Prediction of winter minimum temperature of Kolkata using statistical model

TAPAN KANTI CHAKRABORTY

Regional Meteorological Centre, Kolkata - 700 027, India

(*Received 1 June 2004, Modified 3 June 2005*)

सार – कोलकाता (अलिपुर) के न्यूनतम तापमान का वास्तविक पूर्वानुमान 12 घंटे पूर्व जारी करने के उद्देश्य से पाँच प्राचलों पर आधारित बहु रैखिक समाश्रयण मॉडल विकसित किया गया है । इसके पूर्व सूचकों का चयन अलिपुर वेधशाला से प्राप्त सतह आँकड़ों तथा मौसम कार्यालय कोलकाता के निम्न स्तर के पवन आँकड़ों के आधार पर किया गया है । यह मॉडल 237 दिनों के (वर्ष 1997–2000 की अवधि के जनवरी एवं फरवरी माह के) आँकड़ों के नमूनों तथा काफी लंबी अवधि (वर्ष 1988–2004) के न्यूनतम तापमान के आँकड़ों मे स्थिरता की जाँच के आधार पर विकसित किया गया है । इस मॉडल की जाँच 178 दिनों के आँकड़ों के स्वतंत्र नमूने के आधार पर की गई है । इस मॉडल की क्षमता की जाँच सांख्यिकीय आँकड़ों के आधार पर की गई है और इसे सकारात्मक पाया गया है। इस मॉडल का उपयोग मौसम पूर्वानुमानकर्त्ता द्वारा न्यूनतम तापमान के पूर्वानुमान का आकलन करने के लिए किया जा सकता है और यदि बादल तथा पवन की गति के रूख में बाद में परिवर्तन संभावित हो तो उसमें सुधार किया जा सकता है ।

ABSTRACT. Five parameter multiple linear regression model for objective forecasting of minimum temperature of Kolkata (Alipore) with 12 hours lead period has been developed. The predictors are chosen from the available surface data of Alipore observatory and low level wind data of M. O. Kolkata. Model has been developed from data sample comprising of 237 days (in January and February, period: 1997 – 2000) after stationarity test of minimum temperature data of much longer period (1988–2004). The model is tested with independent sample of 178 days. Efficiencies of the model have been tested with statistical skill score and found to be positive. The model can be used by the forecaster for assessing prediction minimum temperature and modify if cloud cover and wind flow pattern are expected to change subsequently.

Key words – Skill score, Multiple correlation coefficient, Variance explained.

1. Introduction

Temperature of a place mainly depends on a number of meteorological parameters *viz.*, net solar radiation received from the sun, cold/warm advection of air mass, cloud cover, local effects such as nature of soil, orography and topography of that place. Kolkata located on the southern part of Gangetic West Bengal often experiences changes of air masses – south/southeasterly warm and moist airmass from the Bay of Bengal and west/northwest cold and dry airmass from the land areas of Bihar and Jharkhand round the year. During winter months (January and February) series of Western Disturbance (WD) from NW India move across the country in a NE/E'ly direction and some of these extend as far as Kolkata causing light rain, cold spell, fog and haze in its neighbourhood and paralyse normal life. Sometimes minimum temperature suddenly drops down by six degrees Celsius or more. Such a condition however does not continue for long and local weather forecast of minimum temperature based on synoptic condition and trend values of some meteorological parameters may not yield good result on such occasions.

Many authors have carried out studies on minimum and maximum temperature. WMO (1966), Kendal and Stuart (1968), Panofsky and Brier (1968) provide detailed description of various forecasting techniques including the widely used statistical technique of multiple linear regression analysis. Klien and Hammons (1975) suggested a large number of meteorological parameters as the potential predictors for the prediction of maximum and minimum temperatures. Singh and Jaipal (1983), Raj (1989), Charantoris and Liakatas (1990), Vashisth and Pareek (1991), have attempted minimum temperature forecast using various techniques including Markov chains. Attri *et al*. (1995) have used multiple regression method for forecasting minimum temperature at Gangtok in Sikkim. Dimri *et al*. (2002) have attempted minimum temperature forecast on hilly station Manali in Himachal Pradesh using multiple linear regression models for prediction of maximum and minimum temperature during December to February.

Present study is an effort to develop deterministic statistical prediction model for day-to-day forecasting of winter minimum temperature of Kolkata (Alipore) 12 hrs.
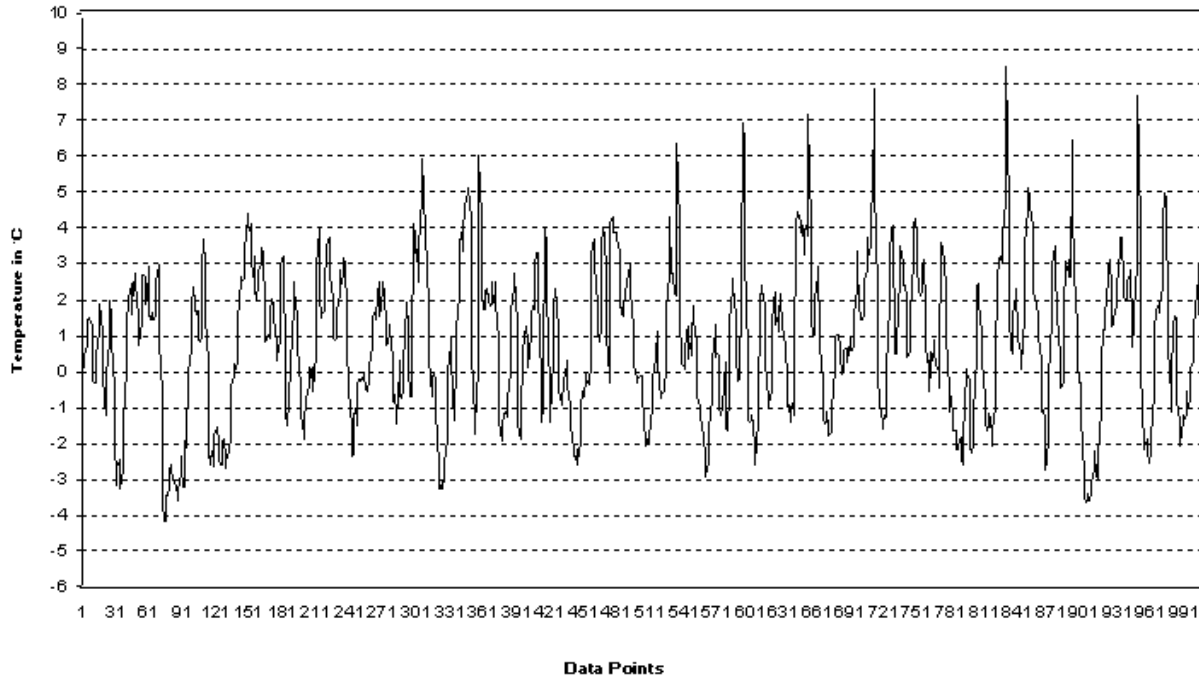
Fig. 1. Variation of daily minimum temperature from normal pentad (JF 1988-2004)

in advance based on data available up to 1730 hr (IST) using multiple linear regression technique.

2.   Data used and processing

Seventeen-year surface data (1988 – 2004) of Alipore Observatory for the months of January and February (JF) has been utilized. Data for the period from 1988 to 2001 has been utilized for developing the forecast model for predicting minimum temperature and the performance of the model has been tested with the independent data set of the remaining period from 2002-04. Thus development period (DP) and the testing period (TP) consist of 830 days and 178 days respectively.

2.1. *Some properties of the predictand*

It is seen from Table 1, minimum temperature of Kolkata of JF during the development period had a mean value of 15.7° C and SD of 3.3° C. The lowest value of minimum temperature of 8.4° C was registered on 14 Jan. 1989 whereas the highest value of 25.1° C was observed on 26 Feb. 2001. Thus range of variation of minimum temperature was found to be of the order of about 17° C. Fig. 1 gives the daily variation of normal minimum temperature based on existing Pentad normal (1931-60) of Kolkata (Alipore). From the graph it is seen that

TABLE 1

Statistics of winter minimum temperature of Kolkata based on data of JF, 1988-2001 and other information on thirty year's normal (1951-1980)

Statistics	Temperature (°C)	Remarks
Mean	15.7	
Median	15.7	
Highest	25.1	26 February 2001
Lowest	8.4	14 January 1989
Range	16.7	
SD	3.3	

variation of temperature from 1988 to 1996 is more or less uniform but it shifts slightly to the higher side from 1997 to 2004. However variation during 1997 to 2004 (from data points 535 onwards) shows some uniformity/stability within it. The effective length of the DP series was reduced to 237 (1997-2000) instead of 830 (1988-2001) on the basis of its stationarity, randomness and persistence.

TABLE 2

Minimum temperature forecast model : Equation, predictor and reduction of variance

S. No.	Predictor	Legend	Unit	Time hrs. (IST)	Mean	SD	Correlation	VE (%)	MCC
1	T_{min}	Previous day minimum temperature	°C	0830	16.0	3.1	+0.83	69.8	0.94
2	T_{max}	Maximum day temperature on previous day	°C	1730	26.9	3.1	+0.75	10.0	
3	Meridional component	Upper wind at 0.9 km on previous day	kmph	1730	-8.3	12.8	+0.49	4.7	
4	RH $_{min}$	Minimum day humidity on previous day	%	1730	45.1	11.5	+0.11	2.6	
5	N	Total cloud cover on previous day	okta	1730	2.4	2.6	+0.23	1.1	

VE (Variance Explained), MCC (Multiple Correlation Coefficient)

2.2. *Introduction of predictors*

The independent meteorological parameters (predictors) were identified in such a way that they are correlated and can explain most of the variance of minimum temperature (predictand). In other words all the chosen parameters that were found to be very much significant in qualifying Fisher's *F* - test at 95% confidence level were considered for the final multiple regression equation for predicting minimum temperature. They are

(*i*) X_1 Minimum temperature record at 0830 hrs. (IST) of previous day

(*ii*) X_2 Maximum temperature record at 1730 hrs. (IST) of previous day

(*iii*) X_3 Meridional component of wind at 0.9 km at 1730 hrs. (IST) of previous day

(*iv*) X_4 Minimum relative humidity record (generally around 1430 hrs) of previous day

(*v*) X_5 Total cloud cover record at 1730 hrs. (IST) of previous day

3. Methodology

3.1. *Multiple linear regression equation*

The dependent predictand y_i is fitted by a multiple linear regression model with p nos. of independent predictors is given by

$$y_i = \beta_0 + \sum_{j=1}^{p} \beta_j x_{ij} + e_i$$

where x_{ij}'s are independent predictor variables with $j = 1,2,3,.........p$ and $\beta_0, \beta_1,........\beta_p$ are the unknown coefficients of the independent variables and e_i is an error term associated with ith of n events. The prediction equation is given by

$$\tilde{y}_i = \tilde{\beta}_0 + \sum_{j=1}^{p} \tilde{\beta}_j x_{ij}$$

where $\tilde{y}_i$ is the predicted value (estimated value) of y_i, such that $\beta_0, \beta_1,........\beta_p$ minimizes the sum of squares of net errors, or in other words

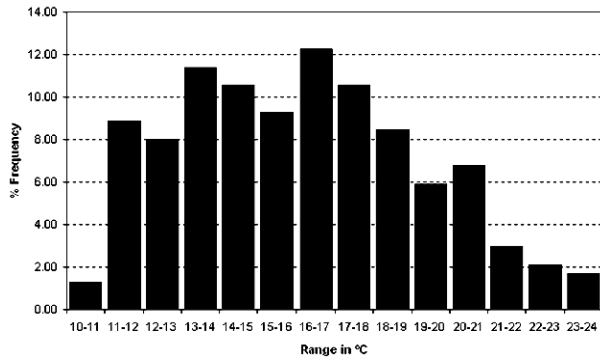$$\sum_{i=1}^{n} (y_i - \tilde{y}_i)^2 = \sum_{i=1}^{n} e_i^2 \text{ is minimum.}$$

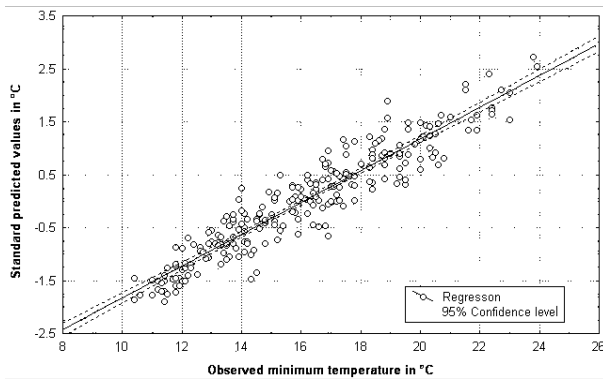**Fig. 2.** Frequency distribution of minimum temperature (JF 1997-2000)

Fig. 3. Standard predicted *vs.* observed minimum temperature (JF 1997-2000)

TABLE 3

Statistics of minimum temperatures of Kolkata based on data of development period (1997 to 2000) and test period (2002 to 2004)

Legend	Development period (1997 to 2000)	Test period (2002 to 2004)
	Observed	Observed
Number of observations	237	178
Mean	16.1	16.1
Median	16.0	16.1
Standard Deviation	3.1	3.4
Highest	23.9	24.1
Lowest	10.4	10.5
Range	13.5	13.6

A positive value of skill score indicates a better performance of the model over the persistence, whereas negative value of the skill score is clear indications that the model does not have any skill to even match the persistence.

4. Discussion and result

4.1. *Nature of distribution*

Nature of distribution of minimum temperature and its relationship with the potential predictors was carried out using developmental sample. Frequency distribution of minimum temperature is given in Fig. 2, which indicates that it is positively skewed in nature. Further studies showed that minimum temperature at Kolkata (Alipore) generally occurs between 0600 and 0700 hrs (IST) in the morning.

4.2. *Relation with potential predictors*

Correlation of minimum temperature with potential predictors, *viz.*, minimum temperature one day before, maximum temperature, meridional wind component at 1730 hrs. of previous day, minimum day humidity on previous day and total cloud cover have been shown in Table 2. It is seen from the table that minimum temperature is highly correlated with maximum and minimum temperature of previous day. It showed very

3.2. *Skill score*

Skill score is a statistical tool to determine whether predicted values of dependent variable based on statistical method is superior and is defined as

$$\left[1 - \left(\frac{\text{RMSE}_m}{\text{RMSE}_p}\right)^2\right] \times 100\%$$

where RMSE_m and RMSE_p are respectively the root mean square error of the model prediction and persistence: 24 hr. change on observed values of dependent variable.

TABLE 4

Error analysis and skill score for minimum forecast temperature (12 hr forecast)

Error range (°C)	Development data (JF 1997-2000)		Test data (JF 2002 to 2004)	
	No. of cases	% Frequency	No. of cases	% Frequency
0.0 – 1.0	160	67.51	129	72.47
1.1 – 2.0	61	25.74	33	18.54
2.1 – 3.0	16	6.75	12	6.74
3.1 – 4.0	0	0.00	4	2.25
Total	237	100	178	100
ABS	0.86		0.88	
RMS	1.08		1.14	
P-ABS	1.39		1.28	
P-RMS	1.78		1.70	
SKILL (%)	63		55	
CC	0.94		0.93	

ABS, absolute error, RMS, Root mean square error, P-ABS, Absolute error assuming persistence, P-RMS, RMS error assuming persistence; SKILL, Skill score; CC, Correlation coefficient.

TABLE 5

Error analysis in the case of persistent change in observed minimum temperature in 24 hrs

Error Range	Development data (JF 1997-2000)		Test data (JF 2002-2004)	
	No. of cases	% Frequency	No. of cases	% Frequency
0.0 – 1.0	117	49.37	96	53.93
1.1 – 2.0	57	24.05	44	24.72
2.1 – 3.0	46	19.41	24	13.48
3.1 – 4.0	11	4.64	9	5.06
> 4.0	6	2.53	5	2.81
Total	237	100	178	100

good correlation only with the meridional component of low-level wind at 0.9 km but not with the zonal component.

4.3. *Regression model*

Analysis of data following stepwise screening procedure led to the minimum temperature prediction model as given below :

$$Y = (0.3130)*X_1 + (0.6467)*X_2 + (0.0394)*X_3 + (0.0631)*X_4 + (0.1446)*X_5 - 9.1806$$

4.4. *Comparison of standard predicted values with observed minimum temperature of development period and testing of model with test period data*

Fig. 3 presents the scatter diagram between standard predicted values of minimum temperature with the observed minimum temperature during the development period (JF 1997-2000). From the figure it is seen that most of the data points are very close to the regression line within about ± 1° C at 95% confidence level. Model developed for minimum temperature was tested with the test data (JF 2002-04). Basic statistics of observed and predicted minimum temperature of the model and the test period data set are given in Table 3 for comparison of mean, variance, highest, lowest value and range of minimum temperature. Range of observed minimum temperature during the test period is slightly more than that of the development period.

4.5. *Error analysis*

Error analysis of forecast values of minimum temperature for both developmental sample and the independent data set is given in Table 4. From the table it is seen that ± 72%, ± 91% and ± 98% of 12 hour forecast minimum temperature values are correct within ± 1° C, ± 2° C and ± 3° C respectively in the test sample.

4.6. *Persistence of minimum temperature*

Persistence of minimum temperature is also studied to see its effect on the model predictions and results are presented in Table 5. It shows change in observed minimum temperature in next 24 hour for both developmental sample and test data in different temperature range. It is seen that the model could inherit a variable nature even with a variation of ± 4° C or more in persistence of minimum temperature.
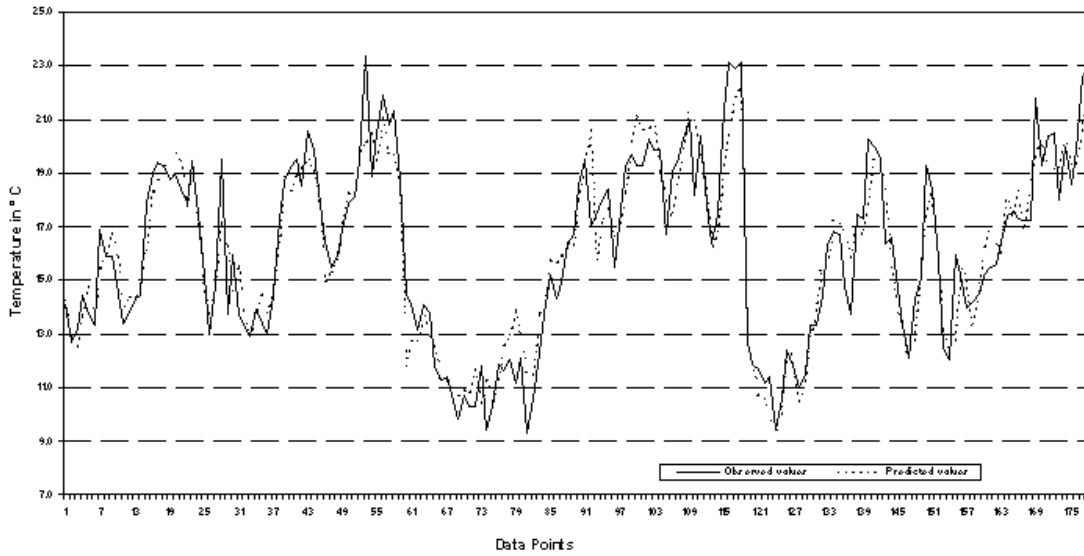
Fig. 4. Comparison of observed and predicted minimum temperature during JF (2002-04)

TABLE 6

Cases of predicted and observed minimum temperatures for error greater than 2º C during test period JF 2002-04

S. No.	Date	Observed values (ºC)	Predicted values (ºC)	Error (Obs.- Pred.) (ºC)	Cloud cover (N)		Wind at 0.9 km.				Weather
					Preday 1730 hrs (IST)	Day 0830 hrs (IST)	Preday 1730 hrs (IST)		Day 0530 hrs (IST)		Fog or mist at 0830 hrs (IST)
							ddd (Deg)	ff (kmph)	ddd (Deg)	ff (kmph)	
1.	28 Jan 2002	19.5	17.1	2.4	2	8	175	19	200	15	Haze
2.	29 Jan 2002	13.7	16.2	-2.5	3	3	270	29	306	08	Mist / Fog
3.	22 Feb 2002	23.4	20.2	3.2	1	6	270	12	338	15	Haze
4.	01 Jan 2003	14.4	11.8	2.6	0	8	345	10	272	23	Mist
5.	20 Jan 2003	11.2	13.8	-2.6	0	0	072	05	309	21	Mist
6.	22 Jan 2003	9.3	11.6	-2.3	0	0	312	27	311	36	Haze
7.	02 Feb 2003	17.0	20.6	-3.6	7	2	101	07	017	19	Haze
8.	20 Feb 2003	18.2	21.0	-2.8	5	0	192	13	157	17	Haze
9.	25 Feb 2003	21.4	19.0	2.4	0	6	313	11	313	21	Mist
10.	26 Feb 2003	23.1	20.7	2.4	1	2	294	14	252	23	Nil
11.	19 Jan 2004	19.7	15.8	3.9	0	4	301	14	268	41	Mist
12.	22 Jan 2004	20.3	17.8	2.5	5	4	325	06	225	11	Mist
13.	01 Feb 2004	19.3	17.1	2.2	7	1	225	08	224	20	Haze
14.	06 Feb 2004	16.0	12.5	3.5	0	7	310	14	272	12	Haze / rain
15.	20 Feb 2004	21.8	19.7	2.1	0	0	230	18	250	11	Nil
16.	28 Feb 2004	22.6	20.3	2.3	0	3	260	11	290	20	Nil

4.7. *Comparison of observed and predicted minimum temperature during test period (JF 2002 – 04)*

Minimum temperature forecast based on regression model along with its observed values for JF 2002-04 are shown in Fig. 4. The model responds well to the variation of minimum temperature, especially when sharp fall is expected.

The predicted minimum temperature carried absolute error of more than 2° C on 16 occasions, details are provided in Table 6. The comparison of data reveals that :

(*i*)   Observed minimum temperatures were higher than the predicted values due to covering of the sky by clouds after 1730 hrs. (IST) on previous night. This was also due to sudden change of wind from North West to SE bringing more moisture for development of clouds, fog and mist in the morning.

(*ii*)   Observed minimum temperatures were less than the predicted value when overcast sky condition became clear and wind direction also changed to NW from SE.

(*iii*) Two potential predictors – total cloud cover and meridional component of low level wind at 0.9 km can modify the minimum temperature values due to their marked variations after 1730 hrs and during the night. But the forecaster on the basis of analysis of 1200 UTC surface and upper air charts can modify values of these two predictors before making the final prediction at 2100 hrs (IST) to minimize the error.

5.   Conclusion

The following salient features have been revealed by the multiple linear regression models for predicting winter minimum temperature of Kolkata (Alipore) during January and February.

(*i*)   Multiple linear regression models can be used for day-to-day forecast of minimum temperature with 91% accuracy just after obtaining 1730 hrs (IST) observation. The values of predicted and observed minimum temperature during the test period are in good agreement to each other.

(*ii*) It can be easily handled by the forecaster as the model is PC compatible

(*iii*) Sharp fall of minimum temperature can be well predicted using potential predictor values of Kolkata. These predictors are pre-determined from this regression model using computer simulation.

(*iv*) Modification if any needed may be considered when cloud amount, low level wind flow is expected to change significantly during previous night based on satellite imagery and other synoptic situations before local forecast is issued at 2100 hrs (IST). This will minimize error in the prediction of minimum temperature and reduce subjectivity also.

References

Attri, S. D., Pandey, A. B. and Dubey, D. P., 1995, "Forecasting of minimum temperature over Gangtok", *Mausam*, **46**, 1, 63-68.

Charantoris, T. and Liakatas, A., 1990, "Study of minimum temperatures employing Markov chains", *Mausam*, **41**, 1, 69-74.

Dimri, A. P., Mohanty, U. C., Madan, O. P. and Ravi, N., 2002, "Statistical model-based forecast of minimum and maximum temperatures at Manali", *Current Science*, **82**, 8, 997-1003.

Kendall, M. G. and Stuart, A., 1968, "The Advanced Theory of Statistics", Griffin, London, (Vol.II), 278-374.

Klein, W. H. and Hammons, G. A., 1975, "Maximum-minimum temperature forecast based on model output statistics", *Mon. Wea. Rev*., **103**, 796-806.

Panofsky, H. A. and Brier, G. W., 1968, "Some applications of Statistics to meteorology", University Park, Pennsylvania, 117-118 and 191-208.

Raj, Y. E. A., 1989, "Prediction of winter minimum temperature of Pune by analogue and regression method", *Mausam*, **40**, 2, 175-180.

Singh, D. and Jaipal, 1983, "On forecasting night minimum temperature over New Delhi", *Mausam*, **34**, 2, 185-188.

Vashisth, R. C. and Pareek, R. S., 1991, "Forecasting of minimum temperature over New Delhi", *Mausam*, **42**, 1, 113-116.

WMO, 1966, "Statistical analysis and prognosis in Meteorology", Technical Note No. 71.