International Journal of Engineering Sciences & Research Technology

(A Peer Reviewed Online Journal)
Impact Factor: 5.164

IJESRT

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

AN ARTIFICIAL NEURAL NETWORK BASED DEVELOPMENT OF CLOUD BURST FORECASTING MODEL BY USING TIME-SERIES DATA

Neha Rana*1 & Er. Shivani Rana2

*1&2Himachal Pradesh Technical University, Hamirpur, Himachal Pradesh

DOI: https://doi.org/10.29121/ijesrt.v9.i10.2020.14

ABSTRACT

Over the last decades, weather forecasting models using the numerical data of various previous years have been improving steadily to provide a more predictable and accurate model. However, the atmosphere conditions are a highly chaotic system and always vary with time and weather forecasts are a major benefit for society and sustainable development. So, in this research we focus to develop a model using the optimized Artificial Neural Network (ANN) for prediction of cloud bursting in India and developed model in known as Cloud Burst Foresting (CBF) model for forecasting of rainfall or cloud burst based on the previous record of the bursting in different state. Here the concept of Particle Swarm Optimization (PSO) is used as an optimization technique in pre-processing step to separate the previous year rainfalls data into two categories such minimum and maximum rainfalls. Basically, PSO separate the cloud burst recorded data using a novel fitness function that help to train the CBF model accuracy and if training is better, then the prediction accuracy will be high. By utilizing the concept of optimized ANN, the prediction accuracy is high in terms of percentage of correct predication and with minimum percentage of incorrect prediction. At last, the performance of the model is calculated to validate the proposed CBF model and this shows that it is possible to use ANN as a machine learning technique in order to estimate future cloud burst forecast uncertainty from past forecasts data of bursting. The main constraint in the performance of our proposed CBF model seems to be the number of past forecasts available for training the ANN.

KEYWORDS: Weather Forecasting, Cloud Burst Foresting, PSO, Machine learning, ANN.

1. INTRODUCTION

Weather forecasting and climatology is important since it helps determine future climate expectations [1]. Through the use of latitude, one can determine the likelihood of snow and hail reaching the surface. Researchers can also be able to identify the thermal energy from the sun that is accessible to a region using the concept of weather forecasting model [2]. It is the scientific study of climates based on the previous records of weather, which is defined as the mean weather conditions over a period of time. Historical weather forecasting focuses primarily on climate changes throughout history and the effects of the climate on people and events over time. Weather simply refers to the condition of air on the earth at a given place and time [3]. It is a continuous, data-intensive, multidimensional, dynamic and chaotic process. Forecasting is the process of estimation in unknown situations from the historical data. Weather forecasting is one of the most scientifically and technologically challenging problems around the world in the last century [4]. Block diagram of tradition weather forecasting model is shown in the Fig. 1.

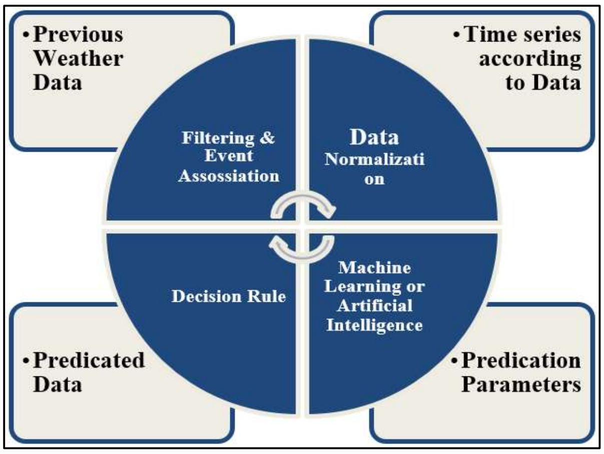


Fig 1: Block Diagram of Tradition Weather Forecasting Model

Weather forecasting entails predicting how the present state of the atmosphere will change. Present weather conditions are obtained by ground observations, observations from ships, observation from aircraft, radio sounds, Doppler radar and satellites [5], [6]. This information is sent to meteorological centers where the data are collected, analyzed and made into a variety of charts, maps and graphs. Modern high-speed computers transfer the many thousands of observations onto surface and upper-air maps. Weather forecasts provide critical information about future weather. There are various techniques involved in weather forecasting, from relatively simple observation of the sky to highly complex computerized mathematical models [7]. Weather prediction could be one day/one week or a few months ahead. The accuracy of weather forecasts however, falls significantly beyond a week. Weather forecasting remains a complex task for researchers, due to its chaotic and unpredictable nature. The types of weather forecasting are given in below section [8]:

Persistence Forecasting: The simplest method of forecasting the weather is persistence forecasting. It relies upon today's conditions to forecast the conditions tomorrow. This can be a valid way of forecasting the weather when it is in a steady state, such as during the summer season in the tropics.

Synoptic Forecasting: This method uses the basic rules for forecasting. Meteorologists take their observations, and apply those rules to make a short-term forecast.

Statistical Forecasting: It works in the basis of records of average temperatures, average rainfall and average snowfall over the years give forecasters an idea of what the weather is "supposed to be like" at a certain time of the year.

Computer Forecasting: Forecasters take their observations and plug the numbers into complicated equations. Several ultra-high-speed computers run these various equations to make computer "models" which give a forecast for the next several days.

Weather forecasting is used in many situations like severe weather alerts and advisories, predicting the behavior of the cloud for air transport, prediction of waterways in a sea, agricultural development and avoiding forest fire [9]. Some importance of weather forecasting is listed as:

- ***Severe weather alerts and advisories***
- ***Predicting the behavior of the cloud or cloud burst prediction***
- ***Prediction of waterways in a sea***
- ***Agricultural development***
- ***Avoiding Forest fire***
- ***Military applications***

Basically in this research, we focus on the development of a Cloud Burst Forecasting (CBF) model using the concept of the Artificial Neural Network (ANN) for the prediction of future chances of cloud bursting in India and the major contributions in this research are listed as:

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

- For the pre-processing of the used dataset of rainfalls in proposed CBF model, Particle Swarm Optimization (PSO) is used with a novel fitness function for the separation of data into two part minimum and maximum rainfalls.
- The concept of ANN is used as a classifier with the optimization algorithm to predict the cloud bursting in India based on the previous time-series recorded data.
- At the Percentage Correct Prediction (PCP) and Percentage Incorrect Prediction (PIP) is calculated to verify the responsiveness of proposed CBF model.

In this section, we introducing the concept of an optimization based pre-processing technique using PSO algorithm and ANN used as a classifier in the proposed CBF model and the remaining paper is systematized as follows. In Sect. 2, survey of weather forecasting or cloud burst prediction model are described and on the Sect. 3, we presents the material and methodology of proposed CBF model. The experimental results of CBF model is presented in Sect. 4 and conclusion with the future possibilities of CBF model is described in the Sect. 5.

2. RELATED WORK

In this section, existing weather forecasting or cloud bursting model are analyzed to find out the issues related. In 2018, ***N. R. Deshpande et al.*** was developed a model for Statistical characteristics of cloud burst and mini-cloud burst events during monsoon season in India. A new and broad definition of cloud burst event has been proposed which covers cloud burst events triggered by orography in Himalayan region and cloud burst events in rest of India in which rainfall in 1 hour exceeds 10 cm. A new high rainfall event called MCB has been coined. The analysis of cloud burst events has been carried out using hourly rainfall data of 126 wide spread stations during the period 1969–2015 for the monsoon season to bring out their statistical characteristics. The study brought out that there are around 26 cloud burst events over the low orography region and plains of India. ***Francisco J. Clemente-Castello et al.*** in 2018, presented a model using the concept of MapReduce for Cloud Burst prediction. Authors address the problem of how to estimate the runtime of iterative MapReduce applications in hybrid cloud bursting scenarios where on premise and off-premise monitoring that host a MapReduce environment need to communicate over a weak link. ***Lei Xu et al. in 2019*** designed a probabilistic load forecasting for buildings considering weather forecasting uncertainty and uncertain peak load. This research work, authors presented a new probabilistic building load forecasting model considering uncertainties in weather forecasts and abnormal peak load. The probabilistic normal load forecasting model is built using the ANN and the probabilistic temperature forecasts. The probabilistic abnormal peak load forecasting model consists of two models quantifying the probabilistic occurrence and magnitude of the peak abnormal differential load respectively. The test results show that the ANN deterministic load forecasting model can achieve satisfactory performance with the average mean absolute percentage error (MAPE) of 5.0%. The percentage of error is less as compare to the existing work but need more improvements. In 2018, ***S. Scher and A. Messori*** had conducted a research to design a model for Predicting Weather Forecast Uncertainty with Machine Learning. Authors proposed a new method based on deep learning with artificial convolutional neural networks that is trained on past weather forecasts. Given a new weather situation, it assigns a scalar value of confidence to medium range forecasts initialized from said atmospheric state, indicating whether the predictability is higher or lower than usual for the time of the year. While our method has a lower skill than ensemble weather forecast models in predicting forecast uncertainty, it is computationally very efficient and outperforms a range of alternative methods that do not involve performing numerical forecasts. This shows that it is possible to use machine learning in order to estimate future forecast uncertainty from past forecasts. The main constraint in the performance of our method seems to be the number of past forecasts available for training the machine learning algorithm.

Usually, weather forecasting is done categorically for different geographical locations as it differs for region to region due to the variations in time and other factors but in the last couple of decades, the quality and overall skill of weather forecasting have achieved notable and worthy improvement. A prime reason for the improved weather forecasting skill is attributed to the improvements in Artificial Intelligence (AI) capabilities. With the establishment of different AI methods, significant efforts were undertaken to upgrade from the optimal interpolation techniques to the AI procedures based on the variation methods. AI based weather forecasting is essentially useful and required to save life and property. It helps people to plan their activities and movements ahead, stay safe and not go out if there's something wrong with nature, and get updated with the possibilities of

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

the arrival of storms that cause destruction and havoc. Time series based weather forecasting also help farmers to plan the sowing and harvesting of their crops as even today farming/agriculture sector is greatly dependent on weather and climatic conditions especially in countries like India. Countries spend huge amount of money on weather forecasting as it is highly important for the safety and security of their people and their state. It is also necessary to stay in pace with other countries in determining weather and excelling in the weather forecasting industry.

But due to presence of un-normalized record of previous weather data, various existing model cannot predict accurate condition of weather and it is still a major challenge for researcher and scientist. So pre-processing like filtering of data or normalization of data is one of the most fundamental steps which should be considered during the development of weather forecasting model. In this research work, ANN is used as a classifier to train the model based on the pre-processed data which helps to achieve better prediction rate with minimum error. At the last of model, the performance metrics of proposed work will be calculated to verify the effectiveness of model terms of predication rate and error with simulation time.

3. MATERIAL AND METHOD

In this section of research paper, we explain the methodology of proposed CBF model using the combination of PSO with ANN as classification or prediction algorithms. The designed model in dived into two different parts named as 1st part (Training Part) and 2nd part (Prediction Part). The brief details about both parts are given in the below section of this paper and Fig. 2 show the flowchart of both part of the proposed CBF model.

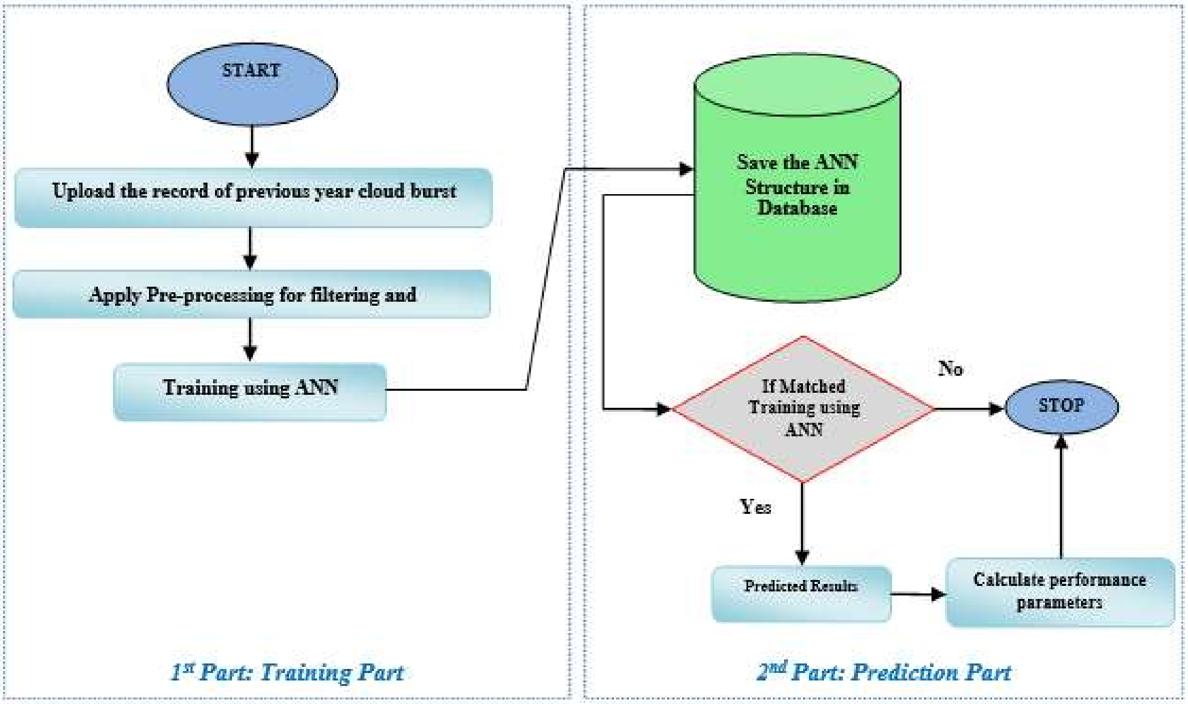


Fig 2: Flowchart of Proposed CBF Model

Above Fig. 2 shows the CBF model flowchart using the PSO with ANN and the problem exiting model is resolve in this research work by using the concept of ANN to train the model based on the normalized previous cloud burst data records. The subsequent steps demonstrate the variety of phases that need to be accomplished:

Step 1: To design the CBF framework we used the concept of GUI in MATLAB software for simulation of proposed CBF model. The developed model is shown in the Fig. 3.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

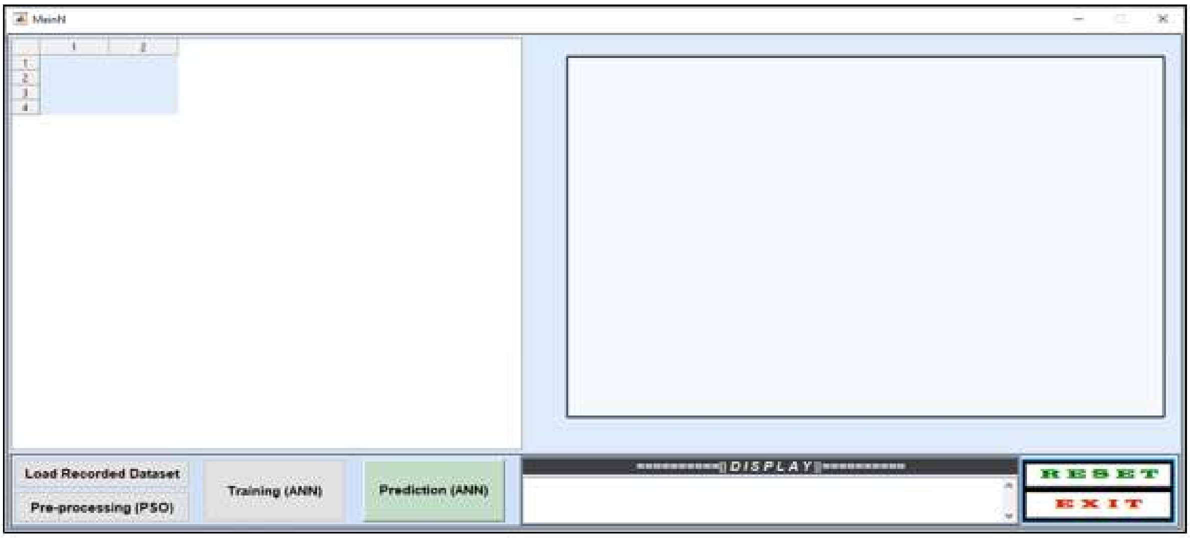


Fig 3: Proposed CBF Model

Step 2: Upload previous recorded weather data regarding the cloud burst for the simulation of proposed CBF model. The uploaded data of previous recorded rainfalls in India's different places or states is shown in the Fig. 4

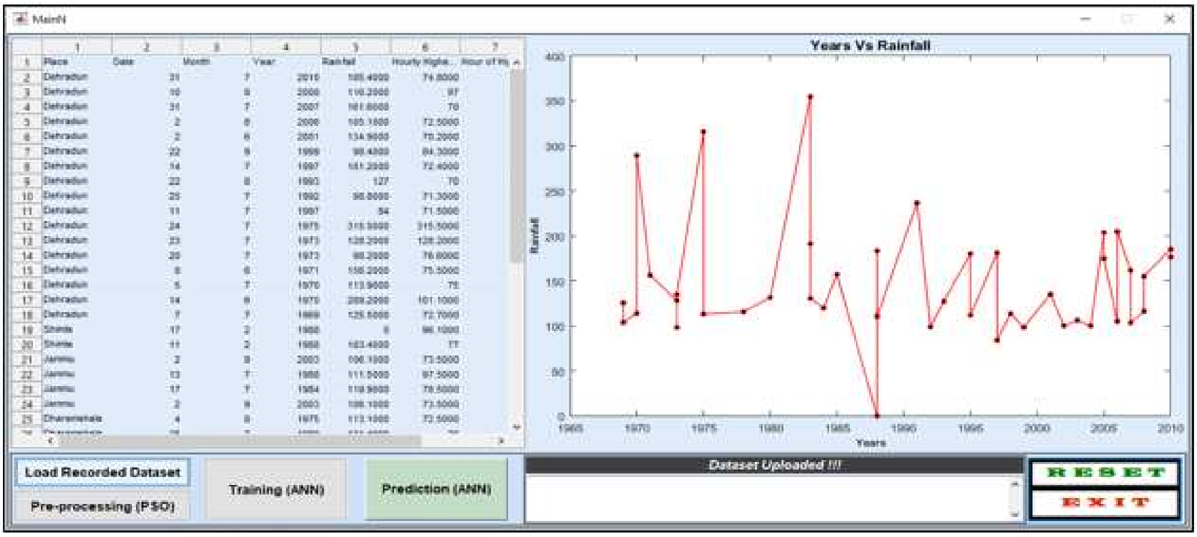


Fig 4: Uploaded Data CBF Model

Step 3: Apply pre-processing on the uploaded cloud burst data to make compatible data for CBF model and swarm based optimization approach will be used to select a better set of the features. The optimized and separated data of previous recorded rainfalls is shown in the Fig. 5

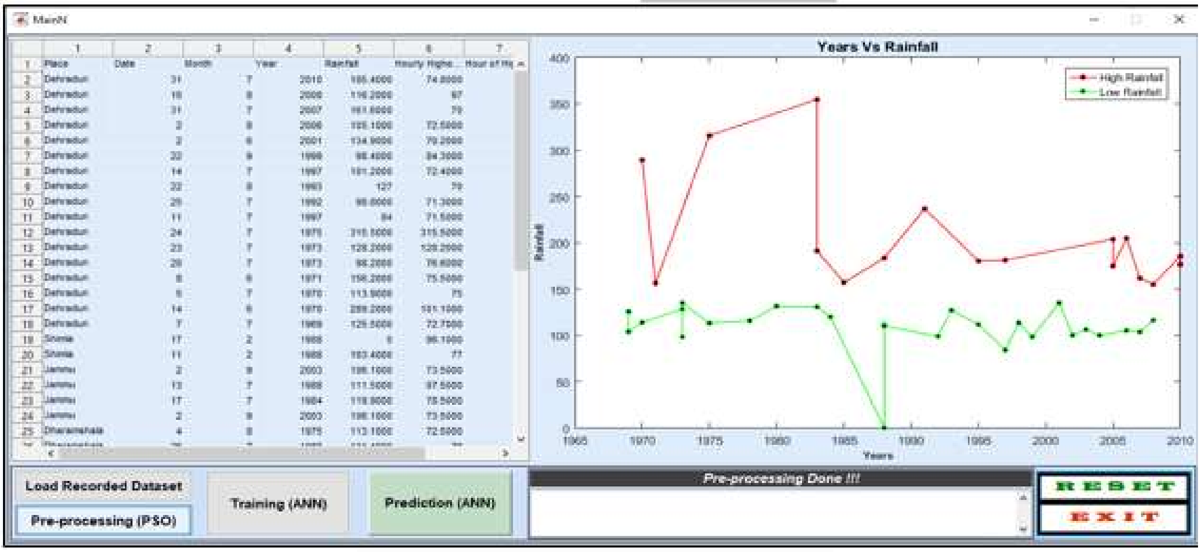


Fig 5: PSO-based pre-processing in CBF Model

Above figure represents the pre-processing of the proposed CBF model using the fitness function of PSO according to the requirement, so we can find out the appropriate set of rainfalls data in two part named as minimum rainfalls and maximum rainfalls. The used fitness function of the proposed CBF model is written below:

$$\boldsymbol{Fitness} = \begin{Bmatrix} F_s = 1(True) & if\ F_s > F_t \\ F_t = 0(False) & otherwise \end{Bmatrix} \quad \textbf{(1)}$$

Where F_s is the selected rainfalls data and F_t is the average of rainfalls in previous years. After the designing of fitness function, we apply the PSO algorithm to the rainfalls data and separate the data based on the minimum and maximum rainfalls using the written algorithm of PSO:

Algorithm 1: PSO-based Pre-processing

Input: Rainfalls Data (R-Data)
Output: Pre-processed Data (P-Data)

```
Strat
Load Dataset as Data
Total Feature Length, L = size (Data)
High-Data = []
Low-Data = []
count1=1
count2=1
For i = 1→L
  If Data (i)>=Average (Data)
    High-Data (count1, 1) = Data (i, 1)
    High-Data (count1, 2) = Data (i, 2)
    High-Data (count1, 3) = Data (i, 3)
    High-Data (count1, 4) = Data (i, 4)
    High-Data (count1, 5) = Data (i, 5)
    High-Data (count1, 6) = Data (i, 6)
    High-Data (count1, 7) = Data (i, 7)
    count1=count1+1
  Else
    Low-Data (count2, 1) = Data (i, 1)
    Low-Data (count2, 2) = Data (i, 2)
```

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

```
        Low-Data (count2, 3) = Data (i, 3)
        Low-Data (count2, 4) = Data (i, 4)
        Low-Data (count2, 5) = Data (i, 5)
        Low-Data (count2, 6) = Data (i, 6)
        Low-Data (count2, 7) = Data (i, 7)
        count2=count2+1
    End –If
End – For
1st For High Rainfall Data
High-PSO-Data = []
[r, c]=size (High-Data)
For si=1→r
   For s=1→c
   Fs=High-Data (si, s)
   Ft=Average (High-Data)
   Fitness = Call fitness function using equation 1
   High-PSO-Data (si, s) = PSO (Fs, Ft, Fitness)
   End – For
End – For
2nd For Low Rainfall Data
Low-PSO-Data = []
[r, c]=size (Low-Data)
For si=1→r
   For s=1→c
   Fs= Low-Data (si, s)
   Ft=Average (Low-Data)
   Fitness = Call fitness function using equation 1
   Low-PSO-Data (si, s) = PSO (Fs, Ft, Fitness)
   End – For
End – For
P-Data = [Low-PSO-Data High-PSO-Data]
Returns: P-Data as a pre-processed rainfall data
End – PSO Algorithm
```

Step 4: Initialize ANN for predication purpose using two phases, namely, training and testing with feed forward back propagation concept. After the training, save the trained structure which used in the cloud burst prediction section to classify the exact weather conditions. The used ANN algorithm is given as:

Algorithm 2: ANN as a classifier

```
Input: Pre-processed Data (P-Data) as Training Data, Group of data in terms of minimum and maximum rainfalls
(Cat = 2) and Neurons (N)
Output: Prediction Results (P-Results)
Strat Training
For i = 1→Cat (2)
    If Data from minimum rainfall record
        Cat (1, i) = P-Data (i)
    Else if Data from maximum rainfall record
        Cat (2, i) = P-Data (i)
    End –If
End – For
Call and set the ANN using P-Data
Set, CBF-NET = NEWFF (P-Data, Cat, N)
CBF-NET = TRAIN (CBF-NET, T, G)
```

RESEARCHERID
THOMSON REUTERS

[Rana *et al.*, 9(10): October, 2020]
IC™ Value: 3.00

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

Start Prediction
Current Data, C-Data = Current year rainfall data
P-Results = Prediction (CBF-NET, C-Data)
Returns: P-Results as a Prediction Results
End – ANN Algorithm

Step 5: At last of simulation, the performance parameters of proposed CBF model is calculated in terms of predication rate and error.

Above flowchart and algorithms are used in the procedural steps of proposed CBF model. By using above procedure we achieve better results which are well described in the next section of this research paper on the behalf of collected dataset from different state of India. The sample of dataset is given in the Table I with their attributes.

Table I: Rainfall Dataset with Attributes

PLACE	Date	Month	Year	Rainfall	Hourly Highest Rainfall	Hour of Highest Rainfall	Duration
DEHRADUN	31	7	2010	185.4	74.8	0	1
DEHRADUN	10	8	2008	116.2	97	14	21
SHIMLA	11	2	1988	183.4	77	13	17
JAMMU	2	9	2003	106.1	73.5	0	2
JAMMU	13	7	1988	111.5	97.5	2	11
DHARAMSHALA	4	8	1975	113.1	72.5	21	29
DHARAMSHALA	25	7	1980	131.4	70	0	5

The simulation results of proposed CBF model using PSO with ANN as a classification of prediction algorithm is described in the below section of paper.

4. RESULTS AND ANALYSIS

In this section, we describe the simulation results of the proposed ANN based development of CBF model by using time-series rainfalls data of previous year in India. Basically we calculate the Percentage of Correct Prediction and Incorrect Prediction for developed CBF model and the simulation results are described in the below section of paper. Firstly we describe the results of training section of proposed CBF model using ANN.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

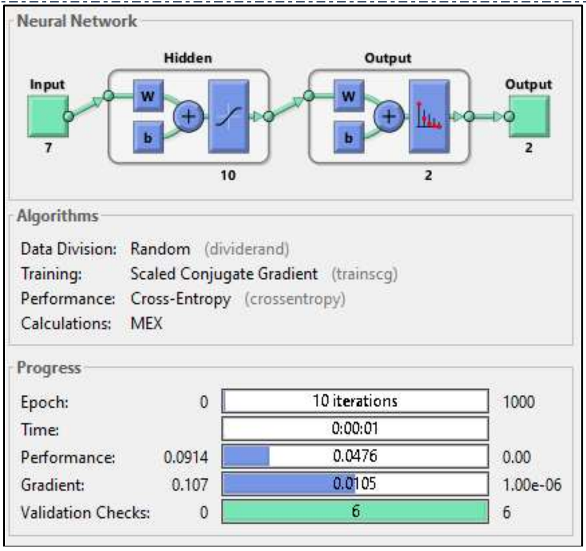
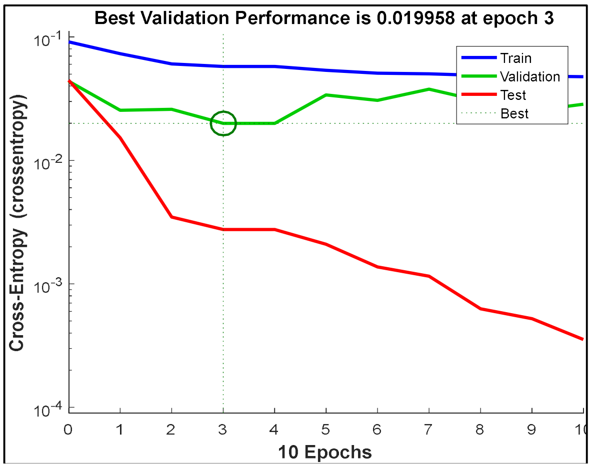


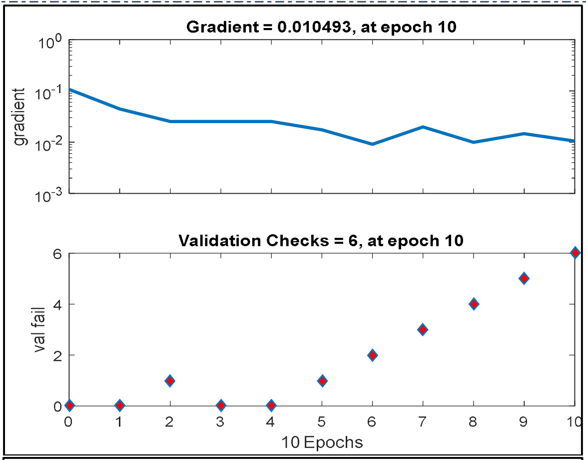
Fig 6: ANN Training Structure (CBF-Net)

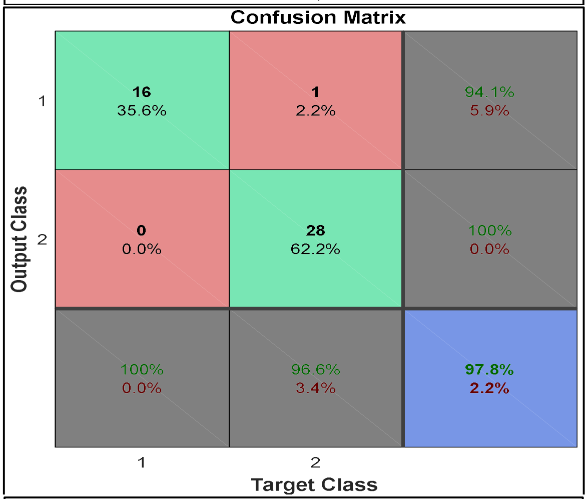
In this figure, training architecture of ANN is shown that have three layers named as, input, hidden and output. In output layer, we passes the Pre-processed rainfalls data as Training Input Data that is obtained after the PSO based optimization. The algorithm of ANN contains some basic information like:

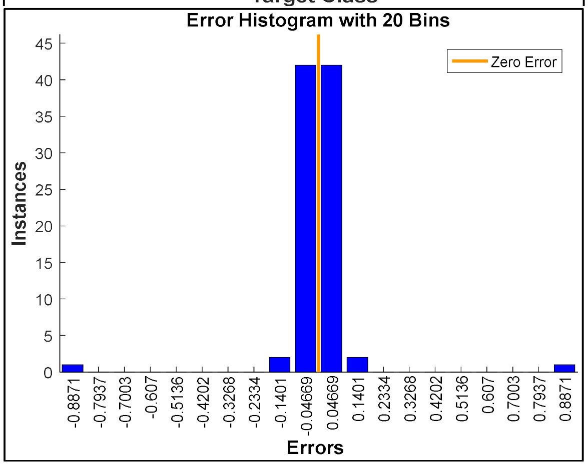
- ***Number of Epochs (E) // Iterations used by ANN***
- ***Number of Neurons (N)***
- ***Performance Parameters: Cross Entropy***
- ***Techniques: Scaled Conjugate Gradient***
- ***Data Division: Random***
- ***Calculation: Mex as transfer function***

The progress of training of the proposed CBF model using ANN is shown in the Fig. 7.




Gradient = 0.010493, at epoch 10
gradient
Validation Checks = 6, at epoch 10
val fail
10 Epochs


Confusion Matrix
Output Class
Target Class
16
35.6%
1
2.2%
94.1%
5.9%
0
0.0%
28
62.2%
100%
0.0%
100%
0.0%
96.6%
3.4%
97.8%
2.2%


Error Histogram with 20 Bins
Zero Error
Instances
Errors

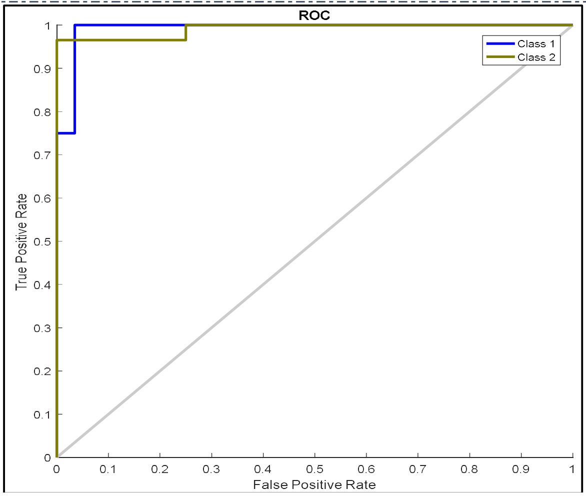


Fig 7: ANN Training Parameters (a) Cross Entropy (b) Training State (c) Confusion Matrix (d) Error Histogram and (e) ROC

The above figures describe the training results of the proposed ANN based development of CBF model by using the PSO as an optimization technique. Where Fig 7 (a) and (b) represents the cross entropy and of the ANN that is used to check the performance of the model with respect to the epochs or iterations and Fig. 7 (c) show the confusion matrix of the training model with error histogram in the Fig. 7 (d). Based on the training of the CBF mode using the ANN, Fig. 7 (e) is generated that show the ROC (Reverse Operating Characteristics). The simulation results of the proposed CBF model is shown in the Table II.

Table II: Percentage of Correct & Incorrect Prediction

Rounds	Correct Prediction	Incorrect Prediction
1	97.8	2.2
2	98.5	1.5
3	97.3	2.7
4	99.7	0.3
5	95.3	4.7
6	98.7	1.3
7	97.8	2.2
8	97.6	2.4
9	98.8	1.2
10	97.2	2.8
Average	**97.87**	**2.14**

The graphical representation of simulation results of the proposed CBF model is shown in the Fig. 8.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

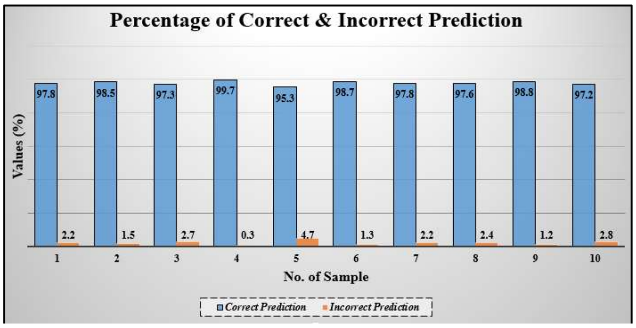


Fig 8: Performance Parameters of CBF Model

Above simulation results based on the prediction accuracy and error of the CBF model show that the model have its own impact in the era of forecasting model using swarm-based metaheuristic technique in with artificial intelligence technique.

5. CONCLUSION AND FUTURE WORK

In this paper, an optimized ANN is used to develop a model of CBF for forecasting of rainfall or cloud burst in India based on the previous record of the bursting in different state. We have evaluated whether it is possible to use machine learning techniques in order to infer cloud burst forecast uncertainty from the error and spread of past ensemble recorded data of forecasts. We showed that this is possible with a method that is based on a PSO-based ANN, which takes as input atmospheric fields and is trained with either the error of past deterministic cloud burst forecasts, or the ensemble spread of past ensemble forecasts. When presented with a new atmospheric field (may be previous or new data), it assigns a scalar value of predictability of the cloud bursting probability. Both our networks are shown to be useful predictors of cloud burst forecast uncertainty with higher percentage of correct predication and with rate of lower percentage of incorrect prediction. In future, proposed CBF model could be designed behalf of the concept of Convolutional Neural Network for the previous year cloud bursting data using the individual swarm intelligence approach to predict the cloud forecasting.

REFERENCES

[1] Narkhedkar, S.G., Morwal, S.B., Padmakumari, B., Deshpande, C.G., Kothawale, D.R., Maheskumar, S. and Kulkarni, J.R. (2015) Rainfall mechanism over the rain shadow region of north peninsular India. Climate Dynamics, 45, 1493–1512.
[2] Ajayamohan, R.S., Merryfield, W.J. and Kharin, V.V. (2010) Increasing trend of synoptic activity and its relationship with extreme rain events over central India. Journal of Climate, 23, 1004–1013.
[3] Das, P.K. (2015) Global warming, glacial lakes and cloud burst events in Garhwal-Kumaon Himalaya: a hypothetical analysis. International Journal of Environmental Science and Technology, 5(4), 697–708.
[4] Dash, S.K., Kulkarni, M.A., Mohanty, U.C. and Prasad, K. (2009) Changes in the characteristics of rain events in India. Journal of Geophysical Research, 114, D10109.
[5] Deshpande, N.R., Kulkarni, A. and Krishna Kumar, K. (2012) Characteristic features of hourly rainfall in India. International Journal of Climatology, 32, 1730–1744.
[6] Dimri, A.P., Chevuturi, A., Niyogi, D., Thayyen, R.J., Ray, K., Tripathi, S.N., Pandey, A.K. and Mohanty, U.C. (2017) Cloudbursts in Indian Himalayas: a review. Earth-Science Reviews, 168, 1–23.
[7] Joseph, S., Sahai, A.K., Sharmila, S., Abhilash, S., Borah, N., Chattopadhyay, R., Pillai, P.A., Rajeevan, M. and Kumar, A. (2015) North Indian heavy rainfall event during June 2013: diagnostics and extended range prediction. Climate Dynamics, 44, 2049–2065.
[8] Khaladkar, R.M., Mahajan Kothawale, D.R. and Rajeevan, M. (2017) monthly, seasonal and annual rainfall time series for all-India, homogeneous regions and meteorological subdivisions: 1871–2016 (Research Report No. RR-138). Chennai: IITM.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

[9] Medina, S., Houze, R.A., Jr., Kumar, A. and Niyogi, D. (2010) summer monsoon convection in the Himalayan region: terrain and land cover effects. Quarterly Journal of the Royal Meteorological Society, 136, 593–616.

[10] Deshpande, N. R., Kothawale, D. R., Kumar, V., & Kulkarni, J. R. (2018). Statistical characteristics of cloud burst and mini-cloud burst events during monsoon season in India. International Journal of Climatology, 38(11), 4172-4188.

[11] Clemente-Castelló, F. J., Nicolae, B., Katrinis, K., Rafique, M. M., Mayo, R., Fernández, J. C., & Loreti, D. (2015, December). Enabling big data analytics in the hybrid cloud using iterative mapreduce. In 2015 IEEE/ACM 8th International Conference on Utility and Cloud Computing (UCC) (pp. 290-299). IEEE.

[12] Xu, L., Wang, S., & Tang, R. (2019). Probabilistic load forecasting for buildings considering weather forecasting uncertainty and uncertain peak load. Applied Energy, 237, 180-195.

[13] Scher, S., & Messori, G. (2018). Predicting weather forecast uncertainty with machine learning. Quarterly Journal of the Royal Meteorological Society, 144(717), 2830-2841.