EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA

DEPARTAMENTO DE MÉTODOS QUANTITATIVOS

A STUDY OF SORGHUM PRODUCTION POTENTIAL IN THE SEMI-ARID TROPIC OF BRAZIL

Fernando Luís Garagorry
Everaldo Rocha Porto
Antonio Williams Moita
Aderaldo de Souza Silva
Homero Chaib Filho

DMQ/A/62

Novembro,1982

A STUDY OF SORGHUM PRODUCTION POTENTIAL IN THE SEMI-ARID TROPIC OF BRAZIL

Fernando Luís Garagorry, Everaldo Rocha Porto,
Antonio Williams Moita, Aderaldo de Souza Silva
and Homero Chaib Filho

Brazilian Agricultural Research Corporation

SUMMARY

The paper presents the application of a computerized agroclimatic system to the study of sorghum production potential in the semi-arid tropic of Brazil. After a brief description of the system, the results obtained for a sample of counties are discussed. In particular, the paper illustrates the use of clustering techniques in connection to agroclimatic zoning. The use of the system in relation to complementary irrigation studies is also stressed.

INTRODUCTION

The work on agroclimatic modeling performed at EMBRAPA has evolved over a period of six years. During all this time, the effort has been centered on the study of the semi-arid tropic of Brazil. The main interest is at the farm level; but, in order to set meaningful priorities for farming systems research, it was found that some work at the county (municipality) level would be required.

The main objectives, at the county level, have beeen the following:

(1) to assess mean productivity and best planting period for several crops (mainly beans, corn, sorghum, cotton and cowpea);

(2) to estimate the amount of excess water that, eventually, could be available to the crops by means of non-traditional irrigation techniques.

A substantial part of the county level studies is performed with a computerized agroclimatic system. In the next section, an outline of the computer system, in its present form, will be given. It evolved from a large program to a system of programs. This gives flexibility to the user, since the system can opperate with different formulas to estimate potential evapotranspiration, and the user can choose a variety of printouts or can perform his own analysis of output values.

Then, some of the results obtained through the system will be presented. They refer to sorghum production potential in 32 counties of 8 states of the semi-arid tropic of Brazil.

Finally, some indications on future work related to studies of sorghum production potential will be presented.

EMBRAPA'S COMPUTERIZED AGROCLIMATIC SYSTEM

General description

The main components of EMBRAPA's computerized agroclimatic system, in its present form, are presented in Fig. 1. In the following, a brief description of these components is given. Temporary disk files, used to pass information from one program to another, are not shown in Fig. 1.

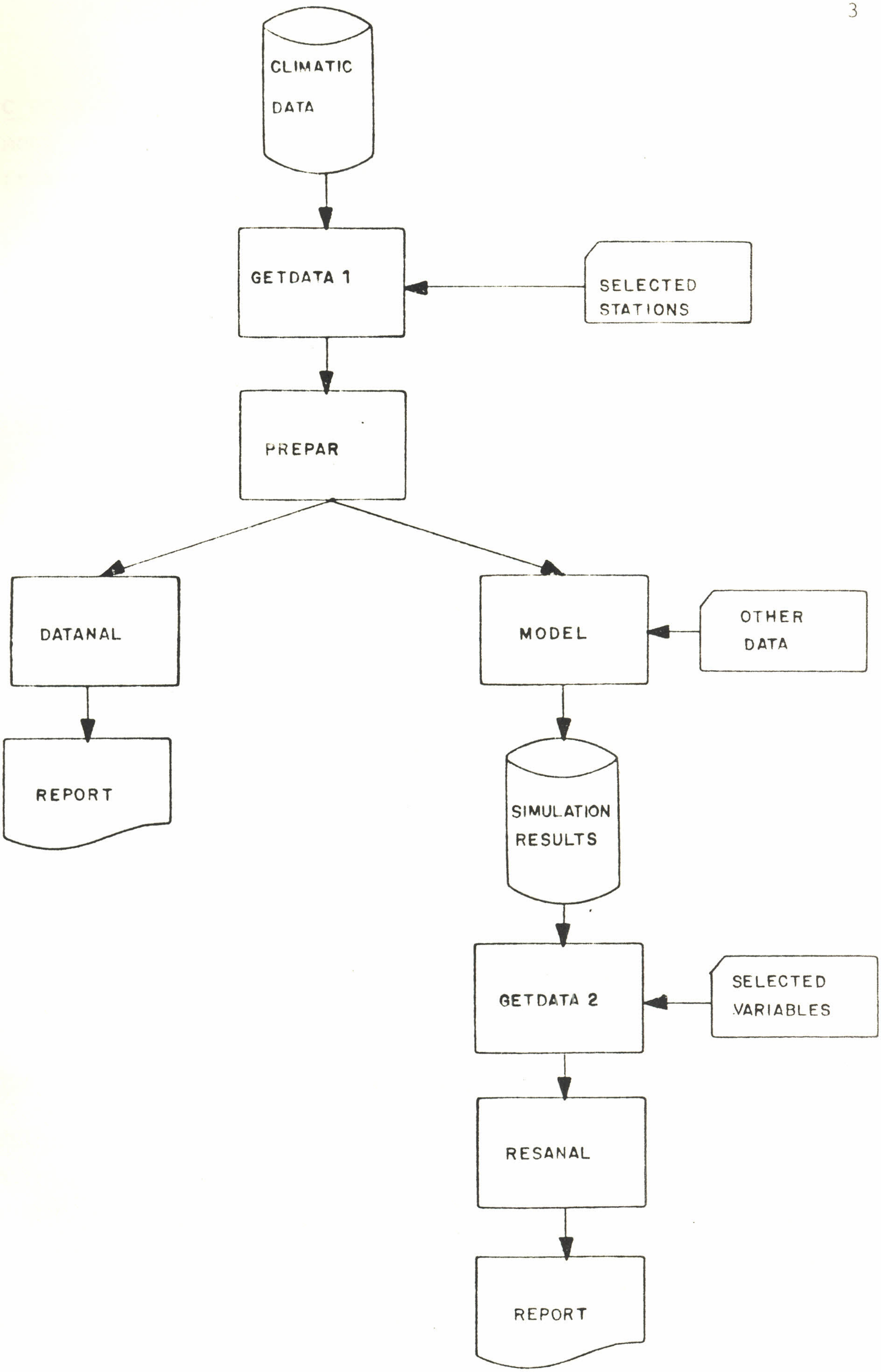


Fig. 1. EMBRAPA's computerized agroclimatic system.

CLIMATIC DATA. A meteorological data base is under development. Although most data are kept on tapes, the operation of the system starts from data placed in a disk file.

GETDATA1. This is a program that retrieves the required weather data and puts them on a temporary disk file.

PREPAR. This name designates, in fact, a small set of programs; they were developed to prepare the climatic data, in order to make easier their handling through the following steps. In particular, potential evapotranspiration is calculated at this stage.

DATANAL. This is a set of programs to analize the climatic data. They are FORTRAN or SAS programs designed to perform, among others, the following tasks: (a) to calculate descriptive statistics (totals, means, standard deviations, medians, modes, etc.); (b) to fit probability distributions and test goodness of fit; (c) to apply cluster analysis, and other classification techniques, so as to group counties according to climatic data ("climatic zoning"). Any other climatic data analysis can be performed by writing the appropriate DATANAL program and, eventually, a PREPAR program if none of the existing ones is suitable for the task.

MODEL. This is a set of simulation programs that perform a water balance and estimate productivity as a function of water stress. They use the climatic data passed by PREPAR and some other data given by the user. These last data refer to the crop and the soil, and also contain some control parameters (initial and final planting day; length, in days, of the simulation step; etc.). In most applications, the simulation step is 5 days, so that the year is divided into 73 planting periods. Some of the programs print a report; but, at present, a version that records the results is used more frequently.

SIMULATION RESULTS. This is a permanent disk file that stores the results obtained from MODEL. In fact, along with the results, the user provided data ("OTHER DATA" in Fig. 1) and some station identification informations are also recorded by the model. In this way, the user will be able to perform any analysis of the results at a later time.

GETDATA2. This is a program that retrieves the required simulation results from the previous file and puts them on a temporary disk file.

RESANAL. This is a set of programs designed to perform, among others, the following tasks: (a) to print simulation results; (b) to apply cluster analysis, and other classification techniques, so as to group counties according to their estimated productivity or water performance with respect to the given crop ("agroclimatic zoning").

Comments

The system is in continuous development along the following lines: (1) more climatic data are being recorded on tapes; (2) more programs are being added to the basic steps shown in Fig. 1.

Over the last years, it evolved from one program that prepared the data, run the water balance and printed the results , to the present situation. It can be said that the main concern has been on modularization, both at system and program levels. Each program of the system must be directed to a specific task; each routine of a program must be small (maximum of one printout page) and perform a simple function.

The model runs a water balance for each planting period of each year for which rainfall data are available. The results are collected into simple statistics (frequencies and means), which are recorded on the SIMULATION RESULTS file (Fig. 1). No rainfall probability distribution is used, since the model works with the original values. The only preparation required is a sum of the climatic data (rainfall and evapotranspiration) corresponding to each period.

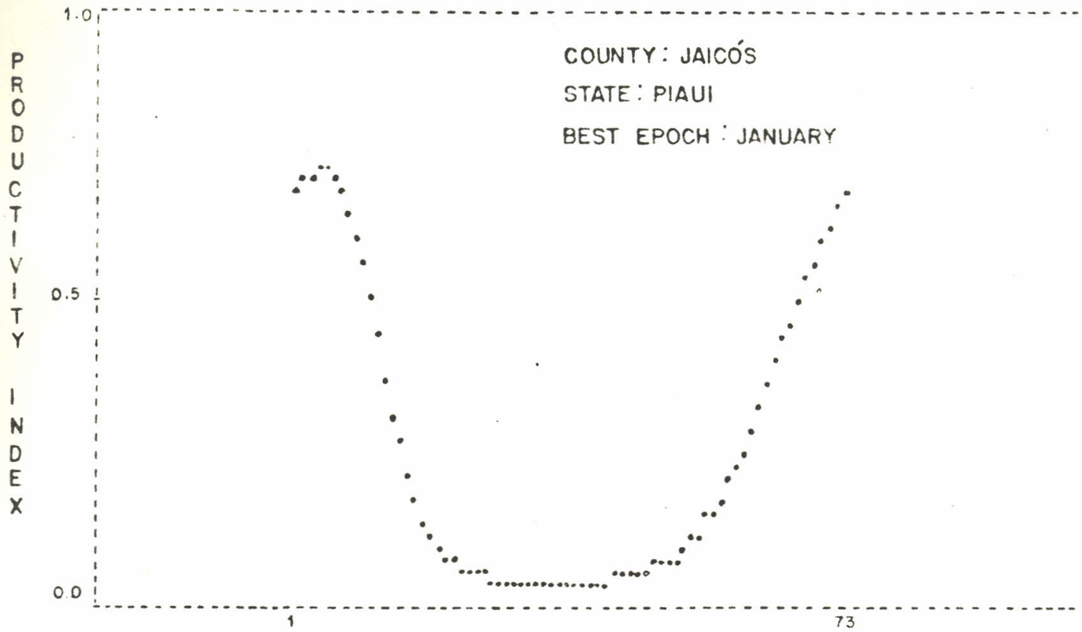
One of the more useful aspects of the system is that it records the simulation results. In this way, the user can apply all sorts of numeric or graphic techniques to analyse the results, and this can be done at any time. This is particularly useful with regard to the zoning work, where the simulation results for several counties must be jointly analysed.

SORGHUM PRODUCTION POTENTIAL IN THE SEMI-ARID TROPIC OF BRAZIL

The computerized system has been applied to the study of sorghum production potential in the semi-arid tropic of Brazil. To illustrate the results, a set of 32 counties has been used. They form an intentional sample, chosen so as to represent the different states and geographic conditions of the region.

Agroclimatic zoning

The model calculates a productivity index for each planting period in each year. This is a number between 0 and 1 , representing the quotient YA/YP (actual yield divided by potential yield). At the end, the program calculates the mean index value for each planting period over the simulation years. Working with a simulation step of 5 days, the model produces, for each county, a vector with 73 components (the productivity indexes) which characterize the county (Fig. 2). Similarly, the model estimates two other vectors of 73 components; one of them gives the mean excess (run-off) water and the other the mean deficit (potential minus actual evapotranspiration), in millimeters, for each planting



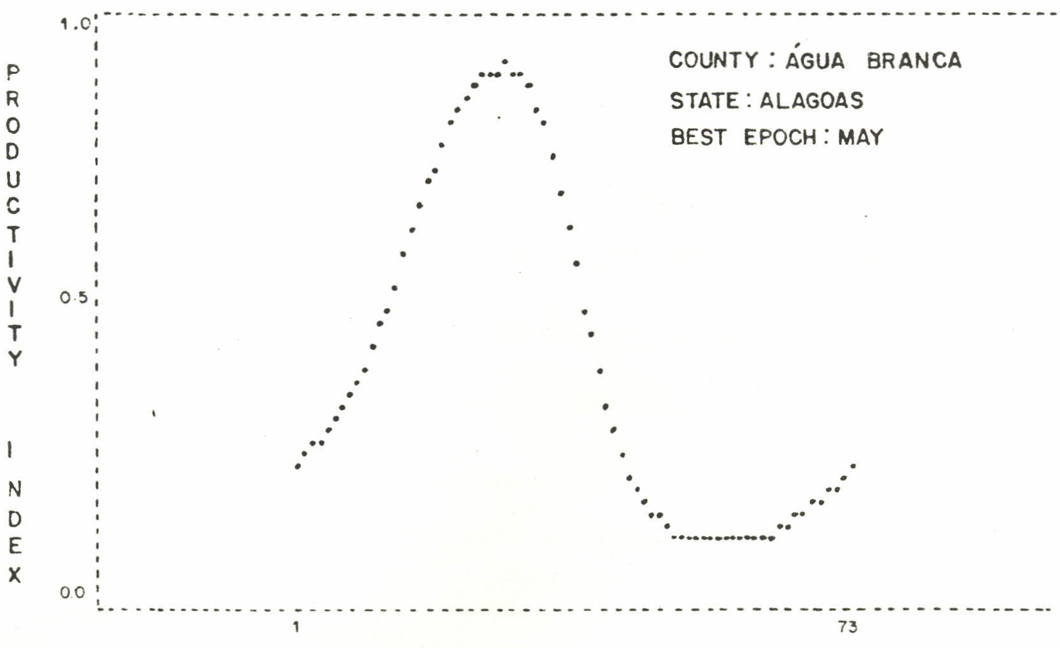


Fig. 2. Mean productivity indexes and best planting epoch for two counties, showing totally different patterns.

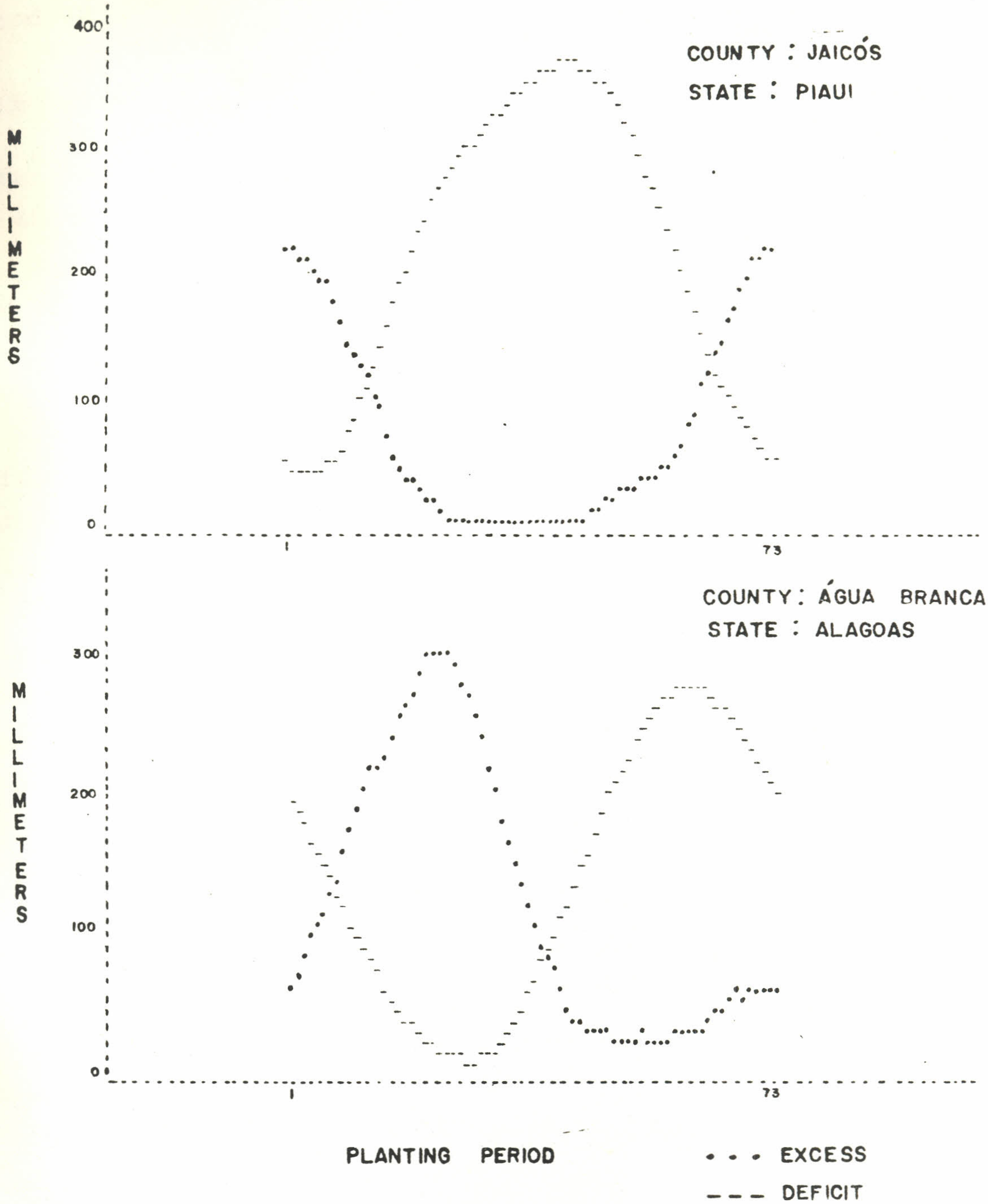


Fig. 3. Excess and deficit for two counties, showing totally different patterns.

period (Fig. 3).

Let $p^1 = (p_1^1, p_2^1, \ldots, p_{73}^1)$ and $p^2 = (p_1^2, p_2^2, \ldots, p_{73}^2)$ be the vectors of productivity indexes for two given counties C_1 and C_2. Let the Euclidean distance between C_1 and C_2 be defined as

$$d\,(C_1, C_2) = \Big(\sum_{i=1}^{73} (p_i^1 - p_i^2)^2 \Big)^{1/2}.$$

Having defined a distance on the set of counties , different clustering techniques can be used in order to find meaningful groups (clusters) of counties (see, for instance, HARTIGAN 1975, JARDINE and SIBSON 1971, SNEATH and SOKAL 1973). In this work, an algorithm developed by GARAGORRY and PALMA (1980), that works under the SAS package, has been used. The algorithm produces a set of clusters which minimizes the maximum diameter of the clusters (as usual, the diameter of a cluster is the maximum distance between any two members of the cluster). The idea, of course, is that the counties in each cluster will have similar productivity patterns. Working on the computer output of the clustering algorithm, it is possible to identify clearly defined groups of counties. Table 1 gives a summary of the results obtained through the application of cluster analysis.

Table 1. Results obtained by cluster analysis

CLUSTER	SORGHUM POTENTIAL	BEST PLANTING EPOCH	NUMBER OF COUNTIES
1	HIGH	JAN	3
2	HIGH	MID-APR/MID-JUN	9
3	FAIR	MID-JAN/MID-MAR	7
4	FAIR	MID-APR/MID-JUN	6
5	FAIR	MID-OCT/MID-FEB	1
6	POOR	-	6

Clusters 2 and 4 constitute a sort of surprise , since there was no information about sorghum planted around May .

This will certainly be the object of further research since, according to the model, in a large proportion of counties, some of them with very high estimated potencial, planting should be done from April to June.

Different distances between two counties were defined using the excess water or the deficit vectors, instead of the productivity index vectors. In all cases, the results oetained through cluster analysis were consistent with the above presentation.

County level recommendations

Zoning means classification or grouping. In this sense, computerized data analysis techniques can be helpful to obtain an initial classification. The final zoning can be obtained by the combination of that classification with other agroclimatic arguments.

On the other hand, independently of the work on agroclimatic zoning, it is convenient to know specific results for each county. In particular, the best planting season, maximum estimated productivity and maximum estimated run-off can be printed for each county. This information is useful to farmers, extension agents and researchers. Table 2 gives a sample of the sort of county level results that can be obtained from the system. Productivity indexes are presented in 10% intervals; the highest interval reached is used, in Table 2, to define the best planting season for each county. Estimated maximum run-off and the period when it occurs are also shown in Table 2.

Table 2. Information on individual counties, ordered by states.(a)

COUNTY	MAXIMUM PRODUCTIVITY INDEX PLANTING SEASON	%INTERVAL	MAXIMUM EXCESS WATER PERIOD	mm
Alagoas				
Água Branca	APR16-JUN09	80- 90	APR21-APR25	303
Delmiro Gouveia	APR16-MAI20	50- 60	FEB25-MAR01	63
Mata Grande	MAI06-MAI25	90-100	APR26-APR30	420
Palmeira dos Índios	APR21-JUN14	90-100	APR21-APR25	399
Pão de Açúcar	APR11-MAI25	60- 70	APR21-APR25	107
Penedo (b)	APR01-JUN19	90-100	APR16-APR20	568
Piranhas	APR26-MAI10	60- 70	APR11-APR15	59
Poço das Trincheiras	MAI11-MAI20	80- 90	APR26-APR30	195
Santana do Ipanema	APR21-JUN09	80- 90	APR21-APR25	275
Bahia				
Barreiras	OCT08-FEB04	60- 70	NOV07-NOV11	301
Irecê	OCT18-NOV21	40- 50	NOV12-NOV16	133
Senhor do Bonfim	APR11-MAI25	60- 70	JAN21-JAN25	117

(a) States names are underlined.
(b) This county does not belong to the SAT of Brazil; it was included for comparison.

In general, a very good agreement was found between the estimates of the maximum productivity index and the maximum excess water. But, in most cases, the period with maximum run-off preceeds the mid-point of the best planting interval. This indicates that it may not be appropriate to equate best planting period with period of maximum run-off, and that the use of a productivity index should be encouraged, although much more research is required in this area.

Excess water studies

A major line of research at CPATSA (EMBRAPA's research center for the semi-arid tropic) is the use of excess water through non-traditional techniques, in order to provide complementary irrigation at critical stages of the crop growth. The model estimates the mean excess water (run-off) for each planting period. Excess water, or run-off, is defined, in the model, as the amount of water that cannot be held in the reservoir; therefore, it refers both to surface run-off and deep infiltration. For that reason, an appropriate technology to retain and use most of the excess water estimated by the model may not be available. Nevertheless, those values give an idea about the potential to introduce, into a given county, the complementary irrigation technology already tested at CPATSA with very encouraging results.

CONCLUSIONS

As stated before, EMBRAPA's computerized agroclimatic system is under continuous improvement. At the same time, several field experiments are being conducted, in order to estimate crop response to water and to study root development.

Several computer experiments are required so as to know, in more detail, the model sensitivity with regard to the following aspects: (1) changes in the crop coefficients (length of the phenological phases, productivity function coefficients, etc.); (2) changes in the crop-soil coefficients (root development). Also, different clustering techniques should be tested in relation to zoning studies.

More importantly, the work with the agroclimatic model should be combined with the use of a more detailed sorghum production model. This will be useful to obtain more insight with regard to sorghum production potential in those counties that already show good prospects with the present model.

REFERENCES

1. GARAGORRY, F.L. & PALMA, O. I. (1980). "Aplicação de programação inteira à construção de conglomerados", Tech. Rep. DMQ/A/47, EMBRAPA, Brasília.

2. HARTIGAN, J. A. (1975). Clustering Algorithms, Wiley.

3. JARDINE. N. & SIBSON. R. (1971). Mathematical Taxonomy, Wiley.

4. SNEATH, P.H.A. & SOKAL, R.R. (1973). Numerical Taxonomy (The Principles and Practice of Numerical Classification), W. H. Freeman and Co., San Francisco.