

*OK - now*

# AgRISTARS

YM-N4-04454  
JSC 18906

A Joint Program for  
Agriculture and  
Resources Inventory  
Surveys Through  
Aerospace  
Remote Sensing

## Yield Model Development

JANUARY 16, 1984

### BRAZIL WHEAT YIELD COVARIANCE MODEL

(E84-10103) BRAZIL WHEAT YIELD COVARIANCE  
MODEL (National Oceanic and Atmospheric  
Administration) 22 p HC A02/MF A01 CSCL 02C

N84-21921

Unclas  
G3/43 00103

SUSAN L. CALLIS  
CLARENCE SAKAMOTO



USDA/NOAA  
CIAD/MODELS BRANCH  
ROOM 200, FEDERAL BLDG  
608 CHERRY STREET  
COLUMBIA, MO 65201



Lyndon B. Johnson Space Center  
Houston, Texas 77058

1. Report No. YM-N4-04454, JSC 18906		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle  Brazil Wheat Yield Covariance Model				5. Report Date January 16, 1984	
				6. Performing Organization Code	
7. Author(s)  Susan L. Callis and Clarence Sakamoto				8. Performing Organization Report No.	
9. Performing Organization Name and Address USDC/NOAA CIAD/Models Branch Room 200, Federal Bldg, 608 Cherry St. Columbia, MO 65201				10. Work Unit No.	
				11. Contract or Grant No.	
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Lyndon B. Johnson Space Center Houston, TX 77058				13. Type of Report and Period Covered	
				14. Sponsoring Agency Code	
15. Supplementary Notes					
16. Abstract <p>A model based on multiple regression was developed to estimate wheat yields for the wheat-growing states of Rio Grande do Sul, Parana, and Santa Catarina in Brazil. The meteorological data of these three states were "pooled" and the years 1972 to 1979 were used to develop the model since there was no technological trend in the yields during these years. Predictor variables were derived from monthly total precipitation, average monthly mean temperature, and average monthly maximum temperature.</p>					
17. Key Words (Suggested by Author(s))  Multiple regression analysis Predictor variables			18. Distribution Statement		
19. Security Classif. (of this report)  Unclassified		20. Security Classif. (of this page)  Unclassified		21. No. of Pages  19	
				22. Price*	

\*For sale by the National Technical Information Service, Springfield, Virginia 22161

BRAZIL WHEAT YIELD COVARIANCE MODEL

by

Susan L. Callis

and

Clarence Sakamoto

AISC Models Branch

January 16, 1984

## INTRODUCTION

The purpose of this study was to select monthly weather variables that could be used to estimate wheat yields for the wheat growing areas of Brazil. Wheat is grown in seven states in Brazil: Rio Grande do Sul, Santa Catarina, Parana, Sao Paulo, Mato Grosso, Goias, and Minas Gerais. Rio Grande do Sul is Brazil's original wheat-producing state; until 1972 it was the country's most important production area. Increasingly, however, Parana, Mato Grosso and Sao Paulo have grown in importance. In 1977, Parana took over the number one spot. Figure 1 shows the wheat-growing areas of Brazil.

Although wheat has been grown since the sixteenth century, Brazil has yet to develop a high-quality wheat variety that produces well under the country's widely varied climatic conditions. Brazil's wheat crops have continuously been plagued with problems resulting in consistently low yields. Frosts and plant disease occasionally reduce yields. The high cost discourages use of fungicides. Expansion of acreage sown to wheat was met with cultivating problems and high costs. New land areas cultivated in wheat are highly acidic and low in fertility. Fertilizer is costly. Late-season rains frequently delay harvest and reduce yield.

The climate of the southern wheat-producing states is "subtropical humid"; rainfall is relatively abundant and well-distributed throughout the year, with slightly more rainfall in the warm months. There is usually no season of drought. The northern states are "semiarid" with a winter dry season and less total annual rainfall. Summers are hot and winters are mild. Parana and southern Sao Paulo are the northern limit for frost occurrence.

Wheat is planted in the months of April through June and is harvested in November and December.

ORIGINAL PAGE IS  
OF POOR QUALITY

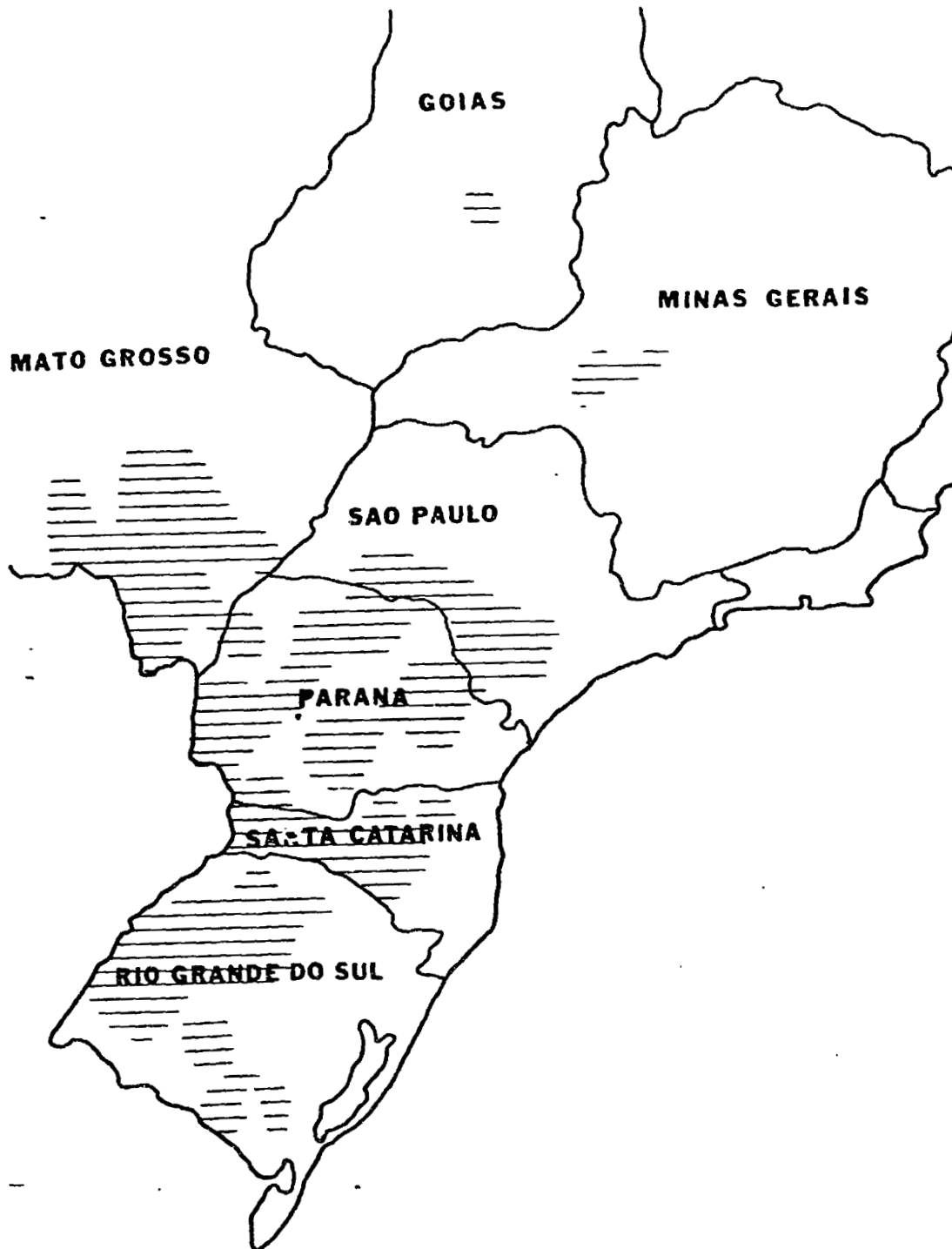


Figure 1. Wheat-growing areas of Brazil.  
(J. McQuigg, R. Willis, 1982, personal  
communication)

## METHOD

Multiple regression analysis of yield with selected agroclimatic indices was used to derive a suitable model. The index P-PET (precipitation minus potential evapotranspiration) was used in the regression equations to represent available soil moisture, monthly precipitation, and monthly maximum temperature.

The regression equation is:

$$\hat{Y} = \alpha + B_1TX_i + B_2R_i + B_3 (P-PET)_i + E$$

where

$\hat{Y}$  = Estimated yield,

$\alpha$  = Constant,

$B_j$  = Coefficients of variables  $j = 1-3$ ,

$TX_i$  = Maximum temperature for month  $i$ ,

$R_i$  = Total precipitation for month  $i$ ,

$(P-PET)_i$  = Precipitation minus PET for month  $i$ , and

$E$  = Unexplained error.

In developing the model, various procedures of the Statistical Analysis System (SAS Institute, Inc., 1979) were used. The procedures used and the operations performed with each are summarized in the Appendix. The selected model had the highest  $R^2$  and included variables that were significant at the 10 per cent level and agronomically meaningful.

## DATA

The Brazil crop data were obtained from the Foreign Agricultural Service (Sam Ruff, Personal Communication, 1982). The data was recorded with year of yield as year of harvest, so the weather influencing the crop occurred during year - 1.

Meteorological data from 1972 through 1977 were used to model because there is no apparent trend in the yield data during this period. Furthermore, 1977 represented the latest available data. Table 1 lists the stations used to derive the meteorological data sets for each state. Figure 2 shows the location of each station.

#### PROCEDURES

Since Rio Grande do Sul's meteorological data has the longest period of record, initial models were developed for this state alone. Various weather variables were tried in regression equations in many different combinations. The results were not good. Several different trends were tried with different regression models, but none were significant at the 10 per cent level. From historical accounts, it is believed that some sort of trend of increased yield began in 1962 because of increased use of fertilizer and more adaptable varieties of wheat. Yet the yield data did not indicate this. It was decided to model for years 1972-1979 for which it was believed no trend existed. Eliminating years of data created the problem of fewer degrees of freedom. However, by combining data for the states of Rio Grande do Sul, Parana, and Santa Catarina, a covariance model could be developed.

Variables used in the regression equations for the covariance model, included "dummy variables" for Parana and Santa Catarina. The "dummy variables" adjust the contributions to yield of both states to a base yield which, in this case, is Rio Grande do Sul's yield. The "dummy variable" for Parana was not significant at the 10 per cent level in the final model. The coefficient for Santa Catarina's "dummy variable" is negative, indicating that its yield is below that of the norm set by Rio Grande do Sul.

<u>STATE</u>	<u>METEOROLOGICAL STATION</u>	<u>WMO NUMBER</u>
Mato Grosso	Campo Grande	83611
	Ponta Pora	83702
	Pres Prudente	83716
Parana	Londrina	83766
	Santa Branca	83810
	Curitiba	83842
	Porto Uniao	83864
Rio Grande do Sul	Ijuí	83805
	Sao Borja	83901
	Passo Fundo	83914
	Julio de Castilhas	83935
	Santa Maria	83936
	Sao Gabriel	83957
	Bage	83980
Santa Catarina	Porto Uniao	83864
	Vacaria	83918
	Sao Joaquim	83920
Sao Paulo	Pres Prudente	83716
	Bauru	83722
	Jau	83723
	Tiete	83777
	Ataliba Leonel	83803
	Tatui	83816

Table 1. Meteorological Stations Used to Derive Data Sets  
for the Brazil Wheat Model.





Figure 2. Location of Meteorological Stations Used to Derive Data Sets for the Brazil Wheat Model.

The following is the selected model:

DUM	Dummy variable to adjust Santa Catarina's yield
TM7	July Mean Temperature
P-PET8	August P-PET
SP-PET8	Squared August P-PET
P-PET9	September P-PET
RDFN11	Deviation from normal of November precipitation

Too high temperatures in July and too much rainfall in November both reduce yield. The negative coefficient for the linear term P-PET8 suggests that in August excess precipitation above demand, PET8, is damaging to yield. This is reasonable in Brazil when in August the crop is in the tillering stage. The quadratic term indicates that increased yield is favorable at some level when PET is higher than precipitation. However, one is cautioned not to extend this interpretation beyond the limits of the data base used. The statistics of the selected model are summarized in Table 2.

The same variables were used in regression equations for only the two states of Rio Grande do Sul and Santa Catarina. The problem encountered was that models with agronomically reasonable variables had too low an  $R^2$ ; models with acceptable  $R^2$  had too many variables for the number of degrees of freedom.

Finally, modeling was attempted for the northern states of Mato Grosso and Sao Paulo. Overlapping plots of weather variables versus yield were made for both states. From these plots, it was determined that modeling for a combination of data for these two states would not be acceptable; their climates are too different. Next, reasonable weather variables were tried in regression for both states separately. No suitable models were derived. For information, plots of yields for both states are shown in Figure 3.

## TEST RESULTS

A jackknife test was run on the selected model. In this test, a year was eliminated from the crop data and the model was used to predict that year's yield. This process was done for each successive year beginning with 1972. The test had to be run separately on each state. The results are printed on Tables 3 through 5 and plotted on Figures 4 through 6.

## APPENDIX

### Definition of Variables

P-PET, precipitation minus potential evapotranspiration, is used a measure of the amount of moisture available for plant growth. Potential evapotranspiration is determined by the procedure developed by Thornthwaite (1948). It requires only temperature:

$$PET = \left( \frac{10T}{I} \right)^a$$

where I = heat index, which is the sum of the 12 monthly indices i,

$$i = \left( \frac{T}{5} \right)^{1.514}$$

T = monthly temperature in °C, and

a = an empirical exponent  $6.75 \times 10^{-7}I^3 - 7.71 \times 10^{-5}I^2 + 1.79 \times 10^{-2}I + 0.49$ .

The duration of daylight is used to adjust potential evapotranspiration as a portion of 12 hours.

### Statistical Analysis System Procedures Used

PROC CORR	Computes correlation coefficients between variables, including Pearson product-moment and weighted product-moment correlation.
PROC PLOT	Graphs one variable against another, producing a printer plot.
PROC STEPWISE	Provides five methods for stepwise regression. Stepwise is useful when selecting variables to be included in a regression model from a collection of independent variables.
PROC STEPWISE FORWARD	Begins by finding the one-variable model that produces the highest $R^2$ . For each of the other independent variables, FORWARD calculates F-statistics reflecting the contribution to the model if the variable were to be included.

#### PROC STEPWISE BACKWARD

Begins by calculating statistics for a model including all the independent variables. The variables are deleted from the model one by one until all the remaining variables produce F-statistics significant at the .10 level.

#### PROC STEPWISE STEPWISE

The stepwise method is a modification of the forward selection technique, differing in that variables already in the model do not necessarily stay there. After a variable is added (as in the forward selection method) the stepwise method looks at all the variables already included in the model and deletes any variable that does not produce an F-statistic significant at the .10 level. Only after this check is made and the necessary deletions accomplished can another variable be added to the model.

#### PROC STEPWISE MAXR

(Maximum  $R^2$  improvement) Unlike the three techniques above, this method does not settle on a single method. Instead it looks for the "best" two-variable model, the "best" three variable model, and so forth.

#### PROC PETM

Uses latitude and mean monthly temperature to calculate Thornthwaite's potential evapotranspiration for each month.

#### PROC ZINDEX

Uses monthly PET's, precipitation, SS (beginning moisture in surface layer), AWCS (available water capacity in surface layer), SU (beginning moisture in the underlying layer), and AWCU (available water capacity in the underlying layer) to calculate Palmer's soil moisture budget, drought index Z, ET, and ET.

PLOT OF YIELD\*YEAR

LEGEND: A = 1 OBS, B = 2 OBS

LEGEND: A = 1 OBS, B = 2 OBS

LEGEND: A = 1 OBS, B = 2 OBS

YIELD  
17.5  
15.0  
12.5  
10.0  
7.5  
5.0  
2.5  
0.0

YIELD  
17.5  
15.0  
12.5  
10.0  
7.5  
5.0  
2.5  
0.0

1961 1967 1973 1979  
YEAR

1961 1967 1973 1979  
YEAR

MATO GROSSO

SAO PAULO

ORIGINAL PAGE 19  
OF POOR QUALITY

Figure 3. Plots of Yields Versus Year for the States of Mato Grosso and Sao Paulo.

BACKWARD ELIMINATION PROCEDURE FOR DEPENDENT VARIABLE YIELD

ALL VARIABLES ENTERED

R SQUARE = 0.66192865    C(P) = 7.00000000

SUM OF SQUARES		C(P)		MEAN SQUARE		PROR>F	
DF						F	
5				21.20184504		5.55	0.0024
17				3.82184587			
23							
STD ERROR		TYPE II SS		PROR>F			
B VALUE							
INTERCEPT	27.91471003			43.54320431		11.39	0.0036
DUMP (Santa Catarina)	-4.90222435			22.28065116		5.83	0.0273
T47	-1.04143789			21.36220762		5.59	0.0302
D PETA	-0.04612431			17.27858241		4.52	0.0484
SO PETA	-0.00016214			29.33485132		7.68	0.0131
P DETA	-0.02743603			70.35548490		19.41	0.0005
DIFFNII	-0.04206852						

VALUES IN THE MODEL ARE SIGNIFICANT AT THE 0.1000 LEVEL.

ORIGINAL PAGE IS  
OF POOR QUALITY

Table 2. Statistics of Brazil Wheat Model





O=OBSERVED YIELD  
P=MODEL'S PREDICTED YIELD

PLOT OF YIELD\*YIELDYR SYMBOL USED IS P  
PLOT OF OBSYIELD\*YIELDYR SYMBOL USED IS O

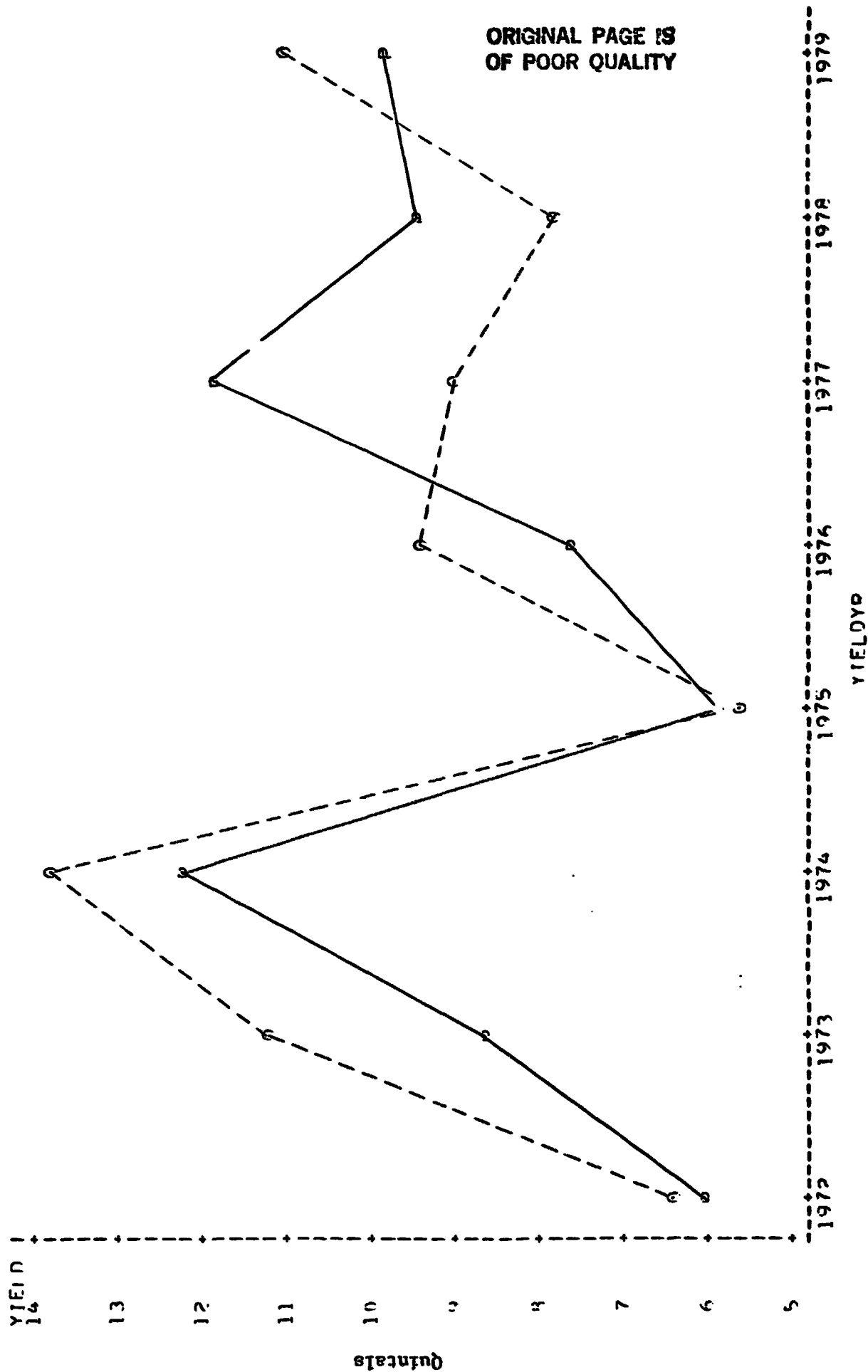


Figure 4. State of Parana Wheat Model.

Q=OBSERVED YIELD  
P=MODEL'S PREDICTED YIELD

HETA3	HTA3	HTA4	HTA5	HETA6	HETA7	CONTR1	CONTR2	CONTR3	CONTR4	CONTR5	CONTR6	CONTR7	BETA1	HETA2
3376	1.73	1.79	1.73	1.73	1.73	3.56	0.712347	7.55088	14	3.32041	33.0919	-6.3795	33.0919	-6.3795
2773	1.73	1.79	1.73	1.73	1.73	11.40	0.666831	3.65706	14	3.92005	31.8241	-5.5776	31.8241	-5.5776
0703	1.72	1.79	1.73	1.73	1.73	10.50	0.608324	1.02987	14	3.96807	27.9412	-4.7060	27.9412	-4.7060
0597	1.72	1.79	1.73	1.73	1.73	9.00	0.631051	0.86645	14	4.54600	28.0310	-4.4158	28.0310	-4.4158
0430	1.72	1.79	1.73	1.73	1.73	4.53	0.665166	0.86604	14	4.37230	28.0549	-5.0135	28.0549	-5.0135
0180	1.72	1.79	1.73	1.73	1.73	12.10	0.764795	2.08209	14	2.73269	14.6674	-2.7415	14.6674	-2.7415
9298	1.72	1.79	1.73	1.73	1.73	4.40	0.641851	0.63813	14	3.91461	27.3324	-4.7451	27.3324	-4.7451
							0.677955	0.87393	14	3.93095	25.8588	-4.4507	25.8588	-4.4507
HETA4	HTA4	HTA5	HTA6	HETA7	CONTR1	CONTR2	CONTR3	CONTR4	CONTR5	CONTR6	CONTR7	BETA1	HETA2	
054569	0.000234444	-0.037211	-0.037211	-0.037211	33.0919	-17.656	-13.400	15.1241	-5.824	-2.1195				
052746	0.000146275	-0.030710	-0.030710	-0.030710	31.8241	-16.244	-17.081	13.7451	-2.2242	-4.5424				
050479	0.000173004	-0.023269	-0.023269	-0.023269	27.9412	-13.911	-5.120	1.7451	-0.4081	0.3645				
044516	0.000145220	-0.024670	-0.024670	-0.024670	24.0310	-13.670	-6.488	3.3519	-3.4046	-0.5531				
053361	0.000131422	-0.025419	-0.025419	-0.025419	24.0544	-13.455	-4.003	1.0304	-1.6719	-1.0410				
053306	0.000102440	-0.022168	-0.022168	-0.022168	14.6674	-12.470	-3.394	1.0685	-0.8096	-1.2567				
044129	0.000155594	-0.026655	-0.026655	-0.026655	27.3324	-15.270	-1.148	0.1053	-0.6761	-1.1019				
038251	0.000132423	-0.024716	-0.024716	-0.024716	25.8588	-12.088	-4.378	1.7348	-3.9234	0.6263				

ORIGINAL PAGE IS  
OF POOR QUALITY

Table 4. Results of Jackknife Test for State of Rio Grande do Sul Wheat Model

Q=OBSERVED YIELD  
P=MODEL'S PREDICTED YIELD

PLOT OF YIELD\*YIELDYR SYMBOL USED IS P  
PLOT OF Q\*YIELD\*YIELDYR SYMBOL USED IS Q

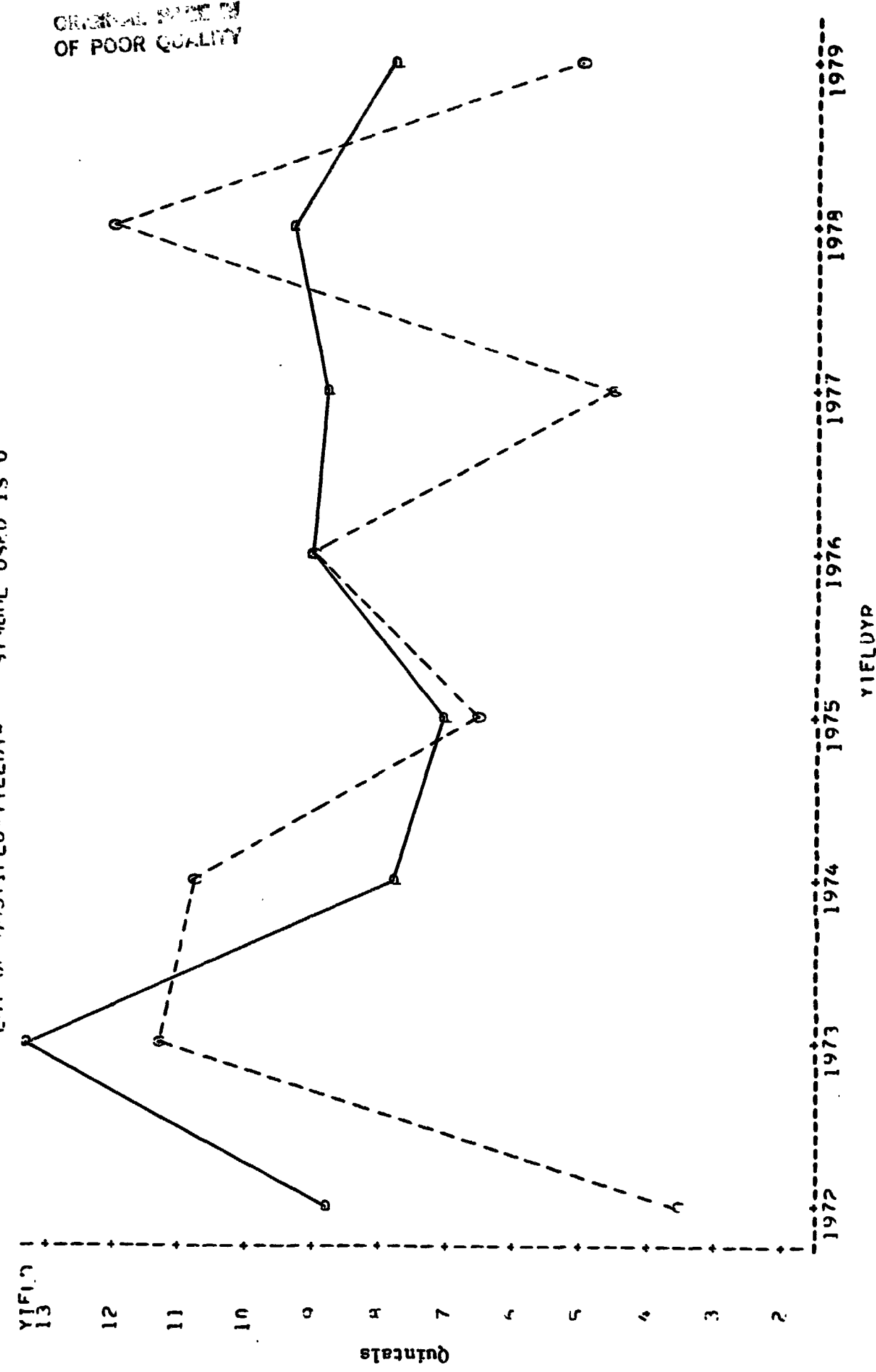


Figure 5. State of Rio Grande do Sul Wheat Model

HETA2	HFGMDLY2	FNMMDLYR	YIELDYR	YIELD	OBSYIELD	PSQ	PIDP	DFRES	SPRES	HETA1
1842269	1973	1979	1972	5.6879	5.10	0.344354	12.8700	15	7.08494	7.35043
145179	1972	1979	1973	7.0093	8.48	0.404981	15.8994	15	6.53424	7.90577
067416	1972	1979	1974	8.8690	8.15	0.343560	1.9440	15	6.01792	8.64346
171277	1972	1979	1975	7.2199	4.50	0.404770	4.5806	15	6.87730	7.61866
451800	1972	1979	1976	3.9227	5.42	0.435712	7.2783	15	2.90850	5.17877
064255	1972	1979	1977	10.7720	3.42	0.731782	0.4588	15	6.15536	9.19480
224322	1972	1979	1978	10.0950	6.15	0.463774	1.6311	15	6.15536	6.22878
	1972	1979	1979	5.4724	6.15	0.459698	1.9774	15	6.15536	6.22878
HETA3	HETA4	HETA5	HETA6	CONRIM1	CONRIP2	CONRIP3	CONRIP4	CONRIP5	CONRIP6	CONRIP7
018688	0.0000541226	-0.0095112	-0.024867	7.35043	2.11910	-4.7694	3.52506	-2.3229	-0.2145	
023050	0.0000323832	-0.007579	-0.019501	7.90577	1.74215	-5.3990	2.16077	-1.1237	1.7234	
021004	0.0000411320	-0.004056	-0.025397	8.64460	0.84630	-1.3503	0.17016	-0.0154	0.5429	
014352	0.0000287391	-0.007513	-0.031867	9.64346	0.25771	-2.2195	0.37803	-0.5847	0.7443	
023002	0.0000559551	-0.007740	-0.030820	7.61866	1.71277	-2.5858	0.59851	-0.1353	-3.2862	
021544	0.0000493990	-0.015570	-0.024336	5.17877	6.28002	-1.0136	0.10936	-0.1138	0.2038	
020427	0.0000305174	-0.009736	-0.031134	9.19480	0.83532	-0.5631	0.03012	-0.5591	1.2570	
005448	0.000012596	-0.011512	-0.026926	6.22878	2.28809	-0.2896	0.00359	-0.9545	-1.7940	

ORIGINAL PAGE 38  
OF POOR QUALITY

Table 5. Results of Jackknife Test for State of Santa Catarina Wheat Model

ORIGINAL PAGE IS  
OF POOR QUALITY

PLOT OF YIELD\*YIELDYR  
PLOT OF OBSERVED YIELD  
P=MODELS PREDICTED YIELD  
SYMBOL USED IS P  
SYMBOL USED IS O

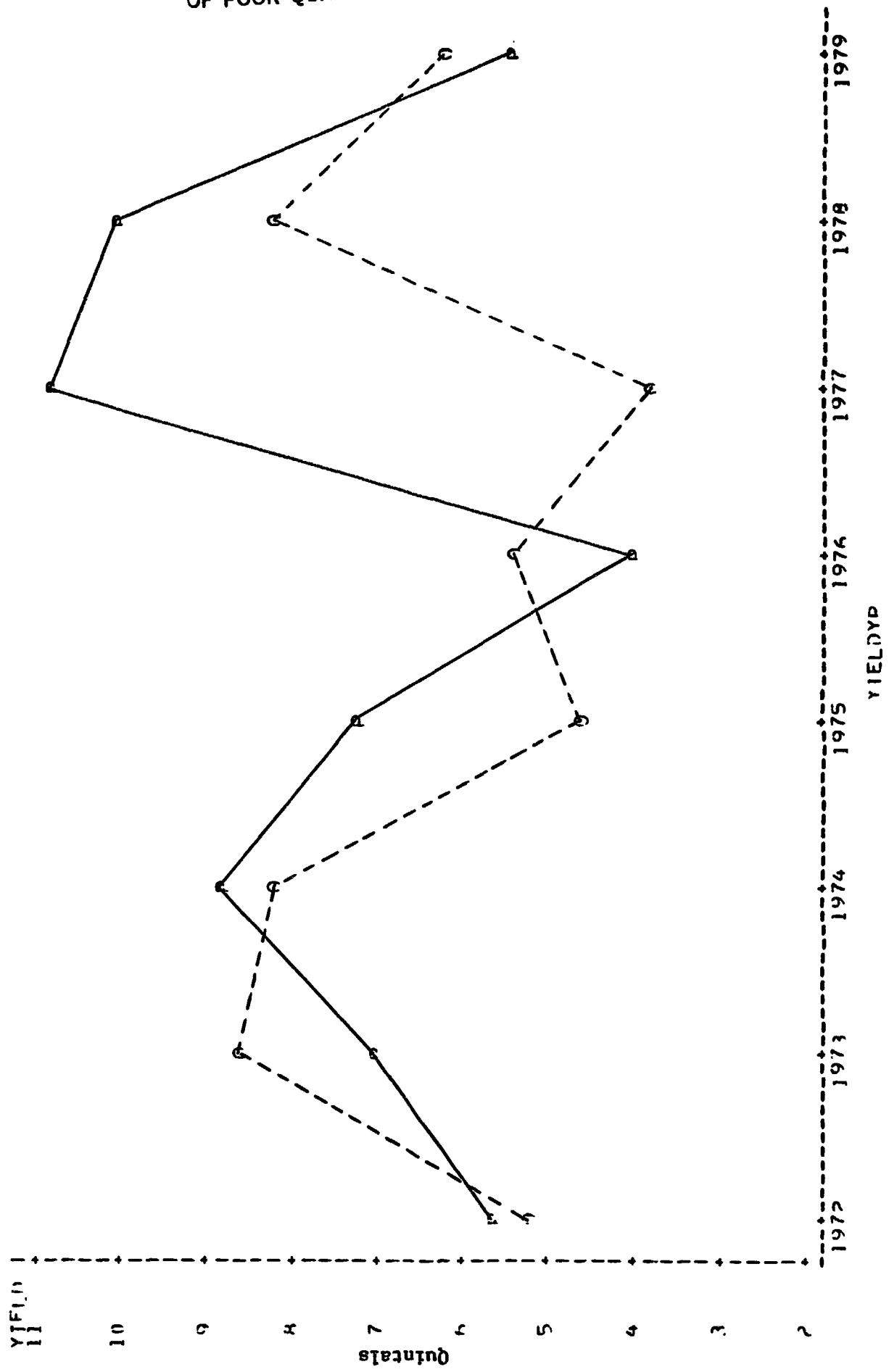


Figure 6. State of Santa Catarina Wheat Model