Application of Pattern Recognition Techniques to Predict Severe Thunderstorms

Himadri Chakrabarty, C. A. Murthy, and Ashish Das Gupta

***Abstract*—Thunderstorm forecasting is a challenging job. Machine learning techniques are being applied nowadays in meteorological fields for prediction purpose. This study presents the application of different machine learning tools based on multiple correlation, Multi-layer Perceptron (MLP), K-nearest neighbor (K-nn) method, and modified K-nn method to predict seasonal severe thunderstorms associated with squall occurring in Kolkata, North-East India. The models are trained and tested with the radiosonde data recorded in the early morning at 00:00UTC. The predictors are moisture difference and dry adiabatic lapse rate at different geopotential heights of the atmosphere. Our aim in this paper is to find how much correctly one can nowcast 10 to 14 hours before the 'occurrence'/ 'no occurrence' of evening squall-storms by using a few upper air diagnostic predictors. Modified K-nn method is found to yield very promising prognostic information with high prediction accuracy. The results indicate that forecasting can be done correctly up to 82.02% both for 'squall-storm/no storm' events, and up to 91.11% for 'squall-storm' events using modified K-nn based approach. In this article, modified K-nn method is proved as the best method in comparison with the other methods for the squall-storm prediction.**

***Index Terms*—Back propagation, K-nearest neighbor, multi-layer perceptron, multiple correlation, squall-storm.**

I. Introduction

Severe thunderstorm is a mesoscale, convective and seasonal atmospheric event. It is associated with squall (very strong wind), thunder, lightning, smart shower, and sometimes with hail. Squall or strong wind is generated from the super cell cumulonimbus clouds or squall line, which is developed from the atmospheric instability condition especially in warm period weather, [1]. Though severe thunderstorms are generally very short-lived phenomena, their effect on human life and property may be devastating on many occasions. Accurate prediction of such severe weather event is necessary and it is a difficult task due to the dynamic nature of atmosphere, [2]. Severe thunderstorm prediction, in a conventional way, generally requires various surface as well as upper air weather data, observed from time-to-time throughout the whole day. Doppler radar and satellite imagery are also necessary for this purpose. Marinaki cited in 2006 that the estimation of atmospheric instability related to thunderstorm generally requires computation based on several thermodynamic parameters, such as, Showalter index [3]; Boyden index [4], which are obtained from the observational weather data.

Our main objective of the present work is to predict squall-thunderstorms. Here, statistical and machine learning (multi-layer perceptron, and K-nearest neighbor classifier) techniques have been applied to forecast the 'occurrence'/ 'no occurrence' of severe thunderstorms at Kolkata (22.3°N/88.3°E), situated in North-East India. The predictors are considered by quantifying the humidity and conditional instability of the atmosphere with the aid of radiosonde data recorded in the morning time at around 5:30 am (00:00UTC) during the period of 40 years from 1969 to 2008 for the months of March, April and May (MAM). These three months are known as the pre-monsoon season in North–East India, and most of the squall-thunderstorms generally occur in this season. The prediction of convective events is usually based on statistical relations between event occurrence (predictand) and various physical variables (predictors), [5]. Upper air vertical moisture difference profile at five different geo-potential heights is considered as five input variables (predictors), and dry adiabatic lapse rates at five different heights are taken as the other five input variables (predictors). The lead time for forecasting here is around 10 to 14 hours. This is a sufficient lead time to alert people from such devastating weather event. The accuracies of the outputs obtained from statistical and neural network models (MLP) and also from K-nn technique have been compared. A modified K-nearest neighbor (modified K-nn) rule has also been tried to be applied for prediction, and it is found to give best results among all these methods.

In the literature, as far as our knowledge goes, there are no papers predicting storms with more than 91% accuracy on the basis of only early morning upper air data, with 10 to 14 hours leading time. This is the main contribution of this paper. There are many research papers on severe thunderstorms, [6]-[8]. Neural network classifiers have been attractive alternatives to conventional classifiers by numerous researchers, [9], and it is studied in the fields of speech and image recognition. Neural Network is a generalization of traditional statistical methods for nonlinear regression and classification, [10]. Literature study shows that weather prediction was done by data mining (K-nn) using historical surface weather parameters such as, rain, wind speed, dew point, temperature, etc., [11]. K-nn techniques were applied by Li *et al.*, 2007 [12] for solar flare forecasting.

Manuscript received February 14, 2013; revised April 25, 2013. This work was supported by University Grants Commission of India.

H. Chakrabarty is with Surendranath College, and he is also with Institute of Radiophysics and Electronics, Calcutta University, Kolkata, India (e-mail: hima.c@rediffmail.com).

C. A. Murthy is with Machine Intelligence Unit, Indian Statistical Institute, Kolkata, India (e-mail: murthy@isical.ac.in).

A. DasGupta is with S. K. Mitra Center for Research in Space Environment, Calcutta University, Kolkata, India (e-mail: adgbkpr@gmail.com).

DOI: 10.7763/IJCTE.2013.V5.810

II. Data

A. Data Collection

All the upper air and surface data during the period of 40 years from 1969 to 2008 were collected from India Meteorological Department, Govt. of India. The data were recorded at 00:00 UTC by radiosonde in the pre-monsoon period for the months of March, April and May (MAM). The data considered for analysis here are both for the days when squall-storms occurred and for some of the days when squall-storms did not occur. The number of MAM squall-storm days in 40 years (1969-2008) is 180. Out of these, 'storm' data of 175 days are available for processing. The rest of the days of March-April-May within this 40 years period are considered as 'no storm' days. The 'no storm' data of 400 days are considered for processing. The whole data set is divided into two parts: training dataset and test dataset. For the training set, 85 days of squall 'storm' data and 84 days of 'no storm' data were considered. In the test set, we shall have 90 (175-85) squall 'storm' days and 316 (400-84) 'no storm' days. Usually, the parameters of the model are calculated by the training set, and accuracy of it is tested by the points in the test set.

B. Data Description

Vertical moisture difference profile and the dry adiabatic lapse rate of the atmosphere are considered as input variables (i.e., predictors), represented by x_i 's. The predictand is the squall-storm, y.

Moisture Difference: The moisture difference has been measured by the difference between dry bulb (T) and dew point temperature (Td) at (i) surface level (x_1), and then at different geo-potential heights of the upper air, such as at (ii) 1000 hpa measured at approximately 75 meters (x_2), (iii) 850 hpa measured at approximately 1500 meters (x_3), (iv) 700 hpa measured at approximately 3100 meters (x_4), and (v) 600 hpa measured at approximately 4500 meters (x_5). So, vertical moisture difference (MD=T-Td) profile indicates the measurement of humidity from the surface to the upper atmosphere of 4.5 kilometers height, from MSL, signifying the amount of saturation in the atmosphere in the morning on the 'squall-storm' days as well as on the 'no storm' days. This moisture forms the thundercloud, [13], if the other atmospheric conditions are suitable.

Adiabatic Lapse Rate: The conditional instability can be evaluated by the adiabatic lapse rate of the atmosphere, [14]. Dry adiabatic lapse rates at different geo-potential heights are determined by dry bulb temperature difference between consecutive two levels (d T / d Z), such as (i) surface and 850 hpa (approximately surface to 1500 meters), denoted by x_6 , (ii) 850 hpa and 700 hpa (approximately 1500 to 3100 meters), denoted by x_7 , (iii) 700 hpa and 600 hpa (approximately 3100 to 4500 meters), denoted by x_8, (iv) 600 hpa and 400 hpa (approximately 4500 to 7500 meters), denoted by x_9, and (v) 400 hpa and 300 hpa (approximately 7500 to 9600 meters), denoted by x_{10}. Lapse rate at these five different above mentioned heights (up to 9.6 kilometers) are considered as the five input parameters (predictors). The more the conditional instability remains in the atmosphere, more moisture would be carried out to the upper atmosphere from the surface level to form thunderclouds, [14].

III. Methodology

A. Statistical Methodology

The statistical methodology adopted here is multiple linear regression technique. Each pair of the total ten predictors (input variables) is correlated. Linear regression equation can be written as,

$$y = a + b_1 x_1 + b_2 x_2 + ... + b_{10} x_{10}.$$

'Occurrence' or 'no occurrence' of the squall-storm is considered as the output of the model or the dependent variable y . It is taken as '1' on those days when squall-storm occurred and '0' on those days when there is no squall-storm. The ten input parameters $x_1, x_2, ...x_{10}$ are taken as independent variables. The task is to estimate (or learn) the parameters $a, b_1, b_2, ..., b_{10}$ from training data on (y , $x_1, x_2, ...x_{10}$), and then to estimate the value of the dependent variable for validation and prediction for every point in test data set. The correlation coefficients between each pair of 11 input variables are used to get the values of b_i 's, [15].

B. Multi-Layer Perceptron

Multi-Layer Perceptron (MLP) network consists of a set of sensory units constituting the input layer, having 11 nodes where the first 10 nodes correspond to 10 predictor weather variables $x_1, x_2, ...x_{10}$, and the eleventh one corresponds to the 'bias' term. The value of the eleventh node is assumed as one, irrespective of 'storm' or 'no storm' days. There may be one or more hidden layers of computation nodes and an output layer having two computation nodes. In the learning phase, the values 1, 0 for nodes 1 and 2 respectively in the output layer would mean that the input is a 'squall-storm' data point, and 0, 1 for nodes 1 and 2 respectively would mean that the observation corresponds to 'no storm' day. A sigmoid function, which is a nonlinear activation function, is widely used as a transfer function. Each unit of each layer is connected to each unit of the next layer by the connection weights. The number of hidden layers is generally taken to be either 1 or 2. More hidden layers indicate more non-linearity of the decision boundary between classes. More nodes in a hidden layer indicate more number of curve segments engulfing a class. MLP model has been studied here by using different architectures.

The different stages of working of MLP:

1) Connection weights are initialized to small random values in the range (-0.5 to 0.5). Threshold value is also assumed. The weight values are modified during back propagation of the learning of the model until the error be minimized. The modified weights are used to validate the

testing datasets. The back propagation method basically uses gradient descent [16] technique for changing the weights. It is used to reduce the possibility of getting stuck in local optimal points or saddle points of the network.

2) Feed Forward stage: In this stage, each node (say *i*) in a layer α containing β nodes is joined to each node (say *j*) in the next layer $(\alpha+1)$ containing τ nodes, with a connection weight represented by $w_{ij}^{(\alpha)}$. If the output from the *i*-th node is y_i then the total input received by the *j*-th node is $\chi_j = \sum_{i=1}^{\beta} y_i w_{ij}^{(\alpha)}$. The output from the *j*-th node is $\frac{1}{1+\exp(-\chi_j)}$. This is valid for every layer.

3) Error: For every point in the training set, the expected output (e_j) is known. For a particular observation, if the actual output value of the *j*-th node in the output layer is o_j , then the error function, which is a mean squared error, [1] for that observation is

$$E = \frac{\sum_{j=1}^{2}(o_j - e_j)^2}{2}$$

. This error is to be minimized during the training phase.

4) Learning: There are two ways of learning the weights of an MLP. They are (i) batch mode learning and (ii) on-line learning. Here, on-line method of learning the weights is followed.

5) Back Propagation of Error: The error is distributed back to the previous layers. Note that the error is a function of every connection weight in the network. Usually, back propagation is done by using gradient descent method by a parameter, λ , which is known as 'learning rate'. Here, its value is 0.01.

6) Updation of weights: The weights are updated. The iterative process continues until the error be minimized to around 0.005 to 0.001. The modified weights are used in test dataset to validate output.

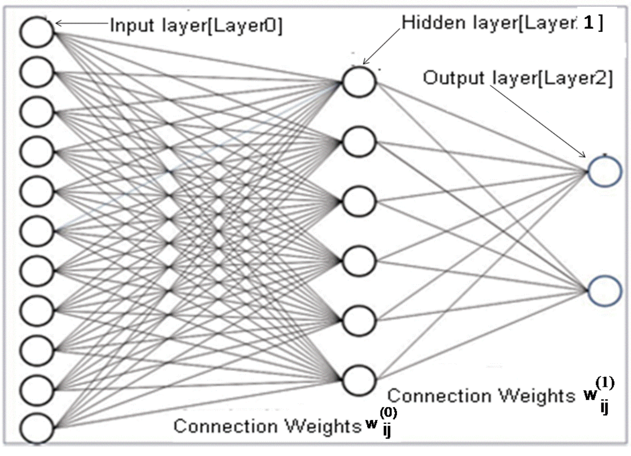

Fig. 1. 3-layered MLP with architecture 11-6-2.

Sometimes, even when the number of iterations becomes a large number or if the classification on the test set be unsatisfactory, the error may not reduce. In such cases, the architecture of MLP is to be changed. So, several 3-layered MLPs and 4-layered MLPs are studied. Three layered MLP consists of input layer, one hidden layer, and output layer. The number of nodes in hidden layer is varied from 3 to 10 to obtain a good classification. The different architectures of 3-layered MLP which were applied here are 11-3-2, 11-4-2, 11-5-2, 11-6-2, 11-7-2, 11-8-2, 11-9-2, and 11-10-2. The architectures of 4 layered MLPs (input layer, two hidden layers, and output layer), which are studied here are 11-4-3-2, 11-4-4-2, and 11-6-6-2.

C. K-nearest Neighbor (K-nn) Method

Yakowitz (1987) [17] extended the K-nearest neighbor method constructing a robust theoretical base for it and introduced it into the successful forecast in the hydrological research, [18]. The 1-nn classifier is an important pattern recognizing method, [19], where the distances of each of the training samples from the test samples are computed. The test samples have the same class label as the representative point nearest to them. K-*nn* is extension of 1-nn. Error bounds for K-*nn* rule are found in literature, [20].

Let $\underline{z}_1, \underline{z}_2, ..., \underline{z}_m$ be a collection of *m* given observations from *c* number of classes. Let $\theta_i \in \{1,2,...,c\}$ denote the class of $\underline{z}_i$, $i = 1,2,...,m$. Let θ_i be known for each *i*. We need to classify a new observation $\underline{z}$ to one of the *c* classes based on $(\underline{z}_i, \theta_i); i = 1,2,...,m$. The usual K-nearest neighbor rule is the following. 1) Choose a positive integer *K*. 2) Find K-nearest neighbors of $\underline{z}$, namely $\underline{e}_1, \underline{e}_2, ..., \underline{e}_K$ among $\{\underline{z}_1, \underline{z}_2, ..., \underline{z}_m\}$. Let K_i of these nearest neighbors belong to class $i, i = 1,2,...,c$. That is $\sum_{i=1}^{c} K_i = K \cdot 3$) Put $\underline{z}$ to class *I*, if $K_i > K_j$, $\forall j \neq i$.

The main problem in applying the rule to any data set is the choice of the value of *K*. Till now, there is no universally acceptable way of choosing the value of *K* for any data set. Here, a modification in the K-nn rule has been made. The suggested procedure to choose *K* in a particular way is described below.

Suggested procedure: 1) Let *N*=3 and *L*=*N*. 2) Find *L* nearest neighbors of $\underline{z}$ among $\{\underline{z}_1, \underline{z}_2, ..., \underline{z}_m\}$. (3) Let L_i of these nearest neighbors belong to class $i, i = 1,2,...,c$. That is $\sum_{i=1}^{c} L_i = L \cdot 4$) Let L_{i1}, L_{i2} denote respectively the maximum and second maximum values among all L_i 's. Let $\delta = L_{i1} - L_{i2} \cdot 5$) If $\delta < N$ then go to 7). 6) Classify $\underline{z}$ to class i1 and stop. 7) If $L < m$ then increase the value of L by 1 and go to 2). 8) If *L*=*m* then decide that no meaningful classification of $\underline{z}$ is possible and stop.

Remarks: 1) The above procedure classifies a point when the difference between the numbers of its neighbors for two most represented classes among the nearest neighbors becomes N. Otherwise it checks one more nearest neighbor. 2) N is taken as 3 in the procedure. The above procedure will yield the nearest neighbor classifier if N=1. 3) N>3 may make many points being unclassified. On the other hand, the value 2 for N may be very small. Thus, the value for N is taken as 3.

For modified K-nn method, five input variables are selected statistically out of these ten parameters, such as $x_1, x_2, ...x_{10}$. These selected five input predictors for modified K-nn are moisture difference at 850 hpa (x_3), at 600 hpa (x_5), and adiabatic lapse rate from 700 hpa to 600 hpa (x_8), from 600 to 400 hpa (x_9), and from 400 to 300 hpa (x_{10}). These five variables are selected on the basis of (i) more difference between the values of means, and (ii) variances of the variables. The similarity between two observation vectors, say,

$\underline{a}' = (a_1, a_2, ..., a_\gamma), \underline{b}' = (b_1, b_2, ..., b_\gamma)$ is defined as $\frac{\sum_{i=1}^{\gamma} a_i b_i}{\sqrt{\sum_{i=1}^{\gamma} a_i^2 \sum_{i=1}^{\gamma} b_i^2}}$. This similarity measure reflects the cosine of the angle between two vectors. The similarity is more if the angle is smaller. For a data point in the test set, its similarity with every point in the training set is calculated using the above said formula. Thus, for every point in the test set, there are 169 similarity values corresponding to 169 points in the training set. K-nearest neighbors in the training set of a point in the test set would mean K points having maximum similarity with respect to the test point. The modified K-nn algorithm is now applied using this similarity measure. When difference in L_i 's becomes 3, the test set point is classified.

IV. Result

The results are shown in Table I and Table II and are described in two ways:

a) Results obtained by the analysis of 10 input predictor weather variables $x_1, x_2, ...x_{10}$.

b) Results obtained by the analysis of 5 input predictor weather variables $x_3, x_5, x_8, x_9, x_{10}$.

It is observed from Table I and Table II that Modified K-nn technique and K-nn classifier yield better results in comparison with MLP and Multiple linear regression. Comparing the results obtained by applying K-nn and Modified K-nn methods on 10 input variables and 5 input variables, it is revealed that the results are better when Modified K-nn is applied on 5 input variables. Modified K-nn technique is the best classifier among all the methods used here, because 82.02% of the observations *(storm/no storm)* in the whole test dataset are properly classified. False alarm rate is 17.98% here. The 'squall-storms' are correctly classified up to 91.11% and 'no storms' are classified correctly up to 79.43% by modified K-nn method. Multiple linear regression technique can classify only 50% of the *storm/ no storm* events. So, this method is not found to provide satisfactory results for these data. Multilayer Perceptron is comparatively better than multiple linear regression.

TABLE I: Results Obtained by Applying Multiple Linear Regression and Multilayer Perceptron (MLP)

Design	No.of misclassified and % of misclassified points for squall storm days for the test set (Number of squall 'Storm' days is 90)	No. of misclassified and % of misclassified points for 'no storm' days for the test set (Number of 'no storm' days is 316.)	Total no. of misclassified and % of misclassified points for the test dataset. (Total size of test dataset is 406)
Multiple linear regression (10variables)	46, 51.1%	162, 51.26%	208, 51.23%
Multiple linear regression (5variables)	44, 48.8%	155, 49.05%	199, 49.01%
MLP (11-5-2)	40, 44.4%	151, 47.78%	191, 47.04%
MLP (11-6-2)	29, 32.2%	126, 39.87%	155, 38.17%
MLP (11-7-2)	46, 51.1%	135, 42.72%	181, 44.58%
MLP (11-8-2)	50, 55.55%	106, 33.54%	156, 38.42%
MLP (11-9-2)	25, 27.77%	155, 49.05%	180, 44.33%
MLP (11-4-3-2)	32, 35.55%	136, 43.03%	168, 41.37%
MLP (11-6-6-2)	29, 32.2%	141, 44.62%	170, 41.87%
MLP (6-4-2)	32, 35.55%	147, 46.5%	179, 44.08%
MLP (6-5-2)	33, 39.9%	135, 42.72%	168, 41.38%
MLP (6-6-2)	16, 17.77%	135, 42.72%	151, 37.19%
MLP (6-7-2)	26, 28.88%	140, 44.3%	166, 40.88%
MLP (6-4-3-2)	53, 58.88%	77, 24.36%	130, 32.01%

Comparing all the results obtained by 3-layered and 4-layered MLPs, the 6-6-2 MLP is better than the other MLPs. This 6-6-2 MLP can classify 82.23% of the 'squall-storm days', 57.28% of the 'no storm' occurrences, and 62.81% of *storm/no storm* correctly. False alarm rate of 6-6-2 MLP network is 37.19%.

Lee *et al.* in 1993 [21] applied decision tree method for the 12-hour forecast of thunderstorms. They used the radiosonde data of significant level temperature, significant level wind up to 100hpa, moisture and instability of the atmosphere, LCL, mixing ratio, lapse rate and potential temperature. The probability of detection is correct up to 81% and false alarm rate is 35%.

It may be mentioned here that the modified K-nn has provided 91.11% true positives for the 'squall-storm' events (which is more than 81% accuracy) with false alarm rate 8.88% (much smaller than 35%). This method classified correctly up to 82.02% of 'storm/ no storm' events. In this study, only two types of upper air weather variables are used for more than 91% correct prediction. But in the study of Lee *et al.* [21], nearly eight types of weather data were required for 81% correct prediction. It may be said from this work that more accurate forecast is obtained by modified K-nn technique using a very few atmospheric features.

TABLE II: RESULTS OBTAINED BY APPLYING K-NEAREST NEIGHBOR RULE (K-NN), AND MODIFIED K-NN RULE

Design	Number and % of misclassified points for squall storm days for the test dataset (Number of 'squall storm' days is 90)	Number and % of misclassified points for 'no storm' days for the test dataset (Number of 'no storm' days is 316.)	Total number and % of misclassified points for the test dataset. (Total size of test dataset is 406)
K-nn with 10 variables, (K=5)	27, 30%	159, 50.31%	186, 45.81%
K-nn with 10 variables, (K=9)	50, 55.55%	91, 28.8%	141, 34.73%
K-nn with 10 variables, (K=11)	33, 36.67%	103, 42.08%	136, 33.49%
K-nn with 5 variables, days (K=5)	10, 11.11%	163, 51.58%	173, 42.61%
K-nn with 5 variables, (K=7)	41, 45.55%	76, 24.05%	117, 28.82%
K-nn with 5 variables, (K=17)	22, 24.44%	90, 28.48%	112, 27.58%
Modified K-nn with 10 variables	11, 12.22% (79 out of 90 = 87.7% properly classified)	90, 28.48% (226 out of 316 = 71.5% properly classified)	101, 24.87% (305 out of 406 = 75.1% properly classified)
Modified K-nn with 5 variables	08, 8.88% (82 out of 90= 91.11% properly classified)	65, 20.57% (251 out of 316=79.4% properly classified)	73, 17.98% (333 out of 406=82.02% properly classified)

V. CLIMATOLOGY

Moisture difference profile and conditional instability of the morning (00:00UTC) time act as two types of important weather predictors from surface to the certain geo-potential heights of the upper atmosphere for the formation of evening squall-storm. Dry adiabatic lapse rate (d T/ d Z) enables us to accurately predict the temperature change of unsaturated air as it moves vertically within the atmosphere, [13]. During the adiabatic process, as an air parcel rises and expands, its temperature drops: if it is unsaturated, the air parcel's relative humidity increases [13]. Expansional cooling is the principal means of cloud formation in the atmosphere, [13]. It is revealed from our study that morning instability in the atmosphere causes the transport of moisture from surface to the upper air during the whole of the day time to form super cell thundercloud for the genesis of severe thunderstorm associated with squall in the evening. Moisture difference (MD=T-T_d) indicates the measure of humidity or saturation of the atmosphere, [13]. More the atmosphere is conditionally unstable, larger is the amount of incursion of moisture from the surface to the upper air. The energy that drives conditional instability is convective available potential energy (CAPE), [14]. The result of the model shows that modified K-nn (82.02% correct predictions) is found to be a more perfect classifier than MLP (62.81% correct predictions) and statistical classifier (50% correct predictions). It is found that modified K-nn method is so efficient classifier that it can perform 91.11% of squall-storm prediction correctly.

VI. DISCUSSION AND CONCLUSION

Generally atmospheric surface parameters, upper air parameters measured by radiosonde, Doppler radar and satellite imageries are required to predict severe storm in a conventional way. Modified K-nn is found to be the best classifier in this study to forecast thunderstorms with a lead time of around 10 to 14 hours. The upper air humidity at 850 hpa and at 600 hpa, and the conditional instability from 700 hpa to 300 hpa of the early morning (00:00UTC) atmosphere are so important parameters that one can predict evening squall-storm only by these two types of weather variables applying modified K-nn method. Two of the upper air morning parameters, humidity and lapse rate play a key role to form thundercloud from the early morning throughout the whole day for the genesis of squall-storm in the evening time. The lead time of 10-14 hours is sufficient for alerting people from this catastrophic weather event. The challenge that has been undertaken for this forecasting work is the proper selection of the machine learning technique to get accurate prediction using only the said two types of input weather variables recorded in the early morning. Experiments were also conducted on the 5 variable data for predicting *storm / no storm* using multiple correlation and MLP, and the corresponding results are found unsatisfactory. Modified K-nn is a new method of classification in Pattern Recognition, and this method has not yet been implemented on other pattern classification data sets. The pattern recognition/machine learning community is not aware of this method. Its theoretical properties are not known. However, its utility in the classification of 'squall-storm' and 'no storm' days is beyond doubt, as more than 82% of *storm/no storm* are accurately classified.

ACKNOWLEDGEMENT

The authors are thankful to India Meteorological Department, Govt. of India for the meteorological data. The corresponding author thanks University Grants Commission of India for providing funds for doing this research work.

REFERENCES

[1] A. Marinaki, M. Spiliotopoulos, and H. Michalopoulou, "Evaluation of atmospheric instability indices in Greece," *Advances in Geosciences*, vol. 7, pp. 131-135, 2006.

[2] Y. Radhika and M. Shashi, "Atmospheric Temperature Prediction using Support Vector Machines," *International Journal of Computer Theory and Engineering*, vol. 1, no. 1, pp. 55-58, 2009.

[3] A. K. Showalter, "A stability index for thunderstorm forecasting," *Bulletin of American Meteorological Society*, vol. 34, pp. 250-252, 1953.

[4] C. J. Boyden, "A simple instability index for use as a synoptic parameter," *Meteorological Magazine*, vol. 92, pp. 198-210, 1963.

[5] Z. Sokol, D. Rezacova, and P. Pesice, "The inclusion of surface synoptic data in very short range predictions of convective event," *Studia Geophysica and Geodaetica*, Springer, Netherlands, vol. 40, issue 2, 1996.

[6] F. H. Ludlam, "Severe Local Storms: A Review," *Meteorological Monographs*, vol. 5, pp. 1-30, *American Meteorological Society*, Sept., 1963.

[7] K. A. Browning, “Airflow and precipitation trajectories within severe local storms which travel to right of the winds,” *Journal of Atmospheric Science*, vol. 21, pp. 634-639, 1964.

[8] R. Rotunno and J. B. Klemp, “The influence of the shear induced pressure gradient on thunderstorm motion,” *Monthly Weather Review*, vol. 110, 136-151, 1982.

[9] C. Y. C. Chung and V. R. Kumar, “Knowledge acquisition using a neural network for weather forecasting knowledge-based system,” *Neural Computing & Applications*, Springer, London, vol. 1, no. 3, pp. 215-223, 1993.

[10] C. Marzban, “Neural Networks for Postprocessing Model Output: ARPS,” *Monthly Weather Review*, American Meteorological Society, vol. 131, pp. 1103-1111, June, 2003.

[11] Z. Jan, M. Abrar, and S. Bashir, “Seasonal to Inter-annual Climate Prediction Using Data Mining KNN Technique,” *Wireless Networks, Information Processing and Systems*, Springer, Berlin, vol. 20, pp. 40-51, 2008

[12] R. Li, H.-N. Wang, H. He, Y.-M. Cui, and Zhan-Le Du, “Support Vector Machine combined with K-Nearest Neighbors for Solar Flare Forecasting,” *Chinese Journal of Astronomy and Astrophysics*, vol. 7, no. 3, pp. 441-447, 2007.

[13] J. M. Moran, M. D. Morgan, and P. M. Pauley, *Meteorology: The Atmosphere and the Science of Weather*, 5^{th} Edition, Prentice Hall, 1997.

[14] H. Volland, *Handbook of Atmospheric Electrodynamics*, vol. 1, 1995.

[15] A. M. Gun, M. K. Gupta, and B. Das Gupta, *Fundamentals of Statistics*, vol. 1, World Press, 2005.

[16] B. Yegnanarayana, *Artificial Neural Networks*, Prentice Hall of India Private Limited, 1999.

[17] S. Yakowitz, “Nearest neighbor method for time-series analysis,” J. Time-series Analysis, vol. 8, 235-247, 1987.

[18] G. Galeati, “A comparison of parametric and non-parametric methods for runoff forecasting,” *Hydrolog. Sci. J.*, vol. 35, pp. 79-94, 1990.

[19] Z. Q. Bian and X. G. Zhang, *Pattern Recognition*, Beijing: TsingHua Univ. Press, 2000.

[20] T. M. Cover, “Estimation by the Nearest Neighbor Rule,” *IEEE Transactions on Information Theory*, vol. IT-14, no. 1, pp. 50-55, 1968.

[21] R. R. Lee and J. E. Passner, “The Development and Verification of TIPS: An Expert System to Forecast Thunderstorm Occurrence,” *Weather and Forecasting*, vol. 8, pp. 271-280, 1993.

Himadri Chakrabarty received Ph.D. (Tech.) degree in radiophysics and electronics from Calcutta University, Kolkata. India. She graduated with Physics (Hons.) and got M.Sc. in Electronic Science and M.Tech. degree in 1992 from Calcutta University, India.

She is an associate professor and Head of the Department of Computer Science, Surendranath College under Calcutta University, India. She was also a teacher fellow in Institute of Radiophysics and Electronics, Calcutta University sponsored by University Grants Commission (UGC) of India. She is carrying on research activities in collaboration with Indian Statistical Institute, Kolkata. Her research field is based on Pattern Recognition, Image Processing, and Data Mining and their applications on Atmospheric Science, Space Science and Communication Engineering. She conducted two Research projects which got financial support from UGC of India. The projects were based on the Application of Artificial Neural Network in Atmospheric Science.

Dr. Chakrabarty is the member of Indian Meteorological Society, India.

C. A. Murthy obtained B. Stat (Hons), M. Stat. and Ph. D. degrees from the Indian Statistical Institute (ISI), India.

He visited the Michigan State University, East Lansing in 1991-92 for six months, and the Pennsylvania State University, University Park for 18 months in 1996 - 97. He is a professor in the Machine Intelligence Unit of Indian Statistical Institute, India. He was the head of the Machine Intelligence Unit, ISI from 2005 to 2010. He is presently the Professor In-charge of the Computer and Communication Sciences Division of ISI. His fields of research interest include Pattern Recognition, Image Processing, Machine Learning, Neural Networks, Fractals, Genetic Algorithms, Wavelets and Data Mining.

Prof. Murthy received the best paper award in 1996 in Computer Science from the Institute of Engineers, India. He received the Vasvik award along with his two colleagues for Electronic Sciences and Technology for the year 1999. He is a fellow of the National Academy of Engineering, India and National Academy of Sciences, India.

Ashish Dasgupta received Ph.D. degree from Calcutta University, India in the year 1971 in the field of Radiophysics and Electronics. He was a NRC/ NAS Senior Resident Research Associate at the then Airforce Geophysics Laboratory, Hanscom AFB, Massachusetts, USA from 1980 to 1983. He was a leader of a group of workers in the area of Radiowave propagation, particularly with reference to communication satellites and GPS.

He served as a professor in Institute of Radiophysics and Electronics of Calcutta University, India from 1987 and retired in 2007 as a professor. He is attached presently as an advisor with the S. K. Mitra Center for Research in Space Environment, University of Calcutta, Kolkata, India. He conducted many projects on Space Research. One of his projects was on the study of Space Weather Environment, sponsored by Indian Space Research Organization (ISRO), Govt. of India. He conducted another project on GPS which was sponsored by Department of Science and Technology (DST), Govt. of India in collaboration with National Physical Laoratory, India. He has about 80 publications in the International journals like “Radio Science,” “Journal of Geophysical Research (JGR)” etc. He is the reviewer of journals “JGR,” “Radio Science” and other international journals.