



BIOMETRY IN AGRICULTURE RESEARCH  
FOR DEVELOPING COUNTRIES

María Cristina Amézquita  
Biometrician //

CIAT

"Biometry is a science that studies the application of mathematical, statistical and computing methods, to the description and analysis of biological phenomena. It is, in summary, the application of the quantitative thought for a better understanding of a live process, random by nature".

Prof. S.C. Pearce, 1983

A biometrician is a statistician, a mathematician, or a statistical biologist who besides possessing good knowledge on quantitative techniques applicable to the study of live processes, has also a clear understanding of the biological problem under concern, and a very good communication capacity.

The author

The success of a support discipline to research, such as Biometry, lies very heavily on a permanent and fruitful communication between the researcher and the biometrician, based on mutual professional respect. It is the only way through which both lines of thought can converge and fully produce for the benefit of agriculture research.

The author

BIOMETRY IN AGRICULTURE RESEARCH  
FOR DEVELOPING COUNTRIES

- CONTENT -

	Page #
1. INTRODUCTION	1
2. THE ROLE OF BIOMETRY IN SUPPORT TO AGRICULTURE RESEARCH	2
2.1 Historic evolution	2
2.2 Its role today	8
3. IS BIOMETRY IN AGRICULTURE RESEARCH INSTITUTIONS SERVING THE DEVELOPING WORLD, COPPING WITH ITS PROPER ROLE? EXISTING CONSTRAINTS AND POSSIBLE IMPROVEMENTS.	10
3.1 Organization of research institutes. Organization of biometry/research data analysis groups.	11
3.2 Research policy and priorities. Responsibilities of the biometrician.	13
3.3 Are administrators of research clear about what they want biometricians for? Are biometricians regarded as scientific partners or only as intelligent analytical tools?	14
3.4 Are the institutions clear about the value of the "Institute information resource" ?	15
3.5 Biometry and Research Data Management Unit. Computing center; Financial and Administrative Systems Unit: A complementary team.	15
3.6 The biometrician personality. Researcher-biometrician communication and attitudes.	16
3.7 Availability of qualified personnel, hardware and software.	17
4. CONCLUSIONS.	19

### ACKNOWLEDGEMENT

To Professor Dr. Paul van der Laan, from the Department of Mathematics and Computing Science, Technical University of Eindhoven, The Netherlands. A very interesting casual conversation with him at lunch gave me the inspiration to write this chapter.

## S U M M A R Y

This is a philosophical paper in nature. It describes the role of Biometry in support to agricultural research with emphasis on the developing world and provides a frame of reference for biometricians and researchers who work in these environments.

In its part I, Introduction, the main objectives of the paper are stated. In its part II, the paper provides a summary of the historical developments of statistics and biometry, of computer hardware and of statistical/data management software in support to agriculture research, in the last two centuries. It then describes the present role of Biometry in agricultural research organizations. In its part III, an analysis is made on whether or not the proper role of Biometry as a support discipline to agricultural research, and the role of the biometrician, as a partner in research teams, are fully operational in agricultural research organizations serving the developing world. Factors that may negatively influence the proper role of Biometry are studied: institutional aspects - policies, organizational structure, status of the biometrician -; availability of qualified personnel in the country; researcher-biometrician communication; and the biometrician personality itself.

After fourteen years of work as a Biometrician at CIAT, (International Center for Tropical Agriculture, located at Cali, Colombia, South America) - one of the 13 International Agricultural Research Centers of the CGIAR System -, and having worked as Consultant on Biometry and research data analysis problems to various National Agricultural Research Institutions in Latin America and Africa, I would like to contribute with this thoughts, product of my day-to-day professional experience, to improve our collaboration with the agricultural research process for the benefit of the under-fed world.

## BIOMETRY IN AGRICULTURE RESEARCH OF DEVELOPING COUNTRIES

### 1. INTRODUCTION

The ultimate goal of agriculture research in developing countries is the expectation that it will contribute to increased production, productivity and quality of specific basic food commodities in a chosen region -the area of impact-, thereby enabling producers and consumers, especially those with limited resources, to increase their purchasing power and improve their nutrition (CIAT, 1989; CGIAR, 1989).

To properly cope with its overall goal, agriculture research makes use of a wide spectrum of disciplines that range from highly specialized laboratory sciences, as virology or biotechnology-utilized for "upstream research"-, to development-oriented research disciplines working on activities such as on-farm testing and technology adoption-- examples of "downstream research". Within this ample range of support disciplines to research, quantitative disciplines as Statistics, Operations research, Mathematical Modelling and Computer Sciences play a very important role: that of providing the tools for a more objective description, analysis and inference on live processes, subject to random variation.

The creation of modern Biometry - due to Sir Ronald Aylmer Fisher (1890 - 1962) with his book "Statistical Methods for Research Workers" published in 1925, reshaped mathematical and statistical reasoning in response to the demands of agriculture research. As very elegantly defined by Prof. S.C. Pearce, 1983: "Biometry is the application of mathematical, statistical and computing methods to the description and analysis of biological phenomena; it is, in summary, the application of the quantitative thought for a better understanding of a live process, random by nature".

After fourteen years of work as a biometrician at CIAT, (International Center for Tropical Agriculture, located at Cali, Colombia, South America) -one of the 13 International Agricultural Research Centers of the CGIAR System-, and having worked as Consultant on Biometry and data analysis problems to various National Agricultural Research Institutions in Latin America and Africa, I intend to put in writing some basic principles, experiences and recommendations concerning the role of Biometry in support to agriculture research in developing countries. These thoughts are a reflection of what I am learning through my daily work and the work of the group of professionals who I lead: The Biometry Unit, at CIAT. It has been indeed very interesting to realize how many similarities exist in the attitudes and practice of Biometry between developing-country institutions and developed-country ones. The papers written by highly respected British and French biometricians, as Profs. D.J. Finney

(1956, 1978, 1981), D.A. Preece (1982) and P. Dagnelie (1982) have been elucidating and inspiring.

This paper also intends to provide a frame of reference for a biometrician working in a developing country and to transmit a central idea to the agriculture research community: The success of a support discipline to research, such as Biometry, very heavily lies on a permanent and fruitful communication -based on mutual professional respect- between the researcher and the biometrician. It is the only way in which both lines of knowledge can converge and fully produce for the benefit of agriculture research.

## 2. THE ROLE OF BIOMETRY IN SUPPORT TO AGRICULTURE RESEARCH

### 2.1 HISTORIC EVOLUTION

#### STATISTICS/BIOMETRY

Steel and Torrie (1980) in their chapter 1 present a brief summary of the history of statistics from the 1st century a.c. up to now. Stiegler (1986) presents a thorough review of the history of statistics before 1900. Gower (1988) reviews the developments of statistics in support to agriculture research in the last two centuries.

Agricultural Science has its origins in the 18th century -in the age of agricultural improvement-, whose impetus came from the Industrial Revolution and the need to feed a rapidly growing urban population. At the end of the 18th century, the importance of chemistry for understanding plant growth was recognized. In the middle of the 19th century, Agriculture, with a firm basis in chemistry, became accepted as an academic subject at the University of Giessen in 1824.

Between 1843 and 1901 J.B. Lawes, an amateur chemist, and J.H. Gilbert, both from Rothamsted, England, embarked on a series of field experiments that established the principles of crop nutrition. Their designs would not be acceptable today: there is no randomization, no replication, no blocking and plots are enormously long; the choices of treatment levels are very irregular. Even with all their imperfections these experiments were useful for demonstrating the effects of different fertilizer combinations to visiting farmers. Lawes and Gilbert have been referred as the fathers of the Scientific Method in Agriculture.

Sir Daniel Hall, appointed as Director of Rothamsted Experimental Station in 1901 is remembered for his uniformity trials work with Mercier (1911). He was the first to realize that experimental data in agriculture should have associated some measure of the size of their errors.

In the USA, at Connecticut Agricultural Experiment Station, founded in 1875, much was done in agricultural experimentation before the end of the 18th century. Systematic Designs were used in France by Cretie de

Palluel (1778) and in Germany and Scandinavia, recognizing the need for replication and some form of balance.

Three special developments in statistics had taken place in the last two centuries:

- a) Development of **Probability Theory**, by Laplace and Fermat, described by Laplace in his "Analytical Theory of Probability" (1812).
- b) Development of the **Theory of Least Squares** for estimation, described by Gauss in his "Theoria Combinationis observationum erroribus Minimis Obnoxiae" (1823).
- c) The work of the **Biometrical School**: Galton, Karl Pearson and Wheldon, covering the notions of correlation and linear regression. In the 19th century, Darwin's work (1809-1882) on evolution was largely biometrical in nature. Mendel, too, with his studies on plant hybrids provided a great advance to biometry. Karl Pearson (1857-1936), a mathematical physicist, and W.S. Gosset (1876-1937), a chemist, student of Pearson, known under the pseudonym of "Student", gave impetus to the statistical science. The former applied his mathematics to evolution following Darwin's work. Founder of the journal Biometrika, he was mainly concerned with large-sample theory. The latter was particularly concerned with small-sample theory. He found exact distributions of the sample standard deviation, of the ratio of the sample mean to the sample standard deviation, and of the correlation coefficient. Today, Student's  $t$  is a basic statistical tool.

Sir R.A. Fisher (1890-1962), influenced by Karl Pearson and Student, gave considerable impetus to the use of statistical procedures in agriculture, biology and genetics. Initially hired as a "young mathematician" by the Director of Rothamsted Experimental Station, John Russell, to analyze 75 years of experimental data, he is considered the creator of the modern Biometry. Through his work at Rothamsted (1919-1933), he laid the foundations of Mathematical Statistics and the modern techniques for the design and analysis of experiments, devising many methods to solve the wide variety of problems that agricultural researchers at Rothamsted were faced with. He created Experimental Design as a statistical discipline; gave a clear understanding to the various sources of variability affecting a given response variable; introduced analysis of variance and multivariate analysis techniques; he used the method of least squares estimation for fitting additive and multiplicative models; he emphasized the fundamental role of replication, randomization, the importance of randomized block experiments, introduced factorial designs, split plot designs and the possibilities of confounding as a way to keep block sizes within acceptable bounds; he made emphasis on sampling methods; on significant tests to assess whether or not a set of data supports what may seem to suggest; he introduced variance components as estimates of the variance associated with random effects. His book "Statistical Methods for Research Workers", published in 1925, made the new methods available to biologists.



"Almost all of what is today the philosophy and practice of Biometry and applied statistics derives from ideas originated by Fisher" (Finney and Yates, 1981).

Frank Yates, British Statistician born in 1902, succeeded Fisher at Rothamsted Experimental Station. He was the Head of the Department of Statistics between 1933 and 1968 and developed Fisher's ideas further: incomplete blocks and Lattice designs; partial confounding; Latin Square designs; further developments on sampling methods and survey design; analysis of multi-site and multi-year trials. M.S. Bartlett, together with Fisher, made a tremendous contribution to the theory of multivariate analysis.

Fundamental development of Biometry - as an application of mathematics and statistics to biological problems - took place at Rothamsted, UK. However, after 1933 due to Sir R.A. Fisher's influence, many statisticians and research workers from other parts of the world started to contribute: Mahalanobis in India, founder of the Indian Statistical Institute in Calcutta in 1932 is a good example. Four main schools have contributed to the development of Biometry after Sir R.A. Fisher: The British, the USA, the French and the Russian, with statisticians from the universities and from private enterprises. ; George W. Snedecor from Iowa State University at Ames in USA, with his great contributions "Statistical Methods" (1937) and "Methods of Statistical Analysis" (1939); R.G.D. Steel, from North Carolina State University at Raleigh; J.H. Torrie, from University of Wisconsin; S.R. Searle from Cornell University with his contribution "Linear Models" (1971); W. G. Cochran, Prof. of Statistics at Johns Hopkins University and Gertrude Cox, from the Department of Statistics at North Carolina State University, with their very practical book "Experimental Design" (1956); E.A. Cornish, Chief of the CSIRO Division of Mathematical Statistics in Australia. Excellent examples of other contributions are the many textbooks that are now available everywhere in the world on the general theory of statistical methods, in multivariate analysis, experimental design, non-parametric techniques, linear models and recently on Mixed Model Theory. Modern statistics, product of the 20th century has been defined as: "Statistics is the science, pure and applied of creating, developing, and applying techniques such as the uncertainty of the inductive inferences may be evaluated" (Steel and Torrie, 1980).

#### HARDWARE DEVELOPMENTS

While modern biometrical methods are mainly a product of the first half of the 20th century, developments in computer hardware and software are mainly a product of the second half of our century.

Electronic Computer technology starts in 1939 with the appearance of the first two military-purposes machines -COLOSUS, in the Great Britain, and ENIGMA, in the Nazi Germany- followed by the first commercial-purpose computer, the ENIAC, in USA in 1943.

Technological development of electronic computers from 1939 to our days, is characterized by the tendency to produce faster, more capable machines, always smaller than their predecessors, less expensive, and with more and more processing and telecommunication power. Mainframe computers have been classified in four main types:

a) "First-generation Computers" (1939- 1958), characterized by the use of vacuum tubes as primary source of memory; use of the binary system; small-memory capacity; large-sized and extremely heavy; b) "Second-generation computers" (1957-1964), characterized by the use of electronic transistors as a substitution of the vacuum tube as primary source of memory (1 transistor = 200 tubes); higher memory capacity, higher processing speed and smaller sized; c) "Third-generation computers" (1964-1980), based on electronic integrated circuits as a replacement of the transistor (1 microminiaturized circuit = 1000 transistors), of very high operational capacity, higher memory capacity, teleprocessing possibilities, and use of an operating system "as the software who administers the use of computer resources"; and d) "Fourth- generation computers" (1980 to our days), based on dense packages of electronic integrated circuits; with very high memory capacity and processing speed, small-sized, light-weighted and with networking possibilities. (IBM, 1984: "Innovations in Computer Technology).

A new era of computer technology started with the appearance of the Personal Computer, a product of our last decade. Computing technology, previously possessed by an elite discipline -composed by some privileged system's engineers with deep knowledge on the operating systems and on more or less difficult computer languages, and by some dedicated "amateurs" such as statisticians, mathematicians, physicists or engineers of other kinds- became available to everyone at his/her desk. User-friendly operating systems were developed for the personal computer (DOS, MS-DOS, CPM, UNIX); new computer languages appeared (BASIC, PASCAL, C) and many software applications were developed to satisfy the wide range of needs of the microcomputer user community: word-processing packages; electronic spreadsheets; database managers, graphic packages and many specific-purpose applications, such as financial packages, administrative applications, statistical packages, etc.

A wide range of models of personal computers has been developed. In IBM micros, for example, the initial IBM PC, with 64 K of memory and 10M bytes of hard disk, followed by the IBM XT, and IBM AT, were very practical for many uses. Lately, the family of IBM Personal System/2 is available all over the world. From Model 25, with 8086 microprocessor, 640K of memory and 20MB of fixed disk to Model 80, with 80386 microprocessor, 1 to 4MB of memory and 44, 70, 115 or 314MB of fixed disk. The 80486 microprocessor is now available and offers higher speed and processing capacity. (IBM Systems and Products guide; 1988).

The appearance of the personal computer has undoubtedly created the need for a new administrative structure in data processing

departments. From a centralized structure around a mainframe, an organized decentralization was needed. In order to have more control on the use of the very powerful computer resources made available to the microcomputer users, the LAN's (Local Area Networks) philosophy appeared on 1984. The possibility for interconnecting different micros to share software resources and information is a great advance. Mixed Networks, where mainframes, or minis or micros can act as SERVERS and multiple micros can be served by them is another recent technological advance. (Pardo, R., 1990).

## SOFTWARE DEVELOPMENTS

Software applications in agriculture research are mainly a product of the second half of our century. Their development has taken place as a consequence of the incredibly fast technological developments in computer equipment.

The first comprehensive statistical package for the mainframe was BMDP (Biomedical Data Processing Program), developed by -Dixon, from Statistical Software, Los Angeles, California, whose first version was released in 1960. SPSS (Statistical Package for the Social Sciences), was initially developed for the mainframe by the University of California in 1966, and latter supported by SPSS Inc., Chicago, Il. SPSS/PC is now available.

Rothamsted Experiment Station acquired its first electronic computer in 1954 (Gower, 1988). Gower says: "Originally, separate programs were written to cover very restricted types of analysis in the following areas: analysis of designed experiments, analysis of surveys, curve fitting, multivariate analysis and multiple regression. In all these areas it was soon found that the way-forward was to standardize on data handling and, so far as possible, subsume many types of analysis as special cases of some more general approach. This pressure to unify statistical methods led to statistical research and in due course to the development of the GENSTAT program, released in 1972". In 1972, Nelder and Wedderburn, under the Society's Working Party for Statistical Computing, from Rothamsted, developed the statistical package GLIM (Generalized Linear Interactive Modelling) -whose facilities are also in GENSTAT-.

Simultaneously in the USA, the first version of SAS, the most widely-used statistical package nowadays in USA and Canada, and very extensively used in Europe, Asia, Latinamerica, Australia and New Zealand, was released in 1972. The creation of SAS is attributed to Drs. Anthony J. Barr and James H. Goodnight, from the Department of Statistics, North Carolina State University at Raleigh, NC, USA. SAS, initially born as "Statistical Analysis System", was conceived as an integrated system for data management and statistical analysis. By combining statistical versatility with extensive capabilities for data manipulation and report writing, SAS in its first version, SAS/72,

provided already a total system to help statisticians and data analysts solve their statistical computing problems. Statistical procedures (PROC's) included, covered a wide range of descriptive statistical analysis and inferential statistical analysis -both univariate as multivariate. After four years of work, the creators of SAS left the Department of Statistics at NCSU and created SAS Institute, Inc. as a commercial software producer company, having as its first members Drs. A. J. Barr, J.H. Goodnight, J.P. Sall and Jane T. Helwig. The second version of SAS, SAS/76, was released in 1976 providing additional statistical analysis capabilities and interfaces with other statistical packages for mainframes such as BMDP, SPSS, and OSIRIS. An International SAS User's Group (SUGI) was created and started to meet every year in different cities of the USA: SUGI2 took place at New Orleans in January 1977, and SUGI15 took place at Nashville in April 1990. From 1979 to 1985, SAS was enriched with new applications outside the field of statistics: SAS/ETS (Econometric and Time Series) and SAS/GRAPH (a set of procedures for graphical aids in two and three dimensions), released in 1980; SAS/FSP (a full-screen product facility for data entry, modification and retrieval), released in 1982; SAS/OR (Operational Research procedures), and SAS/AF (for applications development) released in 1983; SAS/QC (for Quality Control), SAS/IML (Interactive Matrix Language), and SAS/SHARE (for multi-user access to SAS datasets) were latter released. From 1985 to 1990 SAS Institute decided to serve not only the mainframes and minicomputer users, but also the microcomputer user's community. From then on, the name "SAS" was kept without any meaning. The first version of SAS for the Personal Computer -SAS/PC version 6- was released in 1985, with three products: BASICS, STATS and GRAPH. At present, SAS/FSP and SAS/ETS are also available for microcomputers. SAS Institute also supports today a database management software -SYSTEM 2000- fully compatible with all SAS products, and offers software interfaces with major database management software packages supported by other companies, such as: ADABAS, to IDMS/R, DATACOM/DB, IMS-DL/I, DMI, DB2 and SQL/DB. Today the 1990 SAS System -written in C-Language- is defined by SAS Institute as "a software system for data analysis that provides data analysts with one system to meet all their computing and analysis needs": a programming language; information storage and retrieval; file handling; report writing tools, statistical analysis, econometric, financial and forecasting tools, graphic aids and database management. (SAS User's guides version 5, version 6, version 6.03)

Apart from the main mainframe statistical and data management software packages just mentioned, there are thousands of microcomputer statistical packages available in the market. The annual publication "Data Sources -software" by UNISYS, in its 1988 edition, lists more than 5000 different microcomputer software packages on statistical and mathematical applications that work on different operating systems.

## 2.2 THE ROLE OF BIOMETRY TODAY

Biometry has been defined as the science that studies the application of mathematical statistical, and computing methods, to describe, quantify and analyze live processes whether they be human, animal or vegetable. As an application of statistics, it deals with biological problems in which, for the individual observation, laws of cause and effect are not apparent to the observer. Therefore, an objective approach is needed.

A biometrician is a statistician, a mathematician, or a statistical biologist who besides possessing good knowledge on quantitative techniques applicable to the study of live processes, has also a clear understanding of the biological problem under concern. A high percentage of his/her daily work is constituted by continuous dialogue with the researcher, so as to first understand what the researcher desires to learn and then be able to offer and share ideas, suggestions, and alternatives for solving the formulated research problem. Its wide relation with the entire research process requires biometricians to have a personality that is open, communicative, flexible, and ready to learn more each day from the research biologist. "The scientific research process is of a multidisciplinary nature, where minds of different abilities can contribute in endorsing greater generality and applicability of the results. And the biometrician mind is one of these" (Professor S.C. Pearce, 1983).

In an agriculture research institute, the "quantitative thought" of the biometrician must be an integral part of the research process and must be included from the planning and design stages of the research project to the final interpretation and publication of the results. These stages cover: the precise definition of the objectives; the determination of area and conditions for generalization of the results; the choosing of experimental factors and their levels, and the selection of appropriate experimental designs -in the case of designed experiments-; the selection of response variables needed to test the given set of hypothesis; the determination of the most appropriate plan of statistical analysis; and finally, the publication of results. In many research institutions, the biometrician is also responsible for the execution by means of the computer, of the statistical analysis agreed upon in the planning stage. The results often offer more light into the problem and a better understanding of the phenomenon. Depending on the objectives and complexity of each project, Biometric data analysis support varies from standard data management and statistical analysis, to a more specialized support that requires the development of more sophisticated datahandling techniques for specific statistical analysis methodology, or simulation. An agricultural research project could be a diagnostic study in which the theory and applications of sample survey design and analysis is needed; or a designed experiment -either a field experiment or a simulation experiment- in which the theory and applications of experimental design is essential; or a validation

study carried-out through case-studies or via a collection of observations; or a mathematical model for prediction purposes, that describes the interrelation between the various components of a biological system. Although the most frequent involvement for a biometrician in developing countries is in the design and analysis of field experiments and sample surveys, more and more opportunities for collaboration are now open in the other areas.

Following the previous discussion, the type of experimental projects in which a biometrician is asked to provide advise on, can be classified by their nature, in two large classes: "Controlled experimentation" and "Non-controlled experimentation". The first, "Controlled experimentation", refers to designed experiments -field, laboratory or simulation experiments- with controlled experimental factors chosen a-priori, that require orthodox techniques of statistical analysis. It comprises individual experiments which tend to test either one or more hypotheses in specific environments and time, or series of experiments, repeated in space or time, tending to prove one or more hypotheses in a wide range of environments (as in the case of International Networks). Depending on the type of crop, or system of crops, individual experiments as well as series of experiments may require: a) only one time evaluation -as in the case of crops having short vegetative cycles, where evaluation is done at harvest time-, or b) periodic evaluations -as is the case of perennial crops, forestry research, fruit-trees research, pastures research, etc., where it is necessary to consider season, year, and age of the crop as some of the additional sources of variability affecting the response variable. The second group of projects referred as "Non- controlled experimentation", includes technical and socio-economic diagnostic studies, and technology-validation studies, carried-out through sample surveys, case-studies, or by means of a description of a collection of observations in time and/or space, where the emphasis of the statistical analysis falls more on parameters estimation than on hypotheses testing.

The contribution of Biometry to agriculture research covers a wide range of activities which will be classified into five main groups: a) advise and support to the researcher in the various stages of the research process: conceptualization, planning, implementation, data analysis and final publication of research results; b) training in statistical/ mathematical methods and data analysis; c) participation in collaborative methodological projects; d) involvement in applied statistical research and finally, e) participation in the conceptualization of research information systems to solve the needs for storage and retrieval of summarized research data. The first two activities represent the day-to-day work of a biometrician. The third activity, -participation in collaborative research projects-, requires an special effort from the biometrician to gain the authority and credibility to add his/ her specific way of thinking into a new piece of research finding. The biometrician contribution in this context is mainly achieved through

methodological studies that can offer orientation on the efficiency, accuracy and goodness of research practices. Also through the analysis of long-term accumulated information carried-out to evaluate the effectiveness of a given research strategy. However more difficult, it is one of the most challenging and satisfactory activities. The fourth activity -involvement in applied statistical research-, is referred as relevant research to solve the day-to-day problems of research workers: what techniques have been the most effective research tools; what designs are most appropriate for a given research; what sources of variation are relevant for future research planning on a specific subject; the strategy of using multivariate methods for a large number of variables rather than ordinary univariate analysis; suggestions on modifications on a given experimental design to provide better extrapolation capacity to a specific research problem, taking in consideration environmental conditions of the region over which the results are expected to be generalized. The fifth activity, -participation in the conceptualization of information systems to solve the needs for storage and retrieval of summarized research data-, is in many cases as demanding of statistical experience as of experience in the command of database management software. The agricultural researcher -who has a clear objective for the information system as its potential user-, the biometrician -who has a clear understanding of the data and statistical summaries-, and the system's analyst, as the designer and implementor of the information system, will conform an ideal multi-disciplinary team.

An institution whose policies provide the biometricians with the proper atmosphere for expressing themselves in this manner, will undoubtedly gain a lot in terms of a more objective understanding of the phenomenon under study and a better extrapolation capacity of its research results.

### **3. IS BIOMETRY IN DEVELOPING COUNTRIES COPING WITH ITS PROPER ROLE? EXISTING CONSTRAINTS AND POSSIBLE IMPROVEMENTS.**

Let us summarize the proper role of Biometry in agricultural research. Biometricians today should be deeply involved in research cooperation with other disciplines: in quantitative planning of research, in presenting new approaches to the problem solution, in displaying initiative in data analysis projects that will answer relevant research questions, in the conceptualization of research information systems, and in providing training to their colleagues from other disciplines, in statistical methods and data analysis. They must be encouraged to publish as partners in research teams or individually. As Finney (1978) says, both biometricians and researchers must interpret their role in the agricultural research process as one of mutual and permanent collaboration.

Specific services offered by a Biometry/research data management Unit within a research institution are:

1. Statistical consulting services: in research planning, design, implementation, data analysis, generalization capacity of research results and final publication.
2. Data processing/analysis services, or Statistical Computing services.
3. Training in statistical and mathematical methods/data analysis/software use.
4. Collaboration in the conceptualization of research information systems, in close cooperation with system's analysts and the research team involved.
5. Collaborative methodological projects with researchers from other disciplines.
6. Work on applied statistical research, when there is a need for it.

Priorities among these functions depend on the institute culture, research needs, and biometrician-researcher capacity to work in research teams. The right balance must be searched.

Most agricultural scientists in developing countries now recognize this role as a very important component of research progress. The biometrician support with research planning, data analysis and involvement in collaborative projects with other disciplines is in many cases accepted, encouraged and recognized. However, there are situations in which this is not entirely true. There are factors that affect negatively the full expression of the potential contribution of Biometry to research: institutional organization, culture and policies; availability of qualified personnel in the country; communication problems between the biometrician and the researchers; and very often, the biometrician personality itself is a barrier against a fruitful use of quantitative disciplines for the benefit of agriculture research. It is our purpose to analyze some of these factors, in the case of agricultural research institutions working for developing countries.

### 3.1 ORGANIZATION OF RESEARCH INSTITUTES. ORGANIZATION OF BIOMETRY/RESEARCH DATA ANALYSIS GROUPS.

Let us discuss first in very general terms the organization of International Agricultural Research Centers of the CGIAR System, and of national agriculture research institutions of developing countries, to latter discuss the organization of biometry/research data management groups within them.

According to research organization, International Centers, commodity-oriented by nature, can be classified in two groups:

- a) A centralized research Center with a strong headquarters, and various relatively small groups of researchers located in other



countries with responsibility for specific projects for a given region. CIAT and IRRI can be considered examples of this organization.

- b) A decentralized-research Center, with less strong headquarters and many independent and self-sufficient regional projects located in other countries, which are responsible for commodity-oriented research for a specific region. It is usually arranged that regional project researchers may benefit from the housing institution central facilities. The housing institution can be another International Center or the national research institution in the country. CIP and CIMMYT can be considered examples of this organization.

National agriculture research institutions in developing countries can be classified in two categories according to their organizational structure:

- a) A system of research centers, with a strong headquarters -at or nearby the capital city- responsible for central planning, administration and specific research, and many other independent research centers each with one or more experimental stations assigned to them. Each Center has its mandate in terms of research products, ecosystems or regional area within the country. This organization is typical of institutions serving most of Latinamerican countries. Examples are EMBRAPA in Brazil, INTA in Argentina, INIFAP in Mexico, ICA in Colombia, FONAIAP in Venezuela, INIAP in Panama among others.
- b) A set of independent agriculture research institutes within the country, each responsible for conducting research on a discipline or group of disciplines (medicine and veterinary sciences, crop protection, etc.) or on a group of products (cereals, vegetables, tuber-crops, industrial products). This organization is typical of ex-french colonial countries, as Madagascar, Zayre, Burundi, Rwanda, which have inherited the French Agriculture research philosophy. However, the french system is now tending towards a more unified approach. CIRAD (Centre Internationale pour la recherche agricole et le developement) created in 1970 as the International Center for agriculture research and development for France and the French colonies, has brought under one umbrella the 11 independent agriculture research institutes existing in France. Some francophone African countries are working on the same line. That is the case of the FOFIFA (Madagascar National Agriculture Research Institution, its English translation from the malgache language), created in 1974, which has unified agriculture research in the country.

Depending on the different structures being discussed, the organization of Biometry/research data management groups must be thought to best cope with needs. The most common situation in Latinamerican National institutions is to find a more or less well-provided central Biometry or Data Processing Unit at Headquarters, with appropriate hardware, software and capable professionals, leaving Regional Centers almost totally unprovided:

either with one statistician (or "amateur" statistician) and few assistants in large Regional Centers, or with nobody at all in smaller Centers.

In the case of centralized centers, a unit at headquarters is useful; however the Unit leader should be aware of the need to assign a team made-up by biometrician/analyst to support at least the most important research programs. In this way, the biometricians and colleagues are given the opportunity to learn the fundamentals of research problems faced by their users and can become an integral part of the research team. In the case of decentralized research Centers, a central Biometry/ research data management group only at Headquarters is not effective: personal communication is very difficult; the biometrician will be forced to keep contact with researchers by mail or telephone. Results will be delayed. So, it would be necessary to organize well-provided Biometry/research data management groups in all important centers, with a central coordination from headquarters. This is beneficial as it helps to keep biometricians in touch about seminars, new statistical techniques, new software, common problems, policies, etc. In the case of independent institutes spread around the country, it is even more necessary.

In summary, after discussing possible institutional organizations, it becomes apparent that the number of biometricians/data analysts/programmers should respond to research needs, organizational structure and geographic location of research projects outside headquarters. Only one professional with some few assistants-, that is sometimes the case in many institutions-, or a relatively strong central unit at headquarter leaving regional centers totally unprovided, will never be enough to satisfy the needs. To be effective, a multi-skilled team approach with sub-groups at the Regional Centers is necessary.

### 3.2 RESEARCH POLICY AND PRIORITIES. RESPONSIBILITIES OF THE BIOMETRICIAN.

Agricultural research institutions in developing countries can be commodity-oriented, discipline-oriented or a combination of both. Many commodity-oriented national agricultural research institutions are responsible for conducting research on an immense number of crops, of different relative priority in the country's economy, but covering a wide range of agricultural products cultivated in the country. Resources assigned to the various crops are heterogeneous -been influenced by political decisions or availability of funds-. A similar resource-allocation situation may arise when research is conducted under a discipline-oriented scheme. In both cases, however, the area of biometry/data analysis, is generally considered a central service to the entire institution, and very often given a lower relative priority in terms of resource- allocation. In this way, a small group of biometricians/data analysts, if not only one person, are responsible for providing support to a wide variety of research

problems that range from pathology research on small animals at laboratory level, to on-farm research on extensive animal production systems; from short-cycle single crop research, to multiple-crop-associations or rotations including perennial and short-cycle crops; from a single experiment data analysis, to a complex data analysis project including many years, multiple varieties and many experimental sites per ecosystem; from the statistical analysis of electrophoretic diagrams for detection of duplicates in a large germplasm collection, to the statistical analysis of the performance of improved varieties at farm level.

**Problem:** This extreme diversity of research problems with which the biometrician is faced with, makes his/her contribution an ineffective routine. A basic understanding of the biological problem under study is very seldom attained by the biometrician; so, his/her capacity to provide thoughtful and practical research solutions is diminished, if not completely eliminated.

### 3.3 ARE ADMINISTRATORS OF RESEARCH CLEAR ABOUT WHAT THEY WANT BIOMETRICIANS FOR? ARE BIOMETRICIANS REGARDED AS SCIENTIFIC PARTNERS OR ONLY AS INTELLIGENT ANALYTICAL TOOLS?

Very fortunately, more and more research administrators from developing country institutions are becoming aware of the important role of Biometry in support to research. Very often research administrators have been agricultural researchers before having been promoted to an administrative position within the institution or outside their own institution. They have the proper scientific background to appreciate the importance of a consistent and thoughtful scientific research process. However, they very often assume that the researchers themselves should be aware of this, and if they are not, they should take the initiative to look for statistical advise from a biometrician.

**Problem:** What happens if this does not occur? Well, it has been my experience that it is not always easy for a research administrator to establish an institutional policy that reminds researchers of the importance of scientific methodology and accuracy. If there is no such a policy, it is the biometrician responsibility to make the research community aware of this need.

How? Through permanent and positive communication with them and high quality work that shows benefits to the researcher's final output. This task is very satisfying but not an easy one. It takes time and effort to prove oneself continuously in front of the researchers. The biometrician has to demonstrate that a different way of looking at their problem -an analytically oriented way supported by quantitative techniques and a basic understanding of the biological problem- can add a lot of new light to the problem solution. If this is achieved, the biometrician will start being considered as a partner; if not, he or she will continue to be regarded only as an intelligent analytical tool.

There are some cases still in which research administrators are more aware of financial aspects, inter-institutional relations and other

important business, leaving the full responsibility for a proper research conduction to each researcher. If this is the case, the biometrician task is even harder. He or she has to gain credibility to be accepted as a member of multi-disciplinary teams for the conduction of specific research projects; otherwise, the biometrician will gradually start to feel frustrated, inferior and a sense of lack-of-achievement will dominate his/her attitude. If a little timid, he/she will remain doing the same routine day after day, and won't be able to make a positive contribution towards the incorporation of the "quantitative thought" and the more efficient use of information.

#### 3.4 ARE THE INSTITUTIONS CLEAR ABOUT THE VALUE OF THE "INSTITUTE INFORMATION RESOURCE" ?

A lot has been said about the value of information as a decision-making tool. Computer technology -even in developing countries- is available and permits many ways of information interchange, within and outside the institution. Many mainframe and microcomputer statistical, data management, and database management software packages are available; microcomputer networks (LAN's) or mixed networks technology is available.

Problem: However, significant portions of scientific information produced for years by many agricultural research institutions in developing countries still remains untouched - waiting to be cleaned and statistically analyzed.

Why? Perhaps the researchers do not have enough time to devote to this highly time-consuming and delicate task. Perhaps the biometricians do not have the initiative to suggest the researchers a new "data analysis project", whose purpose would be to answer relevant research questions. Perhaps the communication between research administrators, researchers, biometricians and computer specialists has been broken somewhere. Perhaps, although the institute has computers and software, the right persons who could implement the very challenging task of organizing and analyzing a big amount of data, do not exist. In any case, no matter what may be the reasons for not fully utilizing the existing research information for their decision-making, this goal is one of the most important contributions of biometry to agricultural research. Any effort should be made to achieve this objective.

#### 3.5 BIOMETRY AND RESEARCH DATA MANAGEMENT UNIT; COMPUTING CENTER; FINANCIAL AND ADMINISTRATIVE SYSTEMS UNIT. A COMPLEMENTARY TEAM.

For historical reasons -mainly associated with the development of mainframe computing technology in the decades of the 60's and 70's- Biometry/data analysis Units, being the main users of the Institute's Computer facilities, have been made responsible for the provision of computer services to users. This includes administration and control of mainframe computer resources; provision and maintenance of mainframe hardware and software: computer operation service, remote terminal connections, telecommunications, and transcription services.

There are some cases in which even administrative data processing responsibilities lie within the same Unit. In this case, conflict of priorities between administration and research might appear.

The availability of microcomputer technology in the decade of the 80's, increased the number of computer users within the institutes and created new needs for computer services. Although the three functions -Biometry/research data management, administrative/financial data management, and computer services- may be provided by one Department or Unit, a clear distinction of objectives, responsibilities, audience, human resources and reporting lines must be defined. The desired structure very heavily depends on the Institute specific characteristics and needs. It is our own opinion however, that in an agricultural research environment, a Biometry/research data management Unit, given its natural involvement with the research process, must report to the Director of Research. Its precious links with research need to be maintained.

### 3.6 THE BIOMETRICIAN PERSONALITY. RESEARCHER-BIOMETRICIAN COMMUNICATION AND ATTITUDES.

In order for the biometrician to properly deal with the challenge of working in real partnership with other research disciplines, a communicative, flexible, open-minded, and learning attitude is necessary. Apart from a solid background in statistical science, in statistical computing and a basic understanding of the biological problem under study, the successful biometrician needs the right personality.

When appointed at an agriculture research environment, the first step to be followed by the Senior biometrician should be to get an overall understanding of the institute research objectives, program strategies, policies and organizational structure. After that, he or she may promote informal meetings with research teams or with individuals to learn about their needs, and explain the possibilities for mutual cooperation. There are agricultural statisticians who are introverted and solitary; others tend to under-estimate the value of a non-quantitative discipline; others have serious communication problems finding difficult to work in research teams. They should carefully consider the need to modify their attitude before accepting the challenge.

Finney (1968), from Rothamsted Experimental Station, describes the researchers-biometrician relationship, as "variations of the master-servant attitude". This also happens in developing country institutions, although not always. Let's examine when a "master-servant attitude" between researcher and biometrician may arise:

- a) When there is a lack of appreciation from the researcher part on when there is a need for adoption, adaptation or innovation of experimental techniques and data analysis methodologies to better

deal with agricultural research problems of tropical conditions. When this happens, there is a tendency to adopt developed countries, temperate zones, experimental techniques and data analysis methodologies without any question. The biometrician contribution is then regarded as a pure routine. This can be a wrong assumption in many fields of agricultural research for developing countries. For example, beef production research in the tropical savannas ecosystem of Brazil, Colombia and Venezuela, with a carrying capacity of 0.2 animals/ha, with long dry-season periods and mixed animal breeds, needs different considerations for planning, design and data analysis than highly controlled beef production research in England under rye grass and pure animal breeds.

- b) When biometry/research data analysis is looked-at as any other service function within the institution, and is taken for granted, without recognizing its intellectual contribution to the research process.
- c) When the biometrician is a young statistician, with little experience in dealing with biological problems. The researcher then does not look for statistical advice but only for routine data analysis, sometimes telling the biometrician to prove (or try to prove) statistically the conclusions already drawn in his/her own mind. In developing country institutions there is normally at most one Senior biometrician at the Institute, who is in general the Unit Head, responsible for a group of junior biometricians, system's analysts and programmers. At some institutes, the position within the organization is at the same level as that of senior researcher; at other, he/she holds a lower position. Although the Senior biometrician can be involved in many research problems and training, many other projects are delegated to the junior biometricians. If they don't have the right experience, the researcher-biometrician relationship suffers. It is then, the Senior biometrician's responsibility to properly choose and further train and offer professional integration in research teams to his/her own staff.

In a research organization the researchers knowledge in statistics and computers is very variable. Some researchers have very good knowledge on specific applications of statistics to their own field; some have a good general understanding and appreciation of statistical techniques; some others have limited conceptual basis. Some prefer to plan, implement, analyze and publish their results by themselves; some are more willing to work with biometricians/data analyst; and some prefer to leave all data processing/analysis task on the biometrician hands. Within this range, the biometrician has to keep all communication channels open; be appreciative of other disciplines knowledge and try to be always on the user side. Good interdiscipline communication is a key element for success.

### 3.7 AVAILABILITY OF QUALIFIED PERSONNEL, HARDWARE AND SOFTWARE.

With the rapid development in microcomputer technology that has made available to every institution powerful and relatively non-expensive machines -together with the great advances in microcomputer statistical, data management and graphic software packages-, developing country institutions have no limitation nowadays in terms of computer hardware and software for their research data analysis. Software and hardware prices tend to decrease with advances in technology. Site licenses for software are available to cover -in one contract- all the micros spread over the country. Microcomputer maintenance firms proliferate. This is the normal situation in most latinamerican countries. Perhaps the situation in African countries is a little different.

However, the big limitation of national research institutions from Latinamerican and African countries is qualified personnel. Few Latinamerican Universities provide PhD degrees in general. Only Brazil provides a PhD degree in Statistics. Some countries (Argentina, Brazil, Colombia, Chile), provide MSc degrees in Mathematical Statistics and Applied Statistics; but none provide a degree in Biometry as such. A biometrician for a latinamerican country -and presumably for African countries- is not easy to get: the person needs to be either a statistician by profession and further trained "in the job" on the biological sciences, or an Agronomist/biologist with a post-graduate degree on Statistics. The other alternative would be to employ a professional with a degree in Biometry from the USA (North Carolina State University, Cornell, Iowa State, etc.), from England or other European country. However, by qualified personnel I do not only mean professional statisticians: I mean professionals with the right balance required to work in an agricultural research environment.

Once an institution has a good Senior biometrician, it should provide him/her with institutional support and motivation as he/she will be responsible for the employment, training and further professional development junior biometricians for the institute. To be effective, a biometrician, a data analyst, a professional programmer, or a mathematical modeller working in agriculture research, need, apart from their degree a proper exposure and a lot of basic understanding of specific biological sciences. This is normally acquired with experience in the job. To gain this experience, the institution should provide its junior biometricians with the opportunity to learn from the scientists. It is a vicious circle: if the institution does not "train" its young statisticians, mathematicians, or computer specialists, they won't have the right balance to interact with researchers. Without the right balance, they would very seldom be accepted as part of a research team.

So, the solution may be to give the Senior biometrician(s) at the Institute the additional responsibility of training their own staff; facilitate their participation in research projects; send them to the experimental stations; to biology-oriented seminars as well as to

specialized seminars in their own discipline; to facilitate a fruitful interaction between the three key data analysis professionals: "researcher", "biometrician", "computer specialist".

We would like to present in the next chapter, the results of a 1990 survey on "Resources available in the area of Biometry/research data analysis, at national agricultural research/development/ training institutions from Latinamerica and Caribe". These results permit a more accurate quantification of what we have just described.

#### 4. CONCLUSIONS.

Biometricians who work for agricultural research organizations in the developing world have a great challenge today: there is a wide open range of problems to be solved. The opportunities are there: the challenge consists on providing a solid and realistic solution to them. Although modern Biometry as a science was born 65 years ago, it was created as a support discipline to agricultural research carried-out in temperate zones, under high controlled conditions, and as a tool to solve the wide, but specific, range of problems that british researchers use to face at that time. A lot of progress has been made since then. New ideas from statisticians from developed and developing countries have enriched the Biometry Science. Hardware and software resources are not a limiting factor anymore for developing countries. So, what is needed is the right team of professionals with the right balance: solid mathematical and statistical knowledge, solid experience in statistical computing, basic understanding of the biological problem under study, and, which is very important: an open, communicative and flexible personality which can provide a new dimension to the multidisciplinary research team.

Developing countries in the Tropical and Sub-tropical zones, are facing very specific needs in agriculture research: specific climatic conditions, soil variability, multiple-cropping systems, multiple type of associations, a wide variety of production systems that range from extensive agriculture to very intense small-farmer systems; specific animal breeds, plant species, and plant and animal diseases. All this new complex of sources of variability make agricultural research in tropical conditions very different from agricultural research in temperate zones. Innovation and adaptation of experimental methodologies, desing and data analysis -more than adoption of developed countries techniques- is a must.

The challenge is there. But the responsibility to succeed is basically ours: of the biometricians!



## BIBLIOGRAPHY

1. CIAT, 1989. Program Plans and Funding requirements (1989-1993).
2. CGIAR, 1989. 1988-1989 Annual Report.
3. Cochran W.G. and Cox G.M. 1957. Experimental Designs. John Willy & Sons, New York, USA.
4. Dagnelie, P. (1982). La Biometrie dans le pays du Tiers Monde. In: Actes de la 11 Conference Internationale de Biometrie. Versailles, France. I.N.R.A. pp 177-182.
5. Data Sources - Software, 1st edition 1988, Vol. 2. 1988 UNISYS Corporation.
6. Finney, D.J. (1956). The statistician and the planning of field experiments. Journal of the Royal Statistical Society A119, 1-27.
7. Finney, D.J. (1978). Statistics and Statisticians in agriculture research. Journal of Agricultural Sciences. Cambridge. 91, pp 653-659.
8. Finney, D.J. (1981). The misuse of mathematicians, statisticians and computers in agricultural research. Experimental Agriculture, volume 17, pp 345-353.
9. Finney and Yates. 1981. Statistical Methods for Research Workers. Published in 1952.
10. Gower J.C. (1988). Statistics in Agriculture. Journal of Royal Statistical Society. A 151, Part 1, pp 179-200.
11. IBM, 1984: Innovation in IBM Computer Technology.
12. IBM, 1988. Systems and Products guide. (IBM publication No. 800 - 2468.)
13. Pardo, S.R. 1990. Notas del Seminario sobre Redes Locales. Trabajo presentado en: Seminarios de Excelencia. Bogotá, Colombia.
14. Preece, D.A. (1982). La biometrie: pas rites mais Science. In: Actes de la 11 Conference Internationale de Biometrie. Versailles, France. I.N.R.A. pp 189-193.

15. SAS User's guides: 1972, 1976, 1979.  
SAS User's guides: BASIC and STATISTICS, Version 5. Edition 1985.  
SAS guides for Personal Computers. Version 6 Edition 1985.  
SAS guides for Personal Computers. Version 6.03 198
16. Searle, S.R. 1971. Linear Models. John Wiley & Sons, Inc. New York.
17. Snedecor, G.W. and Cochran, W.G. 1937. Statistical Methods. Six edition. The Iowa State University Press.
18. Steel, R.G.D. and Torrie, J.H. 1960. Principles and Procedures of Statistics. McGraw-Hill Book Company.
19. Steel, R.G.D. and Torrie, J.H. (1980). Principles and Procedures of Statistics. A Biometrical Approach. Second Edition. McGraw-Hill Book Company.
20. Stigler, S.M. (1986). The History of Statistics. The Measurement of Uncertainty before 1900. ISBN 0-674-40340-1.