1. Publication Nº	2. Version	3. Date	5. Distribution
INPE-3967-PRE/982		*August, 1986*	☐ Internal ☒ External ☐ Restricted
4. Origin *DME/DAM*	Program *SAFRA*		

6. Key words - selected by the author(s)
CROP YIELD
CLIMATE IMPACT

7. U.D.C.: *551.584.43:633.15*

8. Title *INPE-3967-PRE/982*

CROP YIELD MODELING FOR CLIMATE IMPACT STUDIES
PART I: REASONING AND RESULTS

10. Nº of pages: *12*

11. Last page: *9*

12. Revised by

Jesus Marden dos Santos

13. Authorized by

Marco Antonio Raupp
General Director

9. Authorship *Fausto Carlos de Almeida*

Responsible author

14. Abstract/Notes *The use of regression crop yield models (RCY) for Climate Impact Assessment Studies requires some basic features sometimes not encountered in most RCY models. Many of the RCY models developed have a different aim, i.e., to give a reasonable estimate of the yield of a given crop, for a given location, based on a historical technological trend and suitable chosen meteorological variables. In general they offer good "prediction efficacy", i.e., accuracy and lead time, but should be carefully analized when used for other purposes, such as in Climate Impact Assessment Studies. Depending on what kind of climate scenario we are dealing with, the RCY model chosen has to be able to assimilate the impact of the respective meteorological variables on yield. For this, the model should be representative, i.e., should not depend on the historical time series period we chose to work with. Its regression coefficients should be stable with no change in sign and present a very small (combined) variance, otherwise the climate impact, represented by changes of the meteorological variables being tested, can be masked by its (the regression coefficients) variance.*

15. Remarks *Paper presented at the Laxemburg, Austria, IIASA meeting, on Assessment of Climate Impact on Agriculture in Semiarid and High Altitude Regions (8-12 Oct. 1984).*

TABLE OF CONTENTS

1. INTRODUCTION

Several workers dealing with the food supply and consumption problem have called attention to the high demand for food that the very high population growth rate is creating. This has brought the world to be aware of its finiteness. It is then only a step forward to realize that the only way to bring the problem to an equilibrium stage, without "interfering" in the population growth rate, is to increase food production.

There are two ways to accomplish this. One is to increase the acreage being harvested. The other is to increase the productivity level of the food crops.

Despite the problem of availability of good agricultural land for expansion and/or technology to be absorbed, the solutions to the food versus population problem, from the food increase side, will lead us invariably to a more climate dependent agricultural system. Either because more marginal land is being put into production, or because the higher technologies require better climate management. It is noteworthy to call attention to the effort from the bioagricultural community to develop less weather sensitive crops.

In any event, the concept of "Atmospheric resources management" (Almeida, 1983) has to be taken into account for any productive system.

On the long run, even assuming that we can avoid breaking the balance between food production versus population growth, nevertheless the impact of climate change has to be taken into consideration, at least to guarantee year to year stability on food supply (production + storage).

In studying the impact of climate on food production, it is very important to consider both, the climatic change, or the somewhat long range impact on food production; and the climatic variability, or the somewhat short to medium range impact.

In order to do that, climatic trends and variability have to be distinguished and measured, and models for the assessment of these climate changes and variability on food production have to be chosen.

Here we will try to review the so called climate-crop yield models, analysing them as possible candidate models for the assessment of climate impacts on food production.

2. BRIEF REVIEW OF CROP YIELD MODELS

Crop yield models as they are known today have come about as a way to "predict", with certain lead time, the yield of a certain crop for a specific location, taking into account not only the technological but also the climatic variables.

In this respect, it is an improvement over the more econometric type models solely based on trends which reflect the technological improvements, but not the productivity variation which comes about as a result of the year to year "weather" variability.

In general, statistical crop yield models are based on monthly means or totals of some meteorological variables, chosen as possible candidates regressed against the residuals of the actual yearly yields de-trended by a generally linear technology growth productivity line. These models have the aim to offer good "prediction efficacy", i.e., they should possess good *accuracy*, and also be capable of producing results much before harvesting time (offer good *lead time*). They can be classified as "black box" models, or "physical/empirical" models.

As it is not our aim to offer an exhaustive review of climate-crop yield models, it is sufficient to say that "black-box" type models are usually models in which several meteorological variables are given as input to a stepwise regression analysis scheme, which chooses those which maximize the explained variance. Sometimes these models, although possessing a high R^2 value (percent of the total variability explained by the model), are not biophysically sound. In other words, a variable can appear positively correlated with the yield, when by agronomic reasoning it is known that it should be other way around.

In the more physical models, again, stepwise or multiple regression analysis are used, but to fit previously chosen meteorological variables, which should yield better physically sound relations between yield and climate variables. It is noteworthy to mention here that these models have gained in popularity because of its simplicity and good results, and are today widespread used all over the world. It is important to call attention to the problem of using a model, developed and tested for some type of application, without giving the necessary care to erroneous results which might arise from its improper use, especially when it does not satisfy the basic requirements of the new application.

This is the case when using climate impact assessment models.

3. IMPACT ASSESSMENT MODELS

For this type of application (Climate Impact Assessment), it is necessary that for a given climatic "scenario" the model can respond to the variability of the climate variables being tested, without masking the results because of its statistical poor design and results.

In other words, what it is assumed in using crop yield models, designed for "prediction efficacy", is that they also possess statistical properties to allow yield responses to climate changes and variability to be physically and statistically sound. For a somewhat complete discussion about this problem, the reader is referred to Katz (1981).

4. A MODEL SUGGESTION

In this section, we will introduce a crop yield model suggestion presenting some of its statistical properties and results, to be considered as a candidate for climatic impact assessment studies.

4.1 - DESCRIPTION OF THE MODEL

The model uses a linear trend technology, and the step-wise regression analysis to compose the relations between the yearly yields and some previously chosen climate variables (physical/empirical type model).

The main differences between this suggested model and others existing in the literature are twofold. First, it uses daily variables instead of monthly means or totals. Second, the climate variables to enter the regression analysis are chosen by a method called "critical periods method" (Celaschi and Almeida, 1981; Celaschi, 1983).

In essence it composes from each "climate" daily data, a 10 days moving average, and correlates the resulting time series with the yearly yield series, de-trended from the linear technology line. This procedure will indicate, using the 1% and 5% confidence levels, the critical periods, for the crop being studied, of the climate variables used. This procedure, called the critical periods method, will choose the climate variables to be used in the statistical regression model.

The "critical periods" solely chosen on the basis of correlations between yearly yields and daily mean temperatures and precipitation totals have shown good agreement, when compared to critical periods observed in field experiments for corn, according to the published literature (Matzenauer, 1980; Celaschi, 1983).

4.2 - MODEL ANALYSIS (STATISTICAL)

When writing the statistical crop yield model equation in a generalized manner, i.e.,

$$Y\ (\text{year}) = a + b^{*}\ \text{year} + \Sigma\ C_i\ V_i + \hat{e}\ ,$$

where

Y = yield,

V_i = climatic variable,

C_i = coefficients determining the relations between V_i and Y,

a,b = coefficients determining the technological linear trend,

$\hat{e}$ = residue not explained by the model,

several statistical properties should be analized.

In general, when the aim of the model is "solely" crop prediction, statistics like the R^2 coefficient of determination, and the *lead time* afforded by the model, i.e, the models "prediction efficacy",is all that matters.

For climate impact assessment studies, the importance of other tests should not be overlooked, such as:

1) The physical soundness of the "statistically" derived model, or how the climatic variables chosen to compose the model represent the agricultural known critical periods of the crops being studied.

2) The multicolinearity problem, or how the climatic variables chosen to compose the model are correlated among themselves. This fact is known to cause problems in the statistics of the model, which are very important for climate impact assessment studies (e.g., changes of sign of the coefficients).

3) Stability of the estimations/Beta analysis, or how the variances and covariances of the coefficients, taking into account the statistics of the climatic variables, affect the confidence limits of the yield estimated by the model, therefore putting a high stress on the possibility for the model to respond to climate changes, without masking these changes within its variance (confidence limits).

A complete treatment of this problem, done for corn in five different crop reporting districts, of the State of São Paulo, Brazil, using data ranging from 1950 to 1980, can be found in Celaschi (1983).

Here we will illustrate the problem, showing some statistical results for some crop reporting districts.

4.3 - MODEL RESULTS

In order to illustrate the "prediction efficacy" of the critical periods model, Table 1 presents, as an example, the yield errors, in percent, for two crop reporting districts (CRD), namely, Campinas and Rio Preto.

Both results can be obtained by 05/Jan (Campinas) and 14/Feb (Rio Preto), while the official estimates come about by the end of June. All of the results above were obtained by using the data available up to the preceding year.

TABLE 1

YIELD ERRORS IN PERCENT FOR TWO CRD's IN THE STATE OF SÃO PAULO, BRAZIL

CRD \ YEAR	1976	1977	1978	1979	1980	1981
Campinas	0.7	0.3	-1.8	0.8	-9.7	13.8
Rio Preto	-5.6	-7.5	13.7	-4.2	-6.9	0.7

From Celaschi (1983).

To illustrate the stability of the critical periods model, Figure 1a shows the confidence limits at 95% for the Beta coefficients (Chattersee and Price, 1972), for the crop reporting district of Sorocaba, when compared to the results of a monthly mean type model (Figure 1b) derived from the same reporting district data.

It can be noted that the stability of the coefficients for the critical periods model is better than that of the monthly mean type model. Also noteworthy to be mentioned is the change in sign for the monthly mean model (Figure 1b).

Each point in the graph represents the Beta coefficient for different models, each one composed for time series of variable length, indicated in the figure by the numbers on the right hand corner, i.e., 22 means 22 years of data, 17 means 17 years of data, and so on. Note that the coefficients for the critical periods method (Figure 1a) are much more stable, and that they do not present sign changes as shown for the coefficient of the TDEZ (mean temperature december) monthly mean climate variable (Figure 1b).

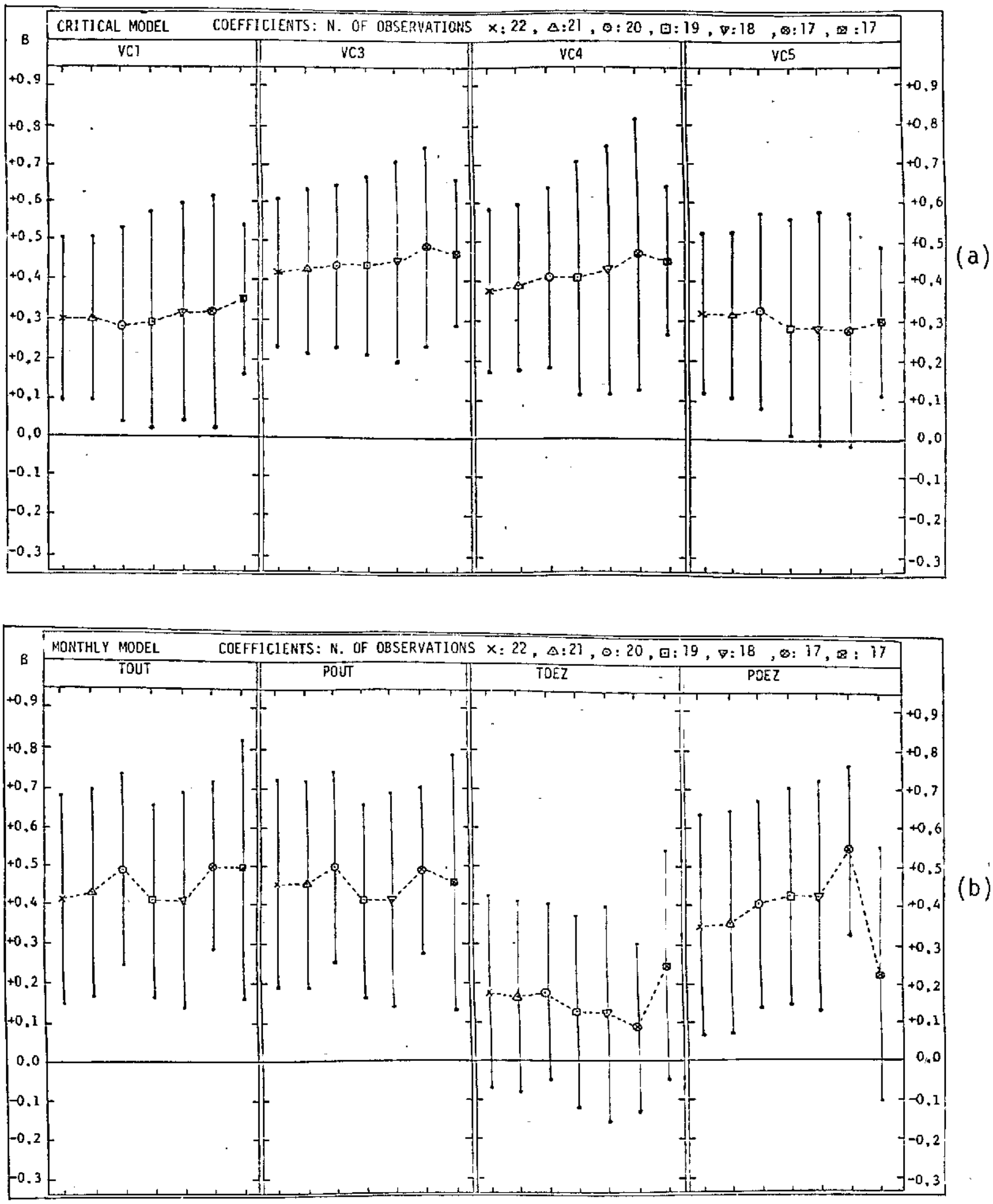

Fig. 1 - Confidence limits at 95% for the Beta coefficient, for the CRD of Sorocaba.

From Celaschi (1983).

5. COMMENTS AND CONCLUSIONS

As per what was presented before, it is suggested that the critical periods, daily data approach crop yield modeling development, leads not only to a model that is comparable to or better than the black box monthly means or total models, but also renders better statistically appropriate results especially for use in climate impact assessment studies.

This model will be used for northeast Brazil climate impact assessment study, to be presented as part II of this paper (Almeida and Sá, 1986).

6. REFERENCES

ALMEIDA, F.C. de *The management of atmospheric resources on food production.* WMO Climate Conf. for Latin America and the Caribbean, Paipa, Colombia, 28 Nov. - 03 Dec. 1983. (INPE-3121-PRE/514) 06 pp.

ALMEIDA, F.C. de; SÁ, L.D.A. *Crop yield modeling for climate impact studies. Part II. An example for NE Brasil.* (to be presented).

CELASCHI, W.; ALMEIDA, F.C. de Um modelo para a estimativa de produtividade de milho no Estado de São Paulo (A model to estimate yield in the State of São Paulo) - II Brazilian Symp. on Agrometeorology, July 10, 1981, Pelotas, RS, Brasil.

CELASCHI, W. *Um modelo para estimação do rendimento do milho aplicado ao Estado de São Paulo (A model for estimating corn yield applied to the State of São Paulo, Brazil).* Ms Thesis, INPE, São José dos Campos, SP, 1983. (INPE-2658-TDL/118) 178 pp.

CHATTERSEE, S.; PRICE, B. Regression analyses by example. New York, NY, John Wiley, 1977.

KATZ, R.W. Assessing the impact of climatic change on food production. Climatic Change 1, 85-96, 1977.

MATZENAUER, R. Evapotranspiração do milho e suas relações com fórmulas e parâmetros (Evapotranspiration of corn and its relations with formulas and parameters). Ms. Thesis, Faculdade de Agronomia, UFRG, Porto Alegre, RS, Brasil, 1980.