Mausam, (1991), **42**, 2, 125-130

551.509.33 : 551.553.21

A power regression model for long range forecast of southwest monsoon rainfall over India

VASANT GOWARIKER*, V. THAPLIYAL**, S. M. KULSHRESTHA†,
G. S. MANDAL†, N. SEN ROY† and D. R. SIKKA††

(*Received 27 February 1991*)

सार – अनेक क्रमचयों तथा संचयों में क्षेत्रीय और भूमंडलीय पूर्ववर्ती संकेतों की तुलना में कई दशकों की दक्षिण-पश्चिम मानसून (जून से सितम्बर) वर्षा के विस्तृत विश्लेषणों के परिणामस्वरूप लेखक इस निष्कर्ष पर पहुंचे हैं कि एक, दो, तीन या चार प्राचलों पर आधारित दीर्घ-अवधि पूर्वानुमान, जिनका कि पहले कई व्यक्तियों ने प्रयास किया, सभी अवसरों पर विश्वसनीय नहीं हो सकते। वास्तव में यही साबित हुआ है। इस शोध-पत्र में वर्णित 16 प्राचलों पर बना प्राचलिक व घात समाश्रयण मॉडल यह सुझाता है कि क्षेत्रीय एवं भूमण्डलीय जलवायुविक प्रणोदनों के अनेक प्राचल एवं परस्पर क्रिया की प्रकृति एक चित्र-यवनिका (टेपेट्री) है जो कि मानसून की गुणवत्ता और मात्रा का निर्धारण करती है। पूर्ववर्ती जलवायु अवस्थाओं और मानसून के मध्य अरैखिक परस्पर-क्रियाओं के विस्तृत विश्लेषणों के परिणामस्वरूप लेखकों ने विभिन्न प्राचलों के संकेतों पर अनुपातिक महत्व की संकल्पना का प्रवर्तन किया है। इसके परिणामस्वरूप घात समाश्रयण मॉडल का विकास किया गया जो कि प्रत्येक प्राचल के प्रभाव की मात्रा को निर्धारित करने में सक्षम है। मॉडल का ब्यौरा दिया गया है। इस मॉडल के आधार पर भारत मौसम विज्ञान विभाग 1988 से 1990 अर्थात् पिछले तीन वर्षों के दौरान समूचे भारत में मानसून वर्षा के प्राचलन दीर्घ-अवधि पूर्वानुमान जारी करता रहा है और ये पूर्वानुमान सही सिद्ध हुए हैं।

ABSTRACT. A detailed analysis of southwest monsoon (June to September) rainfall over India of several decades *vis-a-vis* the regional and global antecedent signals in numerous permutations and combinations has led the authors to conclude that a long range forecast based on one, two, three or four parameters as attempted by several workers in the past, cannot be reliable on all occasions as indeed has proved to be the case. The parametric and power regression models utilizing 16 parameters, described in the present paper, suggest that it is a tapestry of several parameters and interactive nature of the regional and global climatic forcings that govern the quality and quantity of the monsoon. A detailed analysis of non-linear interactions among the antecedent climatic conditions and the monsoon has led the authors to introduce the concept of proportionate weightage to the signals of different parameters. This has led to the development of a power regression model which is able to quantify the effect of each parameter. Details of the model are presented. Based on the model, the India Meteorological Department has been issuing the operational long range forecast of monsoon rainfall over India as a whole during the past 3 years, 1988 to 1990, and these forecasts have proved to be corrrect.

Key words — Southwest monsoon, monsoon forecast, long range forecast, parametric model, power regression model.

1. Introduction

India was the first country to start, a century ago, a systematic development of Long Range Forecast (LRF) techniques for estimating in advance seasonal monsoon rainfall over the country. Noticing the negative association of excessive winter and spring snowfalls in Himalayas with the southwest monsoon (June to September) rainfall over India, Blanford (1884) issued tentative monsoon forecasts from 1882 to 1885 based on the single parameter of Himalayan snow cover. Walker (1908, 1910) introduced the correlation concept in LRF and developed a multiple regression forecast model which used 4 parameters, namely the Himalayan snow accumulation at the end of May, the south American pressure during spring, the Mauritius pressure during May, and the Zanzibar rainfall during April and May. This model was used by the India Meteorological Department (IMD) for issuing operational forecasts from 1909 to 1915. In 1916, Walker (1923, 1924) added one more parameter, namely May rainfall of Ceylon. Many attempts have since been made to identify more parameters. Reviews on the subject are available, *e.g.*, Jagannathan (1960), Thapliyal (1987).

Gowariker *et al.* (1989), hereinafter their work referred to as Paper I, introduced a new LRF technique comprising parametric and power regression models. The models described in Paper I utilize signals from 15 antecedent parameters (since made 16 parameters by adding northern hemisphere surface pressure anomaly) which are significant among the regional and global forcings for the subsequent monsoon. Based on the new model, IMD has been issuing operational LRFs for monsoon rainfall over India as a whole during the past 3 years, 1988 to 1990. Details of the model along with the analysis of its performance are discussed in the present paper.

2. The parametric model

2.1. *Analysis of signals from combinations of a few parameters*

Table 1 shows the year to year performance of southwest monsoon during 1951-1990. In this table, N indicates good monsoon which includes both normal

*Department of Science and Technology, New Delhi. **India Meteorological Department, Pune.
†India Meteorological Department, New Delhi. ††Indian Institute of Tropical Meteorology, Pune.

TABLE 1

Analysis of 16 parameters

		Temperature						Wind			Pressure					Snow cover			
Year	Monsoon condition	*El Nino* (Same year) (—)	*El Nino* (Previous year) (+)	N. I. (March) (+)	E. C. I. (March) (+)	C. I. (May) (+)	N. H. (Winter) (+)	500 hPa (April) (+)	50 hPa (Winter) (—)	10 hPa Zonal (January) (+)	N. H. (January to April) (+)	S. O. I. (Spring) (+)	Darwin (Spring) (—)	Argentina (April) (—)	I. O. E. (January to May) (—)	HIM (January to March) (—)	Eurasian (Previous December) (—)	No. of parameters favourable /Total No. of parameters	% of favourable parameters
1951	D	F	U	U	U	F	U	U			F	U	U	U	F	U		4/13	31
1952	N	F	U	U	U	F	F	U			F	F	U	U	U	F		6/13	46
1953	N	U	U	F	F	F	F	F			F	U	U	U	U	F		7/13	54
1954	N	F	F	F	F	F	U	F			U	F	F	U	F	U		9/13	69
1955	N	F	U	F	F	F	F	F			U	F	F	F	F	F		11/13	85
1956	N	F	U	F	U	F	U	F			U	F	F	F	F	F		9/13	69
1957	N	U	U	U	U	U	U	F			F	U	U	U	U	U		2/13	15
1958	N	F	F	F	F	F	F	F	U	U	F	U	U	F	U	F		10/15	67
1959	N	F	U	F	U	F	F	F	F	F	F	F	F	F	F	U		12/15	80
1960	N	F	U	U	U	U	F	F	U	U	F	F	F	F	F	F		9/15	60
1961	N	F	U	F	F	F	F	U	F	F	F	U	U	F	F	U		10/15	67
1962	N	F	U	U	U	U	F	U	F	U	U	F	U	U	F	F		6/15	40
1963	N	F	U	U	U	U	F	U	F	F	U	F	F	F	F	F		9/15	60
1964	N	F	U	F	F	U	U	F	F	F	F	F	F	U	F	F		11/15	73
1965	D	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U		0/15	0
1966	D	F	F	U	F	U	F	U	F	F	U	U	U	F	U	F		8/15	53
1967	N	F	U	U	U	U	U	F	F	U	F	F	U	U	F	F	F	8/16	50
1968	D	F	U	U	U	U	U	U	U	U	F	F	U	U	F	U	U	4/16	25
1969	N	F	U	F	F	F	U	F	F	F	U	U	U	F	U	F	U	9/16	56
1970	N	F	U	F	F	F	F	F	F	U	U	U	U	F	U	F	F	10/16	63
1971	N	F	U	U	U	U	U	F	F	F	F	F	F	F	F	U	F	10/16	63
1972	D	U	U	U	U	F	U	U	U	U	F	U	U	F	F	U	U	4/16	25
1973	N	F	F	U	F	F	F	F	F	F	F	U	U	F	U	U	U	10/16	63
1974	D	F	U	F	F	U	U	U	U	U	U	F	F	U	F	U	U	6/16	37
1975	N	F	U	U	F	F	F	F	F	F	F	F	F	F	F	U	F	13/16	81
1976	N	U	U	F	F	U	F	F	F	F	F	F	U	U	F	F	U	10/16	63
1977	N	F	F	F	F	U	U	U	F	F	F	U	U	F	U	F	F	10/16	63
1978	N	F	U	U	U	F	F	U	U	F	U	F	U	U	U	F	U	6/16	37
1979	D	F	U	U	U	U	U	U	U	U	U	U	U	U	U	F	U	2/16	13
1980	N	F	U	F	F	F	F	U	F	F	F	U	U	F	U	U	F	10/16	63
1981	N	F	U	F	F	U	F	F	F	U	U	U	U	U	U	F	F	8/16	50
1982	D	U	U	U	F	U	U	U	U	F	F	U	U	U	U	F	U	4/16	25
1983	N	F	F	U	F	F	F	U	F	F	F	U	U	F	U	F	U	10/16	63
1984	N	F	U	U	U	F	F	U	F	U	F	U	U	U	U	F	F	7/16	44
1985	N	F	U	F	F	F	U	U	F	F	F	U	F	F	F	U	F	11/16	69
1986	D	F	U	F	F	U	F	U	U	U	U	U	U	F	F	U	U	6/16	37
1987	D	U	U	F	U	U	F	F	U	U	U	U	U	F	U	F	U	5/16	31
1988	N	F	F	F	F	F	F	U	F	F	F	F	F	U	F	F	F	14/16	87
1989	N	F	U	F	U	F	F	F	U	U	F	F	F	F	F	F	F	12/16	75
1990	N	F	U	U	F	U	F	F	F	F	F	F	U	F	U	F	U	10/16	63

Symbols and abbreviations are explained in the Appendix

TABLE 2

Inferences based on signals from 16 parameters
(Period : 1951 to 1990)

Favourable parameters (%)	No. of occasions	Monsoon condition		Remarks
		Normal	Deficient	
≥80	4	4	0	Rainfall on +ve side of normal
≥70	6	6	0	
≥60	21	21	0	1963 (—2), 1985 (—7)
≥55	22	22	0	1969 (—2)
≥50	26	25	1*	*1966 (—14)
<50	14	5	9	
≤40	12	3	9	
≤30	6	1	5	
≤20	3	1@	2	@1957 (—2)
≤10	1	0	1	
0	1	0	1	

"Normal" and "Deficient" are defined in paragraph 2.1

TABLE 3

List of subscripts indicative of parameters

Subscript of X	Corresponding parameter
X_1	50 hPa wind pattern (Winter)
X_2	Eurasian snow cover (December)
X_3	500 hPa ridge (April)
X_4	Central India temperature (May)
X_5	10 hPa zonal wind pattern (January)
X_6	East coast of India temperature (March)
X_7	Northern hemisphere surface pressure anomaly (January to April)
X_8	Argentina pressure (April)
X_9	Northern hemisphere temperature (January & February)
X_{10}	SOI (Spring)
X_{11}	*El Nino* (Previous year)
X_{12}	North India temperature (March)
X_{13}	Indian ocean equatorial pressure (January to May)
X_{14}	*El-Nino* (Same year)
X_{15}	Himalayan snow cover (January to March)
X_{16}	Darwin pressure (Spring)

and excess rainfalls. Normal rainfall is defined as within ±10% of the long period average value and excess rainfall is defined as more than 110% of the long period average value. D indicates deficient monsoon (<90% of the long period average value). In the same Table, F and U respectively indicate favourable and unfavourable signals for good monsoon rainfall over India.

Firstly, an analysis of the favourable and unfavourable signals obtained from individual parameters was carried out. For this purpose, the correlation coefficients (CCs) between monsoon rainfall and individual parameters were computed; the CCs signs are given in Table 1. A parameter, showing positive CC with rainfall, gives favourable (unfavourable) signal for good monsoon if it has a positive (negative) depature. Reverse is true when a parameter shows a negative CC with rainfall. Following this procedure, F and U signals obtained from different parameters during the past 40 years (1951-1990) are shown in Table 1. This enables us to compare the number of favourable signals in any year with the performance of the corresponding monsoon.

The 16 parameters were then looked at in combinations of 2, 3 and 4 with a view to study the possible inferences regarding the monsoon. For example, it was observed that a combination of two parameters, namely the *El Nino* (current year) and the Darwin pressure which are connected with the ENSO phenomena, seems to be indicative of something interesting. During the past 90 years (1901-1990), these two parameters were together favourable on 37 occasions and out of these, the rainfall was normal on 35 occasions. On the other hand, both the parameters were unfavourable on 13 occasions but rainfall was deficient only on eight occasions. Similarly, a combination of 3 parameters, such as the 3 wind parameters (the 500 hPa ridge, the 50 hPa wind pattern and the 10 hPa wind), indicated that during the past 33 years (1958-1990), all the three wind parameters gave favourable signals on 8 occasions and on all these occasions the monsoon was normal. The three wind parameters were unfavourable on 6 occasions and so was the monsoon on these 6 occasions. Thereafter, combinations of 4 parameters were studied. During the past 40 years (1951-1990), all the four surface temperatures (temperatures of north India, east coast of India, central India and northern hemisphere) indicated favourable signals on 5 occasions and the monsoon was normal on all occasions. However, all the four parameters were unfavourable on 6 occasions but the rainfall was deficient only on 3 occasions.

While the above inferences would seem reasonable, it must be cautioned that in order to draw very definitive conclusions, one may have to have a larger sample size; it needs to be also noted that the available time series for some parameters are truncated and that the measurement and assessment techniques for a parameter like snow cover might require further standardization.

The above analysis would indicate that only in a very selective situation of parametric combinations that there can be a reasonable forecast reliability and that too when some of the most relevant parameters are involved. Such occasions are obviously not easy to come by when year to year operational forecast is required to be given. It was only after a very detailed study of the rainfall of several decades and of relevant parametric signals in several permutations and combinations that led the authors to conclude that a forecast based on one, two, three or just a few parameters, as attempted by several workers in the past, cannot be reliable on all occasions,

TABLE 4

Constants for power regression model

i	α_i	β_i	p_i	C_i
0	0	1		$+0.1303086 \times 10^{+2}$
1	50	10	1.4	$-0.5824688 \times 10^{-1}$
2	0	10	—2.0	$+0.3026653 \times 10^{+2}$
3	0	10	—4.0	$-0.1733019 \times 10^{+1}$
4	0	10	4.0	$+0.6958737 \times 10^{0}$
5	50	10	—0.4	$-0.1956036 \times 10^{+2}$
6	0	10	4.0	$+0.4374519 \times 10^{0}$
7	0	10	4.0	$+0.4687574 \times 10^{-2}$
8	0	100	—3.9	$-0.1662928 \times 10^{+4}$
9	50	10	—4.0	$-0.4617612 \times 10^{+3}$
10	50	10	4.0	$+0.5844691 \times 10^{-1}$
11	50	10	4.0	$-0.4456572 \times 10^{-2}$
12	0	10	4.0	-0.1497441×10^{0}
13	0	1000	3.3	$-0.1183124 \times 10^{+4}$
14	50	10	4.0	$-0.1344925 \times 10^{-1}$
15	50	10	—4.0	$+0.1940398 \times 10^{+5}$
16	0	1000	3.3	$+0.1257985 \times 10^{+4}$

α, β, p, C have been defined in section 3.2.

as indeed has proved to be the case. The basis of our monsoon model utilizing 16 parameters is that it is a tapestry of several parameters and the interactive nature of these global and regional forcings that govern the quality and quantity of Indian monsoon.

2.2. *Analysis of signals from all parameters*

The 16 parameters (*see* Table 1) appear to have physical linkages with the southwest monsoon system. The main feature of the present work lies in a detailed analysis of as many as 16 parameters taken together and relating the results to the behaviour of monsoon. The last column in Table 1 lists the percentage of favourable parameters yearwise. Inferences based on signals from the 16 parameters as obtained during 1951-1990 are given in Table 2. It will be seen that whenever 55% or more parameters were favourable, the subsequent monsoon rainfall was always normal; whenever the favourable factors exceeded 70%, the subsequent rainfall was not only normal but was also on the positive side of the normal. The interesting point is, however, that the situation involving the favourable factors being lower than 55% does not necessarily herald bad monsoon as can be seen from Table 2. In the context of the present work, one is thus talking, as it were, in terms of "sufficient" (and not "necessary") conditions in various sets of situations.

TABLE 5

Comparison of forecasts obtained from power regression model with observed monsoon rainfall over India

Year	Rainfall as percentage of normal	
	Actual	Computed (based on the model)
1958	110	110
1959	114	117
1960	101	104
1961	122	120
1962	97	97
1963	98	97
1964	110	111
1965	82	82
1966	87	90
1967	100	102
1968	90	85
1969	100	98
1970	112	110
1971	104	99
1972	76	76
1973	108	108
1974	88	92
1975	115	118
1976	102	102
1977	104	102
1978	109	107
1979	81	82
1980	104	108
1981	100	104
1982	85	91
1983	113	113
1984	96	100
1985	93	99
1986	87	88
1987	81	79
1988	119	113
1989	101	101
1990	106	102

3. The power regression model

3.1. *The premise*

The basic philosophy of the power regression model stems from the observation that even in the midst of adverse parametric conditions, Indian monsoon can still turn out to be normal. At any rate, the adverse poising of any climatic factor, no matter how significant or important it is viewed to be, need not create a unidimensional impression that the monsoon is going to be bad. We have developed a method of quantifying the effect of each of the factors — whether unfavourable or favourable — towards the quantified forecast of Indian monsoon by assigning due weightages to signals from different parameters. The power regression model based on the 16 global and regional parameters is described below :

3.2. *The details*

As mentioned above, power regression model acknowledges the non-linear interaction of different climatic forcings with the Indian monsoon system. Accordingly, the sign and the magnitude of the influence of each predictor on monsoon rainfall over India have been determined and used for developing the power regression model which takes the following form :

$$R = C_0 + \sum_{i=1}^{16} C_i \, X_i^{\,p_i} \qquad (1)$$

where, R is the southwest monsoon season's total rainfall over India as a whole; X denotes the parameters listed in Table 3; and Cs and ps are model constants.

For determining the exact form of the power regression model, data of past 31 years (1958-1988) have been used. The values of Cs and ps have been determined by using step by step iteration together with the usual method of least square fitting. The iterations have been performed with an increment of 0.1 which has been found effective for determining model constants. The final form of power regression model is given below :

$$\frac{R + \alpha_0}{\beta_0} = C_0 + \sum_{i=1}^{16} C_i \left(\frac{X_i + \alpha_i}{\beta_i}\right)^{p_i} \qquad (2)$$

where, R is monsoon rainfall over India as a percentage of the long period average, Xs are different parameters defined in Table 3 and αs, βs, ps and Cs are model constants and are given in Table 4.

To determine the best power regression model, several experiments have been performed. In the first stage, the best suited order of the parameters has been determined. The power regression model given in Eqn. (2) has been developed by arranging the parameters in order of their decreasing correlation coefficients with the monsoon rainfall over India. The 1989 and 1990 forecasts, for instance, based on the present model were 101% and 102% of the normal, respectively; the actual rainfall was 101% and 106% of the normal respectively. The values based on the model and those observed for 1958 to 1990 are given in Table 5. These show that the forecast rainfall figures lie roughly within $\pm4\%$ of the actual values. The model forecasts have been able to explain 94% variance of the observed rainfall.

4. Conclusions

4.1. A detailed study of monsoon rainfall *vis-a-vis* the regional and global climatological signals in several permutations and combinations has led the authors to conclude that a long range forecast based on one, two, or just a few parameters, as attempted by several workers in the past, cannot be reliable on all occasions as indeed has proved to be the case.

4.2. The parametric and power regression models utilizing 16 parameters suggest that it is the tapestry of several relevant climatic forcings that govern the quality and quantity of the Indian southwest monsoon.

4.3. A detailed analysis of non-linear interactions among the antecedent climatic conditions and the monsoon has led the authors to assign proportionate weightage to the signals of 16 parameters and use these weighted signals in a power regression model for quantifying the forecast.

4.4. The mathematical form of the present mode suggests that at any rate, the adverse poising of any parameter, no matter how significant or important it is viewed to be, need not create a unidimensional impression that the monsoon is going to be bad.

4.5. Forecasts issued by the India Meteorological Department from 1988 to 1990 were based on the authors' work and these forecasts have proved to be correct.

Acknowledgements

The enthusiastic support of the officers and members of staff of IMD's Long Range Forecast Unit at Pune is thankfully acknowledged.

References

Blanford, H.F., 1884, On the connection of the Himalayan snow with dry winds and seasons of droughts in India, *Proc. Roy. Soc. London*, **37**, pp. 3-22.

Gowariker, V., Thapliyal, V., Sarker, R.P., Mandal, G.S. and Sikka, D.R., 1989, Parametric and power regression models — New approach to long range forecasting, *Mausam*, **40**, 2, pp. 115-122.

Jagannathan, P., 1960, Seasonal forecasting in India — A review, Meteorological Office, Poona (India), 67 pp.

Thapliyal, V., 1987, Prediction of Indian monsoon variability evaluation and prospects including development of a new model, *Climate of China and Global Climate*, edited by D.Ye, C.Fu, J. Chao and J. Yoshino, China Ocean Press, Beijing, pp. 377-416.

Walker, G.T., 1908, Correlation in Seasonal Variation of Climate (Introduction), Mem. India Met. Dep., 20, Part VI, 117-124.

Walker, G.T., 1910, Correlation in Seasonal Variations of Weather II, Mem. India Met. Dep. XXI, XXII.

Walker, G.T., 1923, Correlation in Seasonal Variation of Weather VIII : A preliminary study of World Weather, Mem. India Met. Dep., 24, pp. 75-131.

Walker, G.T., 1924, Correlation in Seasonal Variability of Weather IX : A further study of World Weather, Mem. India Met. Dep., 24, pp. 275-332.

APPENDIX

Symbols and Abbreviations

CC	—	Correlation coeffcient
CI	—	Central India
D	—	Deficient
ECI	—	East Coast of India
ENSO	—	*El Nino* Southern Oscillation
F	—	Favourable
HIM	—	Himalayan
hPa	—	Hectopascal
IMD	—	India Meteorological Department
IOE	—	Indian Ocean Equatorial
LRF	—	Long Range Forecast
N	—	Normal
NH	—	Northern Hemisphere
NI	—	North India
Paper I	—	Gowariker, V., *et al.*, 1989, *Mausam* **40** (3)
SOI	—	Southern Oscillation Index
U	—	Unfavourable