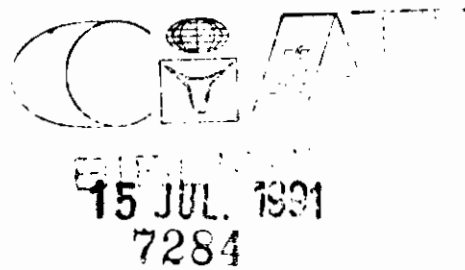




~~INTERNATIONAL NETWORKS FOR VARIETAL~~
EVALUATION: Objectives, information generated and
statistical methods for data analysis



12701

Maria Cristina Amézquita
CIAT, March 1991

ACKNOWLEDGEMENTS

For their very valuable comments on this document, to Drs. Shivajii Pandey, from CIMMYT Maize project for the Andean Zone; to Dr. Federico Cuevas, responsible for the coordination of the IRRI/CIAT rice international network for Latinamerica and Caribe; and to Dr. José M. Toledo, previous Leader of the CIAT Tropical Pastures Program and ex-coordinator of the International network for tropical pastures evaluation.

SUMMARY

This document is intended as a general overview on the role of "International Networks for varietal evaluation", with special emphasis on information generated and statistical methods for analyzing the network information. It discusses objectives, experimental design and information generated by this type of networks (numerals 1 and 2), type of questions that can be answered through the analysis of a network information (numeral 3) and provides a description of the characteristics and use of statistical methods proven useful for answering those questions (numeral 4). The discussion here presented is intended to serve as a frame of reference for agricultural researchers and biometricians working with International Networks in the developing world. For the agricultural researcher, as a way of providing him/her with ideas on how to best utilize the Network's information resource to answer relevant research questions. For the biometrician, as a way of offering a very challenging opportunity to apply quantitative methods to the better understanding of the behavior of newly generated germplasm across a wide range of environmental conditions.

C O N T E N T

	Page No.
Summary	
1. Background	1
2. Objectives, Experimental Design and Information Generated.	2
3. Why a Multilocational Analysis	8
4. Useful Statistical Methods for data analysis	9
5. Concluding Remarks	17
References	18

**INTERNATIONAL NETWORKS FOR VARIETAL
EVALUATION: Objectives, information generated and
statistical methods for data analysis**

Maria Cristina Amézquita
Biometrician, CIAT
March, 1991

1. BACKGROUND

International cooperation in agricultural research is not new and is rapidly increasing as its benefits are realized. Agricultural scientists work in partnership's on a regional or global scale to reduce costs, avoid duplication of efforts and accelerate transfer of technology to farmers. Plucknett and Smith (1984) present a very thorough historical review of networks in agriculture research, from the colonial period -when British, Dutch and French established networks of research stations in their territories to increase the flow of export crops- up to our days, when there are more than 100 international agricultural research networks currently operating. They describe four types of agriculture research networks: a) Informal networks of scientists, to exchange ideas and information; b) Formal networks to exchange technology and information between participants and a central node; c) Formal networks in which participants interact with each other as well as with the central node; and d) More complex networks in which participants set-up subnetworks to focus on a particular problem. International networks for varietal evaluation are generally set-up in the way described by b) or c); cooperative effort involves a 2-way flow of information and germplasm, and entail a commitment of resources from participant institutions.

Important efforts have been made to financially support the initiation and effectiveness of many international agricultural research networks. Let us cite some examples: The US Department of Agriculture (USDA) established in the 1950's the first USDA-coordinated multinational network: the International Stem Rust Nursery, to find resistant varieties of wheat to a new race (15B) of wheat stem rust in the USA. By 1970, 150 wheat breeders from 40 countries in the Americas, Europe and Middle East had joined this network, since they realized that the problem could be extended to their countries (Plucknett and Smith, 1984). The important contribution of the United Nations Development Program in Agriculture (UNDP), whose first global scientific project was initiated in January 1970, in collaboration with CIMMYT, to help develop high-protein maize, a food crop important to millions of people in the developing world. UNDP has continued supporting many agricultural research projects of the International Agricultural Research Centers of the CGIAR System (Consultative Group for International Research) (Mashler, 1985). The International Development Research Center of Canada (IDRC) is another example of an important aid agency.

During the 1960's, 1970's and 1980's, the International Agriculture Research Centers of the CGIAR system, have acted as catalysts for many of the international agriculture research networks of the Third World, in particular for the International networks for varietal evaluation. For example: IRRI established in 1963, with UNDP financial support, the International Rice Testing Program (IRTP), which now counts with more than 800 scientists from 75 countries in Asia, Africa, Latinamerica, Ocenia and Europe (Seshu, 1984). CIMMYT organized in 1964 the International Spring Wheat Field Nursery; other CIMMYT-coordinated international networks on wheat, triticale and barley were established in 1974. CIAT initiated the International Bean Yield and Adaptation Nursery (IBYAN) in 1976 which now counts with participants institutions from 65 countries in all the world (IBYAN reports, CIAT 1976-1986). In 1979, under IDRC financial support, the International Network for Tropical Pastures Evaluation (RIEPT, in Spanish) was created, being CIAT the coordinating institution. The RIEPT now counts with participant institutions from 20 countries in Latinamerica and Caribe (RIEPT reports, CIAT 1982-85, 1985-89). ILCA coordinates the Tripanotolerance Network in Africa, which unites the effort of 20 scientists working on this livestock parasitic disease at other international or national institutions in Africa (Plucknett and Smith, 1984).

2. OBJECTIVES, EXPERIMENTAL DESIGN AND INFORMATION GENERATED

One of the most important activities of the International Centers for Agricultural Research, members of the CGIAR System, is the development or identification of superior germplasm (technology) in close partnership with national agricultural research and development institutions of the developing world. This objective is largely accomplished through the "International networks for varietal evaluation", a multi-institutional effort for the evaluation, selection and distribution of improved germplasm amongst collaborating Institutions. The final responsibility for varietal release as commercial cultivars to their country growers is in the hands of each national institution. (Toledo, 1982; Seshu, 1984).

International networks can be conceived as a germplasm-improvement system combined with a germplasm-delivery system. Network trials help to identify more stable materials for use in breeding programs, but also, they can be a mechanism for final distribution of promising finished lines or promising germplasm collected elsewhere. In a network, the best materials - either produced through breeding or through introductions or collections - are tested for their adaptation and production in a wide range of environments. Varietal evaluation is carried-out in multilocal trials located in representative sites of the target area. Following this line of thought, the central objective of a germplasm evaluation network is to give its members -the participant institutions- reciprocal assistance in two major aspects:

(1) the supply of new germplasm alternatives, from which each region can identify materials to help solve its constraints, and (2) the possibility of attaining -by means of a thorough analysis of the information generated by the network- a precise extrapolation to other environments within the region, on the performance of materials that have shown promise in network sites that are representative of such environments. Thus, the two main products of a Network to be shared among the participant institutions are: improved germplasm and information.

Sites Selection:

In order to provide a network with solid extrapolation capacity to a target region it is necessary to select evaluation sites, evaluation seasons, and evaluation period lengths that represent the existent variability in the region in terms of soil, climate and management practices. The region may be first stratified into major ecosystems and production systems. Subsequently, representative sites must be selected for each ecosystem and production system, in order to cover the actual range of macro-variability.

Let us take as an example, the tropical pastures evaluation network (RIEPT). For the initial selection of network sites in the target region -Tropical America frontierlands from Mexico to Bolivia-, the classification made by Cochrane (1982) in five main ecosystems was used, and representative sites were chosen within each ecosystem (Toledo, 1982). The number of RIEPT sites selected per ecosystem varied with the type of trial (see table 1); more numerous for ERA (small-plot agronomic cutting trials to evaluate monocultures of grasses and legumes for their survival to constraints of soil, climate and biotic factors) and less numerous for ERD (large-plot grazing trials, for the final selection stage, to evaluate grass-legume associations in terms of meat or milk production).

Table 1: Number of RIEPT sites selected per ecosystem for each type of trial

Ecosystem	TYPE OF TRIAL		
	Agronomic evaluation trial (ERB)	Grazing Trial evaluations in small plots (ERC)	Grazing Trial in large plots (ERD)
Well-drained Savannas	46	14	6
Tropical Semievergreen seasonal forest	93	11	20
Tropical rain forest	62	10	10
Total	201	35	36

Source: RIEPT database - State at Dec. 1990
(RIEPT = International Network for the evaluation of tropical pastures)

The experimental design of each type of trial was done taking into consideration the most important sources of variability on pastures performance. For agronomic evaluation trials, (ERA and ERB), soil variation, seasonal variation, and growth habit of the different pasture species under testing, were considered (see Toledo, J.M. (ed), 1982): table 2 illustrates, for example, the dramatic effect of season on grass and legume biomass production under cutting trials. For grazing trials evaluation of legume-grass pastures (ERC and ERD), additional sources of variability were considered, such as among-year variation, season within year variation, variation in agronomic performance between different pasture associations, variations due to the type of experimental animal in terms of breed, age, sex and health conditions, and within-animals variability (see Amézquita, M.C.; 1986): table 3 illustrates the effect of season and among-year variation on weight gain of steers under grazing trials.

Table 2. Pasture trials under cutting. Effect of season on grass and legume biomass production (kg DM/ha, 12 weeks after regrowth).¹

Ecosystem	Season					
	Maximum Rainfall			Minimum Rainfall		
	No. of trials	Mean	CV ² (%)	No. of trials	Mean	CV (%)
Grasses						
Tropical Forest	66	5873	36	66	3571	36
Plains	22	3702	37	21	918	90
Legumes						
Tropical Forest	67	2555	38	68	1684	46
Plains	22	2019	39	23	527	145
Cerrado	10	3271	35	11	377	65

1. SOURCE: Regional Trials "B" (first year of evaluation), conducted on three major ecosystems from tropical America. Original data per trial were taken from RIEPT database, Tropical Pastures Program, CIAT.
2. CV $\sqrt{\text{MSE}/\bar{x}}$, where: MSE = Mean square of the combined error over all trials conducted on this ecosystem. \bar{x} = Mean biomass production for the ecosystem.

Table 3. Pasture trials under grazing. Effect of season and year of evaluation on the production of pastures, the latter expressed as a weight gain of the steers.

Season/year	Number of steers considered	Weight gain ² (grams/animal per day)	CV (%)
Rainy			
1979	73	495	21.4
1980	85	548	14.8
1981	85	478	17.2
1982	40	397	16.4
Dry			
1979	73	140	102.5
1980	85	129	97.3
1981	85	192	64.1
1982	40	252	95.1
Rainy	283	495 a	17.2
Dry	283	175 b	88.3
Years:			
1980	146	407 a	42.3
1981	170	383 ab	32.8
1979	170	376 ab	48.4
1982	80	348 b	47.7

1. The information comes from seven trials on steer weight gain, carried out at Corimagua, Eastern Plains of Colombia, between 1972 and 1982, by the CIAT's Tropical Pastures Program.
2. The averages followed by the same letter do not differ statistically ($P \leq 0.5$). (Tukey test for multiple comparisons).

SOURCE: Tergas et al., 1984a and 1984b.

The previous example illustrates the importance of considering relevant sources of variability for a given crop in the selection of network sites and in the experimental design of trials. This will improve the extrapolation capacity of the network.

Experimental design of network trials:

Network trials can follow an identical experimental design across testing sites (Uniform Networks) or different experimental designs in the various testing sites (Heterogeneous Networks), depending on the specific characteristics and research needs of the particular crop. From the point of view of the statistical analysis of the network information and its storage, it is important to distinguish between these two situations:

- (1) Uniform networks, in which the evaluation of materials follows an experimental design which is identical for all the sites within the network: the same germplasm is evaluated, under similar agronomic and management practices, during similar evaluation times; a common set of response variables is reported for all the sites under study, and standard measuring techniques are used throughout the network. Under these conditions, it is possible to use standard formats for data registration and storage, and to use the same model for performing statistical analysis per site. Likewise, any type of multilocational analysis is feasible.
- (2) Heterogeneous networks: the evaluation of germplasm follows different experimental designs per site. This may be due to variation in the evaluated germplasm between sites, or to the need of applying different management factors to the germplasm according to specific characteristics of the ecosystem, production system, or germplasm types under study; to the need of evaluating the materials under varying times and testing frequencies depending on the environment; or to the fact that plant responses must be expressed through different response variables, relevant to the site's specific ecosystem and production system. Under these conditions, a standard form of data registration, storage, and statistical analysis through the network is no longer possible. The analyses for each site must follow the corresponding experimental design -which is site-specific and must fit the particular needs of each environment-. Multilocational statistical analyses are more difficult since experimental factors vary with evaluation sites. The aim of storing the information produced by a heterogeneous network involves the definition of data registration formats that must be general enough to allow the storage of information which is not standard for all sites within the network.

Let us continue with our previous example, the trials design used by the International Network for Tropical Pastures evaluation (RIEPT) (see Toledo, J.M., 1982; Toledo, J.M. and Schultze-Kraft, R. 1982; Paladines, O. and Lascano, C., 1983; Lascano, C. and Pizarro, E.A., 1986). In the RIEPT, legumes and grasses which survive constraints of soil, climate,

and biotic factors from each ecosystem--material been evaluated in Regional Trials "A" (ERA)--are selected for evaluation in terms of their establishment capacity and seasonal biomass production potential under cutting in Regional Trials "B" (ERB). The ERB's use an experimental design which is identical for all sites: split plots, where the main plot is represented by the ecotype--with a nested arrangement in such a way that ecotypes belonging to the same species are located in contiguous plots--and the sub-plot is represented by the evaluation times of the material under cutting. Thus, the ERB's represent a uniform agronomic evaluation network for legume and grass monocultures evaluation under cutting.

On the other hand, pastures that enter the small-plot grazing trials evaluation stage (500 m², under Regional Trials "C" (ERC)) and large plot grazing trials evaluation stage (> 1 ha under Regional Trials "D" (ERD)), consist of tropical grasses and legumes selected from the ERB and which will be evaluated for their compatibility in grass-legume mixtures under cattle trampling and under different cattle management systems; pasture response is measured in terms of animal production (meat, milk). There is not a unique appropriate experimental design for grazing trials. Depending on the particular ecosystem, even the same pastures under the same treatments can require different management practices and dimensions for the experimental unit; they may require a different number or type of animals per experimental unit, depending on the prevailing ecosystem and production system. Thus, the response variables measured at each site of the network are not necessarily comparable (Amézquita, M.C., 1986). The ERC and the ERD of the RIEPT thus exemplify an heterogeneous network for the evaluation of pastures under grazing.

Information generated

The information generated by a network for varietal evaluation is of two types: a) Data on crop performance across a wide range of environments; and b) Environmental descriptors of the testing sites during the evaluation periods (climate, soil and biotic factors).

Data on crop performance include plant response parameters measured on each experimental plot at each network site. Plant response parameters include morpho-agronomic variables, indicators of plant adaptation to the environment; scores on plant reaction to specific pests and diseases considered relevant to the particular crop and production system, under study; scores on plant tolerance to specific limiting factors in the soil; and final crop yield. With **short-cycle crops** whose final product can be measured at harvest, like beans, cassava, maize, rice, plant response parameters normally require a one-time evaluation during the crop cycle. One single type of trial across network sites can be designed to test the germplasm; environmental sources of variation affecting crop response, like "climate" or "soil" are taken into account by evaluating the germplasm in a wide range of network sites with contrasting climatic and soil conditions. With **perennial crops** whose final product can be measured at harvest, like fruit-trees or forestry research, plant response parameters require periodic evaluations in

time, in order to measure plant performance at different stages of development, under contrasting seasonal conditions. Germplasm can be tested in one single type of trial across a wide range of locations. On **perennial crops whose final product cannot be measured at harvest**, like pastures for example, whose final product is meat or milk, germplasm needs to be tested through many different types of multilocal trials, as exemplified in RIEPT (Toledo, 1982). Each type of trial represents a different selection stage to guarantee that only very promising material are tested under the more advanced and more expensive final trials.

Information Flow:

The information generated by a network represents a very valuable tool to support decisions concerning selection for further evaluation and final release of varieties as commercial cultivars. This is why an effective management of this information is of vital importance.

The responsibility for the centralization and management of the information generated by a network is normally assigned to one of the network member institutions: the **central node**. This institution is chosen based on its overall capacity and resources, including facilities in the areas of computing, data processing and statistical expertise. The Biometry/data analysis Unit, or Data Processing Unit of that institution is naturally involved: it provides human expertise in support to the management and statistical analysis of the network information.

Figure 1 illustrates the steps needed for transforming "data" generated by a network, into "processed information," so it can then be passed to network members to assist them on technical or strategic decision-making. The process here described - which involves data cleaning, data processing, data analyses, storage/retrieval of results and report production - is the responsibility of the network coordinator supported by the Biometry/data analysis Unit of the institution assigned the centralization and management of the network's information. Data generated by network members are first submitted to a cleaning process: correction of syntactical and biological inconsistencies; then, individual analyses are produced per site, aimed at identifying promising germplasm for that specific environment; after selecting a number of sites, based on criteria that match the different objectives of the analysis, multilocal analyses are performed for a target region, country, or ecosystem; the results of these analyses (individual or multilocal) are then shared with all network members.

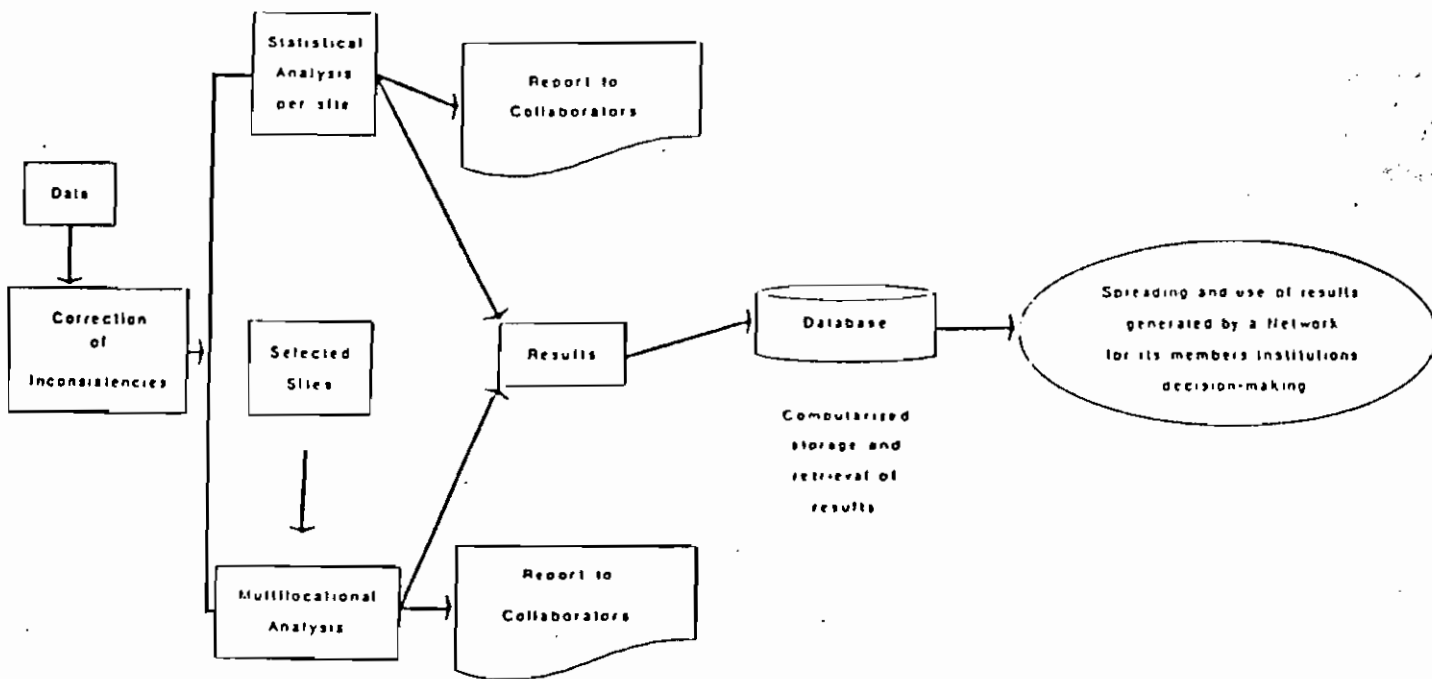


Fig. 1: International Networks for varietal evaluation

- DATA FLOW -

The **network's database** is a computerized storage and retrieval system which contains the information produced by the network. It keeps environmental descriptors per site, and selected results on crop performance analysis per site and per region or ecosystem depending on the network needs. For example, after nine years of RIEPT work, its statistical database stores information on the individual performance, at each network site, of 234 grass and legume ecotypes belonging to a wide variety or species of 6 grass genus and 15 legume genus, evaluated by 18 national institutions throughout the five main ecosystems of Tropical America. (Franco, et al; 1987). However, the creation of a database is only the first step. Its relevance lies on its efficient utilization as a tool for the decision making-process of network members.

3. WHY A MULTILOCATIONAL ANALYSIS.

The multilocal analysis of a germplasm evaluation network can have two purposes: a) to identify more widely adapted materials, for breeding purposes or commercial use, and produce recommendations which can be extrapolated to a target region; and b) to study the influence of specific external conditions, such as soil, climate and biotic factors, on various individual plant characters.

Extrapolation to a target region:

In the first place, a multilocal analysis of a network information is useful in producing a **general description of the ecosystem**. Using the network sites as a representative sample of the ecosystem, the description can be made in terms of soil, climate, and biotic parameters, and of the ecotypes performance parameters through the range of conditions prevailing in the ecosystem. A multilocal analysis can also be performed to **identify materials** showing higher yield potential and adaptability through a wide range of environmental conditions within the ecosystem. An example is the study on the behavior of forage germplasm evaluated by the RIEPT in the savanna and tropical forest ecosystems, based on results on agronomic trials conducted between 1979 and 1982 (see Amézquita, Toledo and Pizarro, 1983). A multilocal analysis can be performed to **identify sub-areas of generalization** within a major ecosystem, in terms of the performance of a given genus, specie or variety. An example is the study "Range of Adaptation of Andropogon gayanus in Tropical America" (see Amézquita, Pizarro and Toledo, 1990), whose purpose was to classify similar environments in terms of plant performance parameters, based on data on 48 RIEPT agronomic cutting trials located between Mexico and Paraguay in which Andropogon gayanus was included. The study of the **extrapolating capacity of a given site** can also be performed using the multilocal data generated by a network. An example is the study carried-out by Cuevas, Amézquita and Rosero (1989), which presents a methodology for evaluating a location site for an international plant breeding program, based on 67 International Rice Testing Trials (IRTP), conducted in Latinamerica for a 6 year-period (1978-84).

Study the influence of specific external factors on plant characteristics

On the other hand, the interest of a multilocal analysis can be to study the influence of external factors (soil, climate, or disease-pressure levels) on one or more plant characteristics. In such a case, the locations selected for analysis play the role of "levels" of the factor whose effect one wishes to measure, and they should be chosen in such a manner that minimal confounding between the factor to be studied and other external factors exist. Different groups of locations can be selected in order to study different factors. The multilocal analysis, in this case, complies with the objective of making inferences regarding the effect of one or more **specific factors** and not of extrapolating over the region.

4. USEFUL STATISTICAL METHODS FOR DATA ANALYSIS

The wide range of statistical methods for descriptive and inferential data analysis can find practical applications in the analysis of data generated by a variety evaluation network. However, following our previous discussion, emphasis will be made on three types of methods: a) reduction of dimensionality techniques, applicable in the presence of a set of highly correlated variables; b) classification techniques; and c) statistical methods for the analysis of genotype adaptability. Also, the use of Mixed Models is of practical relevance for the analysis of models with fixed and random effects; in particular genotype x environment models in which genotype is normally considered a fixed effect and environment a random effect. Also, application of statistical methods for categorical data analysis is of interest for the analysis of binary or multi-category response variables recorded in animal reproductive experiments (conception, abortion, birth events, mortality), or in disease, evaluation trials, where disease reaction is normally a discrete score.

4.1 Principal Components analysis: a reduction of dimensionality technique

Before performing an inferential analysis, it is very important to examine whether there is some structure of correlation among the various environmental parameters measured or among the crop response variables evaluated. Thus, it might be desirable to reduce such a body of information to a lesser number of parameters, independent among themselves, which explain a large part of the variability of the original ones. It is also possible that the range of values of a given parameter may not be sufficient to explain variation in the response variable of interest. In both situations it is necessary to reduce the dimensionality of the problem. In the second case, simply by eliminating parameters - of soil, climate, or biotics - whose range among locations may be so narrow that does not allow the response of the plant to that parameter to be explained. In the first case, in the presence of a structure of correlation among variables, the use of the Principal Components Analysis technique is adequate.

The Principal Components analysis, developed by Pearson (1901) and later by Hotelling (1933), is a multivariable method which examines the relationships among many quantitative variables, and based on these relationships, can produce a reduced set of new variables. Given a body of p numeric variables, the method produces p "principal components" that are linear combinations of the original variables. The coefficients of each linear combination are the eigen-vectors of the correlation matrix. The principal components obtained are normalized variables (with zero mean and standard deviation of 1) and independent among themselves. The first principal component has the maximum variance; the second, the next variance in magnitude; the last, the minimum variance. The number of principal components that one wishes to retain is a subjective decision.

4.2 Cluster Analysis: a classification technique

The purpose of the Cluster Analysis is to place objects into groups or clusters suggested by the data, not defined a priori, such that objects in different Clusters tend to be similar to each other in some sense, and objects in different clusters tend to be dissimilar (Everitt, 1980). A vast number of clustering methods have been developed in several fields, with different definitions of clusters and similarity among objects.

Types of clusters.- There are several types: **Disjoint clusters**, place each object in one and only one cluster. **Hierarchical clusters**, are organized so that one cluster may be entirely contained within another cluster, but no other kind of overlap between clusters is allowed. **Overlapping clusters**, can be constrained to limit the number of objects that belong simultaneously to two clusters; or they can be unconstrained allowing any degree of overlap in cluster membership. Fuzzy clusters, are defined by a probability or grade of membership of each object in each cluster. **Fuzzy clusters** can be disjoint, hierarchical, or overlapping.

Similarity among objects.- An object can be represented by a vector of data values corresponding to the object attributes. A measure of similarity among two objects is a numeric value indicating how "near" or how "similar" are the two objects. Various measures of similarity are used for cluster analysis: the Euclidean distance between the two vectors representing the objects; or a correlation coefficient between the two objects when the object attributes are measured in the same unit, are examples of similarity measures. SAS clustering procedures -PROC CLUSTER, for hierarchical cluster analysis, and PROC FASTCLUS, for disjoint cluster analysis- utilize the Euclidean distance. (See SAS User's Guide: Statistics).

Clustering Methods: Many methods for clustering objects have been developed. They differ in how the distance between two clusters is computed. The clustering methods available are: average linkage, the centroid method, complete linkage, density linkage, maximum likelihood, the flexible-beta method, Mc.Quinty's similarity analysis, the median method, single linkage, two-stage density linkage, and Ward's minimum variance method. Many simulation studies comparing various methods of cluster analysis have been performed. Milligan (1981) presents a review of such studies. In most of these studies, the clustering method with the best overall performance has been either average linkage, or Ward's minimum variance method. In Ward's minimum variance method, the distance between two clusters K and L is defined as follows:

$$D_{KL} = \frac{1}{2} \left\| \bar{X}_K - \bar{X}_L \right\|^2 / (1/n_K + 1/n_L)$$

where, D_{KL} = distance between Clusters K and L

\bar{X}_K, \bar{X}_L = vector of attribute's means for Clusters K and L, respectively

n_K, n_L = number of objects in Clusters K and L, respectively

Number of Clusters: All hierarchical clustering methods are based on the usual agglomerative hierarchical procedure: each object begins in a cluster by itself; the two closest clusters are merged to form a new cluster replacing the two old clusters; merging of the two closest clusters is repeated until only one cluster is left. However, there are no satisfactory methods for determining the optimum number of clusters in the population (Everitt, 1980). There are various criteria, or statistics, useful for estimating the number of clusters; however, these statistics are based on different assumptions related to specific distributions from where the data was sampled. Milligan and Cooper (1983) compared 33 methods for estimating the number of population clusters using four hierarchical clustering methods. The three statistics that performed best in these simulation studies were: a pseudo F-statistics, due to Calinski and Harabasz, (1974), that measures the separation among all the clusters at the current level of the hierarchy; a pseudo t -statistic, due to Duda and Hart (1973), which measures the separation between the two clusters most recently joined; and the Cubic Clustering Criterion, (CCC) due to Sarle (1983). Also the R -statistic, or squared multiple correlation coefficient, which measures the proportion of the total variance among the objects, accounted for by the clusters; and the Semipartial R , measuring the proportion of variance accounted for, resulting from joining the two clusters at a given level in the hierarchy, are very useful. It is always desirable to look at your data graphically to try to identify the number of real clusters. Plots of R against the number of clusters are useful. Also, it may be advisable to look for a consensus among the three best-performed statistics: that is, local peaks of the CCC and pseudo F statistic, combined with a small value of the pseudo t -statistic, and a larger pseudo t for the next cluster fusion. (Massart and Kauffman; 1983).

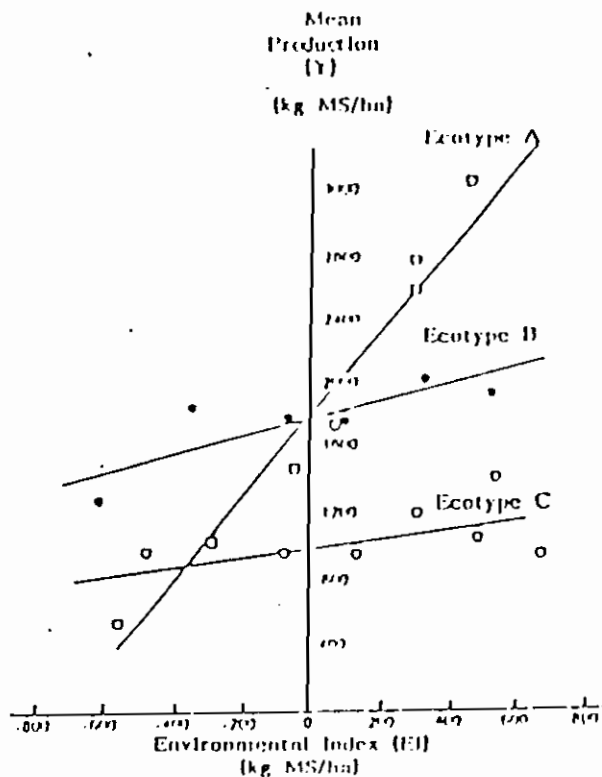
4.3 Statistical methods for the analysis of genotype adaptability

The adaptability of a genotype is defined as "its physiological response to improvements in environmental quality." (Eberhart and Russell, 1966). For environmental quality we understand the combination of soil conditions, climate, pests, diseases, weeds, and aspects of the management of vegetable material (establishment, evaluation techniques, sampling errors, among others). To quantify "environmental quality," several alternatives have been proposed. The most accepted is to express it by means of the overall mean of crop yield in that particular environment. The yield, as a resultant factor of the interaction of soil, climate and biotic plant factors, expresses the potential quality of that environment for the growth of a given genotype of the crop.

In the last 20 years, the literature shows broad revisions of statistical methods available to analyze the adaptability of genotypes, or methods for the analysis of the genotype x environment interaction. The methods used the most can be grouped into three: regression methods, multivariate methods, and geometric methods. In all of them, the environmental index is measured as the average response of genotypes which grow in that environment.

The regression methods (Finlay and Wilkinson, 1963; Eberhart and Russell, 1966) assume that a linear relationship exists between the response of a genotype and the environmental index (see Figure 2). The slope of the linear regression of each genotype represents a first adaptability index of the genotype; the sum of the squares of the variations around the regression represents a second measurement of adaptability. This method has been criticized for three fundamental reasons: (1) the resultant adaptability indexes are very sensitive to extreme values, that is, the index is heavily dependent on the group of locations selected; (2) the two measurements of adaptability - the slope of regression and the SS of the variations around the regression - are not independent; and (3) the response of the genotype is not independent from the "environmental index" estimated as the site mean. Nevertheless, with a good choice of environments according to the objective of the analysis - and with a modification in the calculation of the "environmental index," by means of the exclusion from the overall mean of the specific ecotype under evaluation, this method has shown

Fig. 2: Possible options for the Linear Regression between ecotype productivity (Y) and environmental index (EI).



itself to be useful in the identification of promising material. For example, (see case study of chapter 5) by utilizing this method, the Stylosanthes capitata's 1405, 1315, 1693, 1342, 1728, and 1318 were identified as promising legumes for the plains ecosystem (a well-drained, isohyperthermic savanna); and the Stylosanthes guyanensis 136 and 184 and the Zornia latifolia 728 as promising for the tropical forest ecosystem, through the first analysis of data of the Regional Trials "B" 1979-1982. In fact, these ecotypes have shown excellent performance in their corresponding ecosystems afterwards. (CIAT Tropical Pastures Program Annual Reports 1980, 1981, 1988).

In general, statistical methods to analyze the adaptability of genotypes through a wide range of environmental conditions, assume that the entire set of genotypes is evaluated in the same set of locations during the same years. That is, these methods are valid for balanced datasets. However, Digby (1979) developed a method for adaptability analysis for unbalanced datasets: it is valid when not all the genotypes are evaluated through all the locations or years.

The method is based on the model (non-linear in the parameters)

$$Y_{ij} = \mu_i + b_i A_j + \epsilon_{ij} \quad , \quad \sum_{j=1} A_j = 0$$

yield of variety i in environment j

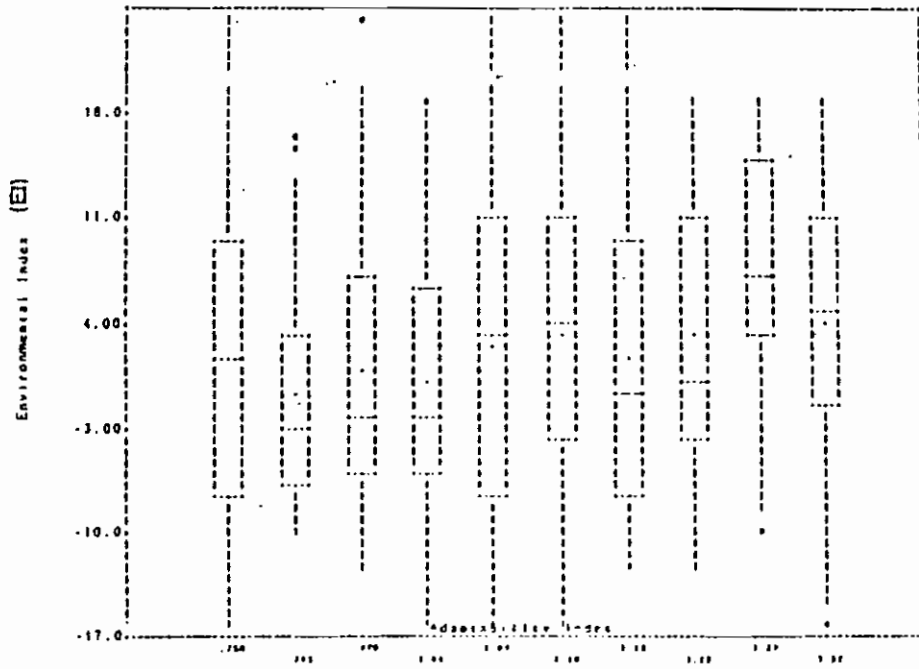
expected mean for variety i in an "average" environment

Adaptability coeff. for variety i

Relative environmental performance with respect to an "average" environment.

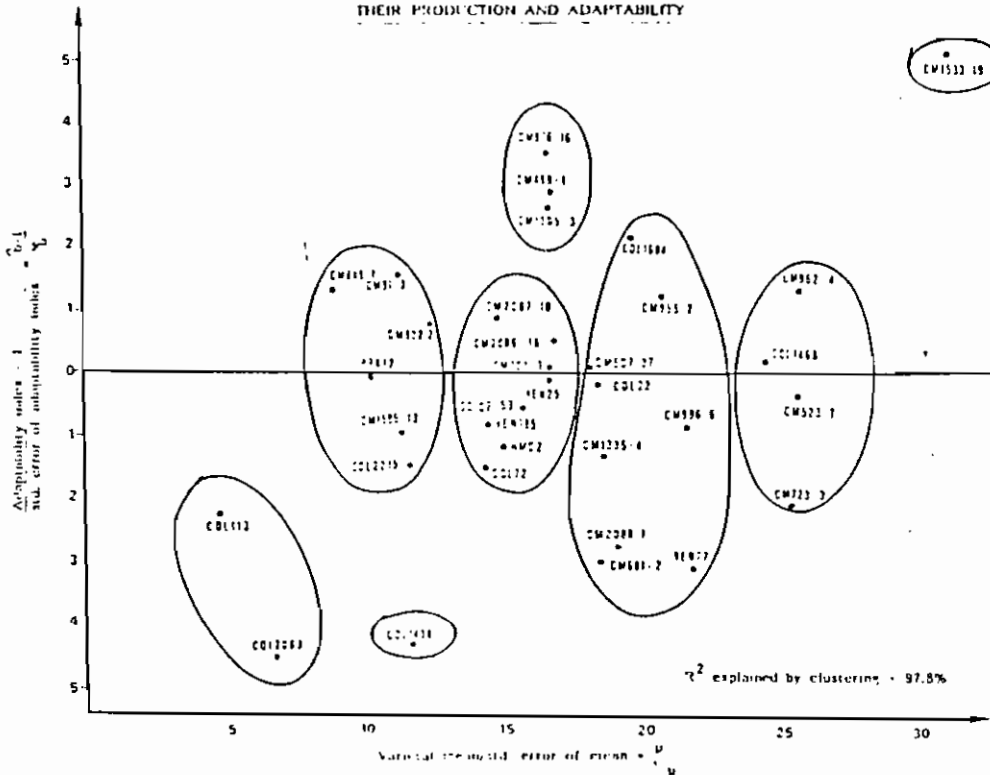
It permits statistical comparisons of varieties tested in different sets of locations and years (as μ_i estimates the expected mean of variety μ_i in an average environment). Also the method allows the comparison of sites where not the same set of varieties were tested (as A_j estimates an "environmental index" expressed as the environment yield potential with respect to the average environment). This method was applied for the adaptability analysis of 34 CIAT cassava varieties evaluated in 53 cassava yield trials conducted between 1979-1986 in five ecozones of Colombia (Hershey, C.; Granados, E. CIAT, work in progress). Based on this method, adaptability indexes were estimated and a classification of the 34 cassava varieties was made, using as classification criteria: $(\mu_i/\text{st. error of } \mu_i)$ and $(b_i/\text{st. error of } b_i)$. Figure 3 shows 10 of the varieties with their adaptability index (horizontal axis) and range of environments where they were tested (vertical axis). Figure 4 shows the 8 variety groups obtained (Ward's Minimum Variance clustering method).

Fig. 3: Adaptability Index for ten cassava varieties and range of environmental conditions (EI) where they were tested.



Source: Hershey, C. and Granados, E. • CIAT (Work in progress).

Fig. 4. CLASSIFICATION OF CASSAVA VARIETIES ACCORDING TO THEIR PRODUCTION AND ADAPTABILITY



Multivariate methods. Two multivariate methods have been mainly used in the analysis of the genotype x environment interaction: the Principal Components Analysis, and the Cluster analysis.

The Principal Components Analysis has been applied either to the set of genotype-response variables, to represent the set of genotypes in a smaller number of dimensions, or to the set of environment-response variables to represent the set of environments in a smaller number of dimensions, or to both (Freeman and Dowker, 1973; Hill and Goodnight, 1981).

An application of the Cluster Analysis method to better understand genotype x environment interaction is to group environments in which the relative position of the tested genotypes is similar; or to group genotypes, using as classification criteria the relative position of the genotype in each one of the environments, so that genotypes belonging to the same resultant group will show similar behavior across environments. In the resulting groups of environments, the genotype x environment interaction is minimal: the SS of genotype x environment interaction in each group is less than the corresponding SS using all the environments (Abou-El-Fittough et. al. 1969; Mungomery et. al. 1974; Byth et. al., 1976; Lin and Thompson, 1975; Fox and Rosiete, 1982). An example of the application of this method is the analysis carried out at CIAT on the data of grain yield (kg/ha) of bean varieties (Phaseolus vulgaris L.) tested in the International Bean Yield and Adaptation Nurseries (IBYAN) over 8 years (1976-1984). (See Voysest, O. and Garcia, J., 1987.) A Cluster Analysis was performed to group 19 locations in which 11 black bean genotypes were tested in the IBYAN 1984 nursery. The relative yield position of each genotype in the locations were used as classification criteria. The resulting six clusters or groups of locations were expected to exhibit minimum genotype x environment interaction. In order to verify the effectiveness of the grouping of locations in terms of the reduction of the genotype x environment SS in each group, an analysis of variance overall locations was performed and the SS for line x location was estimated, without the grouping (top part of table 4) and considering the grouping by decomposing "location" in "group" and "location (group)" (bottom part of table 4). The statistically significant "line x location" F value overall locations, of 2.5, was reduced to non-significant F values of 1.6, 1.0, 1.2 and 0.9 in groups 1 to 4. Groups 5 and 6 did not exhibit a reduction in line x location interaction.

The **Pattern Analysis** developed by the Department of Agriculture, of University of Queensland, Australia in 1977 (See GEBEI Computer Package User's Guide 1988), (GEBEI means "genotype by environment interaction"), is an implementation of the combination of both multivariate methods and analysis of variance for the analysis of varietal performance under a balanced multilocal trial design. The package is a set of eleven FORTRAN 77 sub-routines that perform, the following analyses: a) Combined analysis of variance including year, location, genotype, and replication effects; b) Adaptability analysis based on the regression

Table 4. Effectiveness of grouping of locations (via Cluster Analysis) expressed in terms of the reduction of genotype x environment interaction (see value of F) in the diverse groups.

Sources of variation	ANOVA for yield across location (kg/ha)		
	Df.	F	Prob. F
		<u>Before grouping</u>	
Location	18	132.6	0.0001
Repetition (location)	38		
Line	10	1.3	0.26
Line x location	180	2.5	0.0001
Combined error	380		
Total	626		
		<u>After grouping</u>	
Group	5	290.0	0.00001
Location (group)	13	72.0	0.001
Repetition (location, group)	38		
Line	10	1.3	0.26
Line x group	50	2.7	0.0001
Line x location (within group)	130	1.7	0.0001
Line (Group 1)	30	1.6	0.3
Line (Group 2)	20	1.0	0.5
Line (Group 3)	30	1.2	0.2
Line (Group 4)	20	0.9	0.5
Line (Group 5)	20	3.2	0.0002
Line (Group 6)	20	5.4	0.0001
Combined error	380		
Total	626		

1. Analysis of the International Bean Yield and Adaptation Nurseries (IBYAN). Year 1984, black grain color. Eleven lines tested in 19 locations (Source: Voysest, O. and Garcia, J. 1987).

method; c) Classification of genotypes based on their mean yield in each environment (via Ward's Minimum Variance clustering method) and classification of environments (an environment is a site evaluated in one particular year) in terms of the genotype yields and GxE effects (via Cluster Analysis); d) Principal Component Analysis on genotypes to reduce their dimensionality, i.e. to represent the set of genotypes in a smaller number of dimensions; also Principal Component Analysis on environments to reduce their dimensionality; e) Graphics of the responses of the genotype groups versus environment groups.

The geometric methods have as a basic objective to represent each object (genotype or environment in this case) as a point in a Euclidean space in such a manner that similar objects are represented by nearby points. The configuration of the points is then investigated in order to try to detect some structure. These methods, contrary to the Cluster Analysis, do not form groups of objects. The geometric methods include: Principal Coordinates Analysis (Classical Scaling), Non-metric Multidimensional Scaling, the Biplot Method, and Correspondence Analysis.

The Principal Coordinates method is the only geometric method that has been used to study the adaptability of ecotypes. Given a matrix of distances between points in a Euclidean space, this method (developed by Schoenberg, 1935) determines the coordinates of the points: those are the "Principal Axes" or "Principal Coordinates," independent among themselves. The Principal Coordinates method was used by Crossa, Westcott, and González (1986) to analyze the stability of maize populations of CIMMYT (International Center for Maize and Wheat Improvement), measured according to the behavior of their descendant

varieties, evaluated at 80 locations in a Randomized Complete Block Design with four replicates. The limitation of the Principal Coordinates method is that the computational algorithm is not available yet in statistical packages.

5. CONCLUDING REMARKS

The aim of this paper is to emphasize the value of the information generated by International Networks for varietal evaluation, the type of questions that can be answered through its analysis and the characteristics and use of various statistical methods found useful in the analysis and interpretation of the network information. The discussion here presented is intended to serve as a frame of reference for agricultural researchers and biometricians working with International Networks in the developing world. For the agricultural researcher, as a way of providing him/her with ideas on how to best utilize the Network's information resource to answer relevant research questions. For the biometrician, as a way of offering a very challenging opportunity to apply quantitative methods to the better understanding of the behavior of newly generated germplasm across a wide range of environmental conditions.

REFERENCES

- Amézquita, M.C., Pizarro, E.A. and Toledo, J.M. 1990. Range of Adaptation of Andropogon gayanus. In: Andropogon gayanus Kunth. A grass for tropical acid soils, CIAT, Cali, Colombia. p.37-64.
- Amézquita, M. C. 1986. Consideraciones sobre planeación, diseño y análisis de experimentos de pastoreo. In: Lascano, C. and Pizarro, E. A. (eds.). Evaluación de pasturas con animales: alternativas metodológicas. Red Internacional de Evaluación de Pastos Tropicales, CIAT, Cali, Colombia. p. 13-40.
- Calinski, T. and Harabasz, J. (1974). "A dendrite method for cluster analysis". Communications in Statistics, 3, 1-27.
- CIAT Tropical Pastures Program Annual Reports. 1980, 1981, 1988.
- Cochrane, T.T., 1982. Caracterización agroecológica para el desarrollo de pasturas en suelos ácidos de América Tropical. In: Toledo, J.M. (ed.). Manual para la evaluación agronómica; Red Internacional de Evaluación de Pastos Tropicales. Centro Internacional de Agricultura Tropical (CIAT), Cali, Colombia. p.23-44.
- Crossa, J.; Westcott, B.; and González, C. A. 1986. Yield stability in some CIMMYT maize populations as measured by the experimental varieties. CIMMYT, Mexico. (In preparation.)

- Cuevas-Perez, F., Amézquita, M.C. and Rosero, M. 1989. A methodology for evaluating a location as a selection site for an international plant breeding program. CIAT, Cali, Colombia. *Euphytica* 43: 165-172.
- Digby, P.G.N., 1979. Modified Regression Analysis for incomplete variety x environment data. *J. of Agricultural Sciences Cambridge*, 1979, 93, p. 81-83.
- Duda, R.O. and Hart, P.E. (1973). Pattern classification and Scene analysis. New York: John Wiley and Sons.
- Eberhart, S.A. and Russell, W.A., 1966. Stability parameters for comparing varieties. *Crop Science* 1:135-140.
- Everitt, B.S., 1980. Cluster Analysis, Second edition, London: Heineman Educational Books Ltd.
- Franco, M. A.; Amézquita, M. C.; and Toledo, J. M. 1987. Base de datos estadística-RIEPT: Información disponible y opciones para su utilización. CIAT, Cali, Colombia.
- Freeman, G. H. 1973. Statistical methods for the analysis of genotype-environment interactions. *Heredity* 31:339-354.
- Freeman, G. H. and Dowker, B. D. 1973. The analysis of variation between and within genotypes and environments. *Heredity* 30:97-109.
- Finlay, K.W. and Wilkinson, G.N. 1963. The Analysis in a Plant Breeding Programme. *Aust. J. Agric.*, 1963, Vol. 14, 742-54.
- GEBEI Computer Package Manual 1988. University of Queensland.
- Hill, J. 1975. Genotype-environment interactions: A challenge for plant breeding. *Journal of Agricultural Science* 85:447-493.
- Hill, J. and Goodchild, N. A. 1981. Analysing environments for plant breeding purposes as exemplified by multivariate analysis of long term wheat yields. *Theoretical and Applied Genetics* 59:317-325.
- Hotelling, H. 1933. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology* 24:417-441, 498-520.
- IBYAN reports, CIAT 1976-1986
- IRTP reports: IRRI/CIAT 1977 - 1985
- IBYAN reports: CIAT 1976-1986

- Lascano, C. and Pizarro, E.A. 1986. Evaluación de pasturas con animales. Alternativas Metodológicas. Red Internacional de Evaluación de Pastos Tropicales.
- Mashler, W.T. 1984. "The network approach: United Nations Development Programme in Agriculture". In: Biotechnology in International Agricultural Research. Proceedings of the Inter-Center Seminar on International Agricultural Research Centers (IARCs) and Biotechnology. IRRI, Philippines. 340-344.
- Massart, D.L. and Kauffman, 1983. The interpretation of analytical chemical data by the use of Cluster Analysis. New York: John Wiley and Sons.
- Milligan, G.W. and Cooper, M.C. (1983). "An examination of procedures for determining the number of clusters in a data Set". College of Administrative Science Working paper Series 83-51, Columbus: The Ohio State University. To appear in Psychometrika, 1985.
- Paladines, O. and Lascano, C.. 1983. Germoplasma Forrajero bajo Pastoreo en pequeñas parcelas. Metodología de Evaluación.
- Pearson, K. 1901. On lines and planes of closest fit to systems of points in space. Philosophical Magazine 6(2):559-572.
- Plucknett, D.L. and Smith, N.J.H. 1984. Networking in International Agricultural Research. Science, Vol. 225, 989-993.
- RIEPT reports 1982-1985, 1985-1989. CIAT publications.
- SAS User's Guide Statistics. Version 5. SAS Institute Inc. Raleigh, NC, USA.
- Sarle, W.S. (1983). "The Cubic Clustering Criterion", SAS Technical Report A-108. Cary, N.C.: SAS Institute Inc.
- Schoenberg, I.J. 1935. Remarks to Maurice Frechet's article 'Sur la definition axiomatique d'une classe d'espace distancias vectoriellement applicable sur l'espace d'Hilbert'. Annals of Mathematics (2nd Series), 36, 724-732.
- Seshu, D.V. 1984. "An International Network for Rice Testing and Evaluation". In: An Overview of Upland Rice Research. Proceedings of the 1982 Bouake, Ivory Coast Upland Rice Workshop, IRRI, Philippines. 551-559.
- Silvey, V. 1982. Analysis of crop variety adaptation from performance trials in England and Wales. In: Proceedings in the XIth International Biometric Conference, Toulouse, France. p. 157-163.

- Tergas, L.E. 1984a. El potencial de producción animal de cuatro asociaciones de Andropogon gayanus Kunth con leguminosa en los Llanos Orientales de Colombia. Producción Animal Tropical (Rep. Dominicana). 9:176-186.
- Tergas, L.E. 1984b. Productividad animal de Brachiaria decumbens sola y con pastoreo complementario en Pueraria phaseoloides en los Llanos Orientales de Colombia. Producción Animal Tropical (Rep. Dominicana). (En impresión).
- Toledo, J.M., Amézquita, M.C. and Pizarro, E.A. 1983. Análisis del comportamiento del germoplasma evaluado por la RIEPT en los ecosistemas de Sabana y Bosque Tropical. In: Red Internacional de Evaluación de Pastos Tropicales. Segunda Reunión de la RIEPT, Septiembre 27-29, 1982. pag. 429-447.
- Toledo, J.M. (ed.), 1982. Manual para la evaluación agronómica. Red Internacional de Evaluación de Pastos Tropicales, CIAT publication.
- Toledo, J. M. 1982. Objetivos y organización de la Red Internacional de Evaluación de Pastos Tropicales. In: Toledo, J. M. (ed.). Manual para la evaluación agronómica. Red Internacional de Evaluación de Pastos Tropicales, CIAT, Cali, Colombia. p. 13-21.
- Toledo, J. M. and Schultze-Kraft, R. 1982. Metodología para la evaluación agronómica de pastos tropicales. In: Toledo, J. M. (ed.). Manual para la evaluación agronómica; Red Internacional de Evaluación de Pastos Tropicales. Centro Internacional de Agricultura Tropical (CIAT), Cali, Colombia. p. 91-110.
- Voysest, O. and García, J., 1987. IBYAN report 1984. CIAT publication.
- Westcott, B. (1986a). Some methods of analysing genotype-environment submitted to heredity. Submitted to Journal of Agricultural Sciences. Cambridge.
- Westcott, B. (1986b). A method of assessing the yield stability of crops genotypes. Submitted to Journal of Agricultural Sciences. Cambridge.