Mausam, (1989), 40, 4, 441-446

633.18 : 551.509.313(541.23)

Forecasting rice yields in Gangetic West Bengal using rainfall and agricultural technology

RAJENDRA PRASAD* and S. N. DUDHANE

Meteorological Office, Pune

(*Received 22 May 1987*)

सार — प्रतिवर्ष प्रौद्योगिकीय प्रवृत्ति को अपनाकर और अन्ततः इस क्रम के विस्तार को आरम्भ करते हुए वर्षा के प्रत्यक्ष और व्युत्पन्न प्राचलों का प्रयोग करते हुए, गंगावर्ती पश्चिम बंगाल में चावल की उपज का पूर्वानुमान देने के उद्देश्य से, समाश्रयण निदर्शों (मॉडलों) को विकसित करने के लिए रैखिक बहु-समाश्रयण की परंपरागत पद्धति को क्रमबद्ध रूप में लागू किया गया है। इस अन्तर के विश्लेषण से यह स्पष्ट होता है कि सभी समाश्रयण निदर्श 1 प्रतिशत पर अत्यन्त सार्थक थे। लिए गए नमूने के सभी वर्षों में निदर्शों द्वारा आकलित उपज 10 प्रतिशत के अन्तर्गत थी। समाश्रयण निदर्शों से यह बात स्पष्ट हुई कि 0.90 से अधिक बहु-सहसम्बन्ध गुणांक सहित उपज में कुल अन्तर 80 प्रतिशत से अधिक था।

ABSTRACT. The traditional method of linear multiple regression has been employed systematically to develop regression models to forecast rice yields in Gangetic West Bengal by using direct and derived parameters of rainfall, introducing technological trend and extending the series subsequently every year. The analysis of variance revealed that all regression models were highly significant at 1%. The yields estimated by the models for all sample years lay within $\pm$ 10%. The regression models explained more than 80 per cent of total variation in yields with multiple correlation coefficients exceeding 0.90.

1. Introduction

The impact of weather and climate on food production is of vital importance. According to McQuigg (1975), there are three major sources of variability in yields of grain in a region over a long series of years. These are identified as (*i*) technological change, (*ii*) meteorological variability and (*iii*) random 'noise'. The technological change in yields creeps in due to the recent advances in agricultural technology. A suitable linear time scale dummy variable is introduced in crop yield models to account for the technological trend. Among meteorological parameters, rainfall, temperature of the soil and air, soil moisture etc are some of the important elements which affect crop yields significantly. The impact and contribution of random 'noise' is very little.

An agroclimatic study of the relationship between crop yields and weather parameters is carried out with the help of empirical-statistical multiple regression models. These models are generally employed for making quantitative crop yield forecast on operational basis. Multiple regression models have been developed by Das *et al.* (1971), Sreenivasan *et al.* (1973), Chowdhury *et al.* (1981), Rao *et al.* (1978, 1984) and many others to forecast the yields of principal crops of India. Similar technique has been utilised in the present study to develop linear multiple regression models to forecast yields of rice crop in Gangetic West Bengal where it is generally grown during June to November.

Rainfall is the most important weather parameter that affects rice crop in this region where irrigation plays a minor role. The water requirement of the crop is mainly fulfilled during monsoon season. Monsoon rainfall constitutes about 76% of the annual rainfall and July/August are the rainiest months (IMD 1962). Depressions and low pressure areas (surface and upper air) are the prominent synoptic systems which cause active to vigorous monsoon conditions. In September, monsoon depressions generally intensify into cyclonic storms. In addition to these, other rain producing systems include troughs in mid-latitude westerlies in the middle and upper troposphere which on extending well southwards cause active monsoon conditions (Srinivasan *et al.* 1972). Moreover, the geographic location of Gangetic West Bengal is favourable to the fall of rain from the rain bearing systems traversing over it during monsoon period, which make it, essentially, a rainfed area. The aim of this study is, therefore, to investigate and establish relationship between rainfall and rice yields in this region by using multiple regression technique. Impact of agricultural technology on rice yields has also been examined critically and a modified model has been developed by using technological trend as an independent variable.

2. Data

A data series of 23 years, from 1958 to 1980, has been used in the present study to develop regression models. The average rice production and area figures of ten

*Present affiliation : Meteorological Office, New Delhi.

TABLE 1

Basic and modified models to forecast rice yields in Gangetic West Bengal

Model	Data series	Multiple regression equation	M.C.C.	Percent variation	*F*-value Compu-ted	*F*-value Tabula-ted (1%)
Basic	1958-80	$Y_e = 558.191 + 1.235X_2 + 31.637X_3 + 42.595X_4 + 9.467X_5$ (3.12)* (3.52) (5.36) (2.19)	0.91	83	21.17	4.58
Modified	1958-80	$Y_e = 597.029 + 1.034X_2 + 30.591X_3 - 42.025X_4 - 7.525X_5 - 27.325X_6$ (2.73)* (3.48) (5.38) (1.77) (1.79)	0.92	85	18.66	4.34

Y_e—Estimated yield (kg/ha), X_2—Total rainfall (mm) during 13 to 17 July (Transplantation), X_3—Crop rainy days during 3 to 10 August (Active tillering), X_4—Rainy days during 21 to 28 August (Elongation) and X_5—Restricted rainy days during 11 September to 6 October (Flowering).

*Note : Student's t-values are indicated within brackets.

districts, Murshidabad, Burdwan, Nadia, Birbhum, Bankura, Purulia, Midnapur, Hoogley, 24-Parganas and Howrah were extracted from *Season and Crop Reports* and the *Agricultural Situation in India* published by the State Government and the Economic & Statistical Adviser, Ministry of Agriculture, Government of India respectively. Average of these districts was used in the regression analysis. Similarly, average daily rainfall data based on daily rainfall of seven observatories at Berhanpur, Burdwan, Krishnanagar, Bankura, Purulia, Midnapur and Alipore well distributed over the region under study was used in the regression analysis for the crop growing period from June to October.

The daily and monthly data of direct as well as derived parameters of rainfall was used in the study. The daily derived parameters include :

(1) *Rainy Days (RD)* — Days when the rainfall was equal to or more than 2.5 mm.

(2) *Restricted Rainy Days (RRD)* — Number of rainy days except when preceded by a day having 50 mm or more rainfall.

(3) *Crop Rainy Days (CRD)* — Number of days when the rainfall was 5 mm or more. But a day having rainfall 3.8 mm was considered as a crop rainy day with a gap of two or less non-rainy days.

It is well known that the influence of rainfall on crop yields is closely associated with the role of actual and potential evapotranspiration in a region. Therefore, derived paramcters based on rainfall, actual and potential evapotranspiration were also used in the regression analysis. Of the monthly derived parameters, $P - P_E$ and P/P_E were used, where, P and P_E denote monthly rainfall and potential evapotranspiration respectively. The former shows the excess or deficit rainfall, whereas latter is termed as precipitation efficiency ratio. Monthly P_E values computed by Rao *et al.* (1971) have been utilised in the present study.

Steyaert *et al.* (1981) and Achutuni *et al.* (1982) suggested the use of agroclimatic indices, *viz*., Yield Moisture Index (YMI) and Generalised Monsoon Index (GMI) for the qualitative assessment of drought related food shortages, particularly of rice crop. These indices, monthly as well as cumulative, were used in the regression analysis. Climatic normals of 1901-1950 were utilised to obtain the monthly rainfall weights and the crop coefficients were computed by the method given by Doorenbos and Pruitt (1975).

3. Method

The traditional multiple regression technique has been employed to develop yield forecasting models. Firstly, a basic model has been developed by using direct and derived parameters of rainfall, without introducing technological trend, from a data series of 23 years, from 1958 to 1980.

Secondly, a modified model has been obtained by introducing an assumed technological trend in the basic model keeping other independent variables constant. The development of modified model was intended to improve the forecast of rice yields, particularly higher yields since 1983, by superimposing the impact of agricultural technology in the form of linear time scale dummy variable.

Lastly, modified model has been updated every year by extending the series up to 1981, 1982, 1983, 1984 and 1985 to obtain updated models 1, 2, 3, 4 and 5 respectively. While developing updated models, it was assumed that the same sensitive periods and variables remain significant in the subsequently extended series also. The advantage of developing updated models lies in their ability to ensure the stability of regression coefficients and the validity of regression models as confirmed by Student's '*t*' and '*F*' values respectively. These may suitably be termed as dynamic regression models.

The models have been verified with independent data for the years up to 1985 outside their sample series. The performance of the models has been examined critically by computing percentage departures of estimated and forecast yield figures.

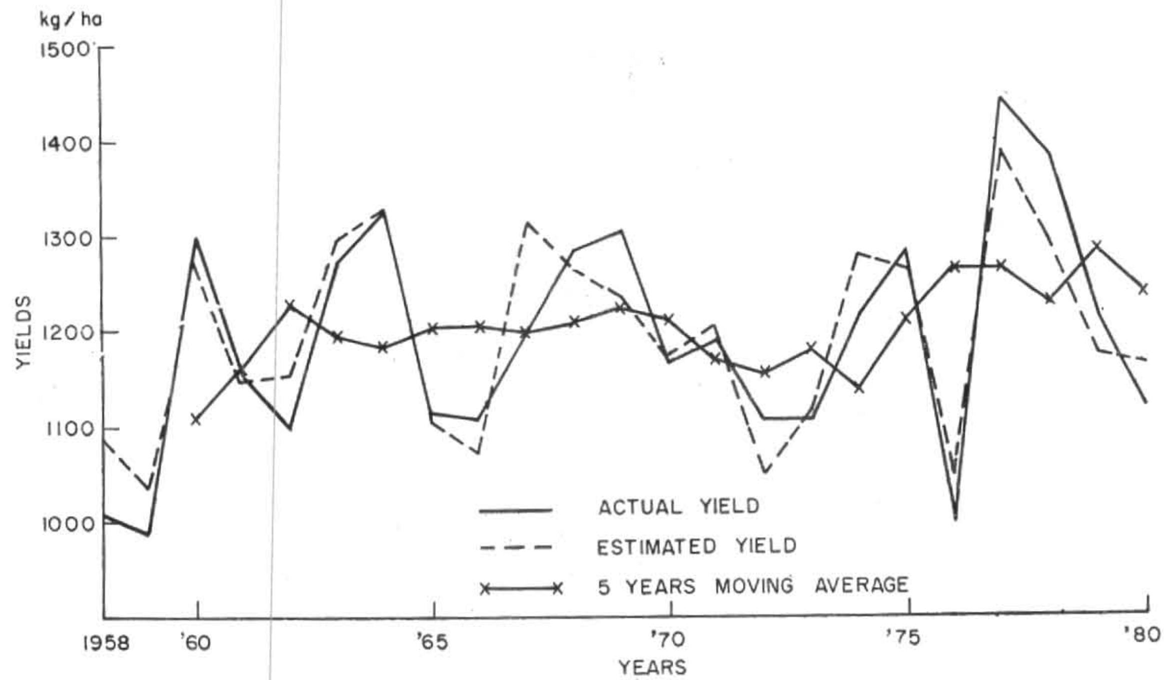

Fig. 1. Actual, 5 years' moving averages and estimated yields (Basic model)

4. Results and discussion

4.1. *Sensitive periods and parameters*

Out of all the optimum periods, the sensitive periods of statistical and phenological significance were selected for the regression analysis. Rainfall during 13 to 17 July, CRD during 3 to 10 August, RD during 21 to 28 August and RRD during 11 September to 6 October emerged significant and useful to the crop. The sensitive periods represent transplantation, active tillering, elongation and flowering stages respectively as per the appropriate crop calendar.

4.2. *Basic model*

The basic model obtained from a data series of 23 years (1958-80) is shown in Table 1. It may be seen that all the independent variables are significant at 5% as confirmed by Student's *t*-test. The model successfully accounts for 83% of total variation in rice yields with a multiple correlation coefficient of 0.91. The analysis of variance proved that the model was highly significant at 1% when its computed *F*-value of 21.17 was compared with the tabulated value of 4.58.

The percentage departures of estimated yields for all sample years lay within ± 10%. The verification of the basic model carried out for the subsequent 5 years outside the sample size up to 1985 is shown in Table 2. The percentage departures of forecast yields for 1981 and 1982 lay within ± 10% whereas they exceeded for remaining years, due apparently to the abnormal increase in reported yields since 1983. The actual and 5 years moving average yields alongwith those estimated by the basic model are shown in Fig. 1.

4.3. *Modified model*

The average of the yields during 1958-80 was 1190 kg/ha with standard deviation and coefficient of variation as 118 and 10 per cent respectively. The abnormal increase of yields in 1983, 1984 and 1985 to 1500, 1595 and 1545 kg/ha respectively was, presumably, due to recent advances in agricultural technology caused by the use of high yielding varieties large scale use of fertilizers, better irrigation facilities, use of insecticides and improved management practices. This is termed as the technological trend and is presumed due to the combined impact of these factors. This necessitated the need to modify basic model by introducing technological trend as an independent linear time scale dummy variable.

A careful examination of 5 years' moving averages in Fig. 1 would indicate that the yields tend to increase gradually since 1975. The trend is mild up to 1980 and rises steadily upwards thereafter. A linear time scale dummy variable was introduced in the basic model to account for technological trend.

A modified model, as shown in Table 1, was obtained with multiple correlation coefficient as 0.92. It accounted for 85% of total variation in yields. RRD and technological trend emerged significant at 10%, whereas other independent variables remained significant at 5% as confirmed by their *t*-values. The analysis of variance was carried out and the modified model was found to be highly significant at 1% as seen from a comparison of the comuted *F*-value of 18.66 with the tabulated value of 4.34.

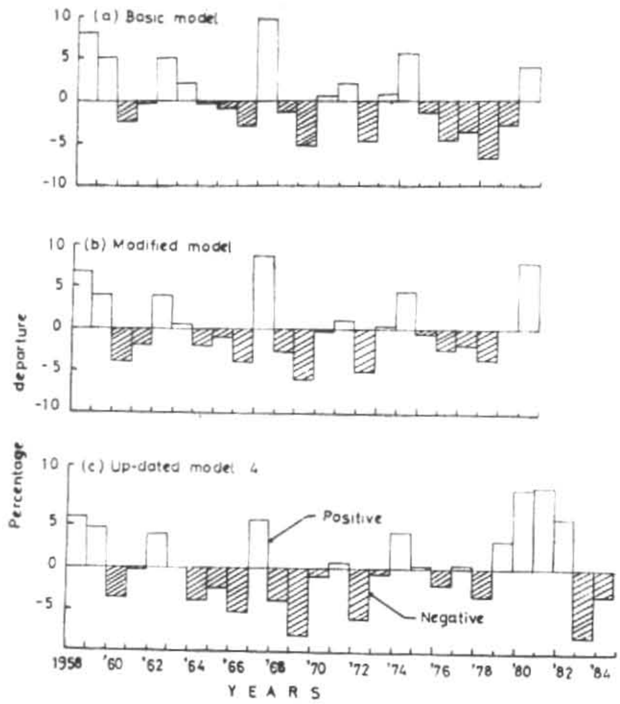

Fig. 2. Percentage departures of yields estimated by (a) Basic model (b) Modified model and (c) Updated model 4

TABLE 2

Percentage departures of rice yields forecast by basic and modified models

	Yield (kg/ha) of years				
	1981	1982	1983	1984	1985
Actual	1257	1228	1500	1595	1545
Basic model					
Forecast	1302	1193	1204	1339	1313
% dep.	+4	—3	—20	—16	—15
Modified model					
Forecast	1376	1313	1336	1490	1489
% dep.	+9	+7	—11	—7	—4

The percentage departures of the yields estimated by the modified model for all sample years were within tolerable limits (± 10%) as shown in Fig. 2. Verification of this model for subsequent 5 years, from 1981 to 1985, outside the sample series showed encouraging results. Percentage departures of forecast yields for all years, except 1983 (—11%), lay within ±10% as shown in Table 2. The model, therefore, tends to forecast higher yields successfully and shows improvement over basic model, as it takes into account the impact of agricultural technology on rice yields in Gangetic West Bengal. Fig. 3 shows the actual yields and those estimated by the modified model.

4.4. *Updated models*

The updated models are shown in Table 3. All updated models, except 3, were able to explain more than 80 per cent of total variation in yields with multiple correlation coefficients exceeding 0.90 and were highly significant at 1% as confirmed by their computed and tabulated *F*-values obtained from the analysis of variance. The yields estimated by them, for all sample years, lay within ±10% as shown in Fig. 2 for updated model 4. Fig. 4 shows the actual yields and those estimated by the updated model 4.

5. Conclusion

(*i*) Multiple regression technique is a formidable tool to develop empirical-statistical crop yield forecasting models using weather and technological parameters. A systematic methodology was adopted to develop basic, modified and updated models to forecast, quantitatively, rice yields in Gangetic West Bengal, from a data series of 23 years (1958-80) of direct and derived parameters of rainfall, with and without technological trend and subsequently extending the series every year up to 1985.

(*ii*) The basic model explained 83 per cent of total variation in yields with multiple correlation coefficient as 0.91. All independent variables emerged significant at 5%.

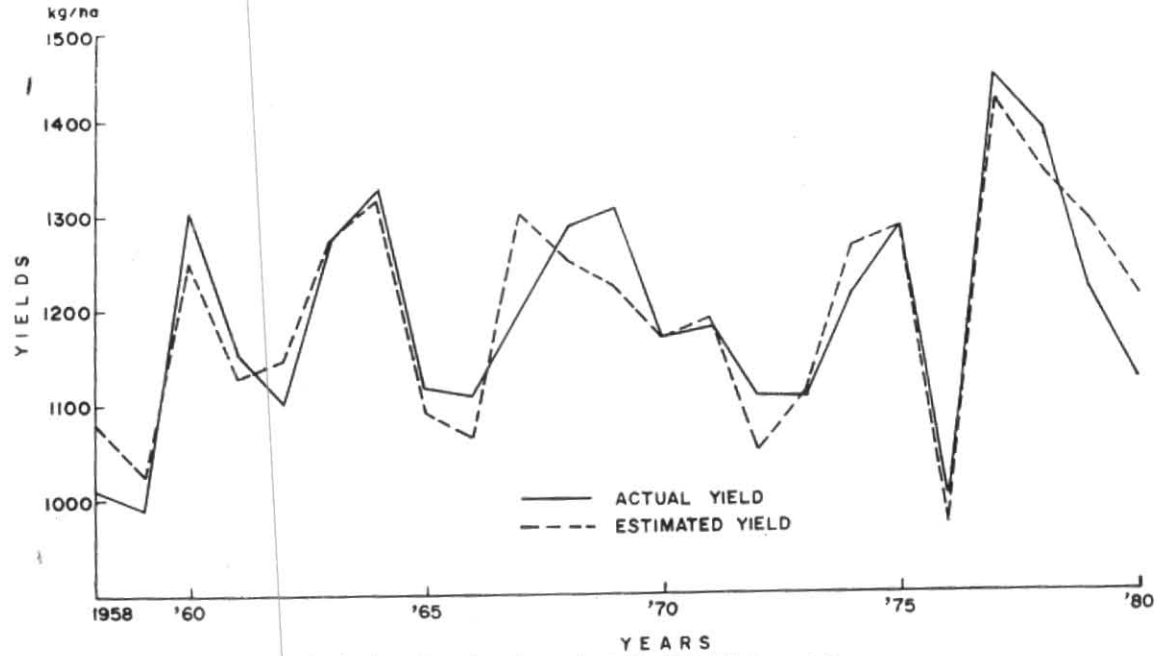

Fig. 3. Actual and estimated yields (Modified model)

TABLE 3

Updated models to forecast rice yields in Gangetic West Bengal

Updated model	Multiple regression equation	M.C.C.	Percent variation	F-value Computed	F-value Tabulated (1 %)
Model 1 (1958-81)	$Y_e=604.661+1.244X_2+27.493X_3+37.914X_4+9.140X_5+13.794X_6$	0.908	82.56	16.98	4.25
Model 2 (1958-82)	$Y_e=604.726+1.244X_2+27.501X_3+37.916X_4+9.132X_5+13.814X_6$	0.909	82.62	18.00	4.17
Model 3 (1958-83)	$Y_e=605.593+1.207X_2+25.350X_3+35.042X_4+10.703X_5+36.217X_6$	0.884	78.22	14.33	4.10
Model 4 (1958-84)	$Y_e=605.789+1.315X_2+23.499X_3+35.470X_4+10.544X_5+42.331X_6$	0.911	83.11	20.60	4.01
Model 5 (1958-85)	$Y_e=586.668+1.33X_2+25.430X_3+37.196X_4+10.328X_5+39.853X_6$	0.925	85.57	25.99	3.99

TABLE 4

Computed Student's t-values of independent variables in updated models

Updated model	X_2	X_3	X_4	X_5	X_6	Degrees of free-dom	Tabu-lated t-value at 5%
1	3.36	3.07	4.93	2.13	1.04	18	2.10
2	3.66	3.27	5.09	2.58	1.22	19	2.09
3	2.88	2.45	3.84	2.48	3.22	20	2.09
4	3.27	2.32	3.90	2.45	4.56	21	2.08
5	3.32	2.64	4.23	2.59	5.40	22	2.07

TABLE 5

Percentage departures of rice yields forecast by updated models

Year	Departure (%) of updated Model 1	Model 2	Model 3	Model 4
1982	—1	—	—	—
1983	—10	—10	—	—
1984	—8	—8	—7	—
1985	—11	—11	—1	+2

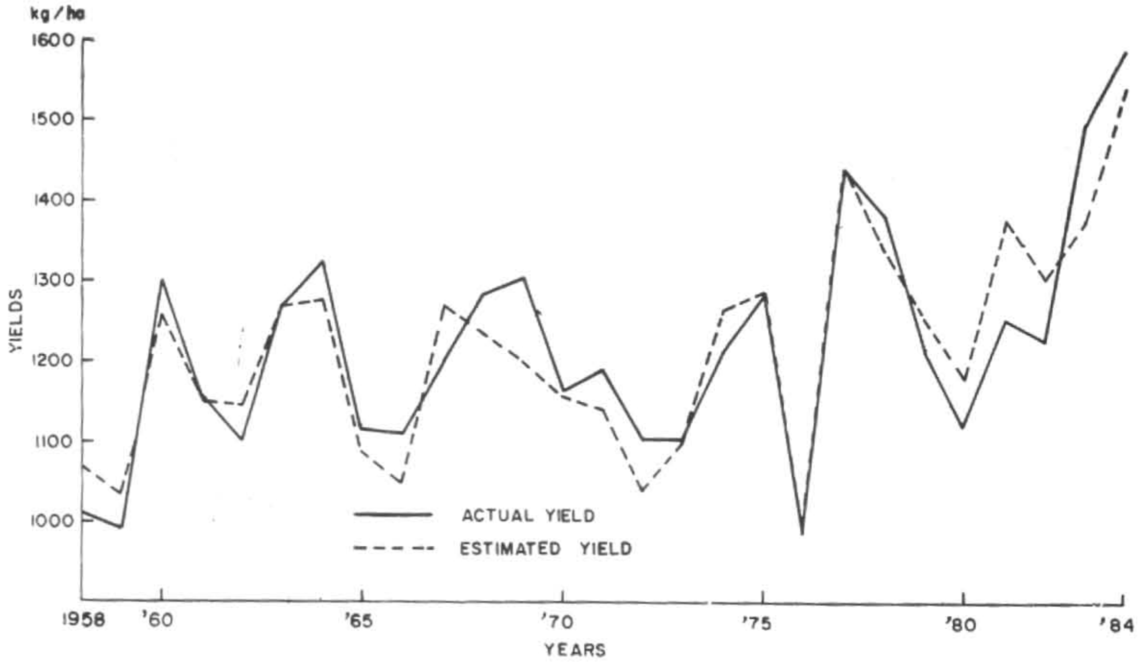

Fig. 4. Actual and estimated yields (Updated model 4)

(*iii*) The modified model, obtained from introducing technological trend in basic model, explained 85 per cent of total variation in yields with multiple correlation coefficient as 0.92. The yields forecast by it lay mostly within ±10%.

(*iv*) All updated models, except for 1958-83, explained more than 80 per cent of total variation in yields with multiple correlation coefficients exceeding 0.90. All independent variables, except technological trend in upated models 1 and 2, emerged significant at 5% and revealed stability of regression coefficients. The updated models forecast rice yields mostly within tolerable limits of statistical significance.

(*v*) The analysis of variance confirmed that all regression models were highly significant at 1%. The models estimated yields within ±10% for all sample years.

(*vi*) The satisfactory performance of updated models suggests that these forecast, successfully, rice yields in Gangetic West Bengal from the direct and derived parameters of rainfall and agricultural technology.

Acknowledgements

The authors wish to express their sincere gratitude to Shri Nootan Das, Deputy Director General of Meteorology, looking after the current duties of Additional Director General of Meteorology (Research) for constant encouragement during this study. Thanks are also due to Shri A.G. Tamboli for typing the manuscript neatly.

References

Achutuni, V.R., Steyaert, L.T. and Sakamoto, C.N., 1982, Agroclimatic Assessment Methods for Drought/Food shortage in South and Southeast Asia—Test and Evaluation : Final Report.

Appa Rao, G., Sarwade, G.S., Jaipal, Sarker, M.B., Joseph Lawrence and Jangle, N.K., 1978, Forecasting rice in India from weather parameters : Marathwada, Rayalseema, Gujarat and Himachal Pradesh, India Met. Dep. Pre-publ. Sci. Rep. No. 78/15.

Appa Rao, G. and Dudhane, S.N., 1984, Weather features associated with abnormal wheat yields of Himachal Pradesh, *Mausam,* **35**, 3, pp. 277-280.

Chowdhury, A. and Sarkar, M.B., 1981, Estimation of rice yield through weather factors in a dry sub-humid region, *Mausam,* **32**, 4, pp. 393-396.

Das, J.C., Mehra, A.K. and Madnani, M.L., 1971, Forecasting the yield of principal crops in India on the basis of weather—Paddy/rice, *Indian J. Met. Geophys.*, **22**, pp. 47-58.

Doorenbos, J. and Pruitt, W.O., 1977, *Guidelines for predicting Crop Water Requirements,* FAO Irrigation and Drainage Paper No. 24, Food and Agricultural Organisation, Rome.

Mc Quigg, J.D., 1975, Economic impacts of weather variability, University of Missouri-Columbia, Department of Atmospheric Science, 256 pp.

India Meteorological Department, 1962, *Mem. India Met. Dep.*, **31**, Pt. III.

Rao, K.N., George, C.J. and Ramasastri, K.S., 1971, Potential Evapotranspiration over India, India Met. Dep. Pre-publ. Sci. Rep. No. 136.

Sreenivasan, P.S. and Banerjee, J.R., 1973, Influence of rainfall on the yield of rainfed rice at Karjat (Colaba District), *Agri. Met.*, **11**, pp. 285-292.

Srinivasan, V., Raman, S. and Mukherjee, S., 1972, *Forecasting Manual,* Part III, 3.6, India Met. Dep.

Steyaert, L.T., Achutuni, V.R. and Atanas Todorov, 1981, Agroclimatic Assessment Methods for Drought/Food shortage in South and Southeast Asia—Proposed Early Warning Program: Final Report.