Meteorol. Appl. 6, 29–38 (1999)

Forecasting of thunderstorms in the pre-monsoon season at Delhi

N Ravi[1], U C Mohanty[1], O P Madan[1] and R K Paliwal[2]

[1]*Centre for Atmospheric Sciences, Indian Institute of Technology, New Delhi 110 016, India*
[2]*National Centre for Medium Range Weather Forecasting, Mausam Bhavan Complex, Lodi Road, New Delhi 110 003, India*

Accurate prediction of thunderstorms during the pre-monsoon season (April–June) in India is essential for human activities such as construction, aviation and agriculture. Two objective forecasting methods are developed using data from May and June for 1985–89. The developed methods are tested with independent data sets of the recent years, namely May and June for the years 1994 and 1995. The first method is based on a graphical technique. Fifteen different types of stability index are used in combinations of different pairs. It is found that Showalter index versus Totals total index and Jefferson's modified index versus George index can cluster cases of occurrence of thunderstorms mixed with a few cases of non-occurrence along a zone. The zones are demarcated and further sub-zones are created for clarity. The probability of occurrence/non-occurrence of thunderstorms in each sub-zone is then calculated. The second approach uses a multiple regression method to predict the occurrence/non-occurrence of thunderstorms. A total of 274 potential predictors are subjected to stepwise screening and nine significant predictors are selected to formulate a multiple regression equation that gives the forecast in probabilistic terms. Out of the two methods tested, it is found that the multiple regression method gives consistently better results with developmental as well as independent data sets; it is a potential method for operational use.

1. Introduction

Of all the phenomena in the atmosphere, there is none that exceeds the thunderstorm in beauty as it builds up from a small detached cumulus into a towering and awe-inspiring cumulonimbus. However, aviators treat thunderstorms with reverence for they are associated with hail, wind shear and turbulence, squalls and at times heavy precipitation. Thunderstorms during the hot and dry pre-monsoon season over north-west India can lead to duststorms, which reduce the visibility to such low values (almost nil) that it is dangerous for any outdoor human activity to take place. Squalls and hail can damage the crops, high-rise structures and uproot trees etc., and the associated turbulence can produce a severe gust load on aircraft which can seriously jeopardise flight safety. The duration of a storm is usually between half an hour to one hour, though on some occasions it may continue for several hours. Thus, prior knowledge of the occurrence of thunderstorms on a time scale varying between an hour to a few hours before the event is very useful for aviation as well as other activities such as transportation, agriculture and construction.

A number of studies using synoptic, synoptic-objective or purely objective techniques for forecasting thunderstorms over an area or over a specific location have been developed. Some such studies are listed below.

- Surendra Kumar (1972) evolved an objective method for forecasting thunderstorms over Delhi by using the convective condensation level, Showalter index, mean mixing ratio at 850, 800, and 700 hPa levels and wind direction as parameters. The technique is graphical in nature and uses parameters based on the experience of the forecaster; hence the method is semi-objective.
- Asoilal (1989) indicated that higher values of the totals total index are associated with more favourable conditions for thunderstorm activity around Jodhpur.
- Lal (1990) in his study on thunderstorms at Lucknow has revealed that a Showalter index of –4 or less, mean relative humidity below the level of 850 hPa of 45% or more and dew point above normal are favourable conditions for the occurrence of thunderstorms. The presence of large-amplitude troughs in mid- and upper-tropospheric westerlies, when they are located around 70° E, are associated with upper-level divergence and hence support the development of thunderstorms.

- Lee & Passner (1993) developed a rule-based expert system to assist military weather forecasters in the prediction of thunderstorms. The knowledge base used by Thunderstorm Intelligence Prediction System (TIPS) was built by using meteorological principles associated with convective activity. TIPS is a single station analysis tool that processes a 1200 UTC sounding to make a 12-hour forecast.
- Collier & Lilley (1994) used instability indices in combination with synoptic, meteorological radar data and satellite-derived information to forecast the initiation of thunderstorms over north-west Europe.
- Reap (1986, 1990) applied screening regression techniques to relate lightning data to large-scale meteorological predictors obtained from numerical forecast models, in order to derive equations for forecasting thunderstorm for different parts of the United States.

The purpose of the present study is to develop a suitable objective technique to forecast the occurrence/non-occurrence of thunderstorms at Delhi during the pre-monsoon season (May and June). The methodology currently in vogue for forecasting pre-monsoon thunderstorms is generally synoptic and hence subjective or semi-objective based on the experience of the forecaster. The synoptic systems in the pre-monsoon season are at times very diffuse and difficult to detect on synoptic charts and hence the forecasting of thunderstorms becomes very difficult. Though satellite pictures/radar reports are very valuable tools, their usefulness depends upon their availability. Furthermore, they often do not give a sufficient lead time for forecasting the thunderstorm in this season. Adequate objective prediction of thunderstorms would lead to better operational planning and precautionary measures by the user agencies. Since the upper-air data of Delhi are usually available 2–3 hours after the time of observation, *viz*. 0000 and 1200 UTC, it was decided to develop forecast techniques that would give the probability of occurrence/non-occurrence of a thunderstorm during the next 12 hours commencing at 0300 UTC and 1500 UTC. Since the frequency of occurrence of thunderstorms is highest towards late afternoon to early night, the initial attempt has been made to develop the 12-hour forecast commencing at 0300 UTC for the pre-monsoon months of May and June.

2. Data and analysis procedure

In India, a thunderstorm is considered to have occurred at an observation station when thunder is heard by the observer; this conforms to international agreement in defining a thunderstorm day. In general, the thunder can be heard over a distance of 30–40 km from the source of the lightning. In the USA and many European countries lightning detectors have been installed and any lightning discharge is counted as an occurrence of a thunderstorm. Reap (1986) defined the report of two or more cloud-to-ground lighting strikes occurring during a 6-hour period within grid blocks that are approximately 47 km on one side as an occurrence. Owing to the absence of these facilities as well as non-availability of weather radar observations, in this study a cumulonimbus cloud, with or without lightning in the vicinity of the observatory, has been considered as an occurrence of a thunderstorm.

During the pre-monsoon season, the occurrence of thunderstorms over north-west (NW) India is mainly synoptic-system based and also depends on a number of dynamical and thermodynamical variables. Therefore, it is felt that a statistical-dynamical technique would meet the location and time specific requirements for forecasts for a metropolitan city like Delhi. In this study, extensive meteorological observations of a location 15 km to the north-east of Delhi (henceforth referred as Delhi) are utilised. The developmental data sets consist of surface as well as upper-air observations of the stations over NW India (such as Delhi, Ambala, Halwara, Patiala and Suratgarh, as shown in Figure 1) for May and June, for the five-year period 1985–89. The stations around Delhi have been selected to cover the movement of synoptic disturbances, which in this season mainly approach from the west. The objective techniques developed for forecasting thunderstorms are tested with independent data sets of May and June for 1994 and 1995.

In order to study the climatology of thunderstorms at the selected place of study, current weather observations taken at hourly intervals at the site have been used. The frequency distribution of thunderstorms at each hour in May and June for Delhi is given in Figure 2. The hourly distribution of thunderstorms indicates that the maximum frequency of occurrence is found during the period 0100–0600 and 1600–2300 local time.

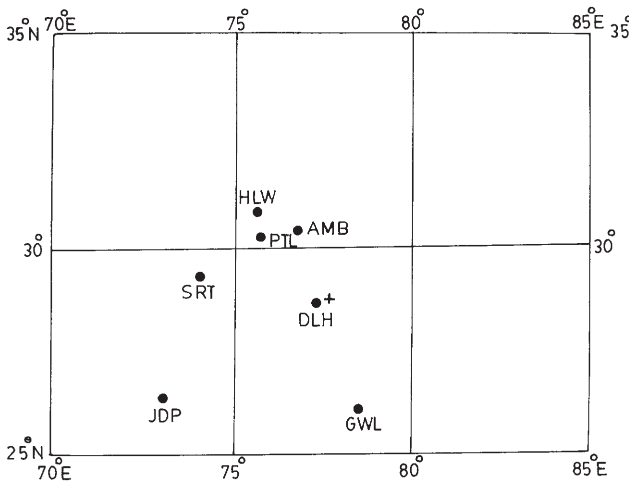

Figure 1. *Locations of meteorological stations from where data has been used in this study: DLH – Delhi, AMB – Ambala, SRT – Suratgarh, HLW – Halwara, GWL – Gwalior, JDP – Jodhpur, PTL – Patiala. Selected place of study is denoted by +.*

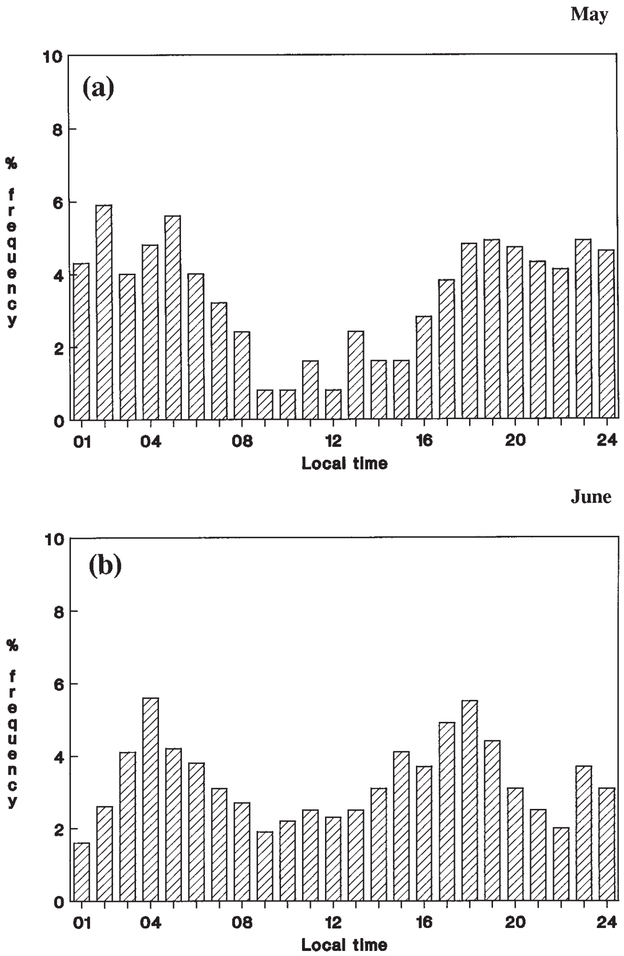

Figure 2. *Hourly distribution of thunderstorms at Delhi for (a) May and (b) June.*

Figure 2(*b*) shows that in June, two maxima occur around 0400 and 1800 local time. It is observed that the frequency of reports of thunderstorms based on thunder heard is half to two-thirds that of the occurrence of thunderclouds. The average number of days of thunderclouds/thunderstorms in May and June is 12.8/6.2 and 15.2/9.2 days respectively. The large increase of thunderstorms in June is due to the arrival of a southwest monsoon current over Delhi by the third week of June.

3. Methodology

During the pre-monsoon season in the north-western parts of India, the thunderstorms are found to be severe and mainly occur in association with synoptic systems. However, in this season the synoptic systems are usually diffuse and hence it becomes difficult to locate them on synoptic charts. Hence forecasting the occurrence of thunderstorms by the synoptic method becomes difficult and unreliable, and one is forced to look at some potential predictors that may help in objectively forecasting thunderstorms. Therefore, a large number of observed meteorological parameters such as temperature, wind, moisture and a number of derived parameters, as given in Table 1, are considered potential predictors for developing methods for forecasting thunderstorms. In this study, two methods, namely (a) graphical and (b) multiple regression, are used for the prediction of pre-monsoon thunderstorms at Delhi.

3.1. Graphical method

For objective weather forecasting techniques for a specific location, stability indices have an important place as an aid to forecasters. A stability index is a single numerical value derived from the vertical distribution of temperature, moisture and wind obtained from upper-air soundings. Since stability and moisture are the two primary factors for the development of thunderstorms, stability indices are derived by considering both dynamic and thermodynamic principles. The objective of the graphical method (based on a scatter/prediction diagram) is to organise the data in such a manner that some kind of relationship can be established between past conditions and the future behaviour of some weather elements (Petterson, 1956). The advantage of this method is its simplicity of application and transparency.

In order to develop an objective method for forecasting thunderstorms, 15 types of stability index are calculated and examined for their ability to predict the occurrence of a thunderstorm at Delhi. The fifteen stability indices used in this study, along with their abbreviations, are given in Table 2. As the methodology for calculating these indices is widely known and found in the literature (e.g. Showalter, 1953; Galway, 1956; George, 1960; Rackliff, 1962; Miller, 1967), the details are not presented here.

It is found from the study of the indices that no single stability index can provide a distinct 'critical' value (the critical value of an index is the threshold value of the index used to decide if the thunderstorm will be realised or not on that particular day) to forecast the occurrence of a thunderstorm. In order to analyse their nature, the mean and standard deviation of the 15 stability indices for the developmental data are calculated and presented in Table 3 for days on which thunderstorms did or did not occur. Here a suitable index is considered to be one whose mean value is distinctly different from occurrence to non-occurrence cases of thunderstorms by at least ±1 standard deviation.

From Table 3, it is noticed that not all the 15 indices are suitable for use alone as a potential index for prediction of occurrence/non-occurrence of thunderstorms. Therefore, using the developmental data sample for May and June together, pairs of indices, one versus the other, are tried in the form of a scatter diagram so that distinct groupings of occurrence and non-occurrence of thunderstorms can be achieved. Thus out of 105 possible pairs, only two pairs, SI versus TTI and TMJ

Table 1. *List of potential predictors used in the present study (the meanings of the abbreviations for the stability indices are given in Table 2). The total number of predictors is 274*

Predictors	Level (hPa)	Time (UTC)	Stations	Number of predictors
Dry bulb and dew point temperatures, maximum and minimum temperatures, dew point depression, total cloud amount, relative humidity, zonal and meridional components of wind	Surface	0000, 1200	DLH, AMB, HLW, JDP, PTL, GWL, SRT	126
24-hour changes in dry bulb, dew point, maximum and minimum temperatures	Surface	0000, 1200	DLH, AMB	16
Difference between maximum and minimum temperatures	Surface		DLH, AMB	2
Dry bulb temperature, saturation mixing ratio, potential, temperature,wind direction, wind speed, zonal and meridional components of wind	1000, 850, 700, 500 300	0000	DLH, JDP	70
Dew point temperature, mixing ratio, wet bulb potential temperature, equivalent potential temperature	1000, 850	0000	DIH, JDP	16
Wind shear	850–700, 850–500, 850–300	0000	DLH, JDP	6
Lapse rate	1000–850, 1000–700, 1000–500, 1000–300, 850–700, 850–500, 850–300	0000	DLH, JDP	14
Advection	1000–850, 850–700, 850–500, 850–300	0000	DLH, JDP	8
Occurrence of thunderstorms on previous day			DLH	1
Stability indices: SI, RI, TMJ, CIIR, K, VTI, CTI, TTI, KMOD, VTIM, CTIM, TTIM, LI, PII, SWEAT		0000	DLH	15

Table 2. *Stability Indices used in the study*

Abbreviation	Index
SI	Showalter index
RI	Rackliff's index
TMJ	Jefferson's modified index
CIIR	Convective index of Reap
K	George index
VTI	Vertical total index
CTI	Cross total index
TTI	Totals total index
KMOD	Modified George index
VTIM	Modified vertical total index
CTIM	Modified cross total index
TTIM	Modified totals total index
LI	Lifted index
PII	Potential instability index
SWEAT	Severe weather threat index

versus K, are able to cluster cases of occurrence, with few cases of non-occurrence along a zone (the region where the majority of the cases of occurrence of thunderstorms are included between the two straight lines and separated from the cases of non-occurrence). For the rest of the 103 pairs, a great deal of overlapping is found, and separation between occurrence/non-occurrence cases do not even satisfy 50% of the cases. On further analysis, it is found that the TMJ versus K pair (Figure 3) can separate the cases of occurrence and non-occurrence better than the SI versus TTI pair (Figure 4). Hence, further verification and discussions are confined to TMJ versus K pair (see Appendix A for the procedure for calculating the TMJ and K indices).

In Figure 3, the zones containing cases of non-occurrence are separated from cases of occurrence by drawing two straight lines, and the equations of these lines are determined. The zone in between these two lines contains a large number of cases of occurrences mixed with a few non-occurrences of thunderstorms. This

Table 3. *Stability Indices - mean and standard deviation (SD)*

Stability index	Occurrence				Non-occurrence			
	May		June		May		June	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
SI	–0.6	3.6	–1.7	4.5	1.9	3.5	0.5	3.9
RACK	29.5	5.6	29.0	3.3	28.0	2.9	27.8	3.3
TMJ	27.4	4.6	30.5	5.8	23.7	4.1	25.8	6.0
CIIR	–5.4	6.7	–8.9	10.4	–3.5	6.3	–8.6	9.0
K	26.9	9.5	35.5	8.9	18.5	10.7	27.7	9.6
VTI	33.0	3.0	29.0	4.7	33.0	2.8	30.9	2.9
CTI	17.1	4.6	19.2	5.2	13.0	4.9	15.5	5.7
TTI	50.1	5.5	48.2	7.5	45.9	6.1	46.4	6.3
KMOD	32.5	8.9	40.7	8.3	23.2	12.1	33.0	8.8
VTIM	34.9	2.9	31.5	4.5	34.8	2.8	32.9	3.0
CTIM	21.0	3.7	22.0	4.6	16.8	4.7	18.8	5.2
TTIM	55.9	4.6	53.5	6.8	51.6	5.9	51.6	6.5
LI	–0.8	3.2	–2.1	4.1	01.9	3.6	0.4	4.2
PII	–0.9	2.6	–0.9	3.1	00.1	2.4	–0.4	3.0
SWEAT	249.3	154.7	253.4	142.7	143.3	101.2	197.8	112.3

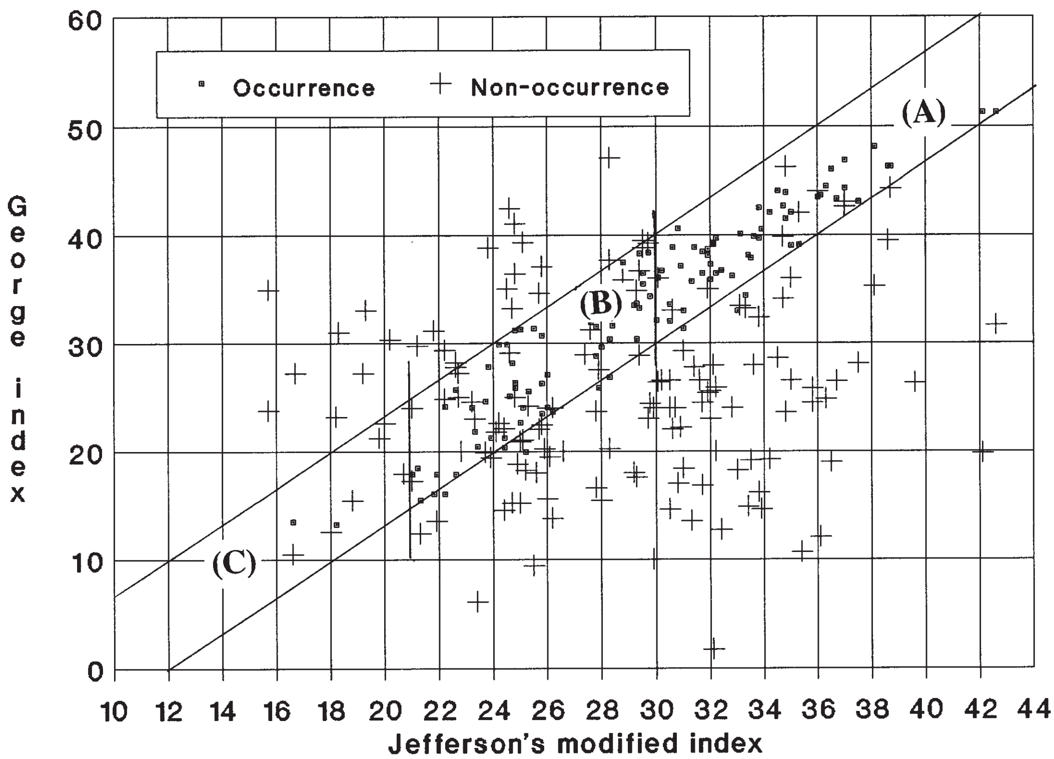

Figure 3. *Scatter diagram for forecasting thunderstorm for TMJ vs K (sub-division* A *has cases mainly of occurrence;* B *has cases of occurrence and non-occurrence mixed;* C *has cases mainly of non-occurrence).*

zone is further subdivided into three sub-divisions and labelled A, B and C, where A has cases mainly of occurrence, C has cases mainly of non-occurrence, and B has a mix of occurrence and non-occurrence cases. From the number of cases of occurrence and non-occurrence of thunderstorms in each sub-division, a probability value can be assigned to each sub-division. The steps to be followed for forecasting of occurrence/non-occurrence of thunderstorms on the basis of the TMJ and K stability indices, along with the probability in each of the three sub-divisions are given in Appendix B. The performance of this objective method with dependent and independent data sets is discussed in section 4.

3.2. Regression method

For the second method, a multiple regression technique is attempted. In the regression model, 9 significant predictors are selected out of the 274 potential predictors listed in Table 1, by use of a stepwise screening procedure (Draper & Smith, 1981). An equation of the type

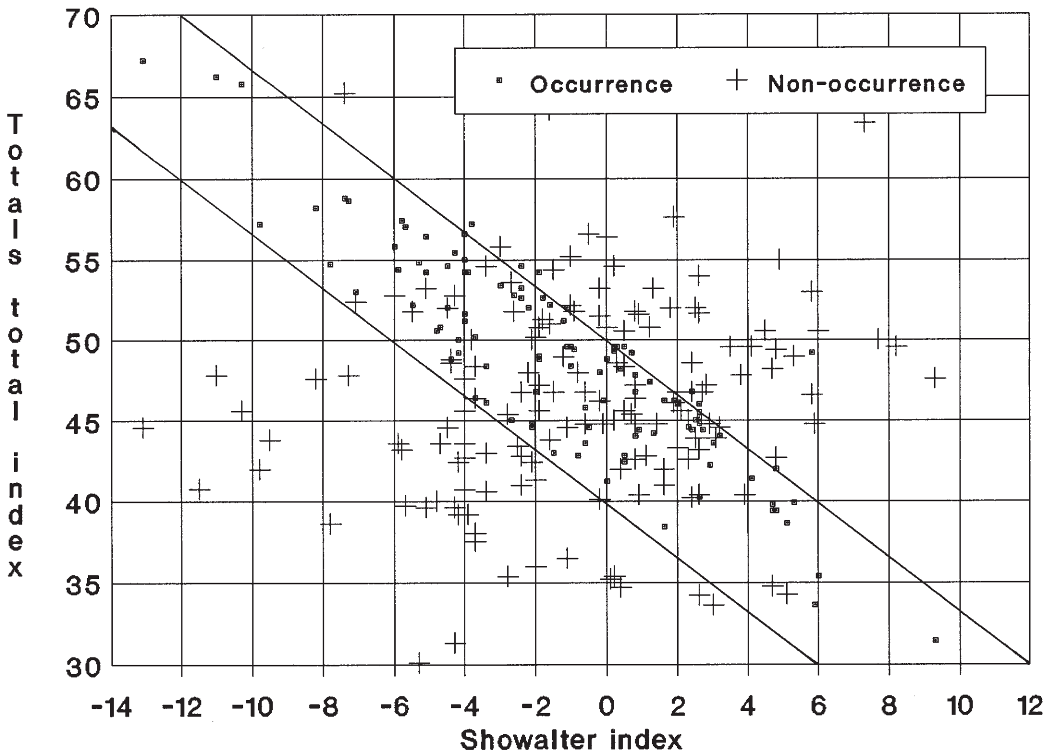

Figure 4. *Scatter diagram for forecasting thunderstorms for SI versus TTI.*

$$Y = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_n x_n$$

is assumed, where Y is the value of the predictand obtained by a linear combination of various predictors $x_1, x_2, \ldots, x_n$ and $b_1, b_2, \ldots, b_n$ are the regression coefficients, with b_0 the regression constant. In the development of the equation, the value of the predictand, Y, is taken as 1 in the event of occurrence of a thunderstorm and 0 if it does not occur.

The stepwise procedure requires a stopping rule. With such a rule, the selection process of the predictors would continue until all candidate predictor variables are included in the regression equation. In this study, selection of predictors is stopped when none of the remaining predictors would reduce the variance by 1% or more.

The selected predictors and the variance explained by each of them is given in Table 4. The selected predictors are mostly upper-air parameters such as dew point temperature, wind direction and speed at Delhi at 0000 UTC, and 24-hour change in dew point temperature at the surface at 1200 UTC. From Table 4 the following should be noted.

(a) The dew point depression at 900 hPa, which is selected as the first potential predictor, is negatively correlated with thunderstorm occurrence. This indicates that the lower values of this predictor (indicative of more moisture in the lower level) favour the occurrence of thunderstorms in the dry pre-monsoon season. It is an established fact that injection of moisture in the lower levels enhances static instability and is favourable for the development of thunderstorms.

(b) The wind direction at 850 hPa is negatively correlated with occurrence, i.e. the north-westerly winds reduce thunderstorm occurrence and south-easterly winds increase it. The south-east wind at 850 hPa over NW India indicates advection of warm and moist air during the passage of a synoptic system.

(c) The saturation mixing ratio at 500 hPa is negatively correlated with occurrence, which indicates that drier air at 500 hPa is favourable for occurrence of thunderstorms.

(d) The 24-hour change in the surface dew point temperature is positively correlated with thunderstorm occurrence. The positive value of the 24-hour change indicates an increase in dew point, which is indicative of an increase in moisture content in the lower troposphere. The buoyancy of the updraft is affected by low-level temperatures and dew point temperature (Colquhoun, 1987). An increase in the dew point of near-surface air increases the updraft buoyancy more than an equivalent increase in temperature.

All the predictors selected are significant and are indicative of occurrence or non-occurrence of thunderstorms.

Table 4. *Regression method for forecasting thunderstorms at Delhi (Multiple correlation coefficient = 0.74)*

Predictor	Level (hPa)	Coefficient	Variance explained (VE)	Cumulative VE	Correlation with predictand
Constant	–	1.4356			
Dew point depression	900	–0.016241	18.9	18.9	–0.435
Meridional wind	300	0.006008	11.0	28.9	0.377
Wind direction	850	–0.001513	8.3	37.2	–0.409
Wind speed	900	0.016035	7.0	44.2	0.284
Wind speed	500	0.007636	3.3	47.5	0.201
24-hour change in dew point	Surface	0.026508	2.8	50.3	0.166
Saturation mixing ratio	500	–0.109190	1.9	50.2	–0.245
Wind direction	700	0.011220	1.7	53.9	–0.273
Dew point temperature	700	0.012015	1.4	55.3	0.122

Using the developed equation, the values of Y are recalculated for all the observations of the development sample. If the recalculated value of Y is less than 0, it is made equal to 0, and if it is more than 1, it is truncated to 1. The values of Y are grouped with an interval of 0.1 and the corresponding number of cases of occurrence/non-occurrence of a thunderstorm and the probability values are calculated. The best-fitting curves to these results are presented in Figure 5. The critical value of 0.4 was used as a cut-off value for deciding occurrence because it was noticed that for Y equal to or more than 0.4, the probability of occurrence of a thunderstorm is 70% or more and for a value of Y less than 0.4 the probability of non-occurrence is higher than that of occurrence. It may be noted that the evaluated value of Y using the regression equation gives the probability of occurrence/non-occurrence of a thunderstorm, which varies between 0 and 1 (or 0 and 100%). However, for the purpose of forecast evaluation, the probability forecast is converted to a categorical forecast; that is, if the value of Y is equal to or more than 0.4, it is assigned the value 1 (forecast of occurrence) and if it is less than 0.4 it is assigned the value 0 (forecast of non-occurrence).

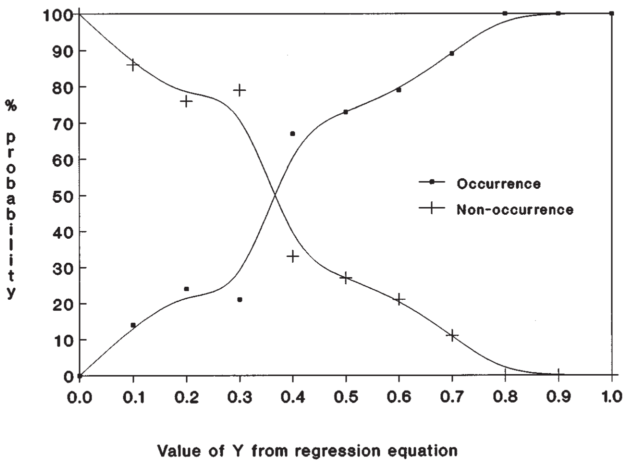

Figure 5. *Probability of occurrence/non-occurrence of thunderstorms by the regression method.*

4. Results and discussions

Using the developmental data sets for May and June for 1985–1989, the graphical and regression methods are applied to generate the prediction graph and coefficients of the equation appropriate to each method. Each method in turn is tested with two independent data sets from May and June for 1994 and 1995. For the purpose of forecast evaluation, a 2 × 2 contingency table is utilised and various verification parameters are calculated (the parameters are listed in Table 5 with the definitions in Appendix C). The performance of each method on developmental as well as independent data sets is discussed below.

4.1. Graphical method

The verification parameters for the developmental as well as the test data sets using the graphical method, TMJ versus K, are presented in Table 6. From this table, it is noticed that the verification of the graphical method with the developmental data set of May gave a *POD* of 0.68, *FAR* of 0.60 and *BIAS* of 1.65. In the case of the independent data sets of May 1994 and 1995, the *POD* varies from 0.75 to 1.00, *FAR* from 0.60 to 0.80 and *BIAS* from 1.78 to 4.70, which are higher than the developmental data set. The verification with developmental as well as independent data sets for June gave *POD* ranging from 0.85 to 1.00, *FAR* from 0.40 to 0.80

Table 5. *Verification parameters*

Abbreviation	Parameter
POD	Probability of detection
FAR	False alarm rate
MR	Miss rate
C-NON	Correct non-occurrence
CSI	Critical success index
TSS	True skill score
HSS	Heidke skill score
BIAS	Bias
PC	Percentage correct

and *BIAS* from 1.56 to 5.00. The skill scores, namely *TSS*, *CSI* and *HSS* in all these cases are found to be small for the developmental and independent data sets. The low values of the skill scores appear to be due to the over-forecasting of the events by this method, as indicated by the higher values of *FAR* and *BIAS.* Considering the *BIAS* alone, though a little overprediction may be preferable so that the occurrences are not missed, the graphical method tends to be oriented towards excessive overprediction.

4.2. Regression method

The multiple regression method, which takes into consideration a large number of potential predictors, has been assessed as an alternative to the graphical method. The skill scores of the regression method for May and June are given in Table 7. From this table, it is noticed that in the case of the developmental data for May, the method is able to predict 92% (*POD* = 0.92) of the total cases of occurrence of thunderstorms. On 38% of the occasions, thunderstorms did not occur though they are forecast (*FAR* = 0.38). The *CSI* is 0.59, the *TSS* is 0.57 and *HSS* is 0.58. The *BIAS* indicates that the method slightly overpredicted thunderstorm occurrence but is certainly less when compared to the graphical method. The testing of the model on independent data sets for May 1994 and 1995, as indicated in Table 7, shows that the model performs much better in 1994 than in 1995 with *POD* of 0.90 and 0.67 respectively. The *FAR* in 1994 is 0.10 whereas in 1995 it is 0.78. The *BIAS* is close to 1.0 in 1994 whereas in 1995 it is oriented slightly towards overprediction.

Furthermore, analysis of Table 7 reveals that in the developmental data for June the multiple regression method yielded a *POD* of 0.84, *FAR* of 0.25 and a near-perfect *BIAS* of 1.12, while in the test data of June 1994, the model slightly under-predicted the occurrence, with a *POD* of 0.67 which is less than that of developmental data set. In the test data of June 1995, though the model slightly overpredicted the occurrence, the overall performance is found to be better.

The *PC* are 79, 88 and 74% for May, and 80, 83 and 87% for June, for the development and the independent test data sets of 1994 and 1995 respectively. Physical scrutiny of the data of 1995 reveals that there were only three cases of occurrence of thunderstorm against an average of 13 days in May, and five cases of actual occurrence in June as compared to the average of 16 days.

A comparison of the verification parameters and skill scores of both methods indicates that the regression

Table 6. *Performance statistics of the graphical method (TMJ vs K) with development and independent data sets*

Parameter	Development data	Test data		Development data	Test data	
	May	May 1994	May 1995	June	June 1994	June 1995
POD	0.68	0.75	1.00	0.85	0.70	1.00
FAR	0.60	0.60	0.80	0.45	0.63	0.80
MR	0.32	0.25	0.00	0.15	0.30	0.00
C-NON	0.53	0.38	0.44	0.40	0.40	0.38
CSI	0.34	0.36	0.18	0.50	0.32	0.20
TSS	0.20	0.14	0.34	0.25	0.10	0.38
HSS	0.17	0.10	0.17	0.24	0.08	0.14
BIAS	1.65	1.78	4.70	1.56	1.90	5.00
PC	57	52	42	61	50	47

Table 7. *Performance statistics of the regression method with development and independent data sets*

arameter	Development data	Test data		Development data	Test data	
	May	May 1994	May 1995	June	June 1994	June 1995
POD	0.92	0.90	0.67	0.84	0.67	0.80
FAR	0.38	0.18	0.78	0.25	0.25	0.43
MR	0.08	0.10	0.33	0.16	0.33	0.20
C-NON	0.65	0.87	0.75	0.76	0.90	0.88
CSI	0.59	0.75	0.20	0.66	0.55	0.50
TSS	0.57	0.77	0.42	0.60	0.57	0.68
HSS	0.58	0.75	0.22	0.60	0.58	0.59
BIAS	1.40	1.10	3.00	1.12	0.89	1.40
PC	79	88	74	80	83	87

method performs better than the graphical method on developmental as well as independent data sets. The *CSI* varies from 0.18 to 0.50, the *TSS* from 0.10 to 0.38 and the *HSS* from 0.08 to 0.24 for the graphical method, whereas for the regression method these are comparatively higher as the *CSI* varies from 0.20 to 0.75, *TSS* from 0.42 to 0.77 and *HSS* from 0.22 to 0.75. The *FAR*, *POD* and *PC* of the regression method are comparatively better than the graphical method.

5. Conclusions

In this study, in order to forecast thunderstorms during the pre-monsoon season at Delhi, two objective methods are developed, and these are evaluated with dependent as well as independent data sets. From the results presented above, the following broad conclusions are drawn.

(a) The graphical method using only stability indices tends to overpredict the occurrence of thunderstorms.
(b) The multiple regression method gives consistently higher skill scores on both the dependent and independent data sets.
(c) The evaluation of both the objective methods indicates that the multiple regression method gives better results than the graphical method.

With the availability of adequate and reasonably accurate numerical model outputs over north west India, other objective methods such as the Perfect Prog Method (PPM) or Model Output Statistics (MOS) could be developed to further improve the prediction of thunderstorms at Delhi.

Acknowledgement

The authors sincerely thank the Directorate of Meteorology, Indian Air Force, and in particular, Air Vice Marshal K. C. Varma, Director, for providing the necessary data and helpful discussions. The authors acknowledge the India Meteorological Department for providing valuable data for this work. Special thanks are owed to the editor and anonymous referees for their valuable comments/suggestions, which significantly improved the manuscript. The work was sponsored by OP & AP of the Aeronautical Research and Development Board under a research grant at IIT, Delhi.

Appendix A. Stability indices used for forecasting thunderstorms at Delhi

(a) Jefferson's modified index (TMJ)

$$TMJ = 1.6\theta_{w850} - T_{500} - 0.5DPD_{700} - 8.0$$

(b) George index (K)

$$K = T_{850} - T_{500} + TD_{850} - DPD_{700}$$

where:

θ_{w850} = wet bulb potential temperature at 850 hPa
T_{500} = dry bulb temperature at 500 hPa
DPD_{700} = dew point depression at 700 hPa
T_{850} = dry bulb temperature at 850 hPa
TD_{850} = dew point temperature at 850 hPa

Appendix B. Procedure for using a graphical method (TMJ versus K) for forecasting thunderstorms

1. From Delhi radiosonde data of 0000 UTC calculate:
   (a) Jefferson's modified index, denoted by X_o
   (b) George's index denoted by Y_o
2. Use the following equations to find Y_U and Y_L:
   $Y_U = 1.7\ X_o - 10.4$
   $Y_L = 1.7\ X_o - 20.4$
3. Check whether Y_o falls within Y_U and Y_L (both values included)
4. If Y_o falls within the limit forecast *YES*, the occurrence of thunderstorms on that particular day between 0900 and 2200 IST. If Y_o does not fall in the limit, forecast *NO*, the non-occurrence of thunderstorms.
5. When Y_o falls within the limit, then:
   (a) If X_o is greater than or equal to 30.0, the probability of occurrence = 1.0 (100 %)
   (b) If X_o is greater than 21.0 and less than 30.0, the probability of occurrence = 0.6 (60%)
   (c) If X_o is less than or equal to 21.0, then probability of non-occurrence = 1.0 (100%)

Appendix C. Verification measures

Observed	Forecast	
	Yes	No
Yes	A	B
No	C	D

Probability of detection (POD)

$$POD = \frac{A}{A+B}$$

False alarm rate (FAR)

$$FAR = \frac{C}{C+A}$$

Miss rate (MR)

$$MR = \frac{B}{B + A}$$

Correct non-nccurrence (C-NON)

$$C\text{-}NON = \frac{D}{D + C}$$

Critical success index (CSI)

$$CSI = \frac{A}{A + B + C}$$

True skill score (TSI)

$$TSI = \frac{A}{A + B} + \frac{D}{D + C} - 1$$

Heidke skill score (HSS)

$$HSS = \frac{2(AD - BC)}{B^2 + C^2 + 2AD + (B + C)(A + D)}$$

Bias (BIAS)

$$BIAS = \frac{A + C}{A + B}$$

Percentage correct (PC)

$$PC = \left(\frac{A + D}{A + B + C + D}\right) \times 100\%$$

References

Asoi Lal. (1989). Forecasting of thunderstorm around Delhi and Jodhpur. *Mausam*, **40:** 267–8.

Collier, C. G. & Lilley, R. B. E. (1994). Forecasting of thunderstorm initiation in NW Europe using thermodynamic indices, satellite and radar data. *Meteorol. Appl.*, **1:** 75–84.

Colquhoun, J. R. (1987). A decision tree method of forecasting thunderstorms, severe thunderstorm and tornadoes. *Wea. Forecasting*, **2:** 337–45.

Draper, N. R. & Smith, H. (1981). Applied Regression Analysis, second edition. Wiley & Sons, New York, 481 pp.

Galway, J. G. (1956). The lifted index as predictor of latent instability. *Bull. Am. Meteorol. Soc.*, **37:** 528–9.

George, J. J. (1960). *Weather Forecasting for Aeronautics*. Academic Press, New York and London, 673 pp.

Lal, R. (1990). Forecasting of severe convective activity over Lucknow in premonsoon season. *Mausam*, **41:** 455–8.

Lee, R. & Passner, J. E. (1993). The development and verification of TIPS: An expert system to forecast thunderstorm occurrence. *Wea. Forecasting*, **8:** 271–80.

Miller, R. C. (1967). Notes on analysis and severe-storm forecasting procedures of the military weather warning center, Technical Report. No. 200, AWS(MAC), USAF, 139 pp.

Petterson, S. (1956). *Weather Analysis and Forecasting*, Mc Graw-Hill, Vol II, 225–37.

Rackliff, P. G. (1962). Application of an instability index to regional forecasting. *Meteorol. Mag.*, **91:** 113–20.

Reap, R. M. (1986). New 6-H thunderstorm probability forecasts for the West. *NOAA, National Weather Service Technical Procedures Bulletin*, No. 362, 6 pp.

Reap, R. M. (1990). Six-hour thunderstorm probability forecasts for Kansas-Oklahoma and northeast United States. *NOAA, National Weather Service Technical Procedures Bulletin*, No. 385, 6 pp.

Showalter, A. K. (1953). A stability index for thunderstorm forecasting. *Bull. Am. Meteorol. Soc.*, **34:** 250–62.

Surendra Kumar (1972). An objective method of forecasting premonsoon thunderstorm/duststorm over Delhi and neighborhood. *Indian J. Met. Geop.*, **23:** 45–50.

Wilks, D. S. (1995). *Statistical Methods in the Atmospheric Sciences An Introduction*. Academic press, New York. 544 pp.