**RAINFALL PREDICTION USING MACHINE LEARNING**

**B. Priyanka1, G. Jhansi2, C. Pavani3, K. Lakshmi4, Ch. SriLakshmi prasanna5**

1BTech, CSE, KV Subba Reddy Engineering College, Kurnool, AP, India

2BTech, CSE, KV Subba Reddy Engineering College, Kurnool, AP, India

3BTech, CSE, KV Subba Reddy Engineering College, Kurnool, AP, India

4BTech, CSE, KV Subba Reddy Engineering College, Kurnool, AP, India

5MTech, CSE, KV Subba Reddy Engineering College, Kurnool, AP, India

**ABSTRACT**

India Meteorological Department has implemented state level medium range rainfall forecast system applying multi model ensemble technique, making use of model outputs of state-of-the-art global models from the five leading global NWP centers. The pre-assigned grid point weights on the basis of anomaly correlation coefficients (CC) between the observed values and forecast values are determined for each constituent model utilizing two season datasets and the multi model ensemble forecasts are generated at the same resolution on a real-time basis. The ensemble forecast fields are then used to prepare forecasts for each state, taking the average value of all grid points falling in a particular district. In this paper, we describe the development strategy of the technique and performance skill of the system during 15 years of rain fall at different states in india. The study demonstrates the potential of the system for predicting future rainfall forecasts for upcoming years and scale over Indian region. District wise performance of the ensemble rainfall forecast reveals that the technique, in general, is capable of providing reasonably good forecast skill over most states of the country, particularly over the states where the monsoon systems are more dominant.

**Keywords:** Multi-Model Ensemble (MME) Forecasting, Rainfall Prediction, Anomaly Correlation Coefficient (CC), State-Level Forecasting, Indian Monsoon Systems

**1.INTRODUCTION**

There has been long demand from the user community for district level quantitative weather forecasts in short to medium range time scale. The quantitative rainfall forecast for smaller spatial distribution such as district level over highly complex inhomogeneous region like India is a very challenging task. For the generation of district level quantitative rainfall forecasts, one has to depend on the forecasts from dynamical Numerical Weather Prediction (NWP) models. During the last two decades, NWP methods have acquired greater skills and are playing an increasingly important role in the operational weather forecasting. But rainfall prediction skill of NWP models is still not adequate to satisfactorily address detailed aspects of Indian summer monsoon. This is because of large spatial and temporal variability of rainfall and some inherent limitations of NWP models. There are various factors like topography, prevailing synoptic situation and its interaction with mesoscale systems, lack of observations, etc., are some of the key factors which pose difficulties for numerical weather prediction of any region, and so Indian region is not an exception. Considering the need of farming sector, India Meteorological Department (IMD) has upgraded the Agro-Meteorological Advisory Service from agro climate zone to district level. As a major step, IMD started issuing district level weather forecasts from 1 June 2008 for meteorological parameters such as rainfall, maximum and minimum temperature, relative humidity, surface wind and cloud octa up to 5 days in quantitative terms (Roy Bhowmik et al 2009). These forecasts are generated through multi-model ensemble (MME) system making use of model outputs of state-ofthe-art three global models from the leading global NWP centres. These forecasts are made available on the national website of IMD (www.imd.gov.in). During summer monsoon 2009, the number of ensemble members is increased from three to five. In the present study, we describe the development strategy of the MME technique, used for high resolution rainfall forecasts over Indian region and demonstrate the prediction skill of the technique during summer monsoon 2009. In our previous study (Roy Bhowmik and Durai 2008, 2010), performance skill of MME at 50 km horizontal resolution for district level short range rainfall forecasts during summer monsoon 2007 was demonstrated from the use of four coarser grid models namely (i) IMD limited area model at 75 km horizontal resolution, (ii) IMD MM5 at 45 km horizontal resolution, (iii) National Centre for Medium Range Weather Forecasting (NCMRWF) MM5 at 30 km resolution, and (iv) NCMRWF T-80 (grid space ∼156 over the tropics). At 50 km resolution, MME could cover only 250 districts. The encouraging results of the study motivated authors for further research to increase the forecast period and model resolution using improved rainfall outputs of state-of-the-art high resolution global models from leading NWP centres to meet the operational requirement of farming community.

**2. METHODOLOGY**

In the present study, we outline the strategic development and implementation of an enhanced **Multi-Model Ensemble (MME) technique** tailored for generating high-resolution rainfall forecasts over the Indian region. This system represents a significant improvement over earlier approaches by integrating advanced data processing methods, high-resolution datasets, and modern computational models. The focus of the study is not only on developing a refined forecasting framework but also on demonstrating its **predictive skill and operational efficiency** under varying climatic conditions, particularly during the crucial summer monsoon period. By incorporating higher spatial granularity and leveraging both conventional and machine learning-based analytical methods, the system aims to deliver **more localized, accurate, and actionable rainfall predictions** that can support planning and preparedness across sectors like agriculture, disaster management, and water resource planning.

### **Advantages of the Proposed System**

The proposed system introduces several **key advantages** that address the limitations of earlier forecasting models and enhance the overall performance and usability of rainfall predictions.

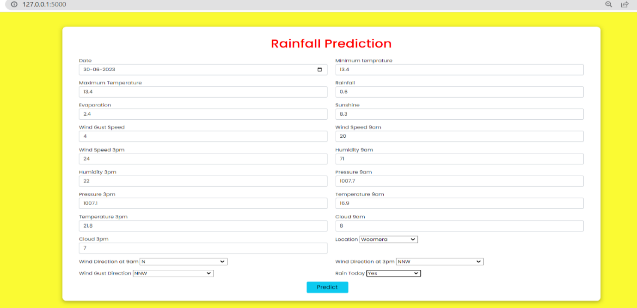
One of the major advantages is the **integration of advanced data visualizations**, which allow for dynamic observation of fluctuations, trends, and anomalies in rainfall data over time and space. Through visual tools such as heat maps, time-series plots, and interactive dashboards, users—including researchers, meteorologists, and decision-makers—can easily identify meaningful patterns in the data. These visual insights play a crucial role in guiding the design of predictive models and refining forecasting strategies, thereby improving the overall understanding of rainfall behavior across different regions.

Another significant benefit of the proposed system is its ability to **predict the amount of rainfall across diverse datasets**, including both structured historical meteorological data and real-time streaming data collected from satellite imagery, ground-based weather stations, and radar sources. This dual compatibility makes the system highly versatile and adaptable for both retrospective analysis and forward-looking forecasting. It ensures that predictions are not limited to one specific format or data source, thereby increasing the robustness and accuracy of the model outputs.

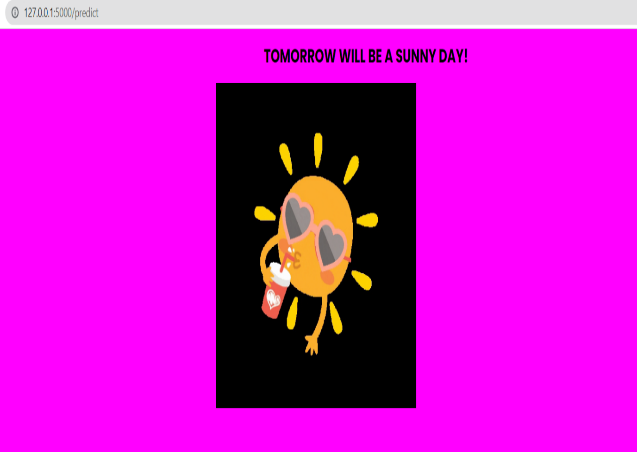
Additionally, the proposed system leverages **machine learning algorithms** that can automatically learn from large volumes of past data to detect complex relationships between atmospheric variables and rainfall occurrences. This leads to more accurate and reliable predictions, even in regions with highly variable or unpredictable rainfall patterns. The use of machine learning also enables continuous improvement, as the models can be retrained with new data to enhance their performance over time.

Furthermore, the system offers **wider geographical coverage** by utilizing high-resolution gridded data, enabling district-level or even sub-district-level forecasting. This granularity is especially beneficial for agricultural and disaster response agencies that require localized information to make timely and effective decisions.

1. **MODELING AND ANALYSIS**
2. In our previous studies conducted by Roy Bhowmik and Durai (2008, 2010), an attempt was made to evaluate the performance of a Multi-Model Ensemble (MME) forecasting system designed for district-level short-range rainfall prediction during the southwest monsoon season of 2007. The MME was configured at a horizontal resolution of 50 km and aimed to improve the accuracy and reliability of rainfall forecasts over the Indian region. This system utilized the combined outputs of four numerical weather prediction (NWP) models with varying horizontal resolutions and dynamical cores. These included: (i) the India Meteorological Department (IMD) Limited Area Model with a spatial resolution of 75 km, (ii) the IMD implementation of the MM5 (Mesoscale Model version 5) operating at a finer resolution of 45 km, (iii) the NCMRWF (National Centre for Medium Range Weather Forecasting) MM5 model at a resolution of 30 km, and (iv) the global T-80 spectral model developed by NCMRWF with a grid spacing of approximately 156 km over tropical regions.
3. The ensemble forecasting approach aimed to utilize the strengths of each model to generate improved rainfall forecasts by minimizing individual model biases and uncertainties. However, despite operating at an enhanced resolution of 50 km, the MME system could only cover approximately 250 districts across the country. Given that India has over 700 districts with diverse geographical and climatological characteristics, this limited coverage significantly constrained the broader application and operational impact of the model. Furthermore, the limited spatial resolution and regional coverage meant that fine-scale rainfall variations and local weather phenomena could not be captured accurately, which is especially critical for regions prone to localized heavy rainfall, flash floods, and agricultural dependence on timely precipitation.
4. **RESULTS AND DISCUSSION**



**Fig 1 : Home Screen**



**Fig 2 : Results**

1. **CONCLUSION**

Various visualizations of data are observed which helps in implementing the approaches for prediction.

• Prediction of amount of rainfall for both the types of dataset.

• Observations indicate machine learning models won’t work well for prediction of rainfall due to Fluctuations in rainfall.

Results of this study showed that all these models, in general, have the capability to capture large scale rainfall features of summer monsoon, such as heavy rainfall belt along the west coast, over the domain of monsoon trough and along the foothills of the Himalayas. It has clearly emerged from the results of the skill score that MME is superior to each member model. For the district level forecast, the procedure has showed appreciable skill to predict occurrence and non-occurrence of rainfall, as well as for the rainfall category of moderate rainfall. But it fails to capture heavy rainfall events. Otherwise, performance of the district level forecast for most of the districts has been fairly good, particularly over the monsoon affected states.

1. **REFERENCES**
2. Thirumalai, Chandrasegar, et al. "Heuristic prediction of rainfall using machine learning techniques." 2017 International Conference on Trends in Electronics and Informatics (ICEI). IEEE, 2017.
3. Geetha, A., and G. M. Nasira. "Data mining for meteorological applications: Decision trees for modeling rainfall prediction." 2014 IEEE International Conference on Computational Intelligence and Computing Research. IEEE, 2014
4. Parmar, Aakash, Kinjal Mistree, and Mithila Sompura. "Machine learning techniques for rainfall prediction: A review." 2017 International Conference on Innovations in information Embedded and Communication Systems. 2017.
5. Dash, Yajnaseni, Saroj K. Mishra, and Bijaya K. Panigrahi. "Rainfall prediction for the Kerala state of India using artificial intelligence approaches." Computers & Electrical Engineering 70 (2018): 66-73.
6. Dr.C K Gomathy, Article: An Effective Innovation Technology In Enhancing Teaching And Learning Of Knowledge Using Ict Methods, International Journal Of Contemporary Research In Computer Science And Technology (Ijcrcst) E-Issn: 2395-5325 Volume3, Issue 4,P.No-10-13, April ’2017.