots, \theta_p$. The site-specific scaling factor is called the index flood, denoted by μ_i at gauging site i. The index flood is usually taken to be the sample mean of annual peak discharges at site i. Thus we can write:

$$Q_i(F) = \mu_i q(F), \; i = 1,2\ldots,N \tag{8}$$

where $Q_i(F)$ is the quantile of non-exceedance probability F at site i.

A standard scaled data approach is the simplest index flood method. This involves dividing each measure by its at-site sample mean of annual peak discharges, and then treating all the scaled data points as if they were observations from the regional frequency distribution. Parameter estimates

are found and the estimated regional flood distribution is then multiplied by the at-site sample mean of the annual peak discharges for the site under investigation.

1.4 Identification of distributions

1.4.1 Identification of distributions using L-moment ratio diagrams

The L-moments ratio diagram is widely used to aid model selection, particularly for skewed distributions. The L-moment ratio diagram can be used to compare the L-skewness – L-kurtosis relations of different distributions and data samples. This gives a visual indication of which distribution may be expected to give a good fit to a data sample or samples. Theoretical curves of the candidate distributions as well as the regional peak number weighted average L-skewness – L-kurtosis space were plotted on the same graph to select the best-fit-distribution based on the proximity of the regional L-skewness – L-kurtosis coordinate point to the theoretical curves.

The proximity of the record length weighted average to a particular candidate distributions theoretical curve in L-skewness – L-kurtosis space has been interpreted as an indication of the appropriateness of that distribution to describe the regional data (Vogel *et al.* 1993; Hosking and Wallis 1995).

To construct an L-moment ratio diagram, it is convenient to have simple explicit expressions for t_4 (L-kurtosis) in terms of t_3 (L-skewness) for the candidate probability distributions. Polynomial approximations of the form

$$t_4 = A_0 + A_1 t_3 + A_2 t_3^{\,2} + ... + A_8 t_3^{\,8} \qquad (9)$$

have been obtained, and the coefficients are given in Table 1 "Overall lower bound" is the lower bound on t_4 for all distributions (Hosking, 1990). For given t_3, the approximations yield values of t_4 that are accurate to within 0.0005 provided that t_3 is in the range -0.9 to +0.9, except that for the generalized extreme-value distribution 0.0005 accuracy is attained only when t_3 is between -0.6 and +0.9.

Table 1: Coefficients of the polynomial approximations of the L-kurtosis – L-skewness relations

	Generalized logistic	Generalized extreme value	Lognormal
A_0	0.16667	0.10701	0.12282
A_1		0.11090	
A_2	0.83333	0.84838	0.77518
A_3		-0.06669	
A_4		0.00567	0.12279
A_5		-0.04208	
A_6		0.03673	-0.13638
A_7			
A_8			0.11368

The blank entries in Table 1 are zero coefficients of the polynomial approximations of the L-kurtosis – L-skewness relations generated by Hosking (1990).

1.4.2 Identification of distributions using minimum L-kurtosis difference criterion

Hosking and Wallis (1997) also proposed a simple but effective approach to fit 3-parameter distribution functions. The approach involves the computation of four L-moments from a given sample. By matching the first three L-moments, a set of 3-parameter distribution functions can be fitted to the sample data. They propose that the distribution with its L-kurtosis value closest to that of the sample value can be taken as the most acceptable distribution functions, which should be used for quantile estimation. In essence, the general idea is to consider the L-kurtosis as an empirical summary measure of the tail weight or shape of the distribution.

1.5 Screening of data using discordancy measure test

Discordancy among the sites is measured in terms of L-moments. This concept has the advantage of testing for homogeneity on the basis of characteristics that are not used in the clustering of sites (Hosking and Wallis, 1997). The objective of screening the data is to check that the data are suitable for carrying out the RFA. Discordancy measure D_i is used to screen out the data from unusual stations whose point sample L-moments are markedly

different from other stations. In this test, L-moment ratio (L-coefficient of variation, L-skewness, and L-kurtosis) of a site is used to describe that site in three-dimensional space. Discordancy measure, D_i, is defined as:

$$D_i = \frac{N}{3(N-1)}\left(\mu_i - \overline{\mu}\right)^T S^{-1}\left(\mu_i - \overline{\mu}\right) \tag{10}$$

where = μ_i^T (L-coefficient of variation, L-skewness, L-kurtosis), $\overline{\mu}$ is the sample mean vector of the annual peak discharges and *S* is the sample covariance matrix.

The site *i* is declared to be discordant, if D_i is greater than the critical value of the discordancy statistic. Generally any site with D > 2.757 can be regarded as discordant (Hosking and Wallis 1993).

2. MATERIALS AND METHODS

2.1 Data

This study utilized data on site characteristics including latitude, longitude, elevation above sea level, catchment area, and annual peak discharge series for the year 1948 through 2005 recorded at 20 gauging sites in Uganda. The data were provided by the Directorate of Water Development, Ministry of Water, Lands and Environment.

2.2 Procedure of analysis

The procedure of analysis followed in this study is outlined by Hosking and Wallis (1997). The main steps were to screen the data, identify homogeneous regions using a hierarchical cluster analysis procedure, choose a 3-parameter frequency distribution using both L-moment ratio diagram and minimum L-kurtosis difference criteria, estimate the parameters of the chosen distribution using the "method of L-moments" by equating the first p sample L-moments of annual peak discharges to the corresponding population quantities and compute exceedance quantile estimates.

The regional flood frequency relationship for estimation of flood quantile of annual peak discharges of various return periods is expressed as:

$$Q_T = \overline{Q} + K_T S\,, \tag{11}$$

where Q_T is the estimated flood quantile of annual peak discharges (in cubic meter per second) for a designated return period (T, years).

$\overline{Q}$ is the mean annual peak discharge which in case of gauged site, is computed from the observed data.

K_T is the frequency factor, a function of the return period and probability distribution.

S is the standard deviation of annual peak discharges.

The regional flood frequency relationship for estimation of floods of various return periods for the gauged catchments is expressed as:

1. Generalized extreme value distribution

$$Q_T = \left[\xi + \frac{\alpha\left\{1 - \left(-\log F\right)^k\right\}}{k} \right] * \overline{Q} \tag{12}$$

$$x_T = \xi + \frac{\alpha\left\{1 - \left(-\log F\right)^k\right\}}{k}$$

2. Generalized logistic distribution

$$Q_T = \left[\xi + \frac{\alpha\left[1 - \left\{\frac{\left(1-F\right)}{F}\right\}^k\right]}{k} \right] * \overline{Q} \tag{13}$$

$$x_T = \xi + \frac{\alpha\left[1 - \left\{\frac{\left(1-F\right)}{F}\right\}^k\right]}{k}$$

Where x_T is the quantile function of the probability distribution.

The return period T is related to the probability of non-exceedance (F) by the relation:

$$F = 1 - \frac{1}{T} \tag{14}$$

where $F = F(x_T)$ is the probability of having an annual peak discharge of magnitude x_T or smaller.

The statistical data analysis tools used were SPSS® 16.0 (SPSS Inc., 2007) and Stata® 11 (StataCorp, 2009). In addition SWP® 2.5 (MacKichan Software Inc., 2002) was used as a supplementary tool mainly for complex mathematical computations.

3. RESULTS AND DISCUSSION

3.1 Descriptive statistics of annual maximum peak discharge by gauging site

Descriptive statistics of annual peak discharge series recorded at the gauging sites are presented in Table 2.

3.2 Delineation of homogeneous regions

The purpose of this step was to form groups of gauging stations that satisfy the homogeneity condition.

3.2.1 Hierarchical cluster analysis procedures

Gauging sites were grouped using hierarchical cluster analysis procedures (average linkage method) based on latitude, longitude, catchment area, elevation above sea level, and annual peak discharge. A key step in a hierarchical clustering is to select a distance measure. The Euclidean distance measure was used. It is computed as:

$$\text{distance(x,y)} = \left\{ \sum_{i=1}^{n} (x_i - y_i)^2 \right\}^{1/2} \tag{15}$$

Table 2: Statistics of annual flood series of various sites in the study

Gauging site	Number of peaks	Elevation (meters, above sea level)	Catchment area (km^2)	Mean annual peak discharge (m^3/sec)	SD of mean peak discharge
Akokorio	30	1042.4	1400.9	40.264	11.345
Aswa	34	914.6	13275.0	399.324	55.317
Enget	21	1076.8	105.2	3.291	1.669
Kafu	50	1040.0	12952.0	1332.908	1270.24
Kagera	9	1133.0	30200.0	273.921	115.892
Kapiri	30	1031.9	14122.9	50.267	15.634
Katonga	27	1130.0	13930.0	7.060	0.936
Kyambura	36	1028.0	660.0	29.886	1.755
Laropi	20	635.0	427131.0	1468.832	358.731
Manafwa	50	1100.2	494.2	39.380	2.858
Masindi	49	1020.6	338465.0	1208.128	353.727
Mpologoma	38	1078.0	3614.0	80.967	11.247
Nkusi	34	1064.0	2839.0	22.353	3.707
Nyagak	27	1443.0	602.0	20.473	1.435
Ora	25	640.2	1750.0	127.120	21.166
Oru	39	887.0	431.0	100.016	7.0404
Panyango	21	611.9	413046.0	1476.525	43.369
Paraa	15	641.0	349207.0	1625.826	235.586
Rwizi	52	1386.0	2070.0	38.562	4.017
Sironko	38	1118.0	265.0	25.025	2.365

Source: Directorate of Water Development, Ministry of Lands, Water and Environment (2006)

Given a distance measure, elements can be combined to build a hierarchy of clusters. The traditional representation of the hierarchy of clusters is a tree data structure (dendrogram) with individual elements at one end and

a single cluster with every element at the other. It shows the relative size of the proximity coefficients at which cases were combined. The bigger the distance coefficient, the more clustering involved combining unlike entities. The result of the grouping and the distance levels at which the rivers occur are clearly illustrated by the dendrogram in Figure 1.

Results from the dendrogram revealed that the 20 gauging sites form two homogeneous regions at third distance level. Region A consists of 16 sites while region B has four sites. The coefficient of variation of the at-site LCv for all sites within a region was computed as a preliminary check to find out whether the proposed region was reasonably homogeneous. A homogeneous region is expected to have a low value of coefficient of variation (less than 0.3). Results show that no region is reasonably homogeneous. However, region B (0.381) is as homogeneous as region A (0.388). However, regions A and B were not further subdivided following from Hosking and Wallis (1997) who state that even in regions with heterogeneity, RFA is more accurate than at-site analysis.

Results showed that gauging sites grouped within a region shared two characteristics in common: lying at nearly the same elevation above sea level and experiencing nearly the same mean annual peak discharge of a site. Generally, region A (except Oru site) constitutes sites lying on highlands of Lake Kyoga drainage and Lake Victoria drainage with a weighted mean annual peak discharge value of 190.986 m^3/sec while region B (except Masindi site) constitute sites lying on lowlands of River Nile drainage with a weighted mean annual peak discharge value of 1371.136 m^3/sec. Therefore, the relief of a gauging site was evident as a crucial factor in forming a homogeneous region as can be seen from the composition of all sites within the R. Nile drainage (except Masindi site).

3.2.2 Screening of data using discordancy measure test

The values of discordancy measure computed indicated that all sites within each homogeneous region are non-discordant (have values below 2.757) and hence, all sites were included into flood frequency analysis. The summary of descriptive L-statistics is shown in Table 3.

Figure 1: Dendrogram using average linkage (within groups) distance method

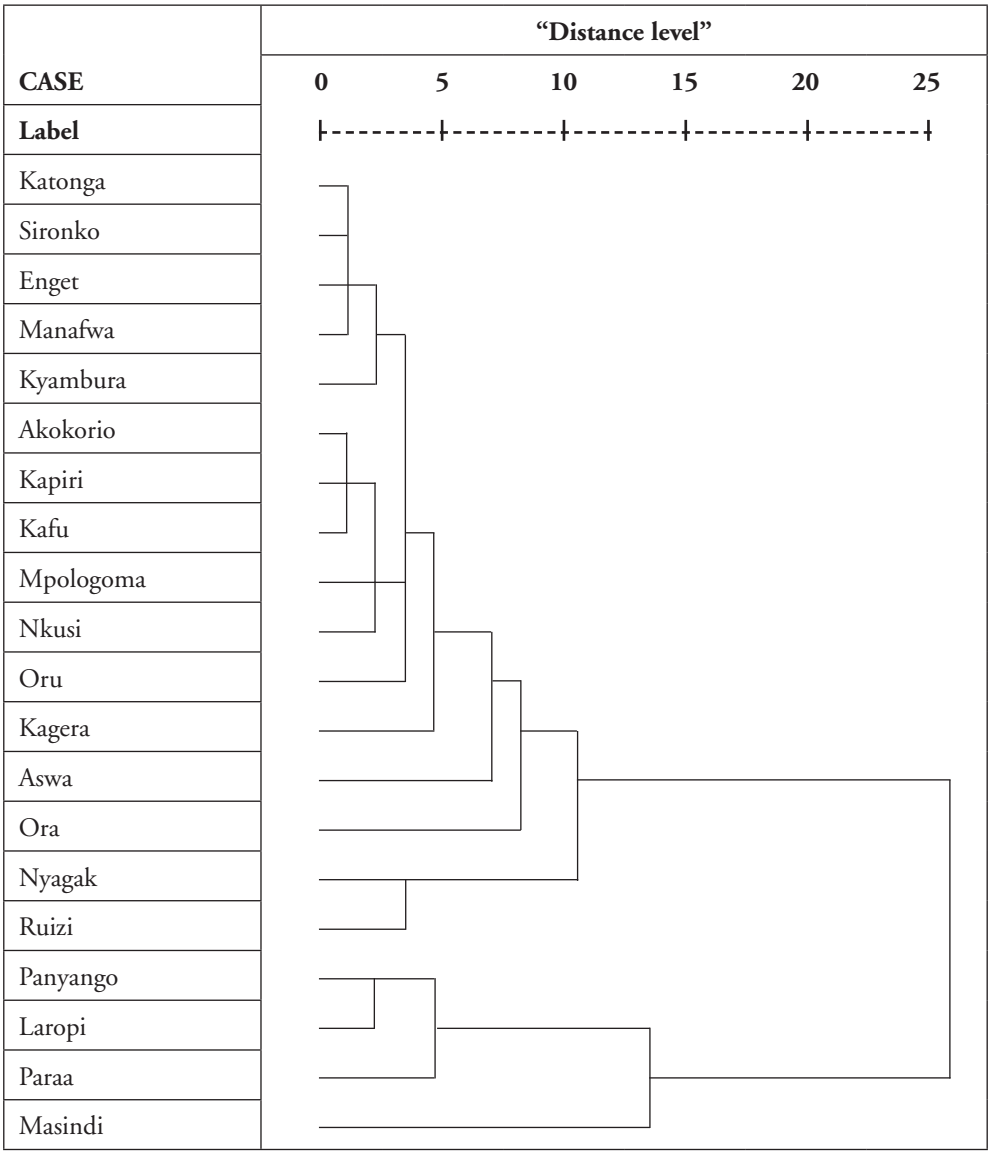


Table 3: Standardized descriptive L-statistics of the annual peak river discharges

Region	Gauging site	L-Coefficient of variation	L-skewness	L-kurtosis	Discordancy measure (D_i)
A	Akokorio	0.648	0.615	0.378	1.432
	Aswa	0.419	0.339	0.207	0.585
	Enget	0.298	-0.025	0.011	1.369
	Kafu	0.361	0.232	0.138	0.183
	Kagera	0.291	0.044	0.094	0.812
	Kapiri	0.698	0.660	0.416	2.109
	Katonga	0.352	0.347	0.258	0.392
	Kyambura	0.200	0.044	0.087	0.481
	Manafwa	0.294	0.086	0.083	0.300
	Mpologoma	0.458	0.299	0.135	0.572
	Nkussi	0.340	0.465	0.507	2.565
	Nyagak	0.194	0.193	0.261	1.126
	Ora	0.399	0.414	0.320	0.634
	Oru	0.247	0.020	0.168	1.781
	Rwizi	0.291	0.184	0.070	1.161
	Sironko	0.315	0.236	0.129	0.700
B	Laropi	0.134	-0.145	0.253	0.893
	Masindi	0.168	0.011	0.061	0.742
	Panyango	0.077	0.070	0.102	0.697
	Paraa	0.081	0.220	0.211	0.763

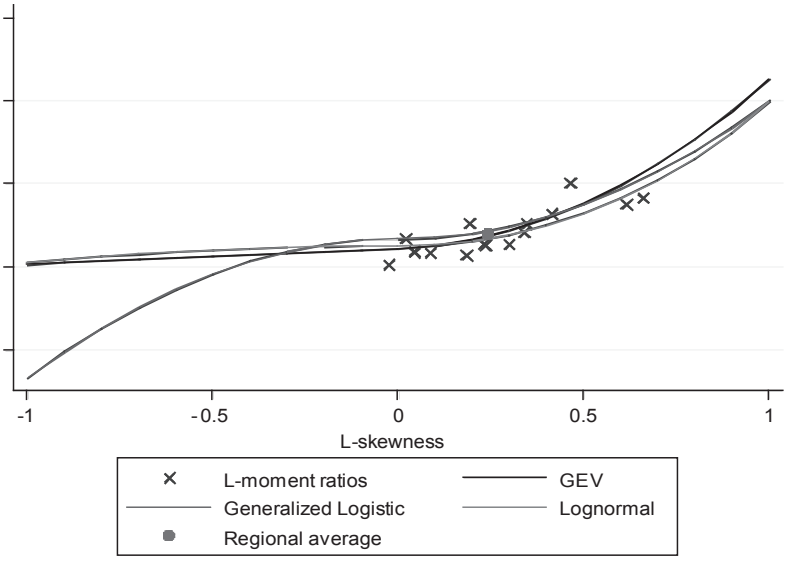
3.3 Determination of the regional probability distribution

Three 3-parameter distributions were fitted to regional data based on L-moment ratio diagram and minimum L-kurtosis difference criterion.

3.3.1 Determination of the frequency distribution based on L-moment ratio diagram

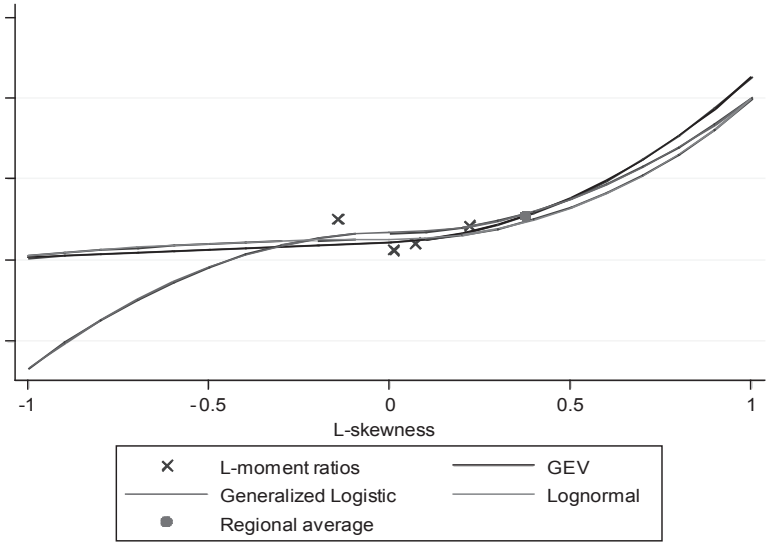
Theoretical curves of the candidate distributions as well as the regional L-skewness and L-kurtosis were plotted on the same graph to select the best-fit-distribution.

Figure 2: L-skewness – L-kurtosis relations of different candidate distributions overlaid on scatter plot of data samples for region A



From Figure 2, it is concluded that generalized extreme value distribution is the most appropriate candidate distribution for data from sites in region A. This result is in agreement with Jaiswal *et al.* (2002), who found out that GEV was the most suitable for the Beas basin and Kumar *et al.* (2003), who also found GEV to be the robust distribution for North Brahmaputra river system.

Figure 3: L-skewness – L-kurtosis relations of different candidate distributions overlaid on scatter plot of data samples for region B



From Figure 3, it is concluded that generalized logistic distribution is the best candidate distribution for data from sites within region B. This distribution was also proposed by Jaiswal *et al.* (2002) and Gelder *et al.* (2002).

3.3.2 Determination of the frequency distribution based on minimum L-kurtosis difference criterion

Four L-moments were computed from the data. By matching the first three L-moments, a set of 3-parameter distribution functions was fitted to the sample annual peak discharges. The distribution with its L-kurtosis value closest to that of the sample value was taken as the most acceptable distribution function, and was used for quantile estimation of annual peak discharges.

Table 4: Minimum L-kurtosis difference criterion based on annual peak discharge data from region A

L-moments of annual peak discharges (m^3/sec)	Sample values	L-moments of fitted distribution functions		
		Generalized Extreme Values	Generalized Logistic	Lognormal
Mean (l_1)	80.17125	80.17125	80.17125	80.17125
Scale (l_2)	30.24394	30.24394	30.24394	30.24394
Skewness (t_3)	0.259563	0.259563	0.259563	0.259563
Kurtosis (t_4)	0.203875	1.6761	0.22281	0.293
L-kurtosis difference	t_4 (sample) - t_4 (distribution)	-1.47223	-0.01894	-0.08913

Results as shown in Table 4 suggest that generalized logistic is the most appropriate distribution which is different from finding based on L-moment ratio diagram, where the generalized extreme value distribution was the most appropriate distribution. Hence, both distributions were fitted to the annual peak discharge data from gauging sites within region A.

Table 5: Minimum L-kurtosis difference criterion based on annual peak discharge data for region B

L-moments of annual peak discharges (m^3/sec)	Sample values	L-moments of fitted distribution functions		
		Generalized Extreme Values	Generalized Logistic	Lognormal
Mean (l_1)	1449.579	1449.579	1449.579	1449.579
Scale (l_2)	161.746	161.746	161.746	161.746
Skewness (t_3)	-0.029	-0.029	-0.029	-0.029
Kurtosis (t_4)	0.157	1.697	0.166	0.293
L-kurtosis difference	t_4 (sample) - t_4 (distribution)	-1.54	-0.009	-0.136

As shown in Figure 2, Figure 3, and Table 5, findings based on L-moment ratio diagram and those based on the minimum L-kurtosis difference both show that generalized logistic distribution is the most appropriate candidate.

3.4 Estimation of distribution parameters and at-site exceedance flood quantiles of annual peak discharges

The quantiles of annual peak discharges that are of interest are those in the extreme right-hand tail of the probability distribution. Consequently, estimates of exceedance quantiles of annual peak discharges, $Q(0.95)$, $Q(0.98)$, $Q(0.99)$, $Q(0.995)$ and $Q(0.999)$ were computed for the probability distribution(s) found appropriate for each gauging site.

Study findings summarized in Table 6 reveal that for the majority of the sites, parameter estimates derived from both GEV and GLO did not show wide variation. However, estimates of exceedance quantile of annual peak discharges derived from GEV were in most cases higher than estimates from GLO at each value of annual exceedance probability.

Determining the probability that a particular event will occur in a particular location within a particular time span is an essential part of hazard prediction. Using estimates of quantiles of annual peak discharges and the observed peak discharges over the period of observation, it was possible to estimate the annual exceedance probability or equivalently the return period for each of the sites in the study. The return period is related to the probability of non-exceedence (F) by the relation,

$$F = 1 - \frac{1}{T},$$

where $F = F(x_T)$ is the probability of having an annual peak discharge of magnitude x_T or smaller. The problem of estimating the return period reduces to the evaluation of the quantile function x_T for a given value of F. In particular, the highest annual peak discharge with a 1 percent probability of occurrence in the next year has a return period of 100 years.

The interpretation of the results is presented for the first two gauging stations Akokorio and Aswa discharges. The interpretation is similar for the other gauging stations.

Table 6: Parameter and quantile estimates of annual peak discharges by gauging site within region A

Site	P.d.f	Parameter			Quantiles (m^3/sec)				
		Location	Scale	Shape	0.95	0.98	0.99	0.995	0.999
Akokorio	GEV	12.506	13.974	-0.593	126.019	227.065	349.122	532.917	1402.607
	GLO	14.315	10.121	-0.615	98.502	178.087	275.618	424.584	1148.891
Aswa	GEV	236.789	180.926	-0.248	1031.225	1427.448	1790.530	2220.558	3553.885
	GLO	311.16	137.44	-0.339	1005.800	1422.400	1830.700	2344.800	4120.400
Enget	GEV	2.718	1.774	0.330	6.077	6.611	6.916	7.158	7.544
	GLO	2.140	1.011	0.025	5.010	5.889	6.528	7.152	8.552
Kafu	GEV	42.671	29.758	-0.095	144.757	183.173	214.268	247.366	332.961
	GLO	39.158	17.184	-0.232	111.747	147.798	180.180	218.001	332.823
Kagera	GEV	190.689	116.896	0.207	449.997	503.535	537.4	566.605	620.080
	GLO	159.971	65.731	-0.044	366.609	438.994	494.714	551.762	690.496
Kapiri	GEV	12.853	16.107	-0.643	156.899	295.603	469.975	741.892	2112.594
	GLO	15.387	12.050	-0.660	124.608	235.355	376.077	597.892	1739.632
Katonga	GEV	4.632	2.647	-0.259	16.468	22.488	28.053	34.696	55.563
	GLO	4.526	1.608	-0.347	12.763	17.772	22.713	28.969	50.790
Kyambura	GEV	25.803	10.106	0.207	48.220	52.848	55.776	58.301	62.924
	GLO	23.147	5.682	-0.044	41.011	47.269	52.086	57.018	69.011

Site	P.d.f	Parameter			Quantiles (m³/sec)				
		Location	**Scale**	**Shape**	**0.95**	**0.98**	**0.99**	**0.995**	**0.999**
Manafwa	GEV	30.851	18.641	0.136	76.419	87.321	94.633	101.261	114.411
	GLO	26.607	10.468	-0.086	61.682	74.991	85.597	96.779	125.341
Mpolog-oma	GEV	45.961	43.245	-0.192	219.169	297.264	365.670	443.595	669.639
	GLO	42.844	25.653	-0.299	163.975	231.735	296.021	374.712	633.661
Nkussi	GEV	14.510	6.221	-0.414	50.860	75.029	100.334	133.972	261.449
	GLO	14.732	4.055	-0.465	40.302	59.282	79.890	108.227	222.458
Nyagak	GEV	17.076	5.534	-0.036	34.418	40.249	44.747	49.342	60.437
	GLO	16.269	3.158	-0.193	28.789	34.583	39.624	45.353	61.956
Ora	GEV	75.782	46.843	-0.349	319.792	464.961	609.195	792.406	1433.583
	GLO	95.366	37.607	-0.414	311.910	459.530	613.310	817.350	1589.700
Oru	GEV	83.967	42.794	0.249	173.792	190.774	201.152	209.844	225.034
	GLO	71.822	24.135	-0.020	145.020	169.503	187.981	206.584	250.579
Rwizi	GEV	27.505	14.983	-0.022	73.487	88.538	100.020	111.636	139.233
	GLO	25.222	8.530	-0.184	58.555	73.731	86.837	101.638	144.074
Sironko	GEV	17.971	10.265	-0.101	53.514	67.041	78.042	89.801	120.425
	GLO	16.788	5.936	-0.236	42.028	54.654	66.031	79.357	120.009

(i) Akokorio discharge The highest annual peak discharge value of 265.835 m^3/sec observed at this gauging site over a period of 30 years is less than the estimated exceedance quantile of the annual peak discharge of magnitude 349.122 m^3/sec corresponding to the annual non-exceedance probability of 0.99 or the return period of 100 years. The results suggest that the gauging site on average will experience the highest annual peak discharge once every 100 years.

(ii) Aswa discharge The highest annual peak discharge value of 1270 m^3/sec observed at this gauging site over a period of 34 years is less than the estimated exceedance quantile of the annual peak discharge of magnitude 1427.448 m^3/sec corresponding to the annual non-exceedance probability of 0.98 or the return period of 50 years. The results suggest that the gauging site on average will experience the highest annual peak discharge once every 50 years.

Estimates of exceedance quantile of annual peak discharges derived from generalized logistic distribution for sites within region B do not show very wide variation. This finding supports the contention that the sites form a homogeneous region.

4. CONCLUSIONS

This study has established that the majority of the gauging sites in Uganda can be regionalized on the basis of relief. Any site within either Lake Victoria or Kyoga drainage could be grouped under highland region while those sites within R. Nile drainage were grouped under lowland region. Other sites outside the three basins were regionalized on the basis of their mean annual peak discharge.

Of the three distributions used in this study and for the same data, LNO gave quantile estimates of annual peak discharges which were highest at each value of annual exceedance probability followed by GEV and then GLO. Generally, parameter and quantiles estimates of annual peak discharges from both GEV and GLO fitted to data from majority of the sites within Lake Victoria or Kyoga drainage did not show very high variations so that either GEV or GLO could be adopted for use.

Table 7: Parameter and quantile estimates of annual peak discharges derived from generalized logistic distribution for gauging sites within region B

Gauging sites	Parameters			Quantiles (m^3/sec)				
	Location	Scale	Shape	0.95	0.98	0.99	0.995	0.999
Laropi	1273.463	167.894	-0.145	1890.116	2151.422	2369.989	2610.172	3267.702
Masindi	975.072	200.583	-0.011	1575.346	1772.657	1920.469	2068.341	2414.437
Panyango	1350.177	104.712	-0.07	1692.568	1818.609	1917.734	2021.086	2280.161
Paraa	1487.928	100.965	-0.22	1906.148	2109.416	2290.211	2499.605	3126.265

5. RECOMMENDATION

The results obtained from this study were based on the data that were available. In most cases the time series available was rather short. The results presented in this study could be improved in future when more data on peak discharges become available.

In this study only the statistical frequency analysis approach was used. RFA requires an assumption of stationarity in which climatic trends or cycles do not affect flood discharges. However, there is clear evidence that this is not the case and as reported by Knox (1993), even modest changes in the climate can result in large changes in flood magnitude. Therefore further research is needed to determine how aspects of climate change can be incorporated into flood frequency analysis for design and planning purposes.

REFERENCES

Adamowski, K., P.J. Pilon, B. Bobée, and Y. Alila (1994). "The Identification of a Regional Distribution for Canada." Paper presented at the International Workshop – "Closing the Gap between Theory and Practice in Urban Rainfall Applications," November 30–December 2, Switzerland.

Burn, D.H. (1990). "Evaluation of Regional Flood Frequency Analysis with a Region of Influence Approach," *Water Resource Research*, vol. 26, no. 10, pp. 2257–2265.

Cavadias, G.S. (1990). "The Canonical Correlation Approach to Regional Flood Estimation," *IAHS Publication*, vol. 191, pp. 171–178.

Cunnane, C. (1989). "Statistical Distributions for Flood Frequency Analysis," *World Meteorological Organization Operational Hydrology*, Report No. 33, WMO-No. 718, Geneva, Switzerland.

Gelder, P. van, and J.K. Vrijling (1997). "A Comparative Study of Different Parameter Estimation Methods for Statistical Distribution Functions in Civil Engineering Applications," *Proceedings of ICOSSAR'97*, pp. 665–668, 7th International Conference on Structural Safety and Reliability, Kyoto.

Gelder, P. Van, M. Pandey, and H. Vrijling (2002). "The Sensitivity of Quantile Estimates to the Distribution Shape," *IABSE Conference Report on Safety, Risk and Reliability: Trends in Engineering*. Malta, pp. 501–506.

Hosking, J.R.M. (1990). "L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics," *Journal of Royal Statistical Society B*, vol. 52, no. 1, pp. 105–124.

Hosking, J.R.M. and J.R. Wallis (1993). "Some Statistics Useful in Regional Frequency Analysis," *Water Resources Research*, vol. 29, no. 2, pp. 271–281.

Hosking, J.R.M. and J.R. Wallis (1995). "A Comparison of Unbiased and Plotting Position Estimators of L-Moments," *Water Resources Research*, vol. 31, no. 8, pp. 2019–2025.

Hosking, J.R.M. and J.R. Wallis (1997). *Regional Frequency Analysis*. Cambridge, UK: Cambridge University Press.

Jaiswal, R.K., S. Mukherjee, D.K. Raju *et al.* (2002). "Forest Fire Risk Zone Mapping from Satellite Imagery and GIS," *International Journal of Applied Earth Observation and Geoinformation*, vol. 4, pp. 1–10.

Jaiswal, R.K., N.K. Goel, P. Singh, and T. Thomas (2003). "L-Moment-based Flood Frequency Modeling," *IE(I) Journal – CV*, vol. 84, no. 2, pp. 6–10.

Knox, J.C. (1993). "Large Increases in Flood Magnitude to Modest Changes in Climate," *Nature*, vol. 361, pp. 430–431.

Kumar, R., C. Chatterjee, N. Panigrahi, B.C. Patwari, and R.D. Singh (2003). "Development of Regional Flood Formulae Using L–moments for Gauged and Ungauged Catchments of North Brahmaputra River System," *Journal of Hydrologic Engineering, American Society of Civil Engineers*, vol. 84, pp. 57–63.

Kumar, R. and C. Chatterjee (2005). "Regional Flood Frequency Analysis using L-moments for North Brahmaputra Region of India," *Journal of Hydrologic Engineering. American Society of Civil Engineers*, vol. 10, no. 1, pp. 1–7.

Kysely, J., J. Piceck, and R. Huth (2005). "Regional Analysis of Extreme Precipitation Events in the Czech Republic," *European Geosciences Union*, vol. 7.

Lecce, S.A. (2000). "Seasonality of Flooding in North Carolina," *Southeasterner Geographer*, vol. XXXX1, no. 2, pp. 168–175.

Lettenmaier, D.P and K.W. Potter (1985). "Testing Flood Estimation Methods using a Regional Flood Generations Model," *Water Resources Research*, vol. 21, no. 12, pp. 1903–1914.

Lettenmaier, D.P., J.R. Wallis, and E.F. Wood (1987). "Effect of Regional Heterogeneity on Flood Frequency Estimation," *Water Resources Research*, vol. 23, no. 2, pp. 313–323.

MacKichan Software Inc. (2001) Scientific Workplace version 2.5. MacKichan Software Inc., Poulsbo, WA.

Nathan, R.J. and T.A. McMahon (1990). "Identification of Homogeneous Regions for the Purposes of Regionalization," *Journal of Hydrology*, vol. 49, pp. 173–192.

NERC (1975). *Flood Studies Report*. Volume I – V. NERC. London: Whitefriars Press.

Royston, P. (1992). "Which Measures of Skewness and Kurtosis are Best?," *Statistics in Medicine*, vol. 11, no. 33, pp. 333–343.

Ruttan, J.A. et al. (2001). *Guidelines on Extreme Flood Analysis*. A document prepared for Alberta Transportation and Civil Engineering Division, Civil Projects Branch. Canada.

Sankarasubramanian, A. and K. Srinivasan (1999). "Investigation and Comparison of Sampling Properties of L-moments and Conventional Moments," *Journal of Hydrology*, vol. 218, no. 1-2, pp. 13–34.

Sneath, P.H.A. and R.R. Sokal (1973). *Numerical Taxonomy*. San Francisco, Ca: W.H. Freeman.

SPSS Inc. (2007) SPSS for Windows, Rel.16. SPSS Inc., Chicago.

StataCorp (2009) Stata Statistical Software: Release 11.0. Stata Corporation, College Station, TX.

Sumioka, S.S., D.L. Kresch, and K.D. Kasnick (1997). "Magnitude and Frequency of Floods in Washington," *WaterResources Investigations Report 97-4277*. Washington: Dept. of Transportation.

Vogel, R.M. and N.M. Fennessey (1993). "L-Moments Diagrams Should Replace Product-Moment Diagrams," *Water Resources Research*, vol. 29, no. 6, pp. 1745–1752.

Vogel, R.M., T.A. McMahon, and F.H.S. Chiew (1993). "Floodflow Frequency Model Selection in Australia," *Journal of Hydrology*, vol. 146, pp. 421-449.

Zrinji, Z. and D.H. Burn (1994). "Flood Frequency Analysis for Ungauged Sites Using a Region of Influence Approach", *Journal of Hydrology*, vol. 153, no. 1-4, p. 153.

3. Hypothèse de constance des coefficients techniques : Mythe ou réalité ?

Achille Pegoue[1]

Résumé

Dans le processus d'élaboration des comptes nationaux, la rareté de l'information conduit souvent à la formulation d'hypothèses économiquement défendables. Si ces hypothèses sont effectivement mises en œuvre, leur vérification, a posteriori, se heurte aux questions statistiques telles que l'agrégation des produits et de branches. Il en est ainsi de l'hypothèse de constance des coefficients techniques qui même si elle est appliquée au niveau des compte de branche reste difficilement vérifiable dans le tableau de synthèse qu'est le tableau des échanges interindustriels aussi bien dans les techniques liant la production et les consommations intermédiaires à prix constants qu'à prix courant. La condition de vérification de l'hypothèse dans un contexte d'agrégation pour les comptes à prix constants est proche du théorème d'agrégation sur les biens de Leontief-Hicks dans le calcul de l'indice de prix alors que l'hypothèse n'est pas envisageable dans les comptes à prix courant, le maintien des taux de valeur ajoutée étant préférable aux distorsions pouvant être engendrées par la prise en compte des effets prix.

Mots clés : *coefficient technique, production, consommation intermédiaire, tableau des échanges interindustriels, comptes nationaux*

Abstract

In the compilation of national accounts, a scarcity of data often leads to a recourse to hypothesis, which in the circumstances is economically defensible. Even though such hypotheses may be effectively implemented, a posteriori verifications usually come up against statistical issues such as aggregation of industries and products. This applies to the constant technical coefficient hypothesis, which is used in industry accounts but is not verifiable in aggregated tables like the interindustry exchange table, which links output techniques to intermediate consumption either in constant or current price. Our finding is that a condition such as the Leontief-Hicks theorem on the aggregation of goods in the computation of price index, should be fulfilled for a posteriori verification of the hypothesis in the framework of constant accounts, while a constant rate of value added should be preferred in the framework of current account.

Key words: *technical coefficient, output, intermediate consumption, interindustry exchange table, national accounts*

1 M. Achille Pegoue est en service au département de comptabilité nationale de l'Institut National de la Statistique du Cameroun depuis 1998 et consultant en comptabilité nationale et en statistique ; apegoue@yahoo.com

1. INTRODUCTION

Une ambition de la comptabilité nationale est de retracer le fonctionnement d'une économie pendant une période donnée aux fins de l'analyse économique, la prise de décision et la définition des politiques. La mise en cohérence d'une masse importante d'informations est alors indispensable dans le cadre soit des tableaux de travail à un niveau relativement désagrégé que constituent les équilibres ressources emplois des produits, les comptes de branches ou les matrices qui-à-qui relatives aux opérations de répartition entre secteurs institutionnels soit des tableaux de synthèse dont le TRE (tableau des ressources et des emplois) et le TCEI (tableau des comptes économiques intégrés) représentent le cadre central. Un des aspects de cette mise en cohérence est la description des relations existant entre la production d'une branche ou d'un secteur d'activité et les consommations intermédiaires (CI) utilisées pour réaliser cette production. Ces relations sont synthétisées dans le tableau des échanges interindustriels (TEI) qui croise les secteurs ou branches d'activité et les produits afin de représenter les disponibilités des CI offertes par l'économie en chaque produit et les utilisations qui sont faites par les secteurs ou branches d'activité. Pour saisir dans leur exhaustivité ces relations lors de l'élaboration des comptes nationaux en année de base, d'énormes moyens humains, matériels et financiers sont déployés pour l'exploitation complète des déclarations statistiques et fiscales (DSF) des entreprises et des comptes administratifs ainsi que pour la réalisation d'enquêtes visant à mesurer l'activité informelle ou les liens entre les maillons d'une filière. Cependant, la faiblesse des systèmes statistiques nationaux (SNS) africains et notamment subsahariens tant en ressources humaines et financières qu'en capacités organisationnelles et gestion des interrelations entre les intervenants des SNS ne permettant pas la pérennité de tels moyens, la carence en information de base lors l'élaboration des comptes en année courante entraine la formulation des hypothèses sur les structures et les évolutions du TEI qu'il importe d'examiner pour mieux comprendre les résultats obtenus. Une de ces hypothèses est la constance des coefficients techniques qui n'échappe pas à l'épineux problème d'agrégation et la question de la prise en compte des effets prix.

Cet article a pour objectif d'attirer l'attention des comptables nationaux et des utilisateurs des données de comptabilité nationale sur l'utilisation abusive de l'hypothèse de constance des coefficients techniques, les précautions à prendre lors de son utilisation et l'interprétation des résultats obtenus lors de son utilisation. En s'appuyant sur quelques exemples, l'article illustre la difficulté de vérification de l'hypothèse de constance des coefficients techniques d'une part à un niveau agrégé pour les volumes et d'autre part sur

les structures lorsque les effets asymétriques des prix sont intégrés. Il met aussi en exergue une condition de vérification de l'hypothèse de constance des coefficients techniques qui rejoint le théorème d'agrégation sur les biens de Leontief-Hicks dans le calcul de l'indice de prix.

Le restant de l'article est organisé ainsi qu'il suit. La deuxième section présente l'hypothèse de constance des coefficients techniques telle qu'elle est mise en œuvre dans les comptes nationaux. La troisième section revient sur les deux méthodes de construction d'un TEI soit en année de base avec la disponibilité des informations élémentaires, soit en année courante avec des hypothèses de structure et d'évolution. La quatrième section examine sur un exemple les difficultés de mise en œuvre de l'hypothèse de constance des coefficients techniques et met en exergue la condition de validité de cette hypothèse dans le cas d'une agrégation des secteurs d'activité ou des branches. La cinquième section illustre les conséquences d'un relâchement non contrôlé de l'hypothèse de constance des coefficients techniques par la prise en compte des effets prix sur la valeur ajoutée. La dernière section conclut et tire quelques perspectives pour des développements futurs.

2. HYPOTHESE DE LA CONSTANCE DES COEFFICIENTS TECHNIQUES

Le compte de production est le premier de la séquence des comptes courants du TCEI. Il décrit le processus de génération de la richesse, ayant pour origine les biens et les services, en soustrayant de la production les entrées intermédiaires (consommations intermédiaires) détruites au cours de cette production. La relation entre les consommations intermédiaires et la production permet de saisir la technique de production représentée par le coefficient technique. Ce dernier est le rapport entre les consommations intermédiaires et la production. L'analyse de création de richesse en biens et services par secteur institutionnel (entendu comme le regroupement selon les objectifs, les fonctions et les comportements économiques semblables d'unités institutionnelles, i.e. des entités économiques capable de détenir des biens et des actifs de leur propre chef) est pertinente pour la traçabilité des facteurs ayant contribué à la génération de cette richesse ainsi que les distribution, redistibution et utilisation du revenu qui s'ensuivent, mais elle ne permet pas de suivre le processus ou la technologie sous-jacente à la production.

Afin de mieux représenter la création de la richesse en biens et services, l'analyse de la production se fait par groupe de producteur homogène i.e.

réalisant le même type de production. Ce regroupement conduit à la formation des branches d'activité dans le cas des unités d'activité économique locale. L'analyse des intrants nécessaires à la production s'effectue par produit. La technologie utilisée dans une branche donnée en un produit donné est mesurée par le rapport entre la consommation intermédiaire en ce produit et la production de la branche. Ce rapport est le coefficient technique.

L'hypothèse de constance des coefficients techniques utilisée lors de l'élaboration des comptes nationaux et notamment du TEI (voir section suivante) postule que ce rapport ne se modifie pas sur une courte ou une moyenne période et donc les coefficients techniques sont identiques entre deux années de base.

3. METHODES CLASSIQUES D'ELABORATION D'UN TEI

Deux méthodes d'élaboration d'un TEI (tableau des échanges interindustriels) peuvent être identifiées : la première est relative à un TEI d'une campagne année de base et la seconde est celle d'un TEI d'une campagne courante.

3.1. TEI d'une année de base

En année de base, le TEI est entièrement refondu par la collecte des données élémentaires sur les utilisations des produits comme entrées intermédiaires au prix de l'année courante dans le processus de production. Cette demande de consommation intermédiaire par les secteurs d'activité ne suffit par pour réaliser un TEI dit équilibré ; il reste à confronter pour chaque produit, la demande totale exprimée par l'ensemble des secteurs dans les comptes des secteurs d'activité ou branches à l'offre totale telle que estimée dans les équilibres des ressources et des emplois - ERE. Cette confrontation délicate est réalisée lors de l'arbitrage du TEI. Elle peut entraîner une révision soit de l'offre en CI du produit soit de la demande de CI.

La révision de l'offre conduit à un déséquilibre des ERE pouvant être corrigé par une modification de la structure des emplois (consommation finale, investissement et exportation) ou par un ajustement des ressources (production, importation ou les autres éléments du prix (marges et impôts sur les produits)). Dans le cas d'un ajustement de la production, les coefficients techniques se trouvent automatiquement modifiés et un nouveau TEI peut être envisagé pour une nouvelle confrontation.

La révision de la demande en CI modifie directement le coefficient technique et affecte la valeur ajoutée ainsi que les autres soldes courants et le besoin/ capacité de financement du compte des secteurs institutionnels (tableau des comptes économiques intégrés – TCEI). Le rejet des soldes ou du coefficient technique conduit à une modification de la production qui déséquilibre l'ERE du produit et engendre éventuellement une nouvelle confrontation.

Ce processus itératif s'arrête lorsque l'offre en CI du produit égalise la demande en CI des secteurs d'activité/branches. Des algorithmes peuvent être développés pour équilibrer un TEI. Le plus utilisé est l'algorithme RAS qui fixe les marges et ventile les écarts en ligne selon le profil ligne et les écarts en colonne selon le profil colonne. Cet algorithme présente l'inconvénient de modifier tous les coefficients techniques simultanément. Une variante de cette algorithme est de fixer certaines valeurs et les éliminant du TEI avant la mise en œuvre de l'algorithme.

3.2. TEI d'une campagne courante

En campagne courante, deux TEI sont élaborés en parallèle : un TEI au prix de l'année précédente (TEI en volume) et un TEI au prix de l'année courante (TEI courant). Le principe général d'élaboration du TEI s'appuie sur deux hypothèses selon le type de TEI. Cette phase s'appelle en général la projection des CI.

Pour le TEI en volume, le produit prenant part à une entrée intermédiaire d'un secteur d'activité garde la même évolution en volume que la production du secteur d'activité. C'est l'hypothèse de constance des coefficients techniques. Elle conduit à des coefficients techniques identiques pour le TEI courant de l'année précédente et le TEI en volume excluant ainsi toute modification du processus de production à court terme (horizon d'une année ou d'un trimestre). Par contre, l'hypothèse de constance des coefficients techniques ne garantit pas l'obtention d'un TEI équilibré pour deux raisons : la structure des CI de la demande peut être modifiée par l'existence/ la disparition des productions secondaires et l'offre des CI peut être affectée par la modification de la structure des emplois ou les autres éléments du prix (marges et impôts sur les produits). Le processus itératif précédemment présenté est mis en œuvre afin de faire converger l'offre et la demande de CI et obtenir un TEI en volume équilibré. Dans certains cas, la construction du TEI en volume peut s'affranchir de l'hypothèse des coefficients techniques notamment lorsqu'un coefficient technique en valeur est disponible ; un indice de prix du produit permet de déflater les CI en valeur et de dériver les CI en volume afin d'obtenir un coefficient technique en volume implicite.

En ce qui concerne le TEI en valeur, partant des CI en volume, l'application des indices de prix des produits permet de déboucher sur des CI en valeur lorsque ces dernières ne sont pas disponibles à partir des données sources. L'utilisation des indices des prix pour inflater les CI modifie profondément le TEI et sa structure, et partant les coefficients techniques, ainsi que les soldes des comptes courants. Pour la validation du TEI, l'arbitrage s'effectue d'abord entre la logique statistique et la cohérence économique avant l'examen des équilibres comptables notamment la confrontation de l'offre et de la demande des CI. La cohérence économique peut conduire à revoir l'hypothèse des coefficients techniques postulée en amont lors de l'élaboration du TEI en volume.

4. AGREGATION DES COEFFICIENTS TECHNIQUES : QUEL NIVEAU DE DETAIL RETENIR ?

4.1. Exemple d'illustration du problème d'agrégation

L'élaboration des nomenclatures (classification) recourt le plus souvent à la mise en œuvre des niveaux hiérarchiques (section/division/groupe/classe ou niveau1/niveau2/niveau3). Les travaux d'élaboration des comptes s'effectuent dans la mesure du possible au niveau le plus détaillé des nomenclatures, les résultats dans les niveaux hiérarchiques supérieurs s'obtenant par agrégation de ces niveaux détaillés. Dans ce contexte, que veut dire l'hypothèse de constance des coefficients techniques des secteurs d'activité ? Le niveau 1 étant celui de publication des données, lorsque l'hypothèse de constance des coefficients techniques est mise en œuvre, les utilisateurs s'attendent à ce que la structure du TEI de l'année courante au prix de l'année précédente (TEI en volume) soit vraisemblablement proche de la structure du TEI de l'année précédente au prix de l'année précédente. Cette condition est rarement vérifiée comme le montre l'exemple suivant illustré dans le tableau 1.

Considérons un secteur d'activité (004 sylviculture et exploitation forestière) qui comprend deux sous secteurs (004001 Sylviculture et 004002 Exploitation forestière) ; pour simplifier nous supposerons que les deux sous-secteurs utilisent en intrants un produit agrégé. A quel coefficient technique du secteur 004 peut-on s'attendre sachant que le calcul des consommations intermédiaires en volume est fait sous l'hypothèse des coefficients techniques constants ? Dans l'exemple 1, nous avons supposé que, pour la production, l'indice de volume de 004001 est de 105 alors que l'indice de volume de 004002 est 95. Si les productions des deux sous-secteurs d'activité avaient un indice de volume identique, le coefficient technique agrégé de 004 n'au-

rait pas changé ; la modification du coefficient technique du 004 qui passe de 0,18 à 0,17 est due au changement de structure (répartition des poids dans le total) de la production des deux sous-secteurs d'activité du secteur 004. De façon formelle, on peut rechercher la condition de constance des coefficients techniques dans le cas d'une agrégation.

Tableau 1 : Illustration du problème d'agrégation des branches en supposant les coefficients techniques constants

Rubrique	Année précédente au prix de l'année précédente	Année courante au prix de l'année précédente
Production de 004	**200**	**200**
dont 004001 Sylviculture	100	105
004002 Exploitation forestière	100	95
CI de 004001	10	10,5
CI de 004002	25	23,75
Total CI	**35**	**34,25**
CT de 004001	0,10	0,10
CT de 004002	0,25	0,25
CT de 004	**0,18**	**0,17**

4.2. Condition de constance de coefficient technique au niveau agrégé

Considérons un secteur d'activité à n sous secteurs {1, 2,…, j, ,…, n}.

Notons :

I_j l'indice de volume du sous-secteur *j*
CI_j les consommations intermédiaires totales utilisées par le sous-secteur *j*
P_j la production du sous-secteur *j*
CT_j le coefficient technique du sous-secteur *j*
crt_n pour indiquer une opération en année *n* au prix de l'année *n*
cst_n pour indiquer une opération en année *n* au prix de l'année *n – 1*

L'hypothèse de constance des coefficients techniques au niveau élémentaire implique

$$CT_j^{crt_{n-1}} = CT_j^{cst_n} = a_j$$

avec $CT_j^{crt_{n-1}} = \frac{CI_j^{crt_{n-1}}}{P_j^{crt_{n-1}}}$. et $CT_j^{cst_n} = \frac{CI_j^{cst_n}}{P_j^{cst_n}}$ (1)

Dans les deux égalités précédentes, la production en volume est liée à la production de l'année précédente par la relation

$$P_j^{cst_n} = I_j \times P_j^{crt_{n-1}} \quad (2)$$

Connaissant la production des sous-secteurs, la production du secteur de l'année précédente (n-1) et la production en volume de l'année courante peuvent être calculées par agrégation respectivement par les formules :

$$P^{crt_{n-1}} = \sum_{j=1}^{n} P_j^{crt_{n-1}} \quad \text{et} \quad P^{cst_n} = \sum_{j=1}^{n} P_j^{cst_n} \quad (3)$$

De même, connaissant les consommations intermédiaires des sous-secteurs, les consommations intermédiaires du secteur de l'année précédente (n-1) et les consommations intermédiaires du secteur en volume de l'année courante peuvent être calculées par agrégation respectivement par les formules :

$$CI^{crt_{n-1}} = \sum_{j=1}^{n} CI_j^{crt_{n-1}} \quad \text{et} \quad CI^{cst_n} = \sum_{j=1}^{n} CI_j^{cst_n} \quad (4)$$

On peut donc calculer les coefficients techniques du secteur qui sont :

$$CT^{crt_{n-1}} = \frac{CI^{crt_{n-1}}}{P^{crt_{n-1}}} \quad \text{et} \quad CT^{cst_n} = \frac{CI^{cst_n}}{P^{cst_n}}$$

En utilisant les égalités de (4), ces coefficients techniques peuvent encore s'écrire :

$$CT^{crt_{n-1}} = \frac{\sum_{j=1}^{n} CI_j^{crt_{n-1}}}{P^{crt_{n-1}}} \quad \text{et} \quad CT^{cst_n} = \frac{\sum_{j=1}^{n} CI_j^{cst_n}}{P^{cst_n}} \quad (5)$$

En déduisant les CI des sous-secteurs des relations données en (1) et en utilisant l'hypothèse de constance des coefficients techniques dans les sous secteurs, les relations données en (5) deviennent :

$$CT^{crt_{n-1}} = \sum_{j=1}^{n} \frac{P_j^{crt_{n-1}}}{P^{crt_{n-1}}} \times a_j \quad \text{et} \quad CT^{cst_n} = \sum_{j=1}^{n} \frac{P_j^{cst_n}}{P^{cst_n}} \times a_j \qquad (6)$$

Ces deux dernières formulations montrent qu'une condition suffisante mais non nécessaire de constance des coefficients techniques au niveau agrégé est l'invariabilité des poids des sous secteurs d'une année à l'autre. Cette invariabilité n'est vraie que si tous les sous-secteurs ont un même indice de volume, i.e. dans la relation (2) I_j = *constante* pour tout sous-secteur j, ce qui est rarement le cas.

Ce résultat montre que, les travaux sur le TEI et notamment la projection des consommations intermédiaires à prix constant se déroulant à un niveau élémentaire, il est illusoire de rechercher la constante des coefficients techniques à un niveau agrégé. L'hypothèse de constance des coefficients à un niveau agrégé se heurte en outre à la condition de départ selon laquelle chaque branche considérée est homogène quant aux processus de production. Ni l'hypothèse ni la condition ne se trouvent satisfaites dans un contexte de pluralité des processus. L'approche de constance des coefficients techniques est fondée sur une espèce de relation linéaire entre intrants et extrants et ne considère pas les autres facteurs de production, dont les modalités d'utilisation influencent assurément la productivité d'une part et la qualité d'autre part. Pour obtenir une constance des coefficients techniques à un niveau agrégé, il faudrait que tous les sous-secteur d'un secteur d'activité aient des indices de volume suffisamment voisins.

5. EFFET PRIX : DES POSSIBLES ABSURDITES

5.1. Illustration du problème des effets prix

Dans l'exemple précédent, la construction d'un TEI en volume en appliquant l'hypothèse de constance des coefficients techniques au niveau détaillé, pouvait conduire à un résultat surprenant traduit par la non-constance des coefficients techniques à un niveau agrégé. Cet effet peut être amplifié lors de la construction du TEI en valeur de l'année courante par la prise en compte des effets prix. Ces derniers peuvent déboucher sur des coefficients techniques supérieurs à l'unité et donc à des valeurs ajoutées négatives comme l'illustre l'exemple suivant.

Dans cet exemple, nous considérons un secteur d'activité à deux sous-secteur ; chaque sous secteur utilise deux produits (CI1 et CI2) en entrée de consommations intermédiaires. Pour déterminer les volumes, nous utilisons l'indice de volume de la production de chacun des sous-secteur (hypothèse de constance des coefficients techniques) ; pour déterminer les valeurs nous inflatons les volumes par les indices de prix correspondants. Le tableau 2 résume ces calculs qui ont été effectués à l'unité près.

Tableau 2 : Illustration du problème des effets prix

	Secteur d'activité 1	**Secteur d'activité 2**	**Secteur d'activité (1 et 2)**
Année n-1			
Production	100	150	250
CI1	60	40	100
CI2	15	80	95
Valeur Ajoutée	25	30	55
Coefficient technique			
CI1	0,60	0,27	0,40
CI2	0,15	0,53	0,38
CI	0,75	0,8	0,78
Année n			
Indice de volume (production)	103	75	
Indice de prix			
Production	102	101	
CI1	140	140	
CI2	130	130	
Volume (Quantité de n aux prix de n-1)			
Production	103	113	216
CI1	62	30	92

	Secteur d'activité 1	Secteur d'activité 2	Secteur d'activité (1 et 2)
CI2	15	60	75
Valeur Ajoutée	26	23	48
Coefficient technique			
CI1	0,60	0,27	0,43
CI2	0,15	0,53	0,35
CI	0,75	0,8	0,78
Valeur (Quantité de n aux prix de n)			
Production	105	114	219
CI1	87	42	129
CI2	20	78	98
Valeur Ajoutée	-2	-6	-8
Coefficient technique			
CI1	0,83	0,37	0,59
CI2	0,19	0,68	0,45
CI	1,02	1,05	1,04

Cet exemple montre qu'il peut arriver qu'en inflatant les volumes de consommations intermédiaires par les indices de prix des produits utilisés en intrants, les valeurs ajoutées soient négatives. L'objectif recherché en utilisant l'indice des prix des consommations intermédiaires est la prise en compte des effets prix dans la modification des structures des coefficients techniques. Ainsi, aux prix de l'année courante, l'hypothèse de constance des coefficients techniques est relâchée afin d'intégrer un effet des prix. Cependant, aucun contrôle ne permet de garantir des résultats plausibles à la méthode par inflation des volumes ainsi utilisée.

Un autre exemple pourrait être utilisé pour montrer un résultat inattendu semblable sur la valeur ajoutée en volume si on utilisait une méthode par déflation des valeurs courantes par des indices de prix.

5.2. Alternative : maintien d'un taux de valeur ajoutée exogène

Cette valeur ajoutée négative du fait de la modification des prix relatifs est économiquement inacceptable. En effet, la prise en compte des évolutions de prix ignore les possibilités de substitution sur les quantités offertes par le marché aux secteurs d'activité.

De façon formelle, les volumes obtenus peuvent être revus afin de tenir compte de l'ajustement par les prix. Une méthodologie rigoureuse de la projection serait ainsi itérative et pourrait modifier les coefficients techniques des volumes avant la confrontation finale dans le TEI.

En pratique, un arbitrage pourrait être fait entre l'indice de volume de la branche et les indices de prix des produits de façon à converger vers un taux de valeur ajoutée exogène.

5.3. Quel indice de prix utiliser ?

Au-delà de la cohérence comptable et des égalités et des conditions mathématiquement vérifiables, le choix de l'indice des prix à utiliser pour déflater les valeurs ou inflater les volumes de consommation intermédiaires n'est pas une question très débattue. En effet, aucun indice parmi les indices de prix disponibles (indice de prix à la consommation finale (IPC), indice de prix producteur (IPP), indice de prix du commerce extérieur (IPCE), indice des cours des matières premières (ICM), etc.) n'est spécifique aux intrants des processus des productions. En effet, l'IPC est relatif essentiellement aux achats des ménages et concerne la vente au détail ; de plus, l'IPC s'élabore à partir de la nomenclature des fonctions de consommation des ménages (COICOP). L'utilisation de l'IPC comme indice de prix des consommations intermédiaires se heurte donc à deux obstacles : (i) les prix relevés pour son calcul sont nécessairement plus élevés que ceux des produits concernés car les entreprises peuvent s'approvisionner à l'amont de la chaîne de distribution, et (ii) la mise sur pied des tables de passage entre les deux nomenclatures (la COICOP et la nomenclature des produits) quelque soit leur finesse n'élimine pas le problème d'agrégation ou d'éclatement.

Un autre élément de différence entre un indice approprié pour la valorisation des consommations intermédiaires et l'IPC est la fiscalité. En effet, la TVA payée par les entreprises sur les intrants est très souvent déductible alors que la consommation finale supporte entièrement la TVA. Quant à l'IPP et l'IPCE, le biais qu'ils présentent dans leur utilisation comme prix de valorisation des consommations intermédiaires est un prix plus faible utilisé

dans le calcul de l'indice ; en effet, les prix utilisés dans le calcul de l'IPP et l'IPCE n'incluent pas les autres éléments du prix d'acquisition à savoir les marges de commerce et de transport et la fiscalité intérieure.

Pratiquement, les comptables nationaux utilisent un indice de prix composite pour valoriser les intrants des branches ou secteurs d'activité. Cet indice de prix dérive de l'équilibre des ressources et des emplois des produits et intègre uniquement la TVA non déductible sur les produits ainsi que les autres éléments du prix d'acquisition. La méthode mise en œuvre pour la construction de cet indice est séquentielle. Elle part d'abord de l'équilibre entre les ressources et les emplois d'un produit à prix de base. Ensuite, les autres éléments du prix d'acquisition sont rajoutés à l'équilibre à prix de base. Finalement, le prix d'acquisition se déduit comme une sommation des différents composants de l'équilibre. C'est cet indice du prix d'acquisition qui est utilisé pour valoriser la demande des branches en produits intermédiaires.

6. CONCLUSION

Cette note visait la mise en exergue des questions et problèmes que peut soulever l'hypothèse de la constance des coefficients techniques. Le premier exemple a permis d'illustrer le fait qu'il peut être illusoire de rechercher, à un niveau agrégé, la vérification de l'hypothèse de constance des coefficients techniques bien qu'elle ait été mise en œuvre sur des sous-secteurs d'activité ; une condition suffisante de vérification de l'hypothèse à un niveau agrégé a été mise en exergue : une évolution similaire des productions des sous-secteurs qui composent le secteur agrégé. Cette condition rejoint le théorème d'agrégation sur les biens de Leontief-Hicks dans le calcul de l'indice de prix. Le second exemple a permis de montrer que l'utilisation non contrôlée des indices de prix peut conduire à une valeur ajoutée (ou un coefficient technique) totalement erronée. Une solution préconisée est la recherche de taux de valeur ajoutée exogène avant la projection des consommations intermédiaires.

D'une façon générale et en guise de recommandations, l'analyse des processus de production peut aller au-delà de l'examen de l'hypothèse de la constance des coefficients techniques par une combinaison plus étendue de facteurs incluant la recherche-développement, le capital et la main-d'œuvre. De même, d'autres études complémentaires peuvent affiner l'analyse de la technologie, notamment la prise en compte des ajustements de productivité et de qualité ainsi que la dynamique des unités de production intégrant les phénomènes de délocalisation, de fusion/scission ou d'intégration verticale/

horizontale. Par ailleurs, la question de l'agrégation dans la mise en œuvre de l'hypothèse de constance des coefficients techniques encourage l'utilisation de données fines.

Certains problèmes liés à l'élaboration du TEI n'ont pas été abordés dans cette note. Un de ceux-ci est le bruit qui existe dans le TEI lorsque l'hypothèse de constance des coefficients techniques est appliquée dans le cas du croisement secteurs d'activité et produits du fait de l'existence des productions secondaires. De même, de manière formelle, les coefficients techniques traduisent un rapport de quantité et non des valeurs monétaires ; la complexité de la mise en œuvre de l'hypothèse de constance des coefficients techniques s'en trouve accrue par la recherche des prix adéquats pour la valorisation des grandeurs en jeu.

BIBLIOGRAPHIE

Braibant, M. (2007). « Le partage volume-prix dans la base 2000 ». INSEE, Banque de France, Trésor Public, Paris, France, N°7.

Branson, W. H. (1972). *Macroeconomic Theory and Policy*. New York : Harper & Row.

Chapron, J.M. et M. Séruzier (1984). *Initiation pratique à la comptabilité nationale selon le nouveau système*. Paris : Masson, 3è édition.

Flouzard, D. (1982). *Analyse économique et comptabilité nationale*. Paris : Masson.

Groupe de travail intersecrétariats sur la comptabilité nationale (1993). *Système de Comptabilité Nationale 1993*. New York : United Nations.

Ngo, T.C. (2002). *Kangaré, conforme au SCN93 et au SEC95*. Paris : INSEE.

Séruzier, M. (1996). *Construire les comptes de la nation*. Paris : Economica.

4. Using Statistics to Assist Decentralization and Local Economic Development in Some West African Countries

Adalbert Nshimyumuremyi[1]

Abstract
Several African countries have adopted decentralization policies that have devolved planning and governance powers from the central to the subnational and lower levels of administration, in order to provide better services to the public. This requires specific information systems and management tools to be put in place to effectively address relevant governance aspects. District-level data are also needed to support national development frameworks such as Poverty Reduction Strategies (PRSs), or to monitor progress toward attaining the Millennium Development Goals (MDGs).

During the 1990s, an innovative quantitative approach called ECOLOC was designed to support the decentralization policy and local economic development in some West African countries. This approach included a statistical methodology aimed at compiling district economic accounts and monitoring indicators by applying at the local level the principles of national accounting (1993 SNA). Social accounting matrices (SAMs) were used as coherence frameworks for local economic accounts. Population census data and vital statistics were complemented by a statistical surveys package to generate basic data required to compile these accounts. Based on the ECOLOC program implementation in four countries (Burkina Faso, Côte d'Ivoire, Mali, and Senegal), this paper shows how the approach helped to improve decentralization and local development policies. However, the sustainability of this approach hinges on the development and implementation of a strategy to mainstream subnational statistical systems into the overall national strategy for the development of statistics (NSDS).

Key words: *Local governance, subnational accounts, statistical surveys, social accounting matrix, statistical systems strategies*

Résumé
Plusieurs pays africains ont adopté les politiques de décentralisation qui ont délégué les responsabilités de planification et de gouvernance du niveau central aux niveaux infranational et plus bas de l'administration, afin de fournir de meilleurs services au public. Ceci exige les systèmes d'information et les outils de gestion spécifiques à mettre en place pour adresser efficacement des aspects appropriés de gouvernance. Les données au niveau des districts sont nécessaires également pour soutenir des cadres de développement national tels que les stratégies

1 Adalbert Nshimyumuremyi is Principal Statistician, Statistical Capacity Building Division, African Development Bank. Email: a.nshimyumuremyi@afdb.org.

de réduction de la pauvreté (SRP), ou pour suivre le progrès vers les objectifs de développement du millénaire (OMD).

Au cours des années 90, une approche quantitative innovatrice appelée ECOLOC a été conçue pour soutenir la politique de décentralisation et le développement économique local dans certains pays de l'Afrique de l'Ouest. Cette approche comprenait une méthodologie statistique ayant pour objectif d'élaborer des comptes économiques et des indicateurs de suivi pour les districts en appliquant au niveau local les principes de la comptabilité nationale (SCN 1993). Des matrices de comptabilité sociale (MCS) ont été utilisées comme cadres de cohérence pour les comptes économiques locaux. Des données de recensement de la population et les statistiques d'état civil ont été complétées par un ensemble d'enquêtes statistiques afin de produire des données de base nécessaires pour élaborer ces comptes. En se basant sur la mise en œuvre du programme ECOLOC dans quatre pays (Burkina Faso, Côte d'Ivoire, Mali, et Sénégal), cet article montre comment l'approche a aidé à améliorer la décentralisation et les politiques de développement local. Cependant, la pérennisation de cette approche repose sur le développement et la mise en œuvre d'une stratégie visant à intégrer les systèmes statistiques infranationaux dans la stratégie nationale pour le développement de la statistique (SNDS).

Mots clés : *Gouvernance locale, comptes infranationaux, enquêtes statistiques, matrices de comptabilité sociale, stratégies des systèmes statistiques*

1. INTRODUCTION

The administrative decentralization process is an approach that most West African countries have adopted, to a greater or a lesser extent, in recent years. The newly created local authorities were given the task of administering their territories, but soon realized that in order to do their job properly, they would need to get involved in local economic development, to gain a better understanding of its workings and interactions with surrounding areas and the country at large.

For these reasons, in 1997 the Municipal Development Program (MDP) and the OECD Club du Sahel and Western Africa Secretariat set up the ECOLOC program: "Managing the Economy Locally in West Africa." Its objectives were to: (i) define the concept of local economy and show that it is possible to analyze and describe its workings, main driving forces, and interactions with the rest of the country and the world at large; (ii) reveal the potential and the obstacles to the development of these local economies

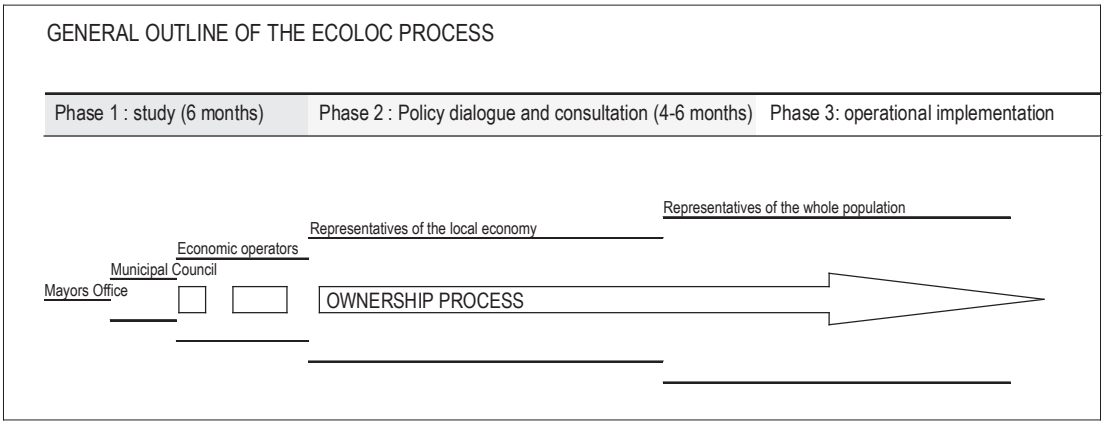
and examine how they could be sustainably revitalized; (iii) define the roles of the various public and private, local and external actors, and set up conditions for more effective dialogue between all parties; (iv) assist in the emergence of a coalition of interests for a strategy of local development; and (v) determine the information and training needs of local actors.

It is this program which is the particular focus of this paper. After this introduction, the second and third sections present some aspects of the methodology for devising a local economy information system. Full details of the ECOLOC process can be found in the methodology handbook and published case-studies. Based on the ECOLOC implementation in four countries (Burkina Faso, Côte d'Ivoire, Mali, and Senegal), the fourth section shows how the approach contributed to improve decentralization and local development policies. The paper ends with a discussion on the sustainability of this approach.

2. THE ECOLOC PROGRAM: "MANAGING THE ECONOMY LOCALLY IN WEST AFRICA"

Each ECOLOC exercise covers a small region comprising a second-tier town (regional capital or town with a population of around 100,000 inhabitants) and its rural area of influence. West Africa has at present about 100 such towns. ECOLOC areas, with total populations of roughly 500,000, are small enough to be studied "exhaustively," but sufficiently differentiated to reveal the complexity of economic and social processes, to help understand the actual behavior of the actors and identify the main determining factors in urban–rural interactions.

An ECOLOC exercise has three phases: (i) a *study phase* lasting about 6 months; (ii) a *policy dialogue and consultation phase* of roughly the same length; and (iii) an *implementation phase* for the local development and economic revitalization strategy, which begins after the consultation phase and lasts for an indefinite period. The study phase leads to the production of a set of coherent retrospective and prospective data on the local economy, its actors, issues, and trends, expressed quantitatively (as urban–rural Social Accounting Matrices [SAMs]), qualitatively, and spatially. The study is steered by a committee initially comprising the local mayor(s) and a few economic agents or representatives of chambers of commerce, gradually extending to other stakeholders. The study phase ends with a public reporting-back of results in the presence of all stakeholders.



The policy dialogue and consultation phase starts as soon as the studies have been reported back. This phase takes place under the auspices of each local authority with the participation of members of the study team and resource persons. The aim is to gradually construct a shared vision of how the local economy works, identifying its strengths and weaknesses. In this way, the participants seek to achieve a broad consensus or coalition of interests around a strategy for local economic development, thereby ensuring that local institutions and operators retain ownership of their own future direction. In this way, ECOLOC can be viewed as an application of the principle of "empowerment by knowledge."[2]

The implementation phase for the local development strategy and local economy revitalization covers a broader spectrum of population, involving every major decision-maker and representatives of local stakeholder groups. These stakeholders comprise local authorities, administration and public services, professional organizations, local and external partners. This phase may involve the creation of associations of municipalities to lobby the national authorities on common issues (such as improving the local network of markets). In the case of border municipalities, it may lead to the creation of consultative bodies with the local authorities in neighboring countries, to discuss common issues, particularly regional planning and the encouragement of regional trade.

2 Source: Club du Sahel/OCDE, MDP. (2001). *Managing the Economy Locally in Africa – ECOLOC Handbook: Assessing Local Economies and their Prospects*. Vol. 1, Summary.

3 QUANTIFYING THE LOCAL ECONOMY: STATISTICAL METHODOLOGY

The ECOLOC program is founded on the principle that for the dialogue between the various components of civil society and their partners to be fruitful, it must be based on quality information available to all. In the ECOLOC exercises, an exhaustive, quantified, and mapped presentation of the local economy is an essential preliminary task. The methodology utilized consists in applying at the local level the principles of national accounting (Mesplé-Somps and Nshimyumuremyi 1999).

Local economic aggregates are defined in a similar way to national accounting aggregates, by Gross Local Product (GLP) and the Gross Local Income (GLI). The GLP measures the value of goods and services produced by the local resident units and which are available for final employment (consumption, investment, etc). The GLI is the sum of incomes received by the resident units and taxes related to the production and the importations, net of subsidies. This aggregate, which takes account of the transfers of incomes (private and public) between the local zone and the outside, is particularly well suited to the measurement of the incomes circulating in a local economy. However, its computation requires the estimation of transfers between the institutional sectors of the local economy and those of the rest of the country, and this is not so easily measurable.

A set of maps, graphs, and indicators of the studied area, designed to be readily accessible to a wide non-specialist audience are grouped into the "*tableau de bord*" of the local economy. It provides graphic views of the study area within the country, whereby economic activities are analyzed according to the following variables: sector, type (modern, informal, etc.), activity complex, location (town, rural area, etc.), public administration, spatial analysis, and outlines of a long-term vision.

The quantified description of the local economy is based on the combination of a partly "modeled" macro approach and a micro approach using targeted observation on the ground. The success of the ECOLOC approach depends to some extent on the capacity of the study team to achieve the right balance and cross-fertilization between the two approaches. The development of the local accounts is carried out in three stages: (i) initial macro framework; (ii) reconciliation of national accounting approaches; and (iii) local SAM (social accounting matrix) construction.

In the first stage, a *macro approach* is used to build an initial picture of the local economy by regionalizing national accounts and using demographic/economic and spatial models (Cour and Snrech 1998; PDM; Club du Sahel/OECD 2001). Assumptions based on the demographic structure, the allocation of resources, and the social data make it possible to regionally spatialize the accounts. This dynamic picture is used as a framework for the studies on the ground, which in turn gradually refine it. The demo-economic and spatial "model" of the local economy that emerges from this iterative process is later used to interpret past performance and sketch out long-term visions, to evaluate the consequences of a given action on the local economy, and to carry out various comparative analyses between the studied area and the rest of the region or country.

The second stage consists in the compilation of existing statistics and conducting ad hoc surveys. This *micro-meso approach*, which consumes most of the study resources, is intended to provide adequate information on topics such as: the household economy; enterprises in the "modern" and "informal" sectors; trading in goods, services, and transfers; private and public capital; and administration and local authorities. Population census and vital statistics are complemented by a statistical surveys package to generate the basic data required to compile the local economic accounts. The information is obtained at the lowest level to allow the establishment of supply and use tables. Data collection tools are prepared using the international classifications (Backiny-Yetna and Gashongore 2001).

In the third phase, *social accounting matrices* (SAMs) are used as coherence frameworks for the different local economic accounts. The various statistical approaches used to build the accounts of industries and institutional sectors invariably lead to differences between the account balances. In order to address this, a systematic review of the coherence of the whole system needs to be carried out through the use of SAMs. It is through balancing the product accounts that the overall SAM final balance is obtained. If the balancing of the entire accounts is not achieved, the least-known variables are used for adjustment (imports and exports of products as well as households' current transfers) (Mesplé-Somps and Nshimyumuremyi 1999).

4. LESSONS DRAWN FROM THE ECOLOC EXERCISES

Since 1997, ECOLOC exercises have been conducted in several West African cities. In Côte d'Ivoire, five case studies were conducted in San Pedro, Daloa, Korhogo, Bondoukou, and Odiénné. In Mali, studies were conducted in

Sikasso and Ségou. In Senegal, ECOLOC studies have been conducted in: Saint-Louis, Dagana, Richall Toll, and Louga. After a successful exercise of the program in Bobo Dioulasso, Burkina Faso has extended the study to eight other sites, namely Banfora, Dedougou, Dori, Kaya, Koudougou, Ouahigouya, Tenkodogo, and Ziniaré. The ECOLOC program has in this way helped local stakeholders to obtain high-quality information appropriate to their needs. In countries where several ECOLOC studies have been conducted (for example, in Burkina Faso), a synthesis allows for a better understanding of poverty in each situation, providing a sound basis for defining poverty reduction programs.

The administrative decentralization process now underway throughout West Africa has not yet extended to a decentralization of information systems, which are still the preserve of central government departments and their external partners. One core principle of the ECOLOC program is that the extreme paucity of information available to local public and private operators hampers the development of the local economy and so perpetuates poverty at the individual level. Saint-Louis in Senegal is a good example of how information from an ECOLOC study can immediately be put to use by the Municipal Council to increase their revenues and negotiate more effectively with higher authorities and external partners.

The ECOLOC approach to territories consisting of one medium-sized town and their essentially rural hinterland demonstrates that "real" households and economic operators live and act in a unified urban–rural context, not in two distinct and separate spaces. The usual conceptual cleavage between rural and urban and between farm and non-farm sectors therefore proves to be artificial and deceptive. As regards rural areas more specifically, the ECOLOC program has confirmed that migration from the poorest and most remote areas to the better-endowed rural areas and towns is one of the keys to sustainable population growth. Such migration is also one of the manifestations of the modernization of farming and of a greater openness to the market economy.

One of the most striking findings of the ECOLOC exercises is the very low rate at which West African urban municipalities tax their local economies. In most cases, local taxes amount to about 0.3 to 0.5 percent of the town's GLP. This is only one-fifth to one-tenth of the rate applied in other parts of the developing world. Land tax is in most places less than 0.1 percent of land value. Current levels of taxation probably need to be tripled, to the benefit of local operators and households alike. But to achieve this, the municipality must be transparent in its workings, manage its affairs efficiently,

and maintain dialogue with the community. The example of Ségou in Mali shows that a process of this kind can be initiated through dialogue.

The ECOLOC exercises also show that it is local authority underspending that holds back local businesses, not overspending. Local governments' inability to spend money is the cause of a vicious cycle, whereby an impoverished urban environment leads to the stagnation or relative regression of the local economy. The lesson that emerges is that the less a local authority takes in taxes, and the less it spends, the poorer everyone is at the end of the day. The ECOLOC studies have revealed that local public spending, which can have a very low "import" content, can be a driver for local business, private investment, and social mobility. One additional franc taken in local authority taxes to boost local public spending can result in increased business that generates at least two francs more in household incomes. As regards raising the necessary funds for public spending, this should gradually be extended from the town to the rest of the local area. Taxation and spending should be organized for the local urban–rural economy as a whole, not for each town and village individually.

5. TAILORING INFORMATION SYSTEMS TO NEEDS: TOWARD THE DEVELOPMENT OF SUBNATIONAL STATISTICAL SYSTEMS

All societies produce the information systems that they need. One of the challenges is to meet the new needs that have arisen from decentralization. This requires a radical reshaping of information systems dating from a time when the central administration held a virtual monopoly on development management. Many of these highly centralized systems broke down during the crisis years of 1980–90, while occasional attempts by the central authorities at reconstruction proved largely unsuccessful.

Starting from the local level has the double advantage of revealing the failings of existing information systems and providing an opportunity to reconstruct new systems that respond to the challenges all the operators face. High-quality information widely shared is an essential precondition for improving governance and living standards for individuals and economic operators. National statistical services and regional and international institutions could benefit from the ECOLOC approach to conduct a more thorough examination of local and/or regional economic accounts.

It is widely recognized that there is a pressing need for the subnational statistical system to provide reliable data to support the decentralization policy and to feed into effective planning and decision-making at the subnational level. In response to this exigency, a variety of approaches have emerged to provide the much-needed data. One approach is to conduct national surveys, the results of which could be combined with census data. A second approach is to conduct censuses by installments, to render the results incremental over time. A third option is a comprehensive and vigorous development of the administrative statistical system (administrative record keeping). There exists a whole range of techniques for small-area estimates plus utilization of complex design techniques, such as multi-phase sample design, in order to obtain reliable data at the lower levels of administration. All these approaches highlight the need for subnational statistical systems to be integrated into the overall national strategies for the development of statistics (NSDS).

The implementation of decentralized statistical systems should not, however, be an end in itself. Its ultimate objective is to improve governance at the local level and to contribute to attaining MDGs. The ECOLOC program, by providing instruments adapted to local government needs, represents a key contribution toward the efficiency of governance at that level.

REFERENCES

Backiny-Yetna, P. and F. Gashongore (2001). *Manuel de procédures de la méthodologie d'élaboration des comptes économiques locaux : Guide du responsable de la MCS*. Abidjan : AFRISTAT/ENSEA.

Backiny-Yetna, P., F. Gashongore., S. Mesplé-Somps, and A. Nshimyumuremyi (2001). *L'élaboration des comptes locaux*. Document technique 1.4, PDM. Paris: Club du Sahel/OECD.

Club du Sahel/OECD, MDP (2001). *Managing the Economy Locally in Africa – ECOLOC Handbook: Assessing Local Economies and their Prospects*. Vol. 1, Summary.

Cour, J-M. and S. Snrech (1998). *Preparing for the Future – A Vision of West Africa in the Year 2020. West Africa Long-Term Perspective Study* (WALTPS). Paris: OECD.

Hussmans, R., F. Mehran, and V. Verma (1990). *Surveys of Economically Active Population, Employment, Unemployment and Underemployment. An ILO Manual of Concepts and Methods*. Geneva: ILO.

Mesplé-Somps, S. and A. Nshimyumuremyi (1999). *Outils d'analyse de l'économie locale*. Abidjan: ENSEA /Paris: DIAL.

Nshimyumuremyi, A. (1998a). *San Pedro: Dynamics and Synergy in the Local Economy. Case Study for the Program "Enhancing Local Economies in West Africa".* Paris: Municipal Development Program and OECD/Club du Sahel.

Nshimyumuremyi, A. (1998b). *Macroeconomic Policy and Poverty Reduction in Côte d'Ivoire*. Case study commissioned by SEPED/BPPS, UNDP, New York.

Nshimyumuremyi, A. (1999). *L'économie locale du département de Daloa, Etude de cas du programme "Relance des économies locales en Afrique de l'Ouest*. Paris : Municipal Development Program, Club du Sahel.

Nshimyumuremyi, A. (2002a). « Approches des grands comptes de l'économie locale ». In Y-A. Fauré and P. Labazee (eds), *Socio-économie des villes africaines : Bobo et Korhogo dans les défis de la décentralisation*. Paris: IRD-Karthala.

Nshimyumuremyi, A. (2002b). « L'action locale des autorités : Le poids de l'État et des municipalités ». In Y-A. Fauré and P. Labazee (eds), *Socio-économie des villes africaines : Bobo et Korhogo dans les défis de la décentralisation*. Paris: IRD-Karthala.

Nshimyumuremyi, A. (2006). "Methodology for Development of Local Economic Accounts," *African Statistical Journal*, vol. 2. Tunis: AfDB.

Séruzier, M. (1996). *Construire les comptes de la nation selon le SCN 1993*. Paris: Economica.

United Nations Statistical Division, Commission of European Community, International Monetary Fund, OECD, and the World Bank (1993). *System of National Accounts 1993* (SCN93).

5. Challenges of Statistical Infrastructure in Africa: A Systemic Approach to Statistical Capacity Development

Michel Mouyelo-Katoula[1] and Adalbert Nshimyumuremyi[2]

Abstract
Despite national and international efforts to develop national statistical systems in Africa, the majority of countries are still characterized by a vicious cycle of weak production and under-utilization of statistics by policymakers and the population at large. African countries are struggling to develop or update their National Strategies for the Development of Statistics (NSDSs) to ensure that statistical requirements in African countries are rooted in wider development strategies, and to build the national statistical system as a whole on a constantly updated and expanded statistical infrastructure base.

This paper highlights the main challenges to developing the necessary "soft" infrastructure, in the context of implementing strategically designed plans of action for the development of statistics. We suggest that the main constraints may be classified under three principal categories: cultural inadequacy; lack of factual synchronicity; and paucity of resources. The paper proposes a systemic approach aimed at developing a truly national statistical culture as the most sustainable foundation for statistical infrastructure development.

Key words: *NSDS, statistical infrastructure, systemic approach*

Résumé
En dépit des efforts nationaux et internationaux de développer les systèmes statistiques nationaux en Afrique, la majorité de pays sont toujours caractérisées par un cycle vicieux de la faible production et de la sous-utilisation des statistiques par les décideurs politiques et la population dans son ensemble. Les pays africains luttent pour développer ou mettre à jour leurs stratégies nationales pour le développement des statistiques (NSDS) afin de s'assurer que des besoins en statistiques dans les pays africains soient enracinées dans des stratégies plus larges de développement, et pour établir le système statistique national dans l'ensemble sur une infrastructure statistique constamment mise à jour et renforcée.

Cet article met en exergue les principaux défis pour développer l'infrastructure "soft" nécessaire, dans le contexte de la mise en œuvre des plans d'action stratégiquement conçus pour le développement des statistiques. Nous proposons que les

1 Michel Mouyelo-Katoula is Global Manager of the International Comparison Program, World Bank. Email: mmouyelokatoula@worldbank.org

2 Adalbert Nshimyumuremyi is Principal Statistician, Statistical Capacity Building Division, African Development Bank. Email: a.nshimyumuremyi@afdb.org.

contraintes principales soient classifiées en trois principales catégories : insuffisance culturelle ; manque de synchronisme effectif ; et manque de ressources. L'article propose une approche systémique visant à développer une culture statistique véritablement nationale comme base la plus soutenable pour le développement des infrastructures statistiques.

Mots clés : *SNDS, infrastructure statistique, approche systémique*

1. INTRODUCTION

Despite efforts made at various levels to develop statistical systems in African countries, there persists a vicious cycle whereby weak production leads to under-utilization of statistics by policymakers, which further undermines production and investment in statistics, and so on. To establish a virtuous cycle, the four major regional organizations in Africa involved in statistical support, as well as PARIS21 and other partner institutions such as the World Bank and bilateral agencies, are assisting African countries to develop or update their National Strategies for the Development of Statistics (NSDSs). The aim is to mainstream statistical requirements into broader development strategies.

The entire process is geared toward improving or laying the foundations for national statistical systems to more effectively inform development policies and programs at the country level. The operational objective is to ensure that the whole national statistical system is based on a statistical infrastructure that will be strengthened and enhanced on a permanent basis. This infrastructural foundation comprises the institutional and organizational structure; status and quality of human resource development and management; methodologies and operational frameworks (including statistical frames, updated registers, database development, management and dissemination); and data-processing equipment, transport equipment and survey material, etc.

While it is acknowledged that "hard" physical infrastructure is as important as "soft" infrastructure,[3] the main objective of this paper is to highlight the main challenges encountered in developing the "soft" type, in the context of the implementation of strategically designed plans of action for the develop-

3 "Hard" infrastructure includes physical equipment and ICT systems for the collection, analysis, dissemination and monitoring of statistical data. "Soft" infrastructure, on the other hand, includes the institutional culture, organizational and methodological frameworks, and strategic planning that facilitate and support statistical systems.

ment of statistics. The paper builds on a survey on the implementation of NSDSs, carried out by the African Development Bank (AfDB) in early 2009.

2. STRUCTURAL FOUNDATIONS OF A NATIONAL STATISTICAL SYSTEM

Although the survey covers all main issues of NSDS implementation, this paper focuses on two aspects of the findings of the study: (i) the institutional and organizational structure and (ii) the conceptual and methodological frameworks and data development. The selection is justified by our understanding that these form the bedrock of statistical infrastructure, without which any other infrastructure asset is a mere empty shell and cannot guarantee the sustainability of the statistical system. This understanding aligns with by the Marrakech Action Plan on Statistics (MAPS), which recommends the mainstreaming of strategic planning of statistical systems. It is further supported by the Reference Regional Strategic Framework for Statistical Capacity Building in Africa (RRSF), which accords a high priority to mainstreaming statistics as a cross-cutting sector in the development process. In fact, this recommendation is listed second on a 12-item list of strategies for the compilation of better statistics for improved development outcomes. The RRSF provides that:

> *"in the formulation of the agenda for development at the sub-national and national levels, priority data needs have to be identified and national capacities have to be built to meet them. These needs relate to benchmarking the existing situation, monitoring progress in the implementation of the development agenda, and evaluating the impact of the agenda on socio-economic conditions. Including a statistical component in the development plans such as the PRSPs and plans to achieve the MDGs, as suggested in MAPS, will help secure political commitment to statistical development and ensure that statistical capacities are built and maintained on a sustainable basis."*

We believe that a properly mainstreamed statistical system is the best statistical advocacy tool. Indeed, as the RRSF makes clear:

> *"Statistical advocacy is about creating statistical awareness (numeracy) and a society that is inclined and able to use statistical information in decision-making; making the general case that statistics is a necessary part of the enabling environment for improving development outcomes; demonstrating*

the use of statistical data for decision-making at the government, business, and community level..."

Therefore, statistical processes should be part and parcel of the whole development agenda.

2.1 Institutional and organizational structure

The study undertaken by the AfDB shows that the weaknesses cited in the evaluation of the Addis Ababa Plan of Actions for Statistical Development in Africa in the 1990s (AAPA) as well as in the preparation process of the RRSF, are yet to be effectively addressed in the processes of NSDS design and implementation. The main weaknesses are: (i) inadequate adherence of the statistical system to the economic, social, and political agendas in most of the countries; (ii) lack of vision and strategic planning of statistical activities; (iii) inadequate funding of statistics by governments and non-sustainable support from development partners; (iv) ineffective coordination of national statistical systems; (v) lack of control over donors' resources; and (vi) under-investment in major factors of statistical production.

According to the study, and as underlined in the RRSF, statistics legislation in most countries should be revised to effectively comply with the UN Fundamental Principles of Official Statistics. The Statistics Act is outdated (having remained unchanged since before 1990) in 21 percent of the countries and requires redrafting, or it should be amended in 34 percent of the countries where it was enacted between 1990 and 2000. It is anticipated that where the Statistics Act is improved/updated, this will make it easier to design or update an NSDS on a regular basis and to ensure that it is implemented according to the set timeframe.

2.2 Conceptual and methodological frameworks and data development

Enhancing statistical infrastructure in a particular country entails developing relevant and robust scientific methods and applying them to the observation, monitoring, and measuring of social and economic phenomena. National Statistics Offices (NSOs) are also expected to adhere to the recommended periodicity in conducting comprehensive surveys aimed at creating and maintaining the sampling frames. This is still a major challenge, as shown by the small proportion of countries (out of a total of 53) achieving the following benchmarks: (i) 12 countries are currently up to date in their program of census taking; (ii) 11 countries plan to conduct an agricultural

census in the timeframe of their NSDS implementation; (iii) 14 undertook a household consumption survey between 1975 and 2007; (iv) 18 completed a health and demographic survey between 1987 and 2007, and 10 intend to do so in the context of their NSDS. This sample of figures speaks volumes about the weakness of the infrastructural foundation upon which to build a national statistical system.

3. MAJOR CHALLENGES

The principal constraints which translate into the "soft" infrastructural weaknesses of the statistical systems in Africa can be classified under three main headings: cultural inadequacy; lack of factual synchronicity; and paucity of intellectual and other human resources.

3.1 Cultural inadequacy

Cultural inadequacy is the discrepancy or disconnect between the multi-dimensional systems of a country, on the one hand, and the prevailing statistical paradigms on the other. It is about the indigenous statistical system failing to empower itself, intellectually, in order to translate the prevailing systems, categories, and other development phenomena into appropriate classifications, concepts, and methods. African statisticians in most countries have limited assurance as to whether their statistical models are a "faithful" numeric representation of the economic and social realities they are supposed to describe and forecast over time and space. They are generally obliged to use statistical classifications that reflect the economic and social categories drawn up by the developed countries. It is these developed countries that tend to have a significant voice in the major expert groups charged with designing international recommendations and norms. Instead of developing a bottom-up approach in African countries, a top-down customization of the recommended standards to the individual country is limited to subdividing the lowest categories of the classification into further detailed subcategories. Such classifications have questionable relevance to reality, as it is perceived by experts in other economic and social fields and the population at large.

3.2 Inadequate factual synchronicity

Inadequate factual synchronicity is another type of discrepancy that constrains the development of national statistical systems in Africa. A statistical system, conceptual framework, methodology or set of data lacks factual synchronicity when there is complete disconnect between the time a phe-

nomenon occurs and the moment when it is statistically observed, recorded, processed, and reported.

3.3 Paucity of intellectual resources

The paucity of intellectual resources is all the more alarming, given the daunting challenges facing NSOs in monitoring and measuring realities, policy implementation, and development actions. The complexities to be observed require an intellectual infrastructure which statistical agencies alone (whether central or sectoral) cannot mobilize or assemble. They require multidisciplinary synergies, which should be organized at the country level as well as at regional and subregional levels.

Bringing together the best intellectual resources in the region will establish a flexible center of excellence for the conceptualization of measurable realities and the design of statistical paradigms specific to Africa. In this way, the bottom-up process of harmonization will be rationalized and enhanced by a top-down system, whereby all major statistical problems facing most of the countries will be addressed at the regional level and thereafter customized to each country's specificities. The statistical intangible infrastructure will be thus created regionally and replicated at the country level according to a regionally adopted research agenda.

4. ENGAGING IN DEVELOPMENT-FOCUSED RESEARCH

Research will entail two complementary lines of action: (i) a more effective participation by Africa as a region in the international statistical debate, and (ii) activities aimed at a better understanding of development issues, developing methodologies, definitions, concepts, classifications and setting standards and norms.

The first line of action will be implemented through a series of interrelated steps[4] that will help to: identify priority themes and needs; create an inventory of existing working groups and establish new ones on important themes not yet covered; enhance the organization of the African statistical system; and streamline knowledge sharing. To the extent possible, this process will take account of the evolving agenda of international debates on statistical development.

4 Michel Mouyelo-Katoula and Thierry Paccoud, "Have Africa's Statistical Voice Heard: How to Prepare Africa's Contribution," *African Statistical Journal*, Vol. 8 – May 2009, African Development Bank.

In respect to the second line of action, we propose a systemic 12-stage approach (see below), based on the truism that social and economic realities to be measured statistically concern living beings and their self-organized systems. Both the living beings and their evolving environment should be studied as *cybernetic systems*, i.e. as a hyper-living-system of interacting elements that engage in various types of exchanges and thus produce information that is their defining and common heritage. This systemic approach will help create intangible infrastructure items that will mainstream statistics both into the development agenda and into development research and analysis. It should be used as an additional advocacy tool to raise or increase awareness about the usefulness and relevance of statistics on the one hand, and to plead for more adequate resources to increase other intangible and physical infrastructure assets on the other.

4.1 Proposing a systemic approach

The challenge to be addressed consists in translating this natural information into statistics, using a systemic approach. Whatever the development phenomenon to be studied and measured in terms of statistics, the following are stages proposed[5]:

1. *Delineation*. The scope, time and space, and nature of the phenomenon must be defined as clearly as possible, to better describe the issues at hand and make it easier to identify the areas where specific expertise will be required.

2. *Expert studies*. The understanding of the issues may require a simple review of existing studies or an optimal mix of expertise to be established (sociologists, geographers, physicians, nutritionists, economists, statisticians, etc.).

3. *Categorization*. Statisticians and other experts will review the findings of the experts to identify categories using the relevant methods, such as segmentation, typology, taxonomy, etc.

4. *Identification of systems and situations*. Any development phenomenon relates to people sharing a reference system that defines the interrelations among them, as well as the interrelations between different portions of their group and their material and non-material environment. This set of interrelations, coded or not, constitutes the prevailing system of

5 M. Mouyelo-Katoula, "Rethinking Statistics for National Development in Africa," *African Statistical Journal*, Vol. 2 – May 2006, African Development Bank.

living conditions. Elements pertaining to the same system or systems are dictated by their own inherent characteristics: by events, necessities, strategies, etc. These overall aspects constitute a situation for an individual or group of individuals at a given point in time and space.

5. *Development of classifications*. Systems and situations as well as other categories identified under item 3 above will lead to statistical classifications related to: individuals and social groups; objectives, principles, beliefs, etc; social links and relationships; spatial partition; time segmentation; functions, social roles, operations; means of action (fungible, semi-durable, durable, etc.); relation to environment; notion of results, principles of satisfaction.

6. *Design of a conceptual framework*. This is a type of intermediate theory that has the potential to connect to all aspects of the study of the phenomenon (e.g., problem definition, purpose, documentary review, methodology, data collection and analysis). It is like a map aimed to give coherence to the empirical knowledge derived under items 1–4.

7. *Development of a methodological framework*. A comprehensive and integrated set of statistical methods will be developed anew or adapted from existing methods, with a view to defining statistical operations to be conducted and data analysis to be carried out to derive proper statistical measurements.

8. *Statistical surveys on systems*. These are benchmark surveys to be conducted according to a periodicity suggested by the natural cycle of the phenomenon under study. While standard benchmark enquiries (population and housing censuses, agricultural censuses, household consumption surveys, demographic and health surveys, etc.) are clearly defined in terms of their objectives, scope, methodology and expected outputs, as well as their periodicity, system-based surveys recommended here are expected to address other developmental aspects which are not commonly investigated under the framework of official statistics. One example is gender statistics, for which we are proposing, in addition to the existing frameworks, another approach aimed at identifying and studying the systemic dimensions that constitute the general framework under which gender dynamics develop (see Box 1 below).

Box 1: Dimensional Aspects of a Systemic Gender Survey

- Contextual dimensions: The space and time boundaries of the framework must be delineated as well as the physical infrastructure supporting processes, activities, and events defining the dynamics;
- Natural resources as an important dimension of gender relations and status;
- Political and judiciary systems, decision-making processes, type of participation, individual and collective obligations, etc.;
- Cosmology and conception of the world/universe. This includes the set of beliefs, standards, and values;
- Sociological bases of gender status: social relations, family status, tradeoff between collective and individual needs;
- Modes and factors of production, productive resources, mechanisms of distribution and exchanges (solidarity, reciprocity);
- Modes of consumption, employment, income and other material possessions;
- Knowledge creation and sharing;
- Health system, reproductive health characteristics.

9. *Statistical surveys on situations, actions, flows, and results*. Situations are to be understood here as any combination of all the things happening in the systems under consideration, as well as all the conditions that exist at a particular time, in a particular place, and that are likely to affect the systems one way or the other. Situations are supposed to evolve more rapidly than systems, and therefore, related surveys will be implemented at a higher frequency rate to be determined according to the types of situations.

10. *Expert analysis and recommendations (sociologists, geographers, physicians, nutritionists, economists, statisticians, etc.)*. The results of the system and situation surveys should be analyzed by a pluridisciplinary team

of experts, some of whom might have contributed to the review of specialized studies recommended under item 2 above. The experts should also recommend follow-up actions for consideration at the appropriate levels.

11. *Inventory of actions implemented as a follow-up to analyses and recommendations*. It is suggested that a review of key activities/events should be carried out after an agreed period of time (resources permitting), to have a sense of how recommended actions or other events have impacted the situations and systems during the period under review. A procedure should be established to properly identify these actions.

12. *Impact assessment of actions*. This stage of the systemic approach is meant to provide information to policymakers and other development stakeholders, including the public at large, on changes that have occurred as a result of combined interactions and actions during a specific period of time.

4.2 Where the proposed approach aligns with the RRSF

The systemic cycle must be considered as a dynamic learning loop whereby constant adjustments are made to the approach itself, building on lessons learned from previous steps in the cycle. It is primarily a development cycle with statistical activities embedded at each stage. This innovative approach has several advantages which can be summarized along the lines of the 12 RRSF strategies:

- *Strategy 1:* Invigorate statistical advocacy by providing sound evidence of the developmental relevance and usefulness of statistics;
- *Strategy 2:* Mainstream statistics as an embedded learning and measurement mechanism in the development process;
- *Strategy 3:* Update the legal and regulatory framework;
- *Strategy 4:* Assess user needs;
- *Strategy 5:* Undertake data development through timely, effective, and integrated use of development data;
- *Strategy 6:* Foster coordination, collaboration, and partnerships among development stakeholders, both domestic and foreign;
- *Strategy 7:* Enhance the statistical infrastructure by creating a knowledge base spanning the whole development realm;
- *Strategy 8:* Harness information and communications technology (ICT);

- *Strategy 9:* Develop human capital through the proposed ever-challenging and cross-sectoral learning process;
- *Strategy 10:* Improve data analysis in a multidisciplinary way, whereby data analysis will be carried out ex-ante before engaging in any new statistical survey, and ex-post using the results of surveys on systems and situations, as well as when making an inventory of actions and undertaking impact assessment;
- *Strategy 11:* Improve data dissemination as statistics will feed into the development processes, which will in turn justify statistical processes;
- *Strategy 12:* Improve funding and sustainability as the proposed statistical approach is expected to benefit from resources available in the various development sectors and segments it will be describing, measuring, and thus helping to monitor.

5. CONCLUSION

This paper has posited that the major challenges facing the development of the "soft" statistical infrastructure in Africa can be classified under three principal categories: cultural inadequacy, lack of factual synchronicity, and paucity of intellectual and other human resources. However, there are other deficiencies which hold back the development of effectual statistical systems in Africa: dependence on donors and hence limited country ownership of statistical programs; insufficient financial resources; lack of coordination between statistical agencies; and the inadequacy of physical infrastructure.

The role of the RRSF and NSDS in addressing these challenges needs to be reinforced through two complementary lines of action: (i) a more effective participation by Africa as a region in the international statistical debate and (ii) activities aimed at understanding development issues, developing methodologies, definitions, concepts, classifications, and setting standards and norms. Development paradigms should be translated into statistics, using a systemic approach. Whatever the development phenomenon to be studied and measured in terms of statistics, the 12 stages proposed in this paper are expected to bring together statisticians and other experts into a multidisciplinary cycle of activities. It is this cycle where empirical observation, conceptualization, measurement, analysis, action, and impact assessment coalesce to generate a dynamic development knowledge base that will ensure the sustainability of other infrastructure assets.

REFERENCES

African Development Bank, UN Economic Commission for Africa, PARIS21, and World Bank (2006). *The Reference Regional Strategic Framework for Statistical Capacity Building in Africa: Better Statistics for Improved Development Outcomes*. Available at: http://www.uneca.org/STATISTICS/docs/rrsf/stat_RRSFDocument_Final.pdf (accessed March 17, 2010).

African Development Bank, Intersect, and PARIS21 (2007). *Mainstreaming Sectoral Statistical Systems in Africa: A Guide to Planning Coordinated National Statistical System*. Available at: http://www.paris21.org/documents/2959.pdf (accessed March 17, 2010).

Ajayi, O., N. Madaya, and A. Nshimyumuremyi (2008). "Mainstreaming Sectoral Statistical Systems into a National Strategy for the Development of Statistics,"*African Statistical Journal*, vol. 7, African Development Bank.

Mouyelo-Katoula, M. (2006). "Rethinking Statistics for National Development in Africa," *African Statistical Journal*, vol. 2, African Development Bank.

Mouyelo-Katoula, M. and A. Nshimyumuremyi (2007). "International Comparison Program for Africa – Towards Economic Convergence Measurement," *African Statistical Journal*, vol. 4, African Development Bank.

Mouyelo-Katoula, M. and A. Nshimyumuremyi (2008). "Towards a Framework for Collaboration among Development Partners in National Accounts Programs in Africa," *African Statistical Journal*, vol. 7, African Development Bank.

Mouyelo-Katoula, M. and T. Paccoud (2009). "Have Africa's Statistical Voice Heard: How to Prepare Africa's Contribution to International Statistical Conferences, Meetings and Working Groups," *African Statistical Journal*, vol. 8, African Development Bank.

PARIS21 (2004). *A Guide to Designing a National Strategy for the Development of Statistics (NSDS),* October. Available at: www.oecd.org/dataoecd/21/57/38308637.pdf.

6. Peer Reviews of African National Statistical Systems

Samuel Blazyk, Guest Charumbira, Lamine Diop, Mary Strode, and Tony Williams[1]

Abstract
The PARIS21 Secretariat has facilitated peer reviews of the National Statistical Systems (NSSs) of seven countries since 2005. The first review of the Ghana statistical system in 2005 was a pilot, followed by reviews in Tanzania and Zambia in 2007. At the time of writing in 2009, peer reviews have been carried out in Burkina Faso, Malawi, Mozambique, and Niger. The peer review findings are treated as confidential but in the case of Tanzania, Malawi, and Mozambique, the host governments have already accepted the reports as public documents for wide dissemination.

This paper sets out the rationale, coverage, and methodology for the peer reviews carried out to date and considers the reporting and dissemination of the findings. Some examples of statistical and organizational good practice that were noted by the peer reviewers are highlighted. The list is not exhaustive but they are included in the interests of sharing experiences.

Key words: *NSDS, rationale, coverage, methodology, good practice, African Charter on Statistics*

Résumé
Depuis 2005, le Secrétariat de PARIS21 a facilité les revues par les pairs des systèmes statistiques nationaux (SSN) de sept pays. La première revue du système statistique du Ghana en 2005 était un exercice pilote, suivi des revues en Tanzanie et en Zambie en 2007. Au moment de la rédaction de cet article en 2009, des revues par les pairs ont été effectués au Burkina Faso, au Malawi, au Mozambique, et au Niger. Les résultats des revues par les pairs sont traités comme confidentiels mais dans le cas de la Tanzanie, du Malawi, et de la Mozambique, les gouvernements hôtes ont déjà accepté que les rapports soient considérés comme des documents publics à disséminer largement.

Cet article présente la logique, l'étendue, et la méthodologie des revues par les pairs effectuées jusqu'ici et les rapports et la diffusion des résultats. Quelques exemples de bonne pratique statistique et d'organisation qui ont été notés par les revues par les pairs sont mis en exergue. La liste n'est pas exhaustive mais ces exemples sont présentés dans l'objectif de partage d'expérience.

1 The co-authors are either members of, or consultants engaged by, the PARIS21 Secretariat, 2 rue André Pascal, 75775 Paris 16, France.

Mots clés : *SNDS, logique, couverture, méthodologie, bonne pratique, Charte africaine sur des statistiques*

1. RATIONALE FOR PEER REVIEWS

The African peer reviews of NSSs were launched in 2003 by the Economic Commission for Africa's Committee on Development Information (CODI). CODI recommended that African countries, supported by PARIS21, should carry out peer reviews to ensure that good practice passes from country to country, based on the first-hand experience of peers, to help accelerate the development of national statistical systems. This responded to the New Partnership for Africa's Development (NEPAD) agenda for greater transparency and openness in governance (including through peer reviews) and drew on the longstanding experiences of peer reviews carried out between members of OECD's Development Assistance Committee (DAC). Subsequently, Eurostat has facilitated peer reviews of EU member states, focused on the adoption of the European Statistical Code of Practice.

CODI recommended that policymakers should be included in the peer reviews of African NSSs for advocacy purposes, to share experiences of policy approaches, to provide their feedback as users of the statistical practices they encounter, and to avoid the peer reviews being associated only with the statistical community.

The focus of the peer reviews is on the governance of the NSS, its organization, strategic planning, service to users, funding and sustainability. In the case of the four peer reviews carried out in 2009, all of this was set against the backdrop of the UN's Fundamental Principles of Official Statistics and on the African Charter on Statistics' best practice principles for professional independence, quality, mandates and resources, dissemination, protection of confidentiality, and coordination and cooperation. The African Charter on Statistics was adopted in Addis Ababa by African Heads of States and Governments on February 4, 2009, while the peer review of Mozambique was underway.

Peer reviews are highly relevant to the design and implementation of National Strategies for the Development of Statistics (NSDSs). The success of each country's NSDS depends on senior managers of statistical systems, as well as on policymakers and decisionmakers, demonstrating commitment to the implementation, which requires a number of years of substantial and often difficult change processes. Peer reviewers share experiences of what

solutions have worked in their own countries and elsewhere in the region; and through their advocacy, peers can help to facilitate change in the processes of government and their relations with cooperating partners. The peer reviews can also help to evaluate the effectiveness of NSDS processes and to ensure that good practice is transmitted from country to country at the level of heads of statistics and senior policymakers.

In the view of the authors, a peer review provides a very valuable input to the process of developing and building support to implement a country's NSDS. The peer review can add value at various stages of the NSDS – for instance, as an input to the design stage in parallel with user consultation exercises; or as a means of reviewing progress during implementation – which has been the case to date for African NSSs.

Part of the unique value of the peer reviews has been the level of support and access afforded by the host countries; the information exchange between the hosts and reviewers; and the inclusion of policymakers as peers. The Government Statistician of Ghana described the peer review of her country as "empowering," while the President of INE Mozambique said that their review will boost advocacy, the self-esteem of NSS staff, and the confidence of users.

The African process differs from the EU peer reviews in that the governance and organization of statistics is the main focus, rather than technical issues of statistical production. The reason for this is that the time and resources available for the peer reviews are limited, and a full technical review would be more resource-intensive. Second, a large proportion of African statistical systems have already undergone or are due to undergo major reforms, and much is to be learned from sharing experiences.

2. TOPICS COVERED BY PEER REVIEWS

Topics for consideration in the peer reviews are decided right at the start. Annex A to this article details potential topics, but all of the reviews so far have assessed:

- Progress of design and implementation of the country's NSDS;
- Statistical governance, legal mandates, and coordination;
- Financial resource availability and predictability, both from governments and cooperating partners; and the extent to which donor funding is focused on national priorities;

- Human resources management; and
- User consultation, prioritization, and satisfaction.

As outlined above, the peer reviews are not intended to make a technical assessment of the quality of official statistics produced by a country; however, an assessment can be made of the capacity to produce, disseminate, and use official statistics. The theoretical underpinning for this is PARIS21's Statistical Capacity Building Indicators, which are based on the IMF's Data Quality Assessment Framework (DQAF), particularly those aspects related to data quality, serviceability, and accessibility. Use of PARIS21's NSDS Essentials checklist provides background information to help to benchmark the quality of NSDS design processes.

3. METHODOLOGY

In the African peer reviews, each country was reviewed by two other countries. For instance, for the peer review of Burkina Faso, the reviewers were from Cameroon and Niger; for Malawi, they came from Mozambique and Tanzania; and for Mozambique, from Malawi and Tanzania. The countries acting as peers provide two reviewers, normally their head of statistics (or deputy head) and a senior policymaker (such as head of the Poverty Reduction Strategy unit, or senior policy user or funder of statistics). Facilitators were provided by the PARIS21 Secretariat.

So far the countries that have participated have either been self-selected or accepted an invitation to participate. Ghana and Tanzania volunteered during the CODI meeting in 2003 and follow-up; the other countries agreed to participate when approached by PARIS21.

The Head Statistician of the country being reviewed hosts the review. The host is responsible for receiving his or her peer reviewers and for arranging a program of meetings to enable them to complete their task. Typically the reviews are carried out over four working days. The host country provides the reviewers with key documents, including: statistical legislation, organizational structure and staffing, strategic plans, funding details, lists of publications, and anything else that the host country thinks might be useful.

Interviews with NSS stakeholders have formed the core of the peer reviews. The teams meet senior managers and staff of the NSS; and with a selection of other producers and key users of official statistics from within Parliament, government, the central bank, private sector, the media and academia; as

well as funders of the NSS, including ministries responsible for finance, and cooperating partners. The interviews are guided by a discussion schedule (see Annex B) but reviewers have the flexibility to deviate from this schedule and to pursue topics in depth. One of the main benefits of the peer reviews has been the opportunity during interviews for peer reviewers to share experiences from their own countries.

4. COSTS

The costs of the peer reviews have been shared by PARIS21 and the host country. This is in addition to the valuable senior staff time given freely by the host governments and reviewers. PARIS21 has funded international travel and per diem expenses, as well as providing the facilitators; while the host countries have provided internal travel and office facilities. Peer reviews are not necessarily directly linked to NSDS processes but, where they are, country costs can be included in NSDS funding.

5. REPORTING AND DISSEMINATING THE FINDINGS

The facilitators accompany the reviewers during their interviews and prepare notes on the findings of each meeting for approval by the review team. The reviewers and the facilitators are professionally bound to retain confidentiality on the issues discussed. To do otherwise might jeopardize the frankness of the peer review process. At the close of the mission, the reviewers generally provide a verbal report with observations and recommendations to the most senior official with responsibility for the statistical system (e.g. the Prime Minister in the case of Mozambique).

A full report is then prepared by the facilitators and agreed with the review team before being submitted to the host government. It is the government that decides if the report can be made public. Any findings that cannot be made public because they are of a sensitive nature – or because they breach the confidentiality of the peer review process – can be provided separately by the peer reviewers to the host government; but this has not been necessary in the case of the reviews conducted so far.

In each of the four peer reviews conducted to date (in 2009), the peer reviewers made observations and recommendations covering aspects such as: governance; strategic planning and implementation; statistical products; advocacy and dissemination; financing; staffing and staff development; and

further sharing of experiences. In the case of Malawi and Mozambique, the host governments have already accepted the reports as public documents to be disseminated widely. It is hoped that this practice will be emulated by all countries. Prior to this latest set of four countries, only the review of Tanzania had to date produced a publicly-available document.

6. SHARING "GOOD PRACTICE" EXAMPLES

Many examples of good practice have been identified by the peer reviewers. The selection below gives examples of the value that the peer reviews can deliver through the sharing of experiences. The exclusion of a country against any particular point does not necessarily mean that it is not applicable to them; rather it indicates that it was not identified as such by the peer reviewers.

Statistical production:

- Quick results from the most recent population censuses (Malawi, Mozambique, Tanzania);
- Re-basing and publishing of a new series of national accounts developed via a Supply-Use Table (Malawi);
- Quality and timeliness of the Consumer Price Index (mentioned by users in Malawi and Zambia);
- Proper account taken of the Fundamental Principles of Official Statistics (Burkina Faso, Niger);
- Publication of a statistical release calendar (Mozambique, planned in Malawi);
- Transformation of production and quality of statistics since semi-autonomous status (noted by users in Tanzania).

Strategic planning/ institutional development:

- Well-developed strategic planning processes (Malawi and Mozambique are both on their third generation of strategic plans);
- Strong leadership from the NSO and regular meetings of the National Statistical Council (Niger, Burkina Faso);
- Legal basis being revised in line with the African Charter on Statistics (Mozambique);
- Semi-autonomous statistical agency and governing board (Tanzania);
- High profile and status of statistics, including level of President of INE and High Council on Statistics is chaired by the Prime Minister (Mozambique).

User focus:
- Training/ workshops for the media (Malawi, Mozambique);
- Innovative advocacy and dissemination work (e.g. Statistics Diary in Mozambique).

Staffing and staff development:
- High retention rates of statistical staff (Malawi and Mozambique);
- Effectiveness and qualifications of staff and increase in staff numbers (Burkina Faso);
- Taking forward training efforts in collaboration with local training institutions (Mozambique, Zambia);
- Existence of a training center for statistics within the NSO (Niger);
- Providing a training resource for the region (EASTC Tanzania);
- Development of staff (INE Mozambique has doubled its number of staff and increased the percentage of professional staff from 8 percent to 30 percent over ten years);
- Existence of a Statistical Association (in Malawi, but not in Zambia or Mozambique).

Cooperating partners:
- Existence of a development partner group for statistics (Tanzania and Zambia; subsequently formed in Malawi);
- Strong support from partners and a desire to create a donor group on statistics (Burkina Faso, Niger);
- Well-developed relationship with development partners (Mozambique, evidenced by pooling of donor funds in a common fund to support statistical development);
- Effective south–south cooperation (e.g. Malawi in its Population and Housing Census);
- Cooperation between countries through AFRISTAT.

7. NEXT STEPS

Following an assessment of the peer review processes and methods conducted so far, other reviews will be planned in Africa and beyond, according to the demand coming from countries. Based on experiences to date, there will be a clear objective to publish the findings with the agreement of the host governments, in the interests of transparency and the sharing of experiences.

It is hoped that as the process of peer review is extended and deepened, the reviews will be linked even more closely to the principles enshrined in the African Charter on Statistics and NEPAD processes.

REFERENCES

Eurostat (2007a). *European Statistical System Code of Practice Peer Reviews: The Peers' Guide*. Luxembourg: Eurostat.

Eurostat (2007b). *European Statistical System Code of Practice Peer Reviews: The National Statistical Institutes' Guide*. Luxembourg: Eurostat.

Eurostat (2008). *Summary of good practices identified during the European Statistics Code of Practice peer reviews carried out during 2006-2008 structured according to the statistics value chain*. Luxembourg: Eurostat.

Fellegi, I.P. and J. Ryten (2000). *A Peer Review of the Swiss Statistical System*. Neuchâtel: Swiss Federal Statistics Office (SFSO).

Laliberté, L. (2002). *Statistical Capacity Building Indicators*. Paris: PARIS21 task team report.

Pagani, Fabrizio (2002). "Peer Review: A Tool for Cooperation and Change." Policy Brief. Paris: OECD.

Strode, M. (2005). *Peer Reviews for Statistical Systems*, Paris: PARIS21.

ANNEX A: POTENTIAL TOPICS FOR PEER REVIEW

1. Financial Resources Availability

a. Adequacy, regularity and reliability
b. Donor dependence
c. Adaptability
d. Utilization
e. Financial reporting
f. Budget preparation and prioritization
g. Donor support & coordination

2. Human Resources Management

a. Staff numbers – qualified and unqualified
b. Staff retention
c. Training programs/strategies in place
d. Salaries and incentives

3. Strategy/Plan Development

a. When prepared
b. How prepared and by whom
c. Implementation progress – what has been achieved / what not achieved
d. Obstacles and possible solutions
e. Progress of legal changes
f. Civil service reform – challenges and opportunities
g. Poverty monitoring plans

4. User Consultation, Prioritization and Satisfaction

a. User consultation processes carried out
b. Prioritization process
c. Dissemination/output plan
d. Links to PRSP and MDG indicators – is there a plan to satisfy needs?
e. User satisfaction, credibility and relevance

f. Publications and website
g. Procedures for dealing with user requests
h. Availability of metadata
i. Quality assurance

5. Statistical System Coordination

a. Is the entire statistical system coordinated by the Government Statistician?
b. How is coordination achieved?
c. Common standards and definitions in place
d. Resolution of conflicting estimates achieved
e. Is there a single source for access to all statistics?

6. Legal Basis

a. Is the law adequate?
b. What revisions are required?
c. Opinion on independence of operation

7. Conclusions and Recommendations

a. Contribution of statistics strategy to the development of the system
b. Improvements recommended
c. Likely support required
d. Obstacles requiring resolution
e. Likely outcome
f. Good practice observed for dissemination
g. Coordination activities

ANNEX B: DISCUSSION GUIDE FOR PEER REVIEWERS

1. Matters of Fact Concerning Respondent

a. Job title and organization
b. Details of their statistical interests
c. Details of the statistics and information they need

2. Opinion of Statistical Products

a. Quality of the information they receive from NSO/NSS
b. Relevance of the information
c. Timeliness of the information

3. Accessibility

a. How is the information obtained from NSO/NSS?
b. How can access be improved?

4. Adequacy of the Institutional Set-up

a. Opinion on the institutional set-up
b. Suitability of the personnel
c. Recommendations for improving personnel weaknesses
d. Adequacy of NSO/NSS resources
e. Who are the potential champions of statistics?

5. Advice

a. What advice would they give to the NSO/NSS?
b. What changes would they recommend?

7. Toward a Vibrant and Effectual Statistics Training Approach

Parin Kurji[1]

Abstract
At the end of a statistics course, if students enquire when the next course is to be scheduled, then the trainer feels that she/he has made a breakthrough! Since 2002, this has been the case at the Faculty of Agriculture, University of Nairobi, where we have been experimenting with an innovative approach to statistics training. The new curriculum integrates the use of computers and the latest developments in education theory, and takes advantage of the expanding network with the user community. Although we offer this as a curriculum for undergraduates, our experience with scientists, postgraduates, meteorologists, and National Statistical Service officers has shown that this approach is equally effective in meeting their training needs. Furthermore, given the unique position that statistics holds in any general training program, the approach can be usefully extended to other subject areas to make higher education generally more effective and relevant.

Key words: *Service courses, integrated curriculum, undergraduate, innovative approach*

Résumé
À la fin d'une formation en statistiques, si les étudiants demandent quand le prochain cours doit être programmé, alors le formateur estime qu'elle/il a eu un impact positif! Depuis 2002, cela a été le cas à la faculté d'agriculture, de l'Université de Nairobi, où nous avons expérimenté une approche innovatrice à la formation statistique. Le nouveau programme d'études intègre l'utilisation des ordinateurs et les derniers développements dans la théorie d'éducation, et tire profit du réseau en expansion de la communauté d'utilisateurs. Bien que nous offrons ceci comme un programme d'études pour les étudiants préparant une licence, notre expérience avec des scientifiques, des universitaires, des météorologistes, et des fonctionnaires des instituts nationaux de statistiques a prouvé que cette approche est également efficace pour satisfaire leurs besoins de formation. En outre, étant donné la position unique que les statistiques tiennent dans n'importe quel programme de formation général, l'approche peut être utilement étendue à d'autres domaines pour rendre l'enseignement supérieur en général plus efficace et approprié.

Mots clés : *Service de formation, programme de formation intégrée, étudiant en licence, approche innovatrice*

1 Head of Biometry, Faculty of Agriculture, College of Agriculture & Veterinary Sciences, University of Nairobi, pkurji@uonbi.ac.ke

1. INTRODUCTION

Traditionally, in many life science degrees in the Sub-Saharan region, one or more statistics courses have been offered at undergraduate level, but more as a sop to tradition than because their importance has been universally recognized. For most students, the statistics course is seen as a necessary hurdle to be overcome and then forgotten. Lack of motivation has worked in both directions: with little or no technology or administrative support, big classes, and students with little or no aptitude or interest in the subject, statistics training has tended to follow a dry "textbook" approach, showing students how to analyze neat, small sets of data from neat, artificially contrived situations. This approach is easy to follow and even easier to assess. The teaching style too, has tended to been standard, using the "talk and chalk" method, with the trainer center-stage and students passively accepting his or her "words of wisdom" – leading to further apathy in class. Unfortunately, students trained in this manner encounter difficulties in abstracting to different situations, which seriously limits their ability to undertake quality research or adapt to the changing demands of the employment market. In mitigation, one has to admit that, until recently, given the scarcity of resources, this was often the only viable approach. In situations where a fresh graduate was employed in a junior capacity, deficiencies in the formal training could be overcome through mentoring and adequate supervision.

The changing demands of the 21st century, however, force us to rethink not only what is taught, but also how it is taught. The focus needs to be on how to obtain, understand and apply knowledge rather than just adding to the student's factual store of knowledge (Crowder *et al.* 1998; Kimani 2002; Kettenring 1995; Markwell 2002). This is true generally, but even more so in the agriculture sector, where the challenges of development, production and employment demand skills and abilities that traditional curricula cannot match. In the words of Professor Gajaraj Dhanarajan (1998),

> *"Educational institutions ... can no longer content themselves with training a labour force for stable industrial or even agricultural jobs, instead, they must train individuals to be innovative, capable of evolving and adapting to a rapidly changing world of work."*

The report by the Vision 2000 Steering Committee (1990) similarly emphasizes the need for graduates with strong interpersonal skills, who are adept at using their talents to meet the changing needs of their professions. H. Lynn Erickson (2006) in her book, *Concept-Based Curriculum and Instruction for the Thinking Classroom* sums this up succinctly: *"effective responses*

to complex problems require the ability to abstract, conceptualize, predict, collaborate, plan, and act accordingly." This could be a description of what training in statistics can do! Properly conceptualized training in statistics can give professionals the tools they require to make evidence-based decisions, while helping them to develop the skills they need in their particular areas of expertise. Moreover, given a situation where secure niches for graduating students no longer exist, such skills are essential to develop a strong, productive workforce.

Until recently, universities had neither the resources nor the capacity to make substantive changes in their approach to statistics training. However, over the last few decades, this situation has changed dramatically. The four areas that can impact statistics training positively are: (i) the availability of and accessibility to computers; (ii) internet facilities to provide virtually unlimited access to information; (iii) developments in statistical methodologies; and (iv) progress in educational research with clearer insights into student learning styles. While there has always been an interest in educational theories, one important milestone was the establishment of the National Science Foundation in the US to facilitate improvement of education in the sciences. The literature abounds with theories on how people learn: for instance, Kearsley (2010) outlines no less than 50 major learning theories in his database. However, from our point of view there are two main points to consider: first, that different people have different learning styles and second, that learning must be organized to cater for all styles in order to achieve maximum impact. The change of focus from that of trainer's convenience to trainee's needs has also ushered in a change in students' attitude to statistics training.

2. APPROACH TO STATISTICS TRAINING

Good statistics practice demands both technical know-how and interpersonal skills. In this respect, strategic planning, execution and critical reflection, using multi-disciplinary teams and good communications skills, are fundamental in applying statistics effectively for problem solving. However, such skills should not be limited to statistics; they can be equally well applied to other fields. At the university level, statistics, by cutting across all disciplines, is ideally suited to provide the lead in the new approach, which is recommended by many leading educators in our field, notably Moore (1997) and Garfield *et al.* (2002) among others. The challenge is to motivate the students by making the statistics courses more interesting, relevant, achievable, and manageable.

The role of undergraduate statistics training is therefore to strengthen the foundation within the context of the bigger picture – the bigger picture being the problem to be solved. The curriculum needs to progress by stages to this "bigger picture." Using a building-block approach and applying modern training methodologies, the student is helped to discover the issues and challenges at each stage, while keeping the ultimate objective firmly in mind. The curriculum naturally divides into three sections: (i) understanding and exploring data; (ii) generalizing the sample results to the population using statistical modeling; and (iii) synthesizing the whole process through course work and by conducting a small project (to be supervised by the relevant discipline faculty and conducted in house or through attachment with a user organization).

3. APPROACH ADOPTED BY THE FACULTY OF AGRICULTURE, UNIVERSITY OF NAIROBI

A study commissioned by The Technical Centre for Agricultural and Rural Cooperation (CTA) in 1995 (Riley 1998) laid the major blame for the poor quality of postgraduate research projects on biometry input. Dialogue with the postgraduate students (class of 2002) at the College of Agriculture and Veterinary Sciences (CAVS) highlighted the need for a better understanding of basic concepts. This was especially true for mature students and those who had come from other institutions, and called for undergraduate statistics training to be strengthened. In response, the Biometry Unit at the College of Agriculture & Veterinary Sciences with the assistance of the Statistical Service Centre, Reading University, UK and the Biometry Support Unit, World Agro-forestry Centre (ICRAF), Kenya, embarked on the ambitious process of overhauling the undergraduate statistics training. The major changes involved reorganization of the curriculum, a student-centered teaching approach, and easier access to quality resources. These three aspects are analyzed below.

3.1 Curriculum

In the first introductory course given to the second-year students, emphasis is on *understanding and exploring data*. It is important for the students to appreciate why and how data are collected, what type of data need to be collected, and the implications this has on the findings of the study. Exploring data in this way introduces the idea of critically examining the collected data for patterns and anomalies before embarking on formal analysis. It

also emphasizes the fact that these data emanate from a sample and will be extrapolated to learn about the population.

An important addition to the curriculum is climatic data, which not only provide an alternative source of monitoring data, but resonate with the current preoccupation with climate variability and change. We discuss below (see section 3.3) resources that are available to facilitate their inclusion in the curriculum.

Inferential statistics is the second statistics course, given in the third year. The idea of sample and population and the standard error as a measure of precision form the central theme in this course. At every level the student is encouraged to reflect on how best to achieve the objectives rather than simply learning different techniques and applying them when specifically asked to do so. The *final undergraduate course*, which is unfortunately an option course, brings together the different aspects of the research process, including practical issues and dissemination of results, and culminates in the final student project.

3.2 Student-centered approach

All training uses the problem-solving, data-based and hands-on approach, enhancing computing proficiency and developing soft skills. The training is student-centered, with practical work involving both individual and group work. Similarly, lecturing is combined with demonstrations, interactive practicals, discussions, and self study. The data used are a combination of simple textbook examples, real datasets from various organizations, and case studies and simulated data from statistical games. The building-block approach ensures that the learning is kept relevant to the student context, becoming progressively more complex but at all times retaining the overall perspective of the problem-solving process.

3.3 Better access to improved resources

Over the years, we have, with support from the University administration, managed to set up an enviable set of resources for both the undergraduate and postgraduate students. These include computer laboratories with free access for all students, free internet connectivity, various statistical software including Excel Add-ons like SSC-Stat, Instat, GenStat, Statistical Games for data simulation, teaching guides for trainers, various CDs with a rich set of materials for reference and learning, and more recently a resource room for seminars and outreach activities. Through the University library, students can now access a number of online journals. One of the most useful resources we have added is an electronic statistics textbook, "CAST for Africa." CAST (Computer Aided Statistics Learning) comprises a series of textbooks with interactive demonstrations, exercises, and examples using African datasets that follow the revised curriculum. It serves as the main statistics textbook for our undergraduate statistics courses and is also a useful resource for postgraduates and scientists who need to refresh their statistics ideas. SSC-Stat, Instat, Teaching guides, including Climatic guides for GenStat and Instat, and other training materials can be accessed from the SSC website, http://www.reading.ac.uk/SSC/ while maintenance work is going on at the Biometry Unit page on the University of Nairobi website. Through assigned work, students are encouraged to make use of all these resources.

With help from our partners, we have implemented this new curriculum with reasonable success. Although no formal evaluation has yet been carried out, the students' demeanor in class, the lively atmosphere during lessons, the enthusiasm with which the students respond, and the positive results in their continuous assessment assignments, all indicate that the students are performing better and are enjoying the learning. The main ideas that we incorporate are detailed below.

The training reflects the real process of a *scientific investigation*, starting with the problem definition, formulation of the objective, leading to what and how to collect the information, exploration of the collected data using appropriate descriptive methods, generalizing the results using appropriate modeling methods if necessary, then interpreting the results, before leading back to the original problem and the identified objectives. The complexity of the problem – the depth and detail at which each stage is discussed – is commensurate with the level of the students.

The emphasis is on the *concept* rather than the technique used. So for example, the focus is not on how to make a two-way table but rather on explaining how a two-way table may be the most appropriate for displaying the data, where the response is categorized by two factors, whereas a scatter diagram would be more appropriate if the data were two quantitative measurements taken on a unit. The students appreciate the importance of appropriateness as opposed to a right or wrong solution.

The approach to the training includes a *variety of student activities*: individual class and take-home assignments, short in-class and longer take-home assignments, snap quizzes, question & answer sessions, peer-judged group presentations, individual and group reports, etc. Students are encouraged to use concept maps and make their own notes in order to discourage rote learning. The emphasis is on good work ethics, including team work, respect, integrity, time management, and proper documentation.

One very important ingredient for the success of this approach is the building of networks within and between universities, and more importantly with the user community. Our collaboration with the national statistics services, research organizations, and meteorological services in the region, through the running of short courses for the scientists in these organizations, has enriched our teaching tremendously. Through this approach, we have real-life scenarios and data to incorporate into our training, while the scientists in these organizations are given the opportunity to learn new skills and retool themselves with new techniques and resources.

The new statistics training builds on the central idea that *while applications are varied*, the concepts are the same. This fact needs to be emphasized and used as a source of strength, especially where students start as a group but go on to specialize in different areas. In a world that is increasingly demanding a multi-disciplinary approach, the earlier that students can learn to accept and work with this reality, the better prepared they will be when they graduate. However, this will only work if we can convert other faculties, scientists,

and users to this way of thinking. One way of gaining their support and participation is to use their data and projects to build the training material, to invite them to attend and give seminars, run short courses and set up discussion groups. The students will then better understand the diversity of problems which they may face and the underlying thread of commonality which points them toward solutions.

There are still areas where more work is needed, the major one being student assessment. Traditional examination methods are not conducive to an approach that emphasizes creative problem-solving and encourages flexibility, discussion, and computer-assisted learning. University regulations stipulate the 70:30 ratio of examination marks to continuous assessment, the latter being in the form of two formal tests. However, we have our own unique approach to student assessment: "exam" is a taboo word until the final two weeks of the semester! The students are encouraged instead to concentrate on understanding, learning, and applying the concepts that have been discussed during lectures. The continuous assessment, which accounts for 30 percent of the final grade, is used to assess students' understanding, clarity of thought and neatness, communication and team-building skills, time management and participation, through their individual tutorial assignments, topic concept maps, group presentations and reports (also peer evaluated), snap quizzes, and class participation. The written final exams are still quite traditional – including multiple-choice questions, calculations, interpretations of outputs, and methodology questions – and quite straightforward. The focus is on the first level, partly because of time constraints and partly because it is what the students expect. There is room for improvement, both in the system of grading which is currently used and in formulating assignments that challenge the bright students without necessarily demotivating the average ones. More mentoring and individual attention needs to be given, but this can only happen if there is an increase in human resources.

As the University moves towards e-learning, it will be possible to encourage students to take more responsibility for their learning by providing them with challenging assignments, including reading assignments, through the course websites. This will also provide the course facilitator with means and time to give prompt, individual feedback – a very necessary part of the learning process.

4. CONCLUSION

The one question that arises is why the focus is on the undergraduate curriculum when the major issue that has bedeviled the statisticians and scientists alike is the poor quality of research. The answer is simple: because this is a critical time – this is the students' first adult exposure to applied statistics. No matter what the schooling has been up to this point, the student realizes that the undergraduate period is the gateway to the real world. The student is therefore open to new experiences, is willing and able to handle the challenges, and is just entering the discipline of choice – so mindsets and attitudes can still be moulded. By giving the right training, we are equipping the students with skills that will be useful in all other courses. Moreover, the students will have a better chance at the postgraduate level if they have received a firm grounding at an undergraduate level.

The Biometry Unit, through its outreach and income-generating unit, BUCS (Biometry Unit Consultancy Services), has continued working within the University and with other universities and organizations in the Sub-Saharan region through seminars and workshops to promote good statistical practice, based on modern effective training approach. It is gratifying to note that statistics training in Africa is slowly rising to the challenge of the continent's development needs; as a consequence, statistics groups are becoming more open and proactive in developing and sharing ideas and resources.

The way forward is not going to be easy or short term. Statisticians providing service training within universities have for too long been on the sidelines – working in their lonely corners, being cornered in odd places to give instant solutions and then blamed for demanding too much or doing too little – to be suddenly accepted as role models by their colleagues. We have to take the lead and woo the students, scientists, and professionals in the field to learn and apply statistics joyfully, in order to make the most of this win–win situation.

REFERENCES

Crowder, L. van, W.I. Lindley, Th. H. Bruening, and N. Doron (1998). "Agricultural Education for Sustainable Rural Development: Challenges for Developing Countries in the 21st Century." Prepared for Extension, Education and Communication Service (SDRE), FAO Research, Extension and Training Division. Published in *The Journal of Agricultural Education and Extension* (1998), vol. 5, no. 2, pp. 71–84. London: Routledge.

Dhanarajan, G. (1998). "International and Inter-Institutional Collaboration in Distance Education." Presented at the "Learning Together: Collaboration in Open Learning" Conference, The John Curtin International Institute Perth, Western Australia, April 20–22, 1998.

Garfield, J., B. Hogg, C. Schau, and D. Whittinghill (2002). "First Courses in Statistical Science: the Status of Educational Reform Efforts." *Journal of Statistical Education*, vol. 10, no. 2. Available at: www.amstat.org/publications/jse/v10n2/garfield.html (accessed on March 15, 2010).

Kearsley, G. (2010). "The Theory Into Practice Database." Retrieved from http://tip.psychology.org (accessed on March 8, 2010).

Kettenring, J.R. (1995) "What Industry Needs," in *Modern Interdisciplinary University Statistics Education: Proceedings of a Symposium*. http://www.nap.edu/openbook.php?record_id=2355&page=5 (accessed on March 8, 2010).

Kimani, G. (2002). "A Tracer Study of University of Nairobi Graduates, 1991–1998." Paper Prepared for CODESRIA's 10th General Assembly on "Africa in the New Millennium", in Kampala, Uganda.

Lynn Erickson, H. (2006). *Concept-Based Curriculum and Instruction for the Thinking Classroom*. London: SAGE Publications.

Markwell, D. (2002). "Undergraduate Education for the 21st Century: Australia at the Crossroads." Trinity Papers Number 20. Available at: http://www.trinity.unimelb.edu.au/publications/trinity_papers/paper20 (accessed on March 15, 2010).

Moore, D.S. (1997). "New Pedagogy and New Content: the Case of Statistics." *International Statistical Review*, 65 (2): 123–165. Available at: http://www.stat.auckland.ac.nz/~iase/publications/isr/97.Moore.pdf (accessed on March 15, 2010).

Riley, J. (1998). *Strengthening Biometry and Statistics in Agricultural Research*. Study Report. Wageningen, The Netherlands: CTA (Technical Centre for Agricultural and Rural Co-operation).

Vision 2000 Steering Committee (1990). Vision 2000: *Quality and Opportunity*. http://thecouncil.on.ca/documents/Vision2000_E.pdf (accessed March 8, 2010).

8. CAST for Africa[1]

Doug Stirling[2], Parin Kurji,[3] and Roger Stern[4]

Abstract
Poor understanding of statistical concepts often hampers agricultural research in Africa and similar problems are experienced by many others who collect and use data. One reason is a traditional approach to statistics training that emphasizes formulae and computation with simple artificial examples and leaves the students unprepared for the complexities of real-life applications. The issue of effective statistics training is not confined to the agricultural sector, nor to Africa. However, improvements have been harder to make in developing countries due to a lack of computers, books, suitable teaching aids, real-life data, and other resources. Now that computers are more readily available, computer-aided learning can be used to overcome some of these constraints and make statistics training more useful and interesting.

A set of e-learning books called CAST was customized to fit a typical statistics course for agricultural students in an African setting. Other e-books were later added to complement courses for meteorologists and staff in National Statistics Offices in Africa. These e-books have been used to support the teaching of statistics in Africa to both university students and researchers and they have received positive feedback from users.

Key words: *e-learning, e-book, statistics training, teaching resources*

Résumé
La faible compréhension des concepts statistiques entrave souvent la recherche agricole en Afrique et des problèmes semblables sont éprouvés par beaucoup d'autres qui collectent et utilisent les données. L'une des raisons concerne une approche traditionnelle à la formation de statistique qui met l'accent sur les formules et le calcul avec des exemples artificiels simples et laisse les étudiants non préparés pour les complexités des applications réelles. La question de la formation efficace en statistiques n'est pas confinée au secteur agricole, ou à l'Afrique. Cependant, il a été plus difficile de faire des améliorations dans les pays en voie de développement à cause du manque d'ordinateurs, de livres, d'aides pédago-

1 ACKNOWLEDGEMENTS: We are grateful to the Rockefeller Foundation, DFID and the UK Meteorological Office for providing funding to enable these customized e-books to be produced and distributed.

2 Doug Stirling is Senior Lecturer, Institute of Fundamental Sciences, Massey University, Palmerston North, New Zealand. Email: d.stirling@massey.ac.nz

3 Parin Kurji is Senior Lecturer, Faculty of Agriculture, University of Nairobi, Kenya. Email: parinkurji@yahoo.com

4 Roger Stern is Chief Biometrician, Statistical Services Centre, University of Reading, Reading, United Kingdom. Email: r.d.stern@reading.ac.uk

giques appropriés, de données réelles, et d'autres ressources. Maintenant que les ordinateurs sont plus aisément disponibles, l'enseignement assisté par ordinateur peut être employé pour surmonter certaines de ces contraintes et pour faire de la formation statistique plus utile et intéressante.

Un ensemble de livres d'étude en ligne appelé CAST a été adapté aux besoins du client pour adapter un cours typique de statistiques pour les étudiants en agriculture dans un contexte africain. D'autres livres en ligne ont été ajoutés plus tard comme complément aux cours en faveur des météorologistes et le personnel dans les instituts nationaux de statistiques en Afrique. Ces livres en ligne ont été employés pour soutenir l'enseignement des statistiques en Afrique à la fois des étudiants d'université et des chercheurs et ont connu une réaction positive des utilisateurs.

Mots clés : *Etude en ligne, livre en ligne, formation statistiques, ressources pédagogiques*

1. INTRODUCTION

The authors have found in their teaching and consulting that agriculture students and researchers often encounter difficulties in understanding statistical concepts and in applying statistical methods appropriately in projects. Reviews and reports have consistently found statistical input in research projects to be inadequate. These problems are not unique to Africa but improvements in the teaching of statistics in the region have lagged behind those in developed countries and this needs to be urgently addressed.

Many undergraduates in applied disciplines view the statistics course as an unwelcome hurdle and an exam to be passed, rather than as an important problem-solving tool that will help them in their careers. The content of courses is partly to blame, since traditional teaching methods for statistics have tended to emphasize proficiency in the more mathematical and mechanical aspects of the subject, to the detriment of a deeper understanding of statistical concepts and reasoning (Moore 1997). Inclusion of real data from the subject area also reinforces the importance of statistics.

Constraints in Africa have exacerbated this problem. Students can afford few, if any, textbooks; consequently the quality of courses is often highly dependent on the approach and notes provided by the lecturer. If the lecturer is not a recently qualified statistician, then his or her approach may rely on a traditional approach that does not take account of modern research into

the effective teaching of statistics (Garfield 1995). A second constraint has been the lack of computers, so that students have been restricted to simple analysis of unrealistically small data sets, which has done little to reinforce the relevance of statistics. Limited numbers of statisticians per institution, large classes, lack of teaching aids, and lack of opportunities for statisticians to interact and keep abreast of developments in methodology and usage, have all contributed to the slow rate of improvement of statistical training in the continent.

However, change has been taking place over the past few years. For example, undergraduates in the Agriculture Faculty at the University of Nairobi had almost no access to computers prior to 2003. Now they can work in a lab of 50 machines that provides internet access, and they can also use computers in their halls of residence or the library. With falling prices, some can even afford to buy their own machines.

Internet access has also provided staff and students with access to a wide range of free resources to support the teaching and learning of statistics. One such resource is the electronic statistics textbook CAST (Computer Assisted Statistics Textbooks). This paper describes how a customized version of CAST was developed for the African continent and how it has become an important resource, not only for lecturers and students in universities, but also for researchers and those who use data in other organizations.

2. BIRTH OF *CAST FOR AFRICA*

Since 2002, with Rockefeller Foundation support, the University of Nairobi has been improving the teaching of statistics for undergraduate students in the Faculty of Agriculture. Two of the authors of this paper (Kurji and Stern) developed a curriculum with a greater focus on real-life data than its predecessor. They realized that this curriculum required resources to support its implementation both in the classroom and for self-study by students. As computers were becoming more common in both areas, a computer-based resource was investigated and, among the many resources that were considered, CAST (Stirling 2001) was found to be the most promising. CAST took a similar data-focused approach and its interactive and dynamic diagrams were particularly effective in explaining statistical concepts.

CAST was initially developed as an e-learning book (e-book) for teaching internal and distance students in New Zealand. It has been incrementally improved by its authors for over a decade. It now contains several e-books

but the initial focus for African use was an introductory e-book for students in biology, agriculture, and health sciences. The structure of this e-book is described in section 3 below.

Unfortunately, the introductory biometrics CAST e-book did not match the order of topics in the curriculum developed for the University of Nairobi, moreover the examples were not topical, being largely drawn from New Zealand, Europe, and the United States. Also, although CAST could be used freely by those with internet access, such access was unreliable in Africa in 2004 and payment was needed at that time for a stand-alone version that could be used offline. (CAST is now free for all users.)

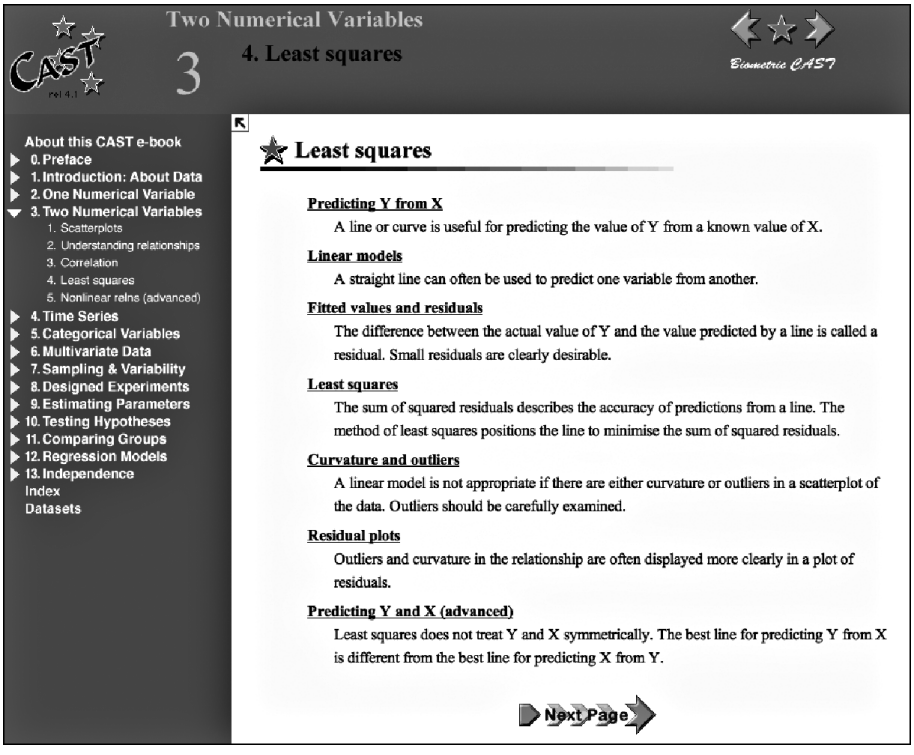
Since the existing CAST biometrics e-book was not ideally suited for use in Africa, discussions between the three authors led to the development of a customized set of e-books, *CAST for Africa*. The content of the new e-books was tailored to the new curriculum; moreover it could be used free of charge by students without internet connections within the University of Nairobi and, more generally, in the whole African continent.

3. THE PUBLIC CAST E-BOOKS

CAST is a collection of electronic statistics textbooks that are accessed on a computer with a web browser such as Internet Explorer, Firefox, Chrome, or Safari. Unlike conventional textbooks, CAST uses dynamic and interactive diagrams ("applets") to explain most concepts. By comparison with static diagrams, these help to retain student interest, are more memorable, and provide more effective explanations of many important statistical concepts (Stirling 2002).

The technology underlying the initial introductory CAST e-book (HTML, Javascript, and Java) made it relatively easy to add variations that were customized for different groups of users so the main public release now contains three introductory e-books (biometric, business, and general) that differ mainly in the data sets and scenarios that are used for motivation and illustration. This also made it easier to create customized e-books for use in Africa. Each introductory e-book contains about 430 pages and includes about 450 dynamic applets.

Figure 1: Pages in a typical CAST section



Three advanced e-books have been added more recently, containing material about multiple regression, experimental design in agriculture, and industrial experiments. An e-book with interactive exercises has also been added (Stirling 2010).

Each e-book is hierarchically structured in a similar way to a conventional textbook with chapters, sections, and pages. Figure 1 lists the pages in one section of the biometric e-book. Clicking the title of any page on the right replaces the section listing with the contents of that page.

3.1. Approach to teaching statistics

In recent years, the emphasis in teaching introductory statistics has moved away from the mathematical underpinnings of the discipline to practical aspects that are important to learners in their later careers. Computers are widely used to apply statistical methods to data, so formulae have become less important than decisions about appropriate analyses and interpretation of computer-generated results. Many learners are primarily interested

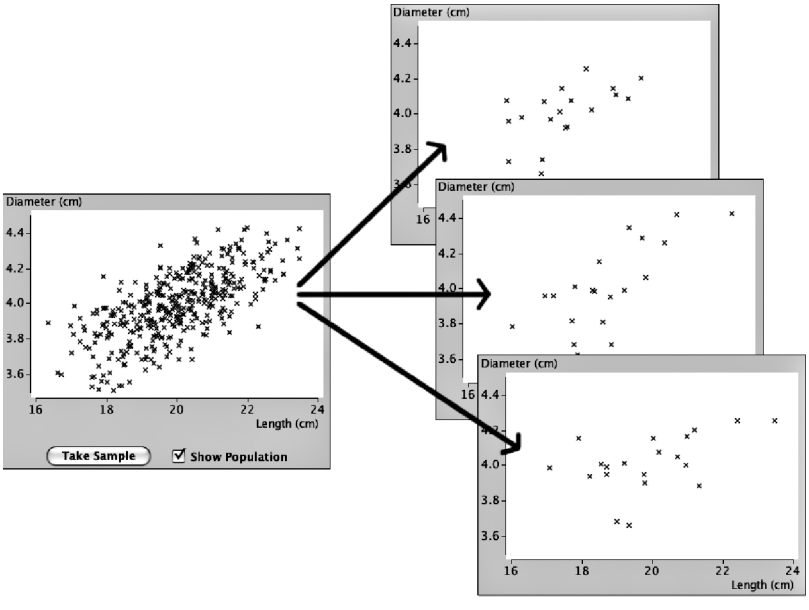
in a specific application area and they study statistics only because they are directed to do so. Hence it is useful if the data sets and scenarios that they encounter when studying statistics are relevant to their own subject areas.

The introductory CAST e-books all adopt a largely non-mathematical approach to statistics and use real-life data extensively for motivation and as examples. Many important statistical ideas and methods do not rely on probability, distributions or inference, so the first third of each introductory CAST e-book is devoted to exploratory and descriptive data analysis. Topics such as transformations, least squares, time series smoothing, and simple analysis of contingency table data are introduced here. Modeling of variability and inference for the main standard data structures are covered in the rest of each e-book. This approach was also consistent with the curriculum being developed for the University of Nairobi.

3.2 Interactive dynamic diagrams

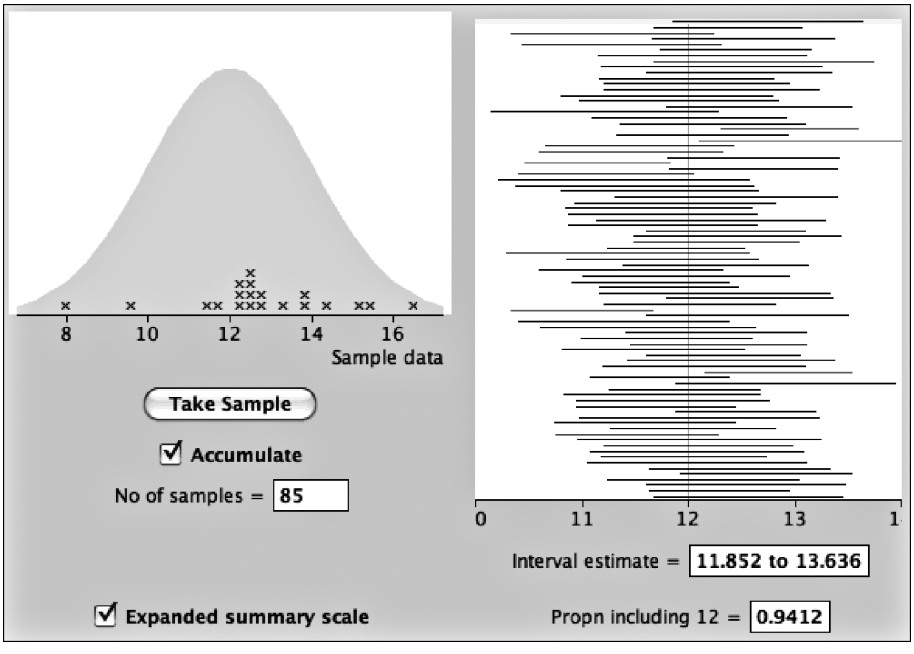
CAST makes extensive use of interactive diagrams to explain and illustrate statistical concepts. Such dynamic diagrams have a better chance of being retained in the student's memory than static diagrams. They provide a form of active learning for self-study (Meyers and Jones 1993) and are also effective in lectures if projection facilities are available, where the animation also provides an interlude from note-taking.

Figure 2: Sampling variability in scatterplots



As an example, the interactive diagram shown in Figure 2 takes random samples from a bivariate population of size 200 to illustrate the idea of sampling variability. The variation in the sample scatterplots is surprising to many learners and is therefore an important concept for them to grasp.

Figure 3: Confidence intervals

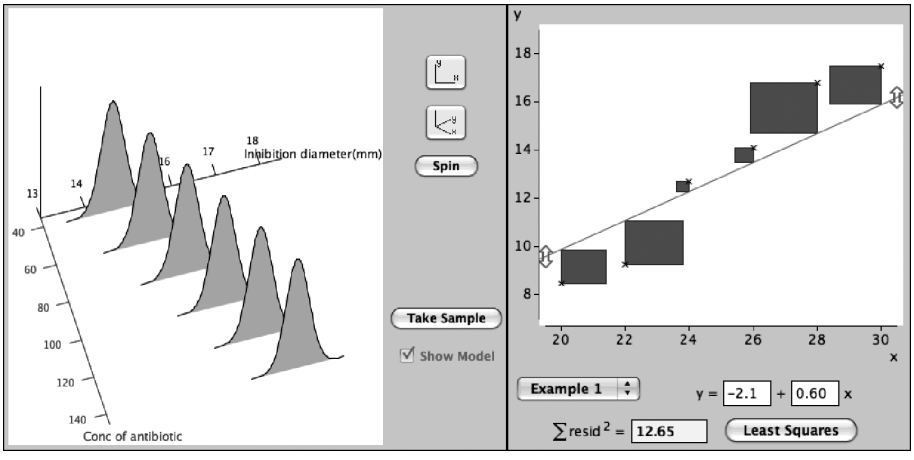


In the chapters that teach inference, the results of repeated simulations are used to accumulate the empirical distribution of sample summaries to illustrate their properties. For example, Figure 3 selects random samples from a normal distribution and accumulates a display of the resulting 95 percent confidence intervals for the population mean. The intervals that do not include the true mean, 12, are displayed in red. From this diagram, students can discover that about 95 percent of such confidence intervals include the true mean, but that 5 percent fail to do so.

Many other diagrams are designed to illustrate a single concept without the use of simulations. For example, in Figure 4, the diagram on the left is three-dimensional and can be rotated with the mouse to give a feel for the linearity and constant standard deviation implied by the simple linear regression model. The diagram on the right allows the line to be dragged with the arrows to illustrate the concept of least squares. The residual sum of squares is represented by the total area of the shaded squares and students

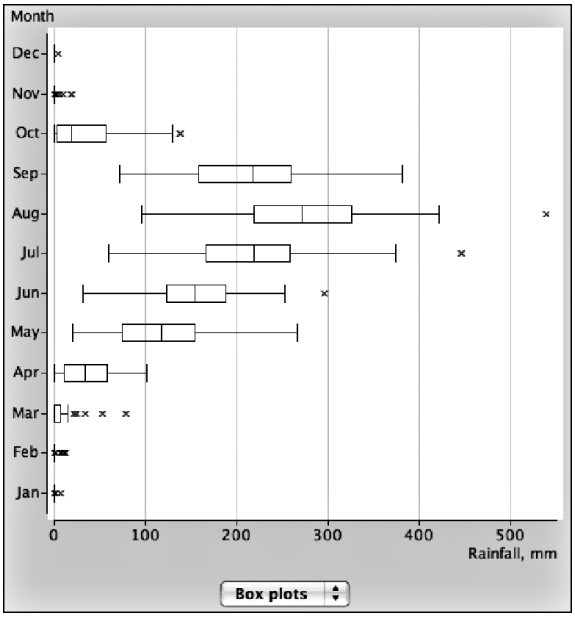
are given the target of minimizing this, with the "Least Squares" button providing the least squares fit of the line.

Figure 4: Explaining regression ideas



4. CUSTOMIZATION FOR AFRICA

Figure 5: Monthly rainfall in Samaru, Nigeria



The CAST introductory biometrics e-book provided an interactive data-based approach that could be used both as a teaching aid in class and as self-study material outside class hours. There were, however, differences between the structure and content of this e-book and the curriculum that had been developed for the Agriculture Faculty at the University of Nairobi. This curriculum involved two courses, the first comprised exploratory material and the second course introduced models for data and inference.

A reorganization of the e-book was needed to make its chapters and sections match the new curriculum, so two customized CAST e-books were created with the same split of topics. Some advanced topics such as transformations were included as optional chapters at the end, whereas other topics, such as time series, that were of lesser relevance in an agricultural context, were given reduced coverage or deleted.

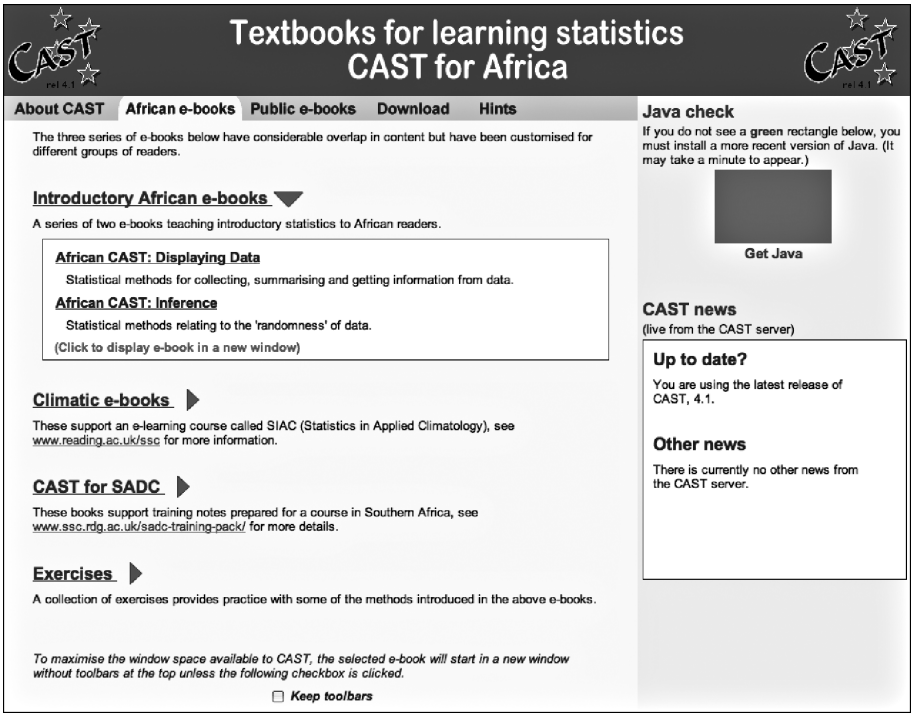
It was also felt that the customized CAST e-books would be more readily accepted if some data sets and scenarios were replaced with ones of greater interest to students in Africa. Africanization of the data sets and scenarios will be an ongoing task, but we do have some African examples in the current release of *CAST for Africa*. For example, the box plots in Figure 5 describe monthly rainfall in Samaru, Northern Nigeria between 1928 and 1983. In some artificial data sets, the context has been changed with the values simply rescaled. We hope to add more such examples as scientists realize the importance of relevant local data both for the students and the scientists themselves.

CAST for Africa was initially developed to teach introductory statistics to university students with a focus on agriculture. However, it was soon discovered that the statistical methods were relevant to teaching introductory statistics in many other disciplines too, so it started to be used in short course training of other groups in Africa. Given the ease with which customized CAST e-books could be created for different applications and markets, two other collections of e-books were developed: *Climatic CAST* is a set of e-books for meteorological officers who are in interested in statistical analysis of climatic data, while *CAST for SADC* is another set of e-books for officers in National Statistical Offices. These were developed with support from various donors and after discussions between the authors and other stakeholders and were customized to include topics of particular relevance. Some new material was specially developed, such as a new section in one climatic e-book about risks and percentiles, which was particularly relevant to the analysis of rainfall data and a new section in one of the SADC e-books about presentation of data in tables.

All of the customized African CAST e-books (*CAST for Africa, Climatic CAST, and CAST for SADC*) are freely accessible with a web browser from: http://cast.massey.ac.nz/african. The main public CAST e-books can be accessed from one tab on this page and the African e-books can be accessed from a different tab (Figure 6). Since all CAST e-books are available from this page, the resource provides a wide range of material covering many application areas via the same web site and supports the current emphasis on a multidisciplinary approach in development work.

The African e-books have also been included in various resource CDs[5] that have recently been developed for scientists, students, and trainers in Africa.

Figure 6: Initial page of CAST for Africa



CAST is free for all to use (under a Creative Commons licence). It can be downloaded (via the Download tab in Figure 6) to a local hard disk or server and run without an internet connection.

5 Biometry Resource Pack (developed with Rockefeller Foundation funding, distributed by BUCS); Biometry & Research Methods Teaching Resource (developed with Rockefeller foundation funding, distributed by ILRI); and Research Methods Resources (distributed by RUFORUM).

5. CONCLUSION

We have demonstrated that free innovative teaching resources can be provided to Africa at a relatively small development cost. *CAST for Africa* was initially used in Kenya at the Faculty of Agriculture in the University of Nairobi and has since been used at the Department of Mathematics and Applied Statistics in University of Maseno. It has also been used as part of various e-learning courses for meteorologists and others in Africa and, through exposure from these courses, at other African universities.

Although there has been no formal evaluation of the educational benefits of using CAST, feedback from student evaluations has been consistently positive about its effectiveness. Students have voiced not only their appreciation for having such a free resource for use in learning, but also their enjoyment from using it. A number of trainers and organizations have expressed interest in using the various e-books in *CAST for Africa* and we hope that future users will contribute African data sets and ideas for further customization.

REFERENCES

Garfield, J. (1995). "How Students Learn Statistics, " *International Statistical Review*, vol. 63, no. 1, pp. 25–34.

Meyers, C. and T.B. Jones (1993). *Promoting Active Learning: Strategies for the College Classroom*. San Francisco: Jossey-Bass Publishers.

Moore, D.S. (1997). "New Pedagogy and New Content: the Case of Statistics," *International Statistical Review*, vol. 65, no. 2, pp. 123–165.

Stirling, D. (2010). "Random Computer-based Exercises for Teaching Statistical Skills and Concepts." In P. Bidgood, N. Hunt and F. Jolliffe (eds), *Assessment Methods in Statistics Education*. Chichester: John Wiley & Sons, pp. 223–233.

Stirling, W.D. (2001). *CAST (Computer Assisted Statistics Textbooks),* http://cast.massey.ac.nz

Stirling, W.D. (2002). "Interactive Content in Web Pages to Teach Statistics." In B. Phillips (ed.), *Proceedings of the Sixth International Conference on Teaching of Statistics*. Cape Town/ Voorburg, The Netherlands: International Statistical Institute.

Ligne éditoriale

Le Journal statistique africain a été établi pour favoriser la compréhension du développement statistique dans la région africaine. Il se concentre sur des questions liées aux statistiques officielles aussi bien que l'application des méthodologies statistiques pour résoudre des problèmes pratiques d'intérêt général pour les statisticiens de métier. L'intérêt particulier est de montrer comment les statistiques peuvent aider à mettre en exergue les problèmes de développement et de politique publique tels que la pauvreté, le genre, l'environnement, l'énergie, le VIH/ SIDA, etc.; le développement de la culture statistique; la prise en compte des questions de développement régional et national; le développement des capacités statistiques et des systèmes statistiques nationaux efficaces; et le développement des statistiques sectorielles comme les statistiques d'éducation, de santé, des statistiques agricoles, etc.

En plus des universitaires et des statisticiens de métier, le Journal devrait revêtir un grand intérêt pour les institutions de la région, notamment les offices nationaux de statistiques, les banques centrales, les instituts de recherche et les organisations économiques sous-régionaux et les agences internationales de développement.

Le Journal constitue un document de recherche et d'information entre les statisticiens et les utilisateurs de l'information statistique, principalement dans la région africaine. Il publie entre autres:

- des articles sur le plaidoyer en matière de statistique qui démontrent le rôle essentiel des statistiques dans la société plutôt que la présentation des outils techniques,
- des articles sur les méthodologies statistiques, avec un accent particulier sur les applications,
- des articles sur les meilleures pratiques et les leçons tirées de la région,
- des avis sur des questions d'intérêt général pour la communauté statistique et les utilisateurs de l'information statistique dans la région africaine,
- des informations et des annonces sur les prochains événements, les conférences, les appels à contribution pour des papiers, et
- les développements statistiques récents et tout autre aspect susceptible d'intéresser la communauté statistique dans la région.

Les articles, qui n'ont pas besoin de contenir du matériel original, devraient intéresser une grande partie des statisticiens professionnels dans la région.

Tous les manuscrits seront passés en revue et évalués sur le contenu, la langue et la présentation.

Editorial Policy

The African Statistical Journal was established to promote the understanding of statistical development in the African region. It focuses on issues related to official statistics as well as the application of statistical methodologies to solve practical problems of general interest to applied statisticians. Of particular interest will be an exposition of: how statistics can help to illuminate development and public policy issues like poverty, gender, environment, energy, HIV/AIDS, etc.; development of statistical literacy; tracking national and regional development agenda; development of statistical capacities and effective national statistical systems; and the development of sectoral statistics e.g. educational statistics, health statistics, agricultural statistics, etc.

In addition to individual academic and practicing statisticians, the Journal should be of great interest to a number of institutions in the region including National Statistical Offices, Central Banks, research and training institutions, subregional economic groupings, and international development agencies.

The Journal serves as a research outlet and information sharing publication among statisticians and users of statistical information mainly in the African region. It publishes, among other things:

- Articles of an expository or review nature that demonstrate the vital role of statistics to society rather than present technical materials;
- Articles on statistical methodologies with a special emphasis on applications;
- Articles about good practices and lessons learned in statistical development in the Africa region;
- Opinions on issues of general interest to the statistical community and users of statistical information in the region;
- Notices and announcements on upcoming events, conferences, calls for papers;
- Recent statistical developments and anything that may be of interest to the statistical community in Africa.

The papers, which need not contain original material, should be of general interest to a wide section of professional statisticians in the region.

All manuscripts will be reviewed and evaluated on content, language and presentation.

Instructions pour la préparation et la soumission de manuscrits

Soumission
Les manuscrits en anglais ou en français doivent être envoyés aux présidents du comité de rédaction par email aux adresses suivantes c.lufumpa@afdb.org et bkiregyera@yahoo.com avec copie à statistics@afdb.org.

Titre
Le titre devrait être bref et détaillé. La page de titre doit inclure le titre du papier, le nom de l'auteur, l'affiliation et l'adresse. L'affiliation et l'adresse doivent figurer comme note de bas de page. Si le manuscrit est produit par des coauteurs, la même information doit être donnée pour les coauteurs.

Résumé, mots clés et remerciements
Un résumé court d'environ 150 mots doit être inclus au début du manuscrit ainsi qu'environ 6 mots clés utilisés dans le manuscrit. Les mots clés ne doivent pas répéter les mots utilisés dans le titre. Les remerciements, s'il y en a, doivent être insérés en bas de la page titre.

Section et numérotation
Les principaux titres doivent être numérotés (par exemple "**1. INTRODUCTION** "). Les sous-titres numérotés (par exemple "**1.1 L'élaboration de SNDS**") peuvent être employés mais par la suite les sous sous-titres ne devraient pas être numérotés. Le corps principal du texte sous forme de paragraphes ne devrait pas être numéroté.

Formatage
Veuillez utiliser le formatage minimal car ceci facilitera l'harmonisation de tous les articles. Garder par défaut le format "normal" (12 pt. Times New Roman) pour le texte principal avec l'espace d'une seule ligne entre les paragraphes. Ne pas appliquer le "corps de texte " en tant que modèle intégré. Les niveaux du titre doivent être facilement identifiables. Nous recommandons les majuscules en gras pour le premier niveau titre dans le texte principal (par exemple "**1. INTRODUCTION**"); ensuite les lettres minuscules en gras pour les sous-sections (par exemple "**1.1 L'élaboration de la SNDS**") et ensuite l'italique en gras sans numérotation (par exemple "***créant une culture de coopération***").Veuillez vous référer au dernier volume du JSA comme guide.

Tables et graphiques
Les tableaux et les graphiques doivent être numérotés et comporter un titre. Ceux-ci devraient être mentionnés (par exemple "voir Tableau 1")

dans le texte par le nombre correspondant, et non par une indication de page ou par d'autres indications telles que "ci-dessous" ou "au-dessus de".

Équations
Toutes les équations dans le papier doivent être numérotées. Les nombres doivent être placés à la droite de l'équation.

Bibliographie
Une liste de références doit être fournie à la fin de l'article (avant les annexes, le cas échánt). Les références doivent être classées par ordre alphabétique selon le nom de l'auteur ou de l'organisation. Là où il y'a plus d'une publication listée pour un auteur, elles doivent être classées chronologiquement (en commençant par les premiers publiés). Les références doivent donner le nom de la'auteur et l'année de publication, le titre du livre, le nom du journal le cas échéant. Utiliser a, b, c, etc. pour séparer les publications du même auteur au cours der la même années. Les titres des journaux et des livres devraient être en italique ; les titres des documents de travail et des raports non publiés devraient être placés dans de doubles guillemets et ne pas être imprimés en italique.

Exemples :

Fantom, N. et N. Watanabe (2008), "Improving the World Bank's Database of Statistical Capacity," *African Statistical Newsletter*, vol. 2, no. 3, pp. 21-22.

Herzog, A. R. and Dielman, L. (1985). "Age Differences in Response Accuracy for Factual Survey Questions," *Journal of Gerontology*, vol. 40, pp. 350-367.

Kish, L. (1988a). "Multipurpose Sample Designs," *Survey Methodology*, vol. 14, no. 3, pp. 19-32.

Kish, L. (1988b). *A Taxonomy of Elusive Populations*, Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 44-46.

World Bank (2006). *Statistical Capacity Improvement in IDA Countries – Progress Report.* Washington DC: The World Bank.

Renvois

Dans le corps principal de l'article, les renvois devraient suivre le modèle de Harvard, par exemple (Kish 1988a ; Herzog et Dielman 1985 : 351). Pour des renvois à trois auteurs ou plus, seulement le premier nom de famille devrait être donné, suivi par *et al.*, bien que les noms de tous les auteurs doivent être fournis dans la Bibliographie elle-même. Les abréviations ibid. et op. cit. ne devraient pas être employées dans le texte ou dans les notes de bas de page.

Guidelines for Manuscript Submission and Preparation

Submissions
Manuscripts in English or French should be sent by email to the Co-Chairpersons, Editorial Board at: c.lufumpa@afdb.org and bkiregyera@yahoo.com with a copy to statistics@afdb.org.

Title
The title should be brief and specific. The title page should include the title, the author's name, affiliation and address. The affiliation and address should be given as a footnote on the title page. If the manuscript is co-authored, the same information should be given for the co-author(s).

Abstract, Key Words, and Acknowledgments
A short abstract of about 150 words must be included at the beginning of the manuscript, together with up to 6 key words used in the manuscript. These key words should not repeat words used in the title. Acknowledgments, if any, should be inserted at the bottom of the title page.

Sections and Numbering
Major headings in the text should be numbered (e.g. "**1. INTRODUCTION**"). Numbered subheadings (e.g. "**1.1 The establishment of the NSDS**") may be used but thereafter sub-subheadings should be unnumbered. Main body text in the form of paragraphs should not be numbered.

Formatting
Please use minimal formatting as this will facilitate harmonization of all the papers. As your default, keep to "normal" (12 pt. Times New Roman) for main text with a single line space between paragraphs. Do not apply "body text" as an inbuilt style. The levels of heading need to be easily identifiable. We recommend all capitals bold for the first level of heading in the main text (e.g. "**1. INTRODUCTION**"); thereafter bold upper and lower case for subheadings (e.g. "**1.1 The establishment of the NSDS**") and unnumbered bold italic (e.g. "***Creating a culture of cooperation***") thereafter. Please refer to the latest volume of the AJS as a guide.

House Style
The Bank's house style is US rather than British spellings (e.g. "organization" not "organisation"; "program" rather than "programme", "analyze" etc.). Use "percent" rather than % or "per cent" and double rather than

single quotation marks. Dates should be US style (e.g. December 11, 1985 not 11 December 1985).

Tables and Figures
Tables and figures should be numbered and given a title. These should be referred to in the text by number (e.g. "See Table 1"), not by page or indications such as "below" or "above".

Equations
Any equations in the paper should be numbered. The numbers should be placed to the right of the equation.

References
A list of references should be given at the end of the paper (to precede the Annexes, if included). The references should be arranged alphabetically by surname/name of organization. Where there is more than one publication listed for an author, order these chronologically (starting with the earliest). The references should give the author's name, year of publication, title of the essay/book, name of journal if applicable. Use a, b, c, etc. to separate publications of the same author in the same year. Titles of journals and books should be in italic; titles of working papers and unpublished reports should be set in double quotation marks and not italicized.

Examples:

Fantom, N. and N. Watanabe (2008), "Improving the World Bank's Database of Statistical Capacity," *African Statistical Newsletter*, vol. 2, no. 3, pp. 21-22.

Herzog, A. R. and Dielman, L. (1985). "Age Differences in Response Accuracy for Factual Survey Questions," *Journal of Gerontology*, vol. 40, pp. 350-367.

Kish, L. (1988a). "Multipurpose Sample Designs," *Survey Methodology*, vol. 14, no. 3, pp. 19-32.

Kish, L. (1988b). *A Taxonomy of Elusive Populations*, Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 44-46.

World Bank (2006). *Statistical Capacity Improvement in IDA Countries – Progress Report*. Washington DC: The World Bank.

Cross References

In the main body of the article, cross-references should be Harvard-style, e.g. (Kish 1988a; Herzog and Dielman 1985: 351). For cross-references to three or more authors, only the first surname should be given, followed by *et al.*, although the names of all the authors must be provided in the References entry itself. Abbreviations ibid. and op. cit. should not be used in the text or in footnotes.

Acknowledgments

The Editorial Board would like to express its appreciation to all authors who submitted papers for publication in this issue of the *African Statistical Journal* and to the following people who assisted with the review and editing of the papers that are published in this volume:

Jeremiah Banda: International Statistical Consultant, Lusaka, Zambia

Barbara Barungi: Lead Economist, AfDB, Tunis, Tunisia

Graham Eele: International Statistical Consultant, UK

Jacky Galpin: Associate Professor, School of Statistics and Actuarial Science, University of the Witwatersrand, South Africa

Michel Mouyelo-Katoula: Global Manager, International Comparison Program, World Bank, Washington DC, USA

Lehana Thabane: Associate Professor, McMaster University, Canada

François Yatta: Coordinator LEDNA, Pretoria, South Africa

Remerciements

Le comité de rédaction voudrait exprimer sa gratitude à tous les auteurs qui ont soumis des articles pour la publication dans cette édition du Journal statistique africain et aux personnes ci-après qui ont assisté pour passer en revue et éditer les articles publiés dans ce volume:

Jeremiah Banda : Consultant International en Statistiques, Lusaka, Zambie

Barbara Barungi : Economiste Spécialiste en Chef, AfDB, Tunis, Tunisie

Graham Eele : Consultant International en Statistiques, Royaume-Uni

Jacky Galpin : Professeur, School of Statistics and Actuarial Science, University de Witwatersrand, Afrique du Sud

Michel Mouyelo-Katoula : Coordinateur Mondial, Programme de Comparaison Internationale, Banque Mondiale, Washington DC, Etats-Unis

Lehana Thabane : Professeur, McMaster University, Canada

François Yatta : Coordinateur LEDNA, Pretoria, Afrique du Sud

Déclaration de Dakar sur le développement statistique

PARIS21 reunion, Dakar, Sénégal, 16-18 novembre 2009

PRÉAMBULE

Lors de la réunion du Consortium de PARIS21 qui s'est tenue à Dakar (Sénégal) du 16 au 18 novembre 2009, les participants ont unanimement réaffirmé qu'un système statistique fiable et efficace constitue un élément fondamental de la bonne gouvernance ; ils ont également rappelé que, pour être en mesure de fournir les données nécessaires au suivi des objectifs du Millénaire pour le développement en 2015, il est indispensable d'agir rapidement. Tout en reconnaissant que de grands progrès ont été accomplis depuis 2000, le Consortium a souligné qu'il reste beaucoup à faire pour concrétiser l'ambition du Plan d'action de Marrakech pour la statistique (MAPS). Des mesures concertées et coordonnées doivent être prises pour utiliser plus efficacement les données statistiques, afin de contribuer à la réduction de la pauvreté et de renforcer et pérenniser les capacités des systèmes statistiques, notamment dans les pays en développement. Le soutien à la compilation, la publication et l'utilisation des statistiques relève d'une responsabilité collective pour que, selon les termes de la Déclaration du Millénaire, « la mondialisation devienne une force positive pour l'humanité entière ».

VALEURS ET PRINCIPES

Le Consortium a appelé tous les partenaires à affirmer que les statistiques officielles sont un bien public et que leur production et leur diffusion sont l'une des missions essentielles de chaque gouvernement. Afin de renforcer leur crédibilité et leur responsabilisation, les gouvernements doivent soutenir et développer leurs systèmes statistiques en appliquant les principes fondamentaux de la statistique officielle définis par les Nations Unies, ainsi que divers autres déclarations et codes adoptés à l'échelon régional. En outre, le Consortium a demandé à tous les acteurs du développement de la statistique de s'engager à respecter les quatre principes suivants :

1. **Tous les pays doivent être encouragés et aidés à déterminer leurs propres priorités en intégrant les besoins des utilisateurs et à définir eux-mêmes leur processus de développement statistique en partant de la collecte jusqu'à la diffusion tout en respectant les normes de qualité internationalement reconnues.**

2. **Tous les partenaires doivent promouvoir le développement de systèmes et méthodes statistiques permettant d'anticiper et de satisfaire les besoins nouveaux en matière de données, d'où qu'ils proviennent.**

3. **Les mesures prises pour améliorer les statistiques doivent avoir pour but de soutenir, renforcer et pérenniser les institutions et offices en charge des systèmes statistiques nationaux.**

4. **Les partenaires du développement doivent contribuer au renforcement des capacités et à la meilleure utilisation des systèmes statistiques des pays en développement, conformément au programme d'action d'Accra.**

APPEL A L'ACTION

Le Consortium appelle tous les partenaires à agir dans cinq domaines clés :

1. **D'ici à 2014, la préparation des plans stratégiques devra avoir laissé place à leur mise en oeuvre, soutenue par des financements durables et des compétences techniques adaptées. De ce fait :**

    a. **tous les pays s'étant engagés à améliorer leur système statistique auront pu mettre en oeuvre leur stratégie nationale de développement de la statistique (SNDS) ;**
    b. **l'assistance au développement sera apportée dans le cadre de SNDS adoptées à l'échelon national, en respectant le principe d'alignement défini dans la Déclaration de Paris sur l'efficacité de l'aide.**

Des progrès significatifs ont été accomplis en termes de renforcement des systèmes statistiques nationaux, particulièrement dans les pays en développement, à la faveur du MAPS et de différents processus régionaux. Il est communément admis qu'il appartient à chaque pays de définir ses propres priorités de développement statistique, par le biais de la préparation d'une SNDS. Mais il est temps aujourd'hui de passer de la préparation à la mise en oeuvre.

2. **Les gouvernements et les partenaires du développement doivent, d'ici à 2014, assurer que tous les grands programmes statistiques mondiaux seront dotés de moyens financiers et techniques durables.**

Le système statistique international repose sur trois grands piliers : les systèmes statistiques nationaux qui sont pilotés, gérés et largement financés par chaque gouvernement ; une gouvernance statistique internationale à l'origine de la fixation de normes ; des programmes statistiques à l'échelle mondiale qui doivent être intégrés le mieux possible dans les programmes nationaux. Ces trois composants sont essentiels et ils devront être soutenus et développés au cours des cinq prochaines années.

3. Afin d'assurer une coordination plus efficace à tous les niveaux d'ici à 2014 :

a. les mécanismes internationaux et régionaux de coordination et de coopération doivent être plus performants et prendre en compte les besoins et priorités des pays en développement et des processus d'intégration ;
b. avec l'assistance des partenaires du développement, les gouvernements doivent constituer des partenariats nationaux pour la statistique, partout où ce type de partenariat n'est pas encore mis en place.

La coordination internationale et nationale des programmes statistiques doit être renforcée, comme il convient d'améliorer les échanges entre les producteurs et principaux utilisateurs de statistiques à tous les niveaux, en s'appuyant dans la mesure du possible sur les institutions, les partenariats et forums existants.

4. D'ici à 2014, les systèmes statistiques de tous les pays devront mieux répondre aux besoins des utilisateurs, et être efficaces, économiquement rentables, et basés sur les résultats, tout en respectant les principes d'indépendance et d'intégrité.

Les responsables des instituts officiels de la statistique ont un rôle capital à jouer pour garantir la qualité et l'intégrité des données qu'ils compilent et qu'ils publient. Ils doivent en outre répondre en toute transparence des ressources qu'ils utilisent ainsi que de leurs actions et décisions. Lorsque l'aide au développement est apportée pour contribuer au renforcement des capacités statistiques, cette responsabilisation est encore plus indispensable.

5. D'ici à 2014, la communauté internationale devra soutenir un programme de recherche et développement visant à moderniser les outils et technologies statistiques, et encourager leur utilisation notamment dans les pays en développement.

Les activités statistiques – à savoir la compilation et la manipulation de grands ensembles de données – ont connu une évolution majeure au cours des dernières années grâce à l'essor des technologies de l'information et de la communication. Toutefois, dans nombre de pays, les procédures statistiques de base n'ont que peu évolué. La mise au point d'outils plus performants est une réelle nécessité, tout particulièrement pour la collecte, l'analyse et la présentation des données. Il est également indispensable de concevoir des méthodes fiables qui soient économiquement supportables par les petits pays.

MISE EN OEUVRE

Afin de concrétiser la mise en oeuvre de la Déclaration de Dakar, le Consortium a proposé le plan d'action suivant :

1. Le secrétariat de PARIS21, en collaboration avec d'autres partenaires, assurera le suivi de la mise en oeuvre de la Déclaration de Dakar. Il élaborera et publiera un rapport annuel sur les progrès accomplis ;

2. Une réunion du Consortium est programmée avant la fin 2014 dans le but de faire le point et de rendre compte des progrès de la mise en oeuvre de la Déclaration ;

3. Tous les membres et partenaires du Consortium s'engagent à poursuivre la promotion de la statistique et à mobiliser des ressources pour mettre en oeuvre la Déclaration ;

4. D'ici la fin 2010, la Déclaration de Dakar sur le développement statistique sera officiellement présentée à tous les organismes et institutions compétents.

Dakar Declaration on the Development of Statistics

PARIS21 Consortium Meeting, in Dakar, Senegal, November 16-18, 2009

PREAMBLE

At the PARIS21 Consortium meeting, held in Dakar, Senegal from November 16 to 18, 2009, the participants unanimously affirmed that an effective and efficient statistical system is an essential element of good governance and that urgent action is still required if the data needed to monitor the Millennium Development Goals are to be provided by 2015. The meeting recognised that a lot has been achieved since 2000, but reaffirmed that much still remains to be done to deliver the vision of the Marrakech Action Plan for Statistics (MAPS). Concerted and co-ordinated actions are required to make more effective use of statistical data to support poverty reduction policies and programs and to strengthen and sustain the capacity of statistical systems especially in developing countries. There is a collective responsibility to support the compilation, publication and use of statistics to ensure that, in the words of the Millennium Declaration "globalization becomes a positive force for all the world's people".

VALUES AND PRINCIPLES

The Consortium called upon all partners to recognise that official statistics are a public good and that their production and dissemination is a core function of all governments. To promote trust and accountability, governments should support and develop their statistical systems in line with the United Nations' Fundamental Principles of Official Statistics as well as various regional declarations and codes. In addition, the Consortium called upon everyone involved in the development of statistics to commit to the following four principles:

1. **All countries must be encouraged and supported to define their own priorities, integrating user needs, and set out their own development pathways for statistics, from collection to dissemination, respecting internationally recognised quality standards.**

2. **All partners should promote the development of statistical systems and methods that anticipate and respond to new and emerging requirements for data at all levels.**

3. **Efforts to improve statistics should support, strengthen and sustain the institutions and agencies that make up national statistical systems.**

4. **Development partners should help strengthen and use developing countries' statistical systems in line with the Accra Agenda for Action.**

CALL TO ACTION

The Consortium called on all partners to take action in five key areas:

1. **By 2014, the focus of attention should have moved from preparing strategic plans to implementing them with sustainable funding and technical capacity, so that:**

    a. **All countries that are committed to improving their statistical systems have been able to put their national strategies for the development of statistics (NSDS) into effect.**
    b. **Development assistance to statistics is being provided within the framework of nationally approved NSDSs, respecting the principle of alignment under the Paris Declaration on Aid Effectiveness.**

Significant and important progress has been made in strengthening statistical systems, especially in developing countries, through MAPS and a number of regional processes. The need for countries to set their own development priorities for their statistical systems through the preparation of NSDSs has been widely accepted; the need now is to move from preparation to implementation.

2. **Governments and development partners should place all essential global statistical programmes on a sustainable financial and technical footing by 2014.**

The international statistical system consists of three main pillars: national statistical systems, which are operated, managed and largely financed by governments; international governance of statistics, including the development of standards; and global statistical programmes, which should be integrated, as far as possible, into national programmes. All three components are essential and must be supported and developed over the next five years.

3. In order to ensure more effective coordination at all levels, by 2014:

a. International and regional mechanisms for coordination and cooperation should work more effectively, taking into account the needs and priorities of developing countries as well as integration processes.
b. Governments, with the assistance of development partners, should establish national partnerships for statistics, where such partnerships do not exist already.

Coordination of statistical programmes, both internationally and nationally, must be strengthened and consultation between statistical producers and key user groups must be improved at all levels, making use of existing institutions, partnerships and forums, wherever possible.

4. By 2014 statistical systems in all countries should better meet the needs of users, operating in an efficient, cost-effective and results-based manner, in line with the core requirements of independence and integrity.

The managers of all official statistical agencies have an essential role to play in ensuring the quality and integrity of the statistics they compile and publish. They must also be open and accountable for the resources they use and their decisions and actions. Where development aid is being provided to help build capacity, the need for accountability is even stronger.

5. By 2014, the international community should support a programme of research and development to modernize statistical tools and technologies and to promote their use, especially in developing countries.

Statistical activities, involving the compilation and manipulation of large data sets have been transformed in recent years by the use of information and communications technology. In many countries, however, basic statistical procedures have changed little in this time. There is a need for the promotion of existing tools as well as the development of better tools at all levels, but especially in the collection of source data and in data analysis and presentation. There is also a need to develop robust methods that are cost-effective in small countries.

THE WAY FORWARD

In order to achieve the vision of the Dakar Declaration, the Consortium proposed that the following actions be taken:

1. The PARIS21 Secretariat, in collaboration with other partners, monitors the implementation of the Dakar Declaration and compiles and publishes an annual report on progress.

2. A follow-up Consortium meeting is convened before the end of 2014 to follow up and report back on progress in implementing the Declaration.

3. All partners and members of the Consortium undertake to continue to advocate for statistics and to mobilise resources to implement the Declaration.

4. By the end of 2010, the Dakar Declaration for the Development of Statistics is presented formally to all relevant processes and institutions.

Déclaration de l'île de Goree : 5ème Symposium sur le Développement de la Statistique en Afrique

19-21 Novembre 2009, Dakar - Sénégal

"Les Technologies de l'Information et de la Communication dans la Diffusion des Données: Rapprocher les Producteurs et Utilisateurs dans les Séries de Recensements de la Population et de l'Habitat des années 2010"

PRÉAMBULE

La Déclaration de L'île de Gorée, une déclaration par les statisticiens africains sur l'île de Gorée, le 21 Novembre 2009, est à la fois significative et symbolique dans le périple de 53 ans de restauration de la statistique sur le continent. Son importance réside dans le fait qu'il marque le 5ème anniversaire du Symposium sur le Développement de la Statistique en Afrique (ASSD) ; un jalon important qui confirme l'achèvement de la phase de fondation de cet important forum qui a mobilisé l'engagement de tous les pays africains à procéder à la Série des Recensements de la Population et de l'Habitat des années 2010 (2010 RHPC) sur une période de cinq ans, un cri de cœur lancé depuis la conférence de Yaoundé en 2005. La plate-forme de l'ASSD ne cesse de gagner du terrain comme en témoigne la participation accrue et l'attraction suscitée par ses programmes pour un meilleur contenu technique.

Le symbolisme de cette déclaration trouve son encrage dans l'histoire paradoxale de l'île de Gorée. L'Ile, de par le statut de son passé inique en tant que antre stratégique de l'esclavage, le lieu ou des africains ont été emmenés de force vers les Amériques pour être réduits en esclaves. Elle est devenue aujourd'hui une étape importante dans le circuit touristique. Gorée représente de nos jours le symbole d'un désir furtif d'une Afrique qui veut prendre le taureau de l'adversité par les cornes. Aussi, la déclaration d'aujourd'hui symbolise le désir discret des africains de sortir des sentiers de la statistique obscure pour entrer dans l'ère mure de la statistique. Il ne pouvait pas y avoir meilleur endroit que la magnifique île de Gorée pour célébrer l'accomplissement de la phase de formation de l'Afrique qui mène vers l'illumination de la statistique et la détermination du continent pour un programme de développement sur la base de faits. Ce qui était autrefois une porte vers l'abîme obscur de l'esclavage est devenu maintenant une fenêtre à travers laquelle les statisticiens africains peuvent regarder vers un avenir radieux de la statistique en Afrique.

DÉCLARATION

Nous, les Directeurs d'Agences Africaines de la Statistique, Statisticiens, Responsables au sein des Agences Statistiques, Partenaires au Développement de la Statistique en Afrique et les Jeunes Statisticiens Africains, réunis sur l'île de Gorée, à Dakar, au Sénégal pour clôturer les travaux du 5ème Symposium sur le Développement de la Statistique en Afrique,

- **Guidés** par le thème, «Les Technologies de l'Information et de la Communication dans la Diffusion des Données : Rapprocher les Fournisseurs et Utilisateurs dans le cadre des Séries de Recensements de la Population et de l'Habitat des années 2010 »
- **Reconnaissant** les Principes et Recommandations de la Série de Recensements de la Population et de l'Habitat des années 2010, et les Dispositions concernant l'Afrique en Annexe,
- **Adhérant** à la mise en œuvre du Cadre Stratégique Régional de Référence en matière de renforcement des capacités africaines et à la Charte Africaine de la Statistique,
- **Conscients** des progrès des Nouvelles Technologies de l'Information et de la Communication et leur rôle dans l'amélioration de l'efficacité des méthodes de travail dans le 21ème siècle,
- **Inspirés** par la réussite de la 57ème Session de l'Institut International de la Statistique (2009) qui s'est tenu pour la première fois en Afrique au Sud du Sahara,
- **Considérant** les résolutions du Symposium sur le Développement de la Statistique en Afrique de Cape Town (2006), la Résolution de Kigali (2007), les Recommandations d'Accra (2008) et la Résolution de Luanda (2009),
- **Considérant** le Plan d'Action de Marrakech pour les Statistiques (2004)
- **Adhérant** à la Déclaration de Dakar sur le Développement de la Statistique,

Déclarons que:

1. Nous devons élaborer un plan rotatif annuel pour mettre en œuvre les recommandations et résolutions des précédents symposiums. Le plan de mise en œuvre rotatif sera élaboré par le Secrétariat du Symposium sur le Développement de la Statistique en Afrique, suivi de manière trimestrielle par les Amis de la CEA (ou un forum équivalent) et évalué lors des réunions ultérieures du Symposium sur le Développement de la Statistique en Afrique,

2. Nous adoptons le Manuel Africain de Traitement de Données (MATD/ADPH) comme plan directeur de traitement des données auquel nous nous référerons toujours comme directives au moment de l'élaboration de nos plans de traitement des données de recensement. Face à l'évolution constante des technologies et des meilleures pratiques, le MATD/ADPH doit être révisé tous les trois ans afin d'en accroître la pertinence,

3. Nous devons nous engager dans la mise en œuvre de la Stratégie de Mobilisation des Ressources pour le Recensement à définir avant la tenue du symposium ASSD de 2010. Cela passe par la signature de partenariat avec les fournisseurs de nouvelles technologies, la définition d'une stratégie de coopération sud-sud par la mutualisation des efforts. L'UNFPA et la BAD doivent prendre l'initiative de mobiliser des ressources pour les activités de recensement de l'Afrique,

4. Nous nous efforcerons à utiliser raisonnablement les Technologies de l'Information et de la Communication les plus efficaces pour planifier les recensements, entreprendre les processus préparatoires de recensement, collecter, traiter, analyser et diffuser les données de recensement. À cette fin, nous invitons les Amis de la CEA (ou un forum équivalent), la Commission de l'Union Africaine, le Secrétariat de l'ASSD, la BAD et la CEA à prendre des dispositions pour faciliter un processus de collaboration pour l'élaboration d'un Manuel de Meilleures Pratiques sur la Technologie de l'Information et de la Communication par le 6ème ASSD,

5. Nous procéderons à la consolidation de nos efforts pour produire des données de qualité et pertinentes qui aident à l'évaluation des cibles des Objectifs du Millénaire pour le Développement et, des documents nationaux de développement notamment le SRP et les plans sectoriels, par l'élaboration et la mise en œuvre des SNDS. Dans le cadre de leurs efforts d'amélioration de la SNDS, les Instituts Nationaux de Statistique (INS), doivent entreprendre des opérations statistiques notamment des enquêtes, des recensements, des sources administratives incluant l'enregistrement des faits d'état civil pour combler les lacunes au niveau des données relatives aux OMD,

6. Nous devons élaborer un Mécanisme d'Appui au Recensement par les Pairs pour rendre effective la résolution prise à Luanda, en Angola, d'établir des mécanismes pour soutenir les pays qui entreprennent des recensements et en particulier ceux qui connaissent des changements politique, économique et social. Le Secrétariat du Symposium sur le

Développement de la Statistique en Afrique (ASSD) devra faire l'ébauche d'un tel cadre,

7. Nous devons investir dans l'avenir en créant des opportunités pour attirer et inspirer la jeunesse africaine à poursuivre des carrières dans les domaines de la statistique et de manière intégrée à travers l'Afrique afin d'éviter les doubles emplois, et encourager le partage des ressources et réduire les écarts de compétences. À cet effet, l'initiative d'ISIbalo Capacity Building Program est à approuver et les partenaires sont invités à mobiliser les ressources nécessaires à la mise en œuvre des programmes,

8. Nous devons déployer des efforts soutenus pour accroître les perspectives de pérennité du Symposium sur le Développement de la Statistique en Afrique. Une Stratégie de Pérennisation de l'ASSD pour le maintien des facteurs critiques de succès et la prise en charge des besoins de financement à long terme doit être élaborée par le Secrétariat de l'ASSD lors de l'ASSD de 2010 (Egypte). L'engagement à long terme des partenaires financiers est d'une importance primordiale ainsi que celui des pays africains. Un plan quinquennal concernant l'organisation tenant compte des responsabilités du pays hôte telles que déclinées dans le Manuel d'Organisation de l'ASSD doit être obtenu auprès de cinq pays à chaque symposium annuel. Nous devons faire le bilan des 5 années d'organisation du symposium et mettre en exergue l'impact dans le développement de la Statistique en Afrique.

Gorée Island Declaration: Fifth Africa Symposium on Statistical Development

November 19-21, 2009, held in Dakar, Senegal

"Information and Communication Technology in Data Dissemination: Bringing Suppliers and Users Closer in the 2010 Round of Population and Housing Censuses"

PREAMBLE

The Gorée Island Declaration, a declaration by African statisticians on the Gorée Island on the 21st of November 2009 is both significant and symbolic in the 53-year rejuvenation journey of statistics on the continent. Its significance is borne out by the fact that it marks the 5th anniversary of the Africa Symposium on Statistical Development (ASSD), a significant milestone that signals the completion of the foundational phase of this important forum which has mobilised all but one African countries commitment to undertake a census in the 2010 Round of Population and Housing Censuses in a space of five years, a far cry from the lamentational tone of the Yaounde conference in 2005. The ASSD platform has gained massive momentum as witnessed by increased participation and gravitation of its programmes towards more technical content.

The symbolism of this Declaration is found in the paradoxical history of the Gorée Island. The Island, from its iniquitous past status as the strategic capital city of slavery when it was used to coercively expel Africans into slavery in Europe and the America's is now a major tourist attraction landmark. Gorée Island has become a symbol of Africa's determination to take the bull of adversity by its horns. Similarly, today's declaration symbolizes Africa's fervent desire to move from statistical obscurity to statistical maturity. There could not be a better place to celebrate the completion of Africa's formative phase towards statistical enlightenment and resolve to a fact-based African development agenda than at the magnificent Gorée Island. What was once a door into the dark abyss of slavery, the "point of no return" in Gorée Island (the opening through which Africans were pushed into ships that would transport them off-shore into a life of indignity) has now become a window through which African statisticians can glimpse a bright future for statistics in Africa.

DECLARATION

We, the Heads of African Statistics Agencies, Statisticians, Officials in Statistics Agencies, Partners in Africa's Statistical Development and Young African Statisticians, gathered on the Gorée Island, Dakar, Senegal to conclude the proceedings of the 5th ASSD,

- **Guided** by the theme, "Information and Communication Technology in Data Dissemination: Bringing Suppliers and Users Closer in the 2010 Round of Population and Housing Censuses"
- **Acknowledging** the Principles and Recommendations for the 2010 Round of Housing and Populations Censuses and the African Addendum to it,
- **Committed** to the implementation of the Reference Regional Strategic Framework for Statistical Capacity Building in Africa (RRSF) and the African Charter on Statistics,
- **Cognisant** of Information and Communication Technology advances and their role in increasing the efficiency of business processes in the 21st century,
- **Inspired** by the successful completion of the 57th Session of the International Statistics Institute (2009) that took place for the first time ever in Sub-Saharan Africa,
- **Noting** the Cape Town ASSD Resolutions (2006), the Kigali Resolution (2007), the Accra Recommendations (2008) and the Luanda Resolution (2009),
- **Noting** the Marrakech Action Plan for Statistics (2004),
- **Committed** to the Dakar Declaration on the Development of Statistics,

Declare that we:

1. Shall develop a rolling plan on an annual basis to implement recommendations and resolutions of preceding symposia. The rolling implementation plan will be developed by the ASSD secretariat, monitored on a quarterly basis by the Friends of the ECA (or equivalent forum) and evaluated in subsequent ASSD gatherings;

2. Adopt the African Data Processing Handbook (ADPH) as our data processing blueprint that we shall always refer to for guidance when developing plans for processing census data. In the face of ever-evolving technology and best practices, the ADPH shall be reviewed on a tri-annual basis to increase its relevance;

3. Shall engage in the implementation of the Census Resource Mobilization Strategy to be developed before the 2010 ASSD. The UNFPA and ADB shall also take the lead in mobilizing resources for Africa's census activities;

4. Shall endeavour to cautiously use the most effective Information and Communication Technologies to plan censuses, undertake census preparatory processes, collect, process, analyse and disseminate census data. To this end, we implore the Friends of the ECA (or equivalent forum), the African Union Commission, the ASSD secretariat, the AfDB, and ECA to collaboratively prepare an Information and Communication Technology Best Practice Handbook by the 6th ASSD;

5. Shall consolidate our efforts to produce quality data that are relevant to assist in the monitoring of the targets of the Millennium Development Goals (MDGs). NSOs, as part of their National Strategies for Development of Statistics (NSDS) improvement efforts shall undertake statistical operations, namely surveys, censuses, administrative sources including civil registration and vital statistics to fill MDGs data gaps;

6. Shall develop a comprehensive African Census Peer Support Framework to give effect to the resolution taken in Luanda, Angola to establish mechanisms to support countries undertaking censuses and especially those in a state of political, social and economic flux. The ASSD Secretariat shall draft such a framework;

7. Shall invest in the future by creating opportunities to motivate and inspire the African youth to pursue careers in statistics-oriented fields and this shall be done in an integrated fashion across Africa to avoid unnecessary duplication, encourage sharing of resources and bridging of skills gaps. To this end, the ISIbalo Capacity Building Programme initiative is applauded and partners are called upon to mobilize the required resources to implement the programmes;

8. Shall embark on an all-out effort to increase the prospects of the ASSD's sustainability. An ASSD Sustainability Strategy lifting critical success factors and long-term funding needs shall be crafted by the ASSD Secretariat by the 2010 ASSD. Long-term commitment from funding partners is of paramount importance; as it is from African countries. A five-year hosting commitment taking into account the host country's responsibilities as spelt out in the Manual for Organising ASSDs shall be obtained from five countries at each and every annual symposium.

Quatrième réunion du Comité africain de coordination statistique

Dakar, Sénégal, 22 novembre 2009

COMMUNIQUE

Le Comité africain de coordination statistique (CACS) a tenu sa quatrième réunion à l'Hôtel Méridien Président à Dakar, au Sénégal, le 22 novembre 2009. Parmi les participants figuraient des représentants de la Banque africaine de développement (BAD), de la Commission de l'union africaine (CUA), de la Commission économique des Nations unies pour l'Afrique (CEA) et d'AFRISTAT.

Le principal objectif de la réunion était de trouver des mécanismes permettant d'établir une plus grande synergie entre les institutions continentales et régionales et de coordonner les programmes et activités relatifs à la statistique, en vue d'accroître et de maintenir l'efficacité de l'aide fournie aux pays africains dans le domaine des statistiques. Le Groupe africain sur la comptabilité nationale (AGNA) a présenté la Stratégie africaine de mise en œuvre du Système de comptes nationaux de 2008.

Les principales conclusions de cette réunion sont présentées ci-après.

- Le CACS a rendu hommage à la BAD pour avoir fait preuve de leadership dans la mobilisation de fonds pour la mise en œuvre du CSRR en Afrique. Il s'agit notamment de sa contribution de 36 millions USD pour la phase II de son Programme de renforcement des capacités statistiques (2009 – 2010), en plus des 23 millions USD qu'elle avait déjà fournis pour la phase I (2004 – 2008).
- Le CACS a noté que bon nombre de pays n'ont pas encore conçu une NSDS conforme aux exigences des directives, et que peu de pays appliquent convenablement leur NSDS. Le CACS a recommandé l'organisation de plus d'activités de sensibilisation dans les pays en 2010. Les institutions concernées doivent être informées à l'avance afin de pouvoir se préparer à y participer.
- Le Comité a rendu hommage à la CUA pour l'adoption de la Charte africaine de la statistique par les chefs d'État et de gouvernement. Pour mettre à profit les avantages comparatifs de chaque institution et programme, les membres du CACS sont convenus que la CUA se chargera de la mise en œuvre de la Charte, tandis que la BAD et la CEA continueront à coordonner le processus d'exécution du CSRR.
- Le CACS a décidé que, autant que faire se peut, la mise en œuvre et les rapports concernant la Charte et le CSRR, ainsi que les activités de StatCom-Africa seront assurés par les mêmes réseaux, les mêmes comités

et les mêmes groupes de travail aux niveaux régional/continental, sous-régional et national. Les mécanismes de mise en œuvre et de coordination harmonisés seront finalisés lors de la réunion d'East London en décembre 2009.

- Il a également été convenu que le CSRR, ainsi que d'autres initiatives relatives au développement de la statistique en Afrique, constituent des instruments de mise en œuvre de la Charte et, en tant que tels, ne devraient pas entrer en conflit avec celle-ci.
- Le CACS a souscrit à l'idée de création du Groupe africain sur la formation statistique et les ressources humaines (AGROST), qui sera désormais l'unique groupe de travail compétent en matière de formation statistique, et qui bénéficiera de l'appui d'un secrétariat permanent basé au Centre africain pour la statistique de la CEA. Il est aussi convenu des prochaines étapes, y compris la préparation d'un programme de travail.
- Le CACS a de nouveau rendu hommage à la BAD pour le bon achèvement du PCI 2005 et lui a recommandé de mobiliser des ressources et de continuer à soutenir les pays pendant le PCI 2011. Le Comité a adopté la proposition de la BAD que le PCI-Afrique 2011 soit mené dans le cadre de groupes économiques sous-régionaux, notamment l'Union du Maghreb arabe (UMA), le Marché commun de l'Afrique de l'Est et de l'Afrique australe (COMESA), la Communauté économique des États de l'Afrique centrale (CEEAC), la Communauté économique des États de l'Afrique de l'Ouest (CEDEAO) et la Communauté de développement de l'Afrique australe (SADC).
- Le CACS a félicité le Groupe africain sur la comptabilité nationale (AGNA), présidé par la BAD, pour la préparation de la Stratégie africaine de mise en œuvre du Système de comptes nationaux de 2008 dans les pays. Le Comité a recommandé que la stratégie soit présentée aux responsables de services nationaux de statistiques lors de la réunion d'East London avant son adoption lors de la réunion de StatCom-Africa en janvier 2010. Il a adopté la proposition que la stratégie soit conjointement préfacée par les présidents de la CUA, de la CEA et de la BAD. Il a en outre recommandé à la CUA de promouvoir cette stratégie au plus haut niveau. Le CACS a félicité la BAD pour son soutien aux activités de l'AGNA et recommandé qu'elle en fasse de même avec les autres groupes de travail.
- La CUA a porté à la connaissance du CACS qu'elle a entamé l'élaboration d'une « Stratégie continentale d'harmonisation des statistiques de l'intégration régionale », grâce à laquelle le Système statistique africain pourra produire des statistiques harmonisées pour soutenir le processus

d'intégration du continent. Le CACS a salué cette initiative de la CUA en promettant de lui apporter tout son soutien.
- Enfin, il a été décidé que les communautés économiques régionales ci-après font désormais partie du CACS : UMA, CEEAC, CEDEAO, COMESA et SADC.

Liste des participants :

BAD :
- Charles Leyeka Lufumpa, Directeur, Département de la statistique
- Abdoulaye Adam, Chef de division par intérim, Division du renforcement des capacités statistiques
- Luc Mbong Mbong, statisticien principal, Division du renforcement des capacités statistiques
- Adalbert Nshimyumuremyi, statisticien principal, Division du renforcement des capacités statistiques
- Stephen Bahemuka, statisticien en chef, Division du renforcement des capacités statistiques

CUA :
- René N. Kouassi, Directeur des affaires économiques
- Thiekoro Doumbia, statisticien, Unité de la statistique

CEA :
- Dimitri Sanga, Directeur par intérim, Centre africain pour la statistique
- Awa Thiongane, conseiller régional en chef, Centre africain pour la statistique

AFRISTAT :
- Birimpo Lompo, Directeur général adjoint

Fourth Meeting of the African Statistical Coordination Committee (ASCC)

Dakar, Senegal, November 22, 2009

I. Preamble

The African Statistical Coordination Committee (ASCC) held its fourth meeting at the Méridien Président Hotel in Dakar, Senegal on November 22, 2009. In attendance were representatives of the African Development Bank (AfDB), AFRISTAT, the African Union Commission (AUC), and the United Nations Economic Commission for Africa (UNECA).

The main objective of this set of meetings is to discuss mechanisms for achieving greater synergy between continental/regional institutions, and coordinating statistical programs and activities within the Reference Regional Strategic Framework for Statistical Capacity Building in Africa (RRSF) with a view to enhancing and sustaining the effectiveness of statistical support to African countries. Specifically, and in line with the recommendations of the first, second, and third meetings, this meeting took stock of progress made by the Working Groups (WGs) set up by the ASCC. In addition, the ASCC considered the report on progress made by the African Group on National Accounts (AGNA), a working group created during the first meeting of the Statistical Commission for Africa's (StatCom-Africa) in 2008. Finally, it considered issues pertaining to other coordination modalities.

II. Reports on the ASCC Working Groups

RRSF Implementation Working Group

The ECA informed the ASCC that the RRSF is being implemented through different statistical initiatives and programs at country, subregional, and regional/continental levels. The implementation report is being jointly prepared by ECA and AfDB. The AfDB has recruited a consultant who is: (i) finalizing the Terms of Reference (ToRs) for the coordinators in order to ensure that the latter are nominated at all levels; (ii) preparing the reporting formats; and (iii) preparing the RRSF implementation report. The RRSF implementation report will be presented at the next StatCom-Africa in January 2010.

The AfDB reported on the implementation of its Phase II statistical capacity-building program in Africa over the next 2 years (2009-2010). This program was approved by the AfDB Board of Directors in November 2008. The Bank's contribution to the Phase II cost is estimated to be up to US$ 36 million. Other contributions are provided by partners.

National Strategies for the Development of Statistics

The ASCC noted the progress made by African countries to design a NSDS, in that 31 countries out of 53 have finalized the preparation of their NSDS. However, there remain a number of countries that have yet to design a NSDS. Most of them have only prepared a statistical master plan focusing on the NSO. There is a need to encourage these countries to design an NSDS that mainstreams sectors and subnational statistical systems. The African Development Bank is leading a working group on mainstreaming subnational statistics into the NSDS. The guidelines will be published by mid 2010. Also the AfDB, FAO, and PARIS21 are jointly working together to prepare a manual on mainstreaming agriculture statistics into the NSDS.

The ASCC also noted the problems known by countries to implement the NSDS. In this regard, the Committee recognized that funding and leadership are key elements of the enabling environment toward a proper implementation of the NSDS by African countries. It was also noted that reliance on donors in financing the NSDS is not sustainable. In 2009, PARIS21, ECA, and the AfDB jointly undertook advocacy missions to four African countries (Angola, Burundi, Côte d'Ivoire, and Democratic Republic of Congo). The ASCC recommended that more joint advocacy missions be organized in 2010. The concerned institutions should be informed ahead of time to enable them to make arrangements to participate.

The African Charter on Statistics

The AUC reported that the African Charter on Statistics had been adopted by the 12th Ordinary Session of the Assembly of Heads of State and Government of the African Union (AU) held in Addis Ababa, Ethiopia, on February 3, 2009. A copy of the Charter has been sent to all AU member states for signature and ratification. Up until the present time, only 10 countries have ratified it; for its entry into force, it has to be ratified by at least 15 countries. The Charter provides a regulatory framework for statistics development in the continent. The AUC representatives recommended that members of the ASCC assist in advocating for the charter signature, ratification and implementation by countries.

The Committee congratulated the AUC on the adoption of the Charter by Heads of State and Government. Leveraging on the comparative advantages of each institution and framework, the ASCC agreed that the AUC would take the lead on the Charter implementation, and that the AfDB and ECA will continue to coordinate the RRSF implementation process.

The Committee also agreed that the RRSF as well as other initiatives in the area of statistical development in Africa are instruments for the implementation of the Charter, and as such there should be no conflict between these activities and the Charter.

The Statistical Training Program for Africa

The ACS reported on the first regional forum on statistical training in Africa that took place in Bujumbura, Burundi, from June 18-19, 2009. The Forum was jointly organized by the UNECA, the AfDB, and PARIS21, in close collaboration with the ACBF, the AUC, the French Ministry for International Cooperation, and the National Institute of Statistics and Economic Studies of Burundi (INSTEEBU). The Forum endorsed the creation of the African Group on Statistical Training and Human Resources (AGROST) as a single working group dealing with statistical training in Africa, to be supported by a permanent secretariat to be based at the African Center for Statistics, UNECA. It also agreed on next steps, including the development of a work program.

Support to fragile states

The AfDB reported that the Bank has committed to provide additional resources to countries emerging from conflicts or that are in a state of fragility. These countries can also use the operation window of the Bank to finance their statistical capacity development.

International Comparison Program for Africa

The ASCC noted that the 2005 ICP final results have been published and disseminated. However, the dissemination process should be enhanced so that these results are widely utilized. In its 39th session, the UN Statistical Commission welcomed the successful completion of the 2005 ICP round and requested that the next round begin in 2011. In preparation for that round and to publish the 2009 and 2010 PPPs for African countries, various activities are currently underway: (i) In view of estimating the 2009 PPPs, data are collected using a reduced geographical coverage with the full and reduced list obtained using the least PPP deviation approach, and (ii) the AfDB participates actively in the harmonization of the Consumer Price Index of the WAEMU countries. The process provided an opportunity to make progress in the integration of ICP and CPI by enlarging the number of products that are common to both statistical operations.

It is envisaged to conduct the ICP-Africa 2011 on the basis of subregional economic groupings: Arab countries (AMU), AFRISTAT (including the two subregions CEEAC and ECOWAS), COMESA and SADC. The item list

needs to be reviewed and updated to take into consideration new products, the implementation structure of the program, and lessons learned from the 2005 round.

Training of data collectors and supervisors, testing of survey instruments and other preparatory activities should be conducted in the third quarter of 2010 so that data collection can start in January 2011 in compliance with the Global ICP data collection program.

Statistical Associations

The Committee noted that a meeting of statistical associations was being organized in Durban alongside the ISI meetings in August 2009. The ASCC recommended that countries should establish strong statistical associations (only 20 countries are reported to have a statistical association to date) before coming up with the regional association. African statisticians should also be encouraged to take up membership of the international statistical associations.

The ASCC was informed that African universities are also in the process of launcing a statistical association. The WG should approach the promoters to avoid confusion and duplication of efforts.

Joint African statistical databases and publications

The ASCC noted with satisfaction the progress made toward joint publications. The African Statistical Yearbook 2009 (ASYB) has been published. Other publications such the MDGs Progress Report are published jointly by the AfDB, the UNECA, and the AUC. The joint database is a long-term project that is under discussion.

Consolidated African statistical program

The ASCC encouraged concerned institutions to provide their 2010 work program and planned meetings (seminars and workshops) to the AfDB by mid-December 2010. The AfDB will consolidate them and propose a Consolidated African Statistical Program to be reviewed by all ASCC members.

III. Reports on the AGNA (StatCom-Africa's Working Group)

The African Group on National Accounts (AGNA), chaired by the African Development Bank, was set up during the first meeting of the Statistical Commission for Africa (StatCom-Africa), which was held from January 21-23, 2008 in Addis Ababa. Its main task is to develop a strategy for the

implementation of 2008 SNA. In order to facilitate the activities of this WG, the AfDB has to date organized five meetings of the AGNA: in Lusaka (April 2008), Tunis (July 2008), Durban (August 2009), Addis Ababa (October 2009), and Dakar (November 2009). In addition, the WG undertook different activities toward designing a strategy and an action plan for the revised 2008 SNA implementation in African countries; this comprises five strategic elements, with a view to optimizing the compilation of National Accounts on the continent. The timeframe of the African Strategy for the Implementation of the 2008 SNA is estimated at 5 years, starting in January 2010 until 2014.

After a presentation of the strategy by representatives of the AGNA, the ASCC congratulated the Working Group for a job well done and recommended the strategy be presented to the heads of NSOs at the East London meeting before its adoption during the StatCom-Africa in January 2010. The ASCC recommended this strategy to be advocated at a highest level by the AUC. The ASCC commended the AfDB for its support to the AGNA activities and recommended that the AfDB supports also the other working groups.

IV. African Regional Integration Statistics

The AUC informed the ASCC that it has embarked on the elaboration of a "Continental Strategy for the Harmonization of Regional Integration Statistics," to endow the African Statistical System with a mechanism to produce harmonized African statistics in support of the regional integration process. Among others, the study has three major objectives: (i) to identify priority areas of the African integration process; (ii) to assess statistical functions within Regional Economic Communities (RECs) as well as to review different initiatives for statistical harmonization both at regional and continental levels; and (iii) to define a global strategy at the continental level involving all actors with a view to progressive and total harmonization of statistics relating to the areas of integration, and the definition of an appropriate coordination mechanism for data requests sent to countries by international organizations. A Peer Review Meeting was organized in Addis Ababa in October 2009, to examine the draft strategy. Another meeting will be organized in East London, South Africa, in December 2009. The ASCC welcomed this initiative and committed to fully support the AUC in the endeavor.

V. Other Coordination Issues

The meeting followed up on the output of the two working groups created last year, focusing on price statistics and financial statistics.

After a presentation by the AUC on a project to create an African Fund for Statistics, the ASCC recommended that this Fund be hosted by the AfDB, due its comparative advantages as a financial institution. The AfDB will share in the East London meeting the study prepared last year on a similar project.

Finally it was decided that the following RECs should become members of the ASCC: AMU, ECCAS, ECOWAS, COMESA, and SADC.

Résolution de East London : Réunion du comité des directeurs généraux des instituts nationaux de statistique africains

Nous, directeurs généraux des Instituts nationaux de statistique des pays africains, réunis, sous la direction de l'Union africaine (UA), de la Banque africaine de développement (BAD), de la Commission économique des Nations Unies pour l'Afrique (CEA), pour discuter de *l'harmonisation et du Renforcement des capacités statistiques en Afrique* à East London (Afrique du Sud), du 10 au 12 décembre 2009 :

CONSIDÉRANT la décision des chefs d'Etat et de Gouvernement de l'Union africaine visant à accélérer le processus d'intégration de l'Afrique par la mise en œuvre de tous leurs engagements et programmes en vue de promouvoir le développement politique, économique, social, culturel et autoentretenu, ainsi que l'intégration des économies africaines ;

CONSIDÉRANT ÉGALEMENT que l'information statistique est essentielle à la prise de décision par les diverses composantes de la société, et en particulier celles les décideurs politiques, des acteurs économiques et sociaux, et qu'elle est par conséquent indispensable pour l'intégration et le développement durable du continent, ainsi que pour le suivi des progrès réalisés dans la réalisation des Objectifs du Millénaire pour le développement (OMD) ;

CONSCIENTS DU FAIT que tous les engagements pour mettre en œuvre les programmes de développement et combattre la pauvreté devraient être basés sur des faits réels qui requièrent par conséquent un système de statistique solide, pertinent, fiable, complet, harmonisé et sensible ;

PRENANT NOTE de la nécessité d'établir un lien entre la production des statistiques, le programme et le calendrier d'intégration ainsi que toutes les priorités continentales et régionales ;

RECONNAISSANT le fait que les statistiques africaines doivent appuyer les politiques nationales ainsi que les politiques sur lesquelles les pays africains se sont engagées aux niveaux régional, continental et international ;

NOTANT que la qualité et la fiabilité des statistiques africaines mises à la disposition du public et des décideurs politiques reposent sur une collabora-

tion efficace entre les fournisseurs, les producteurs, les utilisateurs de données statistiques et les institutions régionales, panafricaines et internationales ;

NOTANT EN OUTRE la nécessité de l'implication continue et exhaustive de toutes les composantes de la gouvernance, à savoir les gouvernements, les parlements, la société civile et le secteur privé ;

RECONNAISSANT que les défis émergents tels que les changements climatiques, la crise alimentaire, la crise financière, la crise énergétique, etc. imposent une nouvelle dimension et une nouvelle manière de penser le développement statistique en Afrique et dans le monde ;

RAPPELANT la décision prise par les chefs d'État et de Gouvernement de l'Union africaine à Addis-Abeba (Ethiopie) en janvier 2009, portant adoption de la Charte africaine de la statistique et invitant les États parties à en accélérer le processus de signature et de ratification pour permettre à la Charte d'entrer en vigueur en tant que cadre réglementaire pour l'harmonisation et le développement de la statistique en Afrique ;

PRENANT NOTE des progrès accomplis par les pays africains dans l'élaboration de statistiques conformément aux campagnes que mènent la Commission de l'Union africaine (CUA), la BAD et la CEA à travers les initiatives telles que la Commission des statistiques pour l'Afrique (StatCom-Afrique) et le Symposium africain sur le développement statistique (ASSD) ;

PRENANT NOTE EN OUTRE de la décision N° EX.CL/DEC.504 (XV) Rev1. prise par le Conseil exécutif de l'Union africaine en juillet 2009, à Syrte, (Libye), sur le rapport de la seconde Conférence conjointe UA / CEA des ministres africains des Finances, de la Planification et du Développement économique, décision donnant mandat à la Commission de l'UA pour travailler en étroite collaboration avec les Etats membres, les Communautés économiques régionales (CER), la BAD, la CEA et la Fondation pour le renforcement des capacités en Afrique (ACBF) en vue d'élaborer une stratégie pour l'harmonisation des statistiques en Afrique, finaliser l'étude technique sur la création d'un Fonds pour le développement de la statistique et préparer une évaluation à mi-parcours des progrès accomplis dans la réalisation des OMD dans notre continent ;

PRENONS LA RESOLUTION DE:

Sur la Charte africaine de la statistique

- Réitérer notre attachement à la Charte africaine des statistiques et nous engageons à prendre toutes les mesures nécessaires pour la faire signer et la ratifier dans les plus brefs délais ;

Sur la stratégie africaine pour l'harmonisation des statistiques de l'intégration africaine

- Adopter la stratégie africaine pour l'harmonisation statistique de l'intégration africaine ;
- Adopter la Stratégie africaine pour la mise en œuvre du Système de comptabilité nationale 2008 (SCN 2008) qui constitue le premier grand pilier de la Stratégie africaine pour l'harmonisation des statistiques de l'intégration africaine et de la création d'un Centre de formation supérieure en comptabilité nationale ;
- Donner mandat à la Commission de l'UA pour présenter la Stratégie africaine pour l'harmonisation des statistiques de l'intégration africaine et la Stratégie africaine pour la mise en œuvre du Système de comptabilité nationale 2008 (SCN 2008) à la prochaine Conférence conjointe de la Commission de l'UA des ministres africains de l'Economie et des Finances et à la Conférence de la CEA des Ministres des Finances, de la planification et du développement économique qui se tiendront au Malawi en mars 2010, pour qu'elles soient recommandées à l'adoption et la prise de décision des Chefs d'Etat et de Gouvernement, au cours du juillet 2010 ;
- Lancer un appel à tous les gouvernements africains pour qu'ils mettent les ressources adéquates et durables à la disposition des INS pour leur permettre de produire les statistiques nécessaires aux politiques nationales et à l'intégration africaine ;
- Nous engager à nous approprier la Stratégie africaine pour l'harmonisation statistique de l'intégration africaine et nous assurer qu'elle est reflétée dans nos stratégies nationales de développement des statistiques (SNDS) ;

Sur la création d'une interaction entre les statisticiens et les décideurs

- Donner mandat à la Commission de l'UA, en collaboration avec la BAD et la CEA, pour créer une plate-forme entre le Système statistique africain et les différentes composantes de la gouvernance, la toute première tâche devant être la convocation d'une réunion avec les parlementaires africains sur la Charte africaine des statistiques, la Stratégie africaine pour l'harmonisation statistique de l'intégration africaine, les changements climatiques et les OMD ;

Sur le renforcement des capacités statistiques

- Fournir les efforts visant à intégrer les initiatives de renforcement des capacités statistiques sur le continent africain.

FAIT à East London, Afrique du Sud, le 12 décembre 2009.

East London Resolution of the Meeting of the Committee of Directors-General of African National Statistical Offices

We, heads of National Statistical Offices (NSOs) of African countries, under the leadership of the African Union (AU), African Development Bank (AfDB), United Nations Economic Commission for Africa (UNECA), meeting under the theme *African Statistical Harmonization and Capacity Building*, in East London, South Africa, on 10-12 December 2009:

CONSIDERING the decision of the Heads of State and Government of the African Union for accelerating Africa's integration process by implementing all their commitments and programmes that aim to promote political, economic, social, cultural and self-sustained development, as well as integration of African economies;

CONSIDERING ALSO that statistical information is vital for decision-making by all components of the society, particularly policy makers as well as economic and social players, and is therefore essential for the continent's integration and sustainable development as well as tracking progress towards the attainment of Millennium Development Goals (MDGs);

AWARE OF THE FACT that all commitments to implement development programmes and combat poverty should be based on clear evidence and therefore require a robust statistical data system which is relevant, reliable, comprehensive, harmonized and responsive;

NOTING the necessity to link the production of statistics to the integration agenda and timelines as well as all continental and regional priorities;

RECOGNISING THE FACT THAT African statistics should serve national policies as well as policies that African countries have committed themselves to at regional, continental and international levels;

NOTING that the quality and accuracy of the African statistics available to the public and to policy makers depends on effective collaboration between statistical data suppliers, producers and users as well as regional, pan-African and international institutions;

NOTING FURTHER a need for continued and exhaustive engagement with all governance segments, namely, governments, parliaments, civil society and private sector.

RECOGNIZING the emerging challenges such as climate change, food crisis, financial crisis, energy crisis, etc. impose a new dimension and a new way of thinking about statistical development in Africa and the world;

RECALLING the Decision taken by the Heads of State and Government of the African Union in Addis Ababa, Ethiopia, in January 2009 adopting the African Charter on Statistics and calling upon State Parties to expedite the signature and ratification process in order for the Charter to enter into force as regulatory framework for statistics harmonization and development in Africa;

NOTING the progress that is being made by African countries in the development of statistics as championed by the African Union Commission (AUC), AfDB and UNECA under initiatives such as the Statistical Commission for Africa (StatCom-Africa) and the Africa Symposium on Statistical Development (ASSD);

NOTING FURTHER the Decision No EX.CL/DEC.504 (XV) Rev1. taken by the Executive Council of the African Union in July 2009, in Sirte, Libya, on the Report of the second AU/UNECA Joint Conference of African Ministers of Finance, Planning and Economic Development, mandating the AUC to work in close collaboration with Member States, Regional Economic Communities (RECs), AfDB, UNECA and African Capacity Building Foundation (ACBF) to develop a strategy for integration, statistics harmonization and finalize the technical study on the creation of a Statistical Fund as well as preparing a mid-term review on progress made towards the attainment of MDGs on our continent;

RESOLVED TO:

On the African Charter on Statistics

- Reiterate our commitment to the African Charter on Statistics and commit to take active steps to have it signed and ratified as expeditiously as possible;

On the Africa Strategy for Statistical Harmonization for African Integration

- Adopt the Africa Strategy for Statistical Harmonization for African Integration;
- Adopt the African Strategy for the Implementation of the 2008 System of National Account (SNA 2008) which constitutes the first key pillar of the African Strategy for Statistical Harmonization for African Integration and the establishment of a Post-Graduate Training Center on National Accounts;
- Mandate AUC to submit the African Strategy for Statistical Harmonization for African Integration and the African Strategy for the Implementation of the 2008 System of National Account (SNA 2008) to the next Joint AUC Conference of African Ministers of Economy and Finance and UNECA Conference of Ministers of Finance, Planning and Economic Development to be held in Malawi in March 2010 for further recommendations to the Summit of Heads of State and Government for adoption and decision taking in July 2010;
- Call on all African governments to avail adequate and sustainable resources to NSOs to enable them to produce statistics required for national policies and for African integration;
- Commit ourselves to take ownership of the African Strategy for Statistical Harmonization for African Integration and to ensure that it is reflected in our National Strategies for the Development of Statistics (NSDS);

On creating an interaction between statisticians and policy-makers

- Mandate the AUC, in collaboration with the AfDB and UNECA, to create a platform between the African Statistical System and the different governance segments, first of which should be convening a meeting with African parliamentarians on African Charter on Statistics, African Strategy for Statistical Harmonization for African Integration, climate change and MDGs;

On statistical capacity building

- Come up with efforts to integrate statistical capacity building initiatives on the African continent.

EFFECTED in East London, South Africa, on 12 December 2009.

Second Session of the Statistical Commission for Africa (StatCom-Africa II) – Conclusions and Recommendations

Addis Ababa, Ethiopia, January 18-21, 2010

Preamble

Acknowledging that the theme of this second session of StatCom-Africa "Harnessing Financial and Technical Resources in Support of the Monitoring of the Millennium Development Goals in African Countries" is timely as we are preparing for the second five-year-review of the MDGs;

Considering reports from seven working groups namely on (i) Data Management, (ii) Development Indicators, (iii) Gender Statistics, (iv) Informal Sector, (v) National Accounts, (vi) Statistical Training, and (vii) Harmonisation of Statistics in Support of Economic Integration;

Considering other ongoing and emerging issues namely: (i) the 2010 Round of Population and Housing Censuses, (ii) Civil Registration and Vital Statistics, (iii) Improving Agricultural Statistics, (iv) Environment Statistics and Climate change, (v) Coordination of Advisory Services to member States;

Having examined reports on (i) the implementation of the Reference Regional Strategic Framework for Statistical Capacity Building in Africa (RRSF) and, (ii) the African Charter for Statistics;

Taking cognizance of a number of programmatic issues including (i) the implementation of the 2008-2009 and the planned 2010-2011 work programmes of the African Centre for Statistics (ACS);

Noting with satisfaction the participation of young statisticians and universities in the discussion of the Commission for the first time;

The second session of the Statistical Commission for Africa concluded and recommended as follows:

A. ON DATA MANAGEMENT

The Statistical Commission for Africa:

a. **Urges ECA** to:

i. organize an expert group to undertake a comprehensive study of the major data management software tools to draw up a comparative table of their features to enable National Statistical Offices (NSOs) to select the appropriate ones for drawing up a composite optimum set of features for an African software system for data management and dissemination to be developed as an open source software under the auspices and full ownership of the African Statistics fraternity;
ii. undertake a survey of database systems and software tools available in countries to determine what databases and systems (if any) they are using with a view to identifying those that need assistance to implement databases;
iii. produce a handbook on data management including guidelines on data management policies and software for analysis, dissemination, archiving, etc.;
iv. involve its sub-regional offices in making requests/recommendations to member States because these sub-regional offices are closer to the countries and will have to harmonise information coming from those member States;
v. involve and collaborate with the regional economic communities (RECs) and other regional institutions in the process of data collection and database updating;

b. **Encourages** professional Statistical Associations to participate in meetings, access and use available data with a view to offering comments and identifying errors, mistakes, and best practices for data management;

c. **Recommends** that the capacity of NSOs be enhanced in the use of geospatial tools to reinforce the combination of geographic and statistical data in development planning and monitoring;

d. **Encourages** ECA to prepare a work programme, in collaboration with United Nations Group of Experts on Geographical Names (UNGEGN) that will lead to authoritative national databases, which will be consolidated into a continental database and feed into the global geographical names database.

B. ON DEVELOPMENT INDICATORS

The Statistical Commission for Africa:

a. **Recommends that:**

i. a study be conducted to ascertain which African countries may not meet some MDG targets, and why they may not meet those MDG targets;
ii. countries, and regional and international organizations encourage and support research and study initiatives on methodologies that would allow to identify development objectives for African countries as well as indicators to measure them beyond 2015;

b. **Calls on** the ECA and its partners to develop a Handbook on the derivation of development indicators from population censuses and household surveys;

c. **Urges** African countries to build national databases which include MDG indicators; set up national focal points and Coordination Committees on MDGs; and promote better involvement of national political authorities in MDG monitoring.

C. ON GENDER STATISTICS

The Statistical Commission for Africa:

a. **Recommends:**

i. the preparation of an African handbook on gender statistics;
ii. the preparation of training material on gender statistics;
iii. the speedy completion of the Compendium on gender statistics; and
iv. the formulation of a regional strategy on gender statistics;

b. **Calls for** the mainstreaming of gender into the themes of all working Groups, and the identification of gender-responsive indicators to monitor Poverty Reduction Strategy Programmes (PRSPs.)

D. ON THE INFORMAL SECTOR

The Statistical Commission for Africa:

a. **Adopts** the Action Plan of the Working Group for improving and monitoring informal sector statistics in Africa including the proposed terms of reference for the period 2010–2012;

b. **Re-affirms** the importance of ECA collaborating with other UN regional commissions on the measurement of the informal economy and urges them to implement an inter-continental project on measuring the informal economy;

c. **Recommends** that identified pilot countries finalize their project on information sector surveys and conduct, in collaboration with AFRISTAT, ECA and ILO, regular employment surveys, preferably on an annual basis.

E. ON NATIONAL ACCOUNTS

The Statistical Commission for Africa:

a. **Adopts** the African strategy for the implementation of the 2008 System of National Accounts (SNA); and

b. **Makes it** the first pillar of the African Strategy for the Harmonization of Statistics in Africa (SHaSA);

c. **Urges** pan-African institutions, regional economic communities and member States to take ownership of the strategy and mobilize the necessary funds for its implementation; and urges Africa's development partners to support its implementation.

F. ON STATISTICAL TRAINING

The Statistical Commission for Africa:

a. **Recommends** that ECA coordinates the various initiatives and establish a programme on statistical training;

b. **Endorses** the inception of the African Group on Statistical Training and Human Resources (AGROST);

c. **Advocates** for the establishment of a permanent secretariat for AGROST at ECA;

d. **Requests** the African Statistical Coordination Committee to explore funding modalities for the permanent secretariat.

G. ON HARMONIZATION OF STATISTICS IN SUPPORT OF ECONOMIC INTEGRATION

The Statistical Commission for Africa:

a. **Recognizes** the important work of AUC to develop SHaSA, especially advocacy and its strong policy commitment at all levels;

b. **Recommends** the establishment of 12 working groups on thematic areas of harmonization, and the assignment of a lead country or institution for each thematic group;

c. **Proposes** that the African Statistical Coordination Committee rationalizes the programmed meetings of heads of NSOs.

H. ON EMERGING AND OTHER ISSUES

Civil Registration and Vital Statistics
The Statistical Commission for Africa:

a. **Recognizes the importance** of Civil Registration and Vital Statistics (CRVS) and the gaps in National Statistical Systems (NSSs) due to the inadequacy of civil registration systems in most African countries;

b. **Recognizes** the limited awareness on, and lack of political support for CRVS among governments and political leaders; and

c. **Supports** the organization of a high-level Ministerial Conference on CRVS.

Agricultural Statistics
The Statistical Commission for Africa:

a. **Recognizes** the crucial role of agricultural statistics in dealing with the challenges of food security and climate change, and observes that the current arrangement for collecting and managing them are inadequate for these crucial roles;

b. **Supports** the Global Strategy to Improve Agricultural And Rural Statistics;

c. **Recommends** that NSDS include provisions for producing agricultural statistics, with technical expertise and funding in the area of agriculture statistics being channelled through the appropriate country institutions that have the infrastructure for producing them.

Environmental Statistics and Climate Change
The Statistical Commission for Africa:

a. **Notes with appreciation** the work undertaken in the area of environment statistics by the United Nations Environment Programme (UNEP), ECA and UNSD to adapt the core list of indicators to the Economic Community of West African States (ECOWAS) subregion for use in national and regional reporting;

b. **Further appreciates** the Biodiversity Indicators Capacity Strengthening workshops organized for selected East and Southern African countries by ECA and UNEP – World Conservation and Monitoring Centre;

c. **Reiterates** the recommendation of StatCom-Africa I calling on ECA to establish a task force on environment statistics to:

    i. Streamline the core list of environmental indicators for Africa;
    ii. Prepare a work programme on environment statistics for the Africa region; and
    iii. Work closely with the Inter Sectoral Working Group on Environment Statistics;

d. **Requests ECA** to provide technical assistance to enable countries to develop their environment statistics and indicators and to continue organizing training workshops in the area of environment statistics and accounts;

e. **Recommends** the strengthening of Statistical Training Centres to ensure that they include environment statistics in their curricula.

Poverty Reduction
The Statistical Commission for Africa:

a. **Adopts** the recommendations of the PRSP group:

    i. To improve collaboration between statisticians and planners, including capacity building activities and use of modern software

ii. To include among its priority the statistics on social protection which are the pillar of PRSP;

b. **Acknowledges** the need to develop sectoral statistics in social fields, as well as for infrastructures;

Geography and Statistics
The Statistical Commission for Africa:

a. **Endorses** the view that geographic information is essential and should be incorporated more into statistical processes and:

   i. **calls on** African countries to ensure that national statistical, planning and cartographic authorities have effective collaboration between them in the development of respective data infrastructures and systems by having National Statistics and Planning Offices represented in the coordination arrangements for the National Spatial Data Infrastructures (NSDI) and National Mapping and Planning Agencies represented in the coordination of National Strategies for the Development of Statistics; and
   ii. **requests ECA**, with technical support from UNSD and the two United Nations Cartographic Conferences to organize a parallel session for cartographers and geoinformation practitioners.

Social Inclusion and Protection
The Statistical Commission for Africa:

a. **Recognizes** the importance of social inclusion/protection as an important strategy in the fight against poverty, and for safeguarding Africa's gains in the MDGs in the face of the global financial crisis (particularly rising food and oil prices), and climate change;

b. **Recommends** that the African Statistical Coordinating Committee (ASCC) undertakes the necessary conceptual/technical review and establish modalities for setting up a working group on social inclusion/protection and report back on progress made to the next session of StatCom-Africa.

Coordination
The Statistical Commission for Africa:

a. **Recommends** that pan African institutions in charge of statistics improve on the coordination of their respective activities;

b. **Acknowledges** that statutory statistical councils/boards constitute major stakeholders in some NSDSs of member States and that these organizations are important players in the enhancement of the development of and coordination of the NSDSs in the countries and agrees that they be represented in StatCom-Africa and other strategic formations in pursuit of statistical excellence in Africa;

c. **Recognizes** the opportunities provided by strategic initiatives within member States such as the 2010 round of population censuses in creating the basis for the implementation of other programmes such as the civil registration, collection of agricultural statistics, employment statistics, etc, and agrees that coordination efforts be mounted in order to create optimum comprehensive statistical programmes in respective states;

d. **Endorses** the view that Africa's input into the work of the sessions of the United Nations Statistical Commission should be coordinated during forthcoming sessions of StatCom-Africa.

Conférence sur « Une Stratégie Mondiale pour l'Amélioration des Statistiques Agricoles : Plan de mise en œuvre pour l'Afrique »

Hammamet, Tunisie, 3-4 février 2010

Actuellement, il y a une proposition d'une Stratégie Mondiale pour l'Amélioration des Statistiques Agricoles et Rurales qui est en cours de développement sous les auspices de la Commission des Statistiques des Nations Unies. Le plan de mise en exécution de cette Stratégie Mondiale implique le développement d'un programme exhaustif de renforcement des capacités en Statistiques Agricoles pour les pays africains. Ceci est en parfaite harmonie avec les objectifs du Plan d'Action de Marrakech pour les Statistiques (MAPS), le Cadre Stratégique de Référence Régional (RRSF) pour le Développement de la Statistique en Afrique et la Stratégie Nationale pour le Développement de la Statistiques (SNDS). Il a l'objectif de fournir une vision pour les systèmes de statistiques nationales et internationales pour la production des informations et données de base permettant de guider la démarche de prise de décision requise pour le 21ème Siècle.

Cette Stratégie Mondiale est ancrée sur :

- Le consentement pour les pays à produire un minimum de données fondamentales répondant à la demande émergeante et que tout le monde promet de fournir annuellement ;
- L'agriculture est intégrée dans les système nationaux de statistiques afin de répondre aux attentes des décideurs et autres utilisateurs de données ; et ces données sont comparables entre les pays et dans la durée ; et
- L'intégration est atteinte par un consentement sur la méthodologie qui inclut le développement d'un Plan d'Echantillonnage Maître pour l'agriculture, la mise en exécution d'un Cadre Intégré d'Enquêtes, et avec des résultats disponibles dans une Base de Données Intégrée.

Une Conférence des Partenaires aux Développement sur la Stratégie Mondiale pour l'Amélioration des Statistiques Agricoles : Plan d'Exécution pour l'Afrique s'est tenue en Tunisie, du 3 au 4 février 2010. La réunion était conjointement organisée par le Groupe de la Banque Africaine de Développement (BAD), l'Organisation des Nations Unies pour l'Alimentation et l'Agriculture (FAO) et la Fondation Bill et Melinda Gates (BMGF). Les objectifs clés de la réunion étaient : (i) de finaliser des indications pour le développement d'une proposition de renforcement des capacités pour l'amélioration des statistiques agricoles en Afrique ; (ii) d'établir une coalition de

donateurs pour constituer des fonds conséquents pour le plan d'action ; (iii) et de s'accorder sur la suite des travaux, les prochaines étapes et le calendrier.

La Conférence a réuni 47 participants venant de l'intérieur comme de l'extérieur de l'Afrique et composés des Partenaires aux Développement, des Centres de Formation, des Institutions sous-régionales et régionales, les Bureaux Nationaux de Statistiques et des Services de Statistiques Agricoles.

Une réunion Post-Conférence de quatre institutions (BAD, FAO, BMGF et CEA des Nations Unies) et les Amis du Bureau de la Commission Statistique des Nations Unies (Ethiopie, Maroc et Ouganda) s'est tenue séparément, a discuter sur la suite des travaux et a recommandé la composition des membres des équipes qui devront préparer la proposition pour chaque composante de la Stratégie Mondiale ainsi que le choix des agences responsables. Cette réunion a aussi arrêté le calendrier et les différents événements.

Il a été convenu que le Plan d'Exécution pour l'Afrique sera développé sous une supervision générale d'une Structure de Gouvernance qui sera définie à travers la Composante «Gouvernance». Le Plan d'Exécution comprendra trois composantes techniques et une Structure Globale. Chaque composante technique sera développée comme une composante indépendante mais entièrement intégrée et complémentaire aux autres composantes techniques sous la Structure Globale. Le rôle de la Structure Globale sera d'assurer la consistance et la cohérence des composantes techniques ainsi que le lien avec le Plan d'Exécution au niveau mondial.

Les composantes seront sous la responsabilité des institutions suivantes :

- Composante Assistance Technique : BAD
- Composante Formation : CEA
- Composante Recherche : FAO
- Structure Globale : FAO
- Composante Gouvernance : BAD

Les autres principales recommandations de la Conférence incluent :

- l'établissement d'un Trust Fund (TF) s'étalant sur cinq ans pour appuyer le Plan d'Exécution du plan pour l'Afrique (ce plan aura une perspective de long terme, de 10 à 15 ans, mais suivra une approche par phase, avec une première phase couvrant les 5 prochaines années (2011-2015) ;
- le TF devra être abrité par la BAD ;

- le TF devra être utilisé pour l'appui des pays en terme d'Assistance Technique dans la mise en exécution des Stratégies Nationales de Développement de la Statistique (SNDS), spécialement pour la composante agriculture, et pas pour conduire des enquêtes ou recensements ;
- il devra y avoir un plaidoyer pour la mobilisation des ressources pour la mise en exécution du plan ; et
- le Comité de Coordination Statistique en Afrique (ASCC), composé de la BAD, la CEA, la Commission de l'Union Africaine, la Fondation de Renforcement des Capacités en Afrique (ACBF) et la Communauté Economique Régionale (RECs), et la FAO devront participer dans la préparation de la proposition pour la mise en exécution du plan ;
- trois organisations prendront la responsabilité pour la mise en exécution du programme : (i) la BAD sera en charge de la mobilisation et du décaissement des ressources et fournira de l'Assistance Technique aux pays ; (ii) la CEA s'occupera de la formation ; (iii) tandis que la FAO (en tant que secrétariat de la Structure Globale) assurera la coordination du programme régional avec les structures de la mise en exécution de la Stratégie Mondiale, fera beaucoup plus le travail technique dont la recherche et le développement des méthodologies ainsi que la production des manuels techniques et guides tout en donnant de l'Assistance Technique aux pays.
- Etc.

Pour ce qui est de la Gouvernance et des trois composantes de la mise en exécution du plan de la Stratégie Globale en Afrique, les équipes suivantes ont été formés et convenues pour la préparation des propositions.

Fonction	**Institution Responsable**	**Membres**
Assistance Technique	BAD	1. Ouganda 2. Consultants responsables 3. AFRISTAT
Formation	CEA	1. Maroc 2. Consultant responsable 3. EASTC 4. Institut de Statistiques et d'Economie Appliquée (Université de Makerere) 5. ENSEA, Abidjan 6. BAD
Recherche	FAO	1. Ethiopie 2. Université de Bologna 3. Tunisie 4. Regional Centre for Mapping of Resources for Development (RCMRD) 5. EC Joint Research Center
Structure Globale	FAO	1. AfDB 2. ECA 3. FAO 4. MGF
Mecha-nisme de Gouver-nance	ASCC	1. CEA 2. CUA 3. ACBF 4. Ethiopie, Ouganda, Maroc 5. FAO 6. Consultants responsables

Conference on "A Global Strategy for Improving Agricultural Statistics: Implementation Plan for Africa"

Hammamet, Tunisia, February 3-4, 2010

There is currently a proposal on a Global Strategy to Improve Agricultural and Rural Statistics, which is being developed under the auspices of the United Nations Statistical Commission (UNSC). The Implementation Plan for Africa of this Global Strategy involves the development of a comprehensive capacity-building program in Agricultural Statistics for African countries. This is fully in line with the objectives of the Marrakech Action Plan for Statistics (MAPS), the Regional Reference Strategic Framework (RRSF) for Statistical Capacity Building in Africa, and National Strategies for the Development of Statistics (NSDS) in seeking to provide a vision for national and international statistical systems that will produce basic data and information to guide decision-making processes in the 21st century.

The Global Strategy is anchored on:

- Countries agreeing upon a minimum set of core data that meet the emerging demands, and that all will pledge to provide annually;
- Agriculture being integrated into the National Statistical Systems in order to meet policy-maker and other data user expectations; and that the data will be comparable across countries and over time; and
- The integration being achieved by an agreed upon suite of methodology that includes the development of a Master Sample Frame for Agriculture, the implementation of an Integrated Survey Framework, and with the results available in an Integrated Data Base.

A Conference of Development Partners on "The Global Strategy for Improving Agricultural Statistics: Implementation Plan for Africa" took place in Tunisia, from February 3-4, 2010. The meeting was co-organized by the African Development Bank Group (AfDB), the Food and Agriculture Organization of the United Nations (FAO), and the Bill and Melinda Gates Foundation (BMGF). The key objectives of the meeting were to: (i) finalize guidelines for developing a capacity-building proposal for improving agricultural statistics in African countries; (ii) build a donor coalition for raising subsequent funding for the action plan; (iii) and agree on the way forward, next steps and calendar.

The Conference was attended by 47 participants composed of key development partners, training institutions, subregional and regional institutions,

National Statistical Offices, and Agricultural Statistical Services within and outside Africa.

A post-conference meeting of four agencies (AfDB, FAO, BMGF, and UNECA) and the Friends of the Chair (FoC) in the UNSC (Ethiopia, Morocco, and Uganda) was held separately and discussed the way forward and recommended the membership of the teams to prepare the proposal for each component of the Global Strategy as well as the choice of the lead agencies. The meeting of agencies and Friends of the Chair in the UNSC (FoC) also decided on the calendar of events.

It was agreed that the Implementation Plan for Africa would be developed under the overall supervision of a governance structure to be defined through the governance component. The Implementation Plan will comprise three technical components and an umbrella framework. Each technical component will be developed as a stand-alone component but fully integrated and complementary to the other technical components under the umbrella framework. The role of the umbrella framework will be to ensure consistency and coherence of the technical components as well as linkages with the overall Implementation Plan of the Global Strategy.

The components will be led by the following Organizations:

- Technical Assistance Component: African Development Bank Group (AfDB)
- Training Component: United Nations Economic Commission for Africa (UNECA)
- Research Component: FAO (Food and Agriculture Organization)
- Umbrella Framework: FAO (Food and Agriculture Organization)
- Governance Component: African Development Bank Group (AfDB).

Other main recommendations of the Conference included:

- Establishment of a five-year Trust Fund (TF) to support the Implementation Plan for Africa (the plan will have a long-term perspective of 10 to 15 years, but will follow a phased approach, with the first phase covering the next 5 years (2011-2015));
- The TF should be housed in the AfDB;
- The TF should be used to support countries in terms of Technical Assistance (TA) in National Strategy for the Development of Statistics (NSDS) implementation, especially for the agricultural component, and not to conduct surveys or censuses;

- There should be advocacy for resource mobilization for the Implementation Plan;
- The African Statistical Coordination Committee (ASCC) (composed of the AfDB, UNECA, AUC, African Capacity Building Foundation (ACBF), and RECs) and the Food and Agricuture Organization (FAO) should participate in the preparation of the proposal for the Implementation Plan;
- Three organizations would take the lead in the Implementation Program: (i) the AfDB would take the lead in mobilization and disbursement of resources, providing coordination and TA to countries; (ii) the UNECA would be responsible for training; (iii) the FAO (as the Secretariat of the umbrella framework) would coordinate the regional program with the Global Strategy implementation structures, do much of the technical work including research and development of methodologies and production of technical manuals and guides, as well as providing TA to countries.

With regard to governance and the three components of the Implementation Plan for the Global Strategy in Africa. the following working teams were formed and agreed to draft the proposals.

Function	Lead agency	Team members
Technical Assistance Component	AfDB	1. Uganda 2. Lead consultants 3. AFRISTAT
Training Component	UNECA	1. Morocco 2. Lead consultant 3. EASTC 4. Institute of Statistics and Applied Economics (Makerere University) 5. ENSEA, Abidjan 6. AfDB
Research Component	FAO	1. Ethiopia 2. University of Bologna 3. Tunisia 4. Regional Center for Mapping of Resources for Development (RCMRD) 5. EC Joint Research Center
Umbrella Framework	FAO	1. AfDB 2. ECA 3. FAO 4. BMGf
Governance Mechanism	ASCC	1. ECA 2. AUC 3. ACBF 4. Ethiopia, Uganda, Morocco 5. FAO 6. Lead consultants

Upcoming Events / Evénements en vue – 2010

Date	Venue / Lieu	Title / Titre	Organisers / Organisateurs
MAI / MAY			
4-6	Helsinki, Finland	European Conference on Quality in Official Statistics / Conférence Européenne sur la qualité des statistiques officielles	Statistics Finland, EUROSTAT
24-25	Washington, USA	Third Global Conference on Agricultural and Rural Household Statistics / Troisième Conférence Mondiale sur les statistiques agricoles et les ménages ruraux	FAO, OECD
JUIN / JUNE			
14-25	Nairobi, Kenya	Launching Regional Workshop of the 2011 round of the International Comparison Program for Africa / Atelier regional de lancement du cycle 2010 du Programme International de Comparaison pour l'Afrique	AfDB
OCTOBRE / OCTOBER			
13-15	Kampala, Uganda	Fifth International Conference on Agriculture Statistics / Cinquième Conférence Internationale sur les Statistiques Agricoles	FAO, AfDB, World Bank, EUROSTAT, UNSD, Uganda
20	New York, USA	World Statistics Day / Journée mondiale de la statistique	UNSD
20-22	Santiago, Chile	International Association for Official Statistics (IAOS) Conference / Conférence de l'Association Internationales pour les Statistiques Officielles	ISI, IAOS, INE-Chile
20-22	Washington, USA	Third ICP 2011 Regional Coordinators' Meetings / Troisième réunion des coordinateurs régionaux du PCI 2011	World Bank
NOVEMBRE / NOVEMBER			
18	All countries / Tous les pays	African Statistics Day / Journée africaine de la statistique	NSOs / INS

Temporary Relocation Agency (TRA) – Agence Temporaire de Relocalisation (ATR)
13 Avenue du Ghana
BP. 323, 1002 Tunis Belvedère
Tunis, Tunisia / Tunisie
Tel: (+216) 71 103 216
Fax: (+216) 71 103 743
Email: statistics@afdb.org / Internet: http://www.afdb.org