METHODS FOR FINDING ERRORS IN PREDICTING MODES OF OSCILLATION

Anuradha M. Dhumale, Sneha Mishra
Department of Computer Science & Engineering , Ajay Kumar Garg Engineering College Ghaziabad
dhumaleanuradha@gmail.com, mishrasneha@gmail.com

Abstract— **The Indian rainfall is very important for agriculture, as it replenishes water sources and nourishes crops all over the country. The Indian monsoon is a seasonal wind pattern characterized by heavy rainfall across the Indian subcontinent. Statistical forecasting plays a pivotal role in anticipating weather patterns, especially in the context of the Indian monsoon and related atmospheric phenomena. The Indian monsoon is a very complex coupled ocean atmospheric system which exhibits variability beyond different time scale, it is very challenging task for making accurate prediction. The modes of oscillation are predicted like intra seasonal and interannual variability is crucial for concerning the behavior of Indian monsoon and/ or allied phenomena. Using different climate models error dynamics involved in forecasting the oscillatory modes of the Indian monsoon is explored. By analyzing error propagation, model biases and unpredictability in initial condition this study talks about how expand over time and affect the accuracy of prediction. The madden -Julian Oscillation (MJO) an the El Nino- Southern Oscillation (ENSO) are the special case which influence monsoon dynamics. Through the help of Linear regression and Fast Fourier Transform and various Machine Learning algorithm the prediction of error dynamics will be processed.**

Keywords— **FFT, LSTM, Loss Function, Regression.**

I. INTRODUCTION

Statistical forecasting is pivotal in enhancing our understanding of the Indian monsoon and related phenomena. By leveraging historical data and employing sophisticated statistical models, meteorologists and researchers can gain valuable insights into the complex dynamics of the monsoon system. This understanding is crucial for predicting seasonal variations, identifying trends, and assessing the risks associated with extreme weather events such as heavy rainfall, droughts, and cyclones. Such forecasts are instrumental in facilitating early warning systems, enabling communities and authorities to take proactive measures to mitigate the impact of adverse weather conditions. Moreover, statistical forecasting aids in informing decision-making processes in sectors like water resource management, agriculture, and disaster preparedness, thus contributing significantly to sustainable development and resilience in the face of climate variability and change. [1] Non-dynamical methods offer an alternative forecasting approach, relying on historical data and statistical patterns to capture the relationships between predictor variables and the target variable. Statistical regression is a common non-dynamical method that uses linear or nonlinear regression techniques to estimate these relationships.

Rainfall forecasting using statistical methods is a valuable approach that controls historical weather data to make predictions about future precipitation patterns. Meteorologists often employ statistical models to analyze past rainfall records and identify patterns or trends that can help predict upcoming weather conditions. These models may include techniques such as time-series analysis, regression analysis, and machine-learning algorithms to capture the relationships between various meteorological variables and rainfall. By statistically analyzing the historical data, forecasters can identify factors that influence rainfall and develop models capable of making predictions based on current atmospheric conditions.

The two approaches have been used previously to predict the error dynamics in the statistical method of the Indian monsoon. The first one is dynamical and the second one is the non-dynamical approach. In the first approach, prediction is done by the computer simulation. The second method used for the error dynamics is the non-dynamical approach. These are nothing but the statistical method. In this statistical method, all the predictions are made based on the previous data set pattern without any evaluation law. (FS) farmer-sidorowich [2] is the technique of this category. (ANN) Artificial neural networks are non-linear techniques that are used for prediction.[2]

Prediction using dynamical law growth rate of error depend on the state of the system. The error grow rapidly in some region of the state space. In some region the errors rate will be decrease. Prediction is not dependent on the state of the system in non-dynamical technique which is not used dynamic evaluation law. Prediction is done in non-dynamical technique with the history of the system weather evaluation law is known or not.

In terms of actual and observed time series in 1999 murphy compared the dynamical and statistical method with the co-relation. The 976 European station observed during 1983 to 1994. In his research he is explore that both statistical and dynamical method have same level of prediction. Statistical method gives the better result for the summertime prediction and dynamical method gives the better result for the winter sess.

II. MOTIVATION

Studying error dynamics in statistical forecasting of the Indian monsoon and allied phenomena can be highly motivating for several reasons:

1. Nonlinear and Chaotic in Nature: The Indian monsoon is known for its nonlinear and chaotic behaviour. The intricate interaction of various atmospheric and oceanic factor make it challenging to develop precise predictive models. The presence of nonlinearities can lead to unexpected errors, especially when forecasting over extended periods.
2. Spatial and Temporal Variability: spatial and temporal variability in rainfall patterns experiences by the Indian subcontinent. In different region climate behaviour are different. And monsoon onset and withdrawal timing can vary widely. Capturing this variability accurately is considered as challenge in forecasting model.
3. Topographical Influences: The diverse topography of India, including mountain ranges and coastal areas, adds complexity to the monsoon dynamics. Localized effects, such as orographic lifting and land-sea interactions, can significantly impact rainfall patterns. Incorporating these influences into forecasting models requires sophisticated spatial modelling techniques.
4. Influence of Monsoon Dynamics on Allied Phenomena: Tropical temperature wind and flood this are the various allied phenomena which are interconnected through Indian Monsoon. Comprehensive understanding of the complex interaction between atmospheric and oceanic processes and the monsoon are require for predicting these above associated events.
5. Limited Data Availability and Quality: For training and validating forecasting model historical data is essential. In india there may be limitation in quality and availability of meteorological data. In data -sparse region this scarcity can affect the accuracy of forecasting models.
6. Extreme Events and Abrupt Changes: The Indian monsoon is prone to extreme events, including heavy rainfall, droughts, and sudden changes in weather patterns. Forecasting such extreme events is challenging, and errors in predicting these abrupt changes can have significant socio-economic impacts.

III. METHODOLOGY

A. Simple linear regression:

For statistical forecasting simple linear regression plays an important role for solving the challenges. Historic data set is used to take climate value or parameter. Simple linear regression is a statistical method used to model the relationship between two variable by fitting a linear equation to observed the data. Based on the independent variable value of dependent variable is help for prediction. To understand the mapping between the independent variable and dependent variable by fitting a straight line to the data this is the goal of the linear regression. The linear equation is the heart of the linear regression. The linear regression is represent as follows:

$$Y=\beta 0+\beta 1X+\epsilon \quad (1)$$

Y is the value we want to predict that is dependent variable. X is the predictor that is independent variable. β0 is the value of Y when X is 0 the intercept. β1 is how much Y changes when X increases by one unit the slope. ϵ represents the difference between the observed and predicted Y the error term. In the context of predicting rainfall, a simple linear regression model is established with the dependent variable being the amount of rainfall and the independent variables encompassing factors like temperature, humidity, and wind speed. The algorithm seeks to identify a linear relationship between these independent variables and the target variable, allowing for the prediction of rainfall amounts based on observed conditions.

Historical data, consisting of past instances of rainfall and corresponding values of the independent variables, is crucial for training the model.

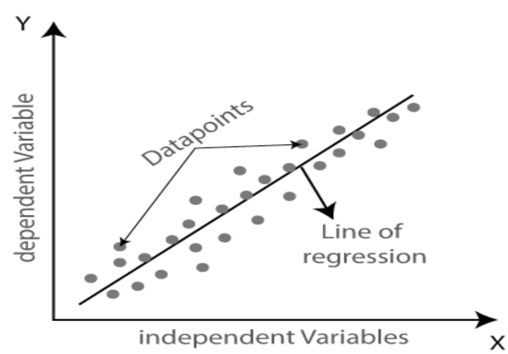

Fig 1: Linear Regression showing dependent variable and an independent variable.

B. Error Calculation:

We decided to minimize the total sum of the squared residual which is the sum of the squared difference between the observed and predicate value. This term is called the sum of squared error.

$$SSE=\sum_{i=1}^{n}(y_i-\widehat{y_i})^2 \quad (2)$$

We can find optimal value for β0 (intercept) and β1 (slope)by minimizing this quantity. The least squares method is used to minimize the SSE and find the best-fit line.

The average of the squared residual is mean squared error (MSE).

$$\text{MSE}= = \frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y_i})^2 \tag{3}$$

In which no of data point is shown by n, number of residual is represented by $(y_i - \widehat{y_i})^2$ the square root of the MSE is the root mean squared error (RMSE) and is often used because it has the same units as the dependent variable:

$$\text{RMSE}= \sqrt{\sum_{i=1}^{n} \frac{(y_i - \widehat{y_i})^2}{n}} \tag{4}$$

RMSE[15] gives a sense of how large the residuals are on average, in terms of the dependent variable's units. This all equations are used to calculate the error dynamics between statistical method.

C. Fast Fourier Transform:

The Fast Fourier Transform (FFT)[16] is a computational algorithm used to efficiently transform a time-domain signal into its frequency-domain representation. It plays a crucial role in various fields, including signal processing, communications, and scientific research. The primary advantage of FFT is its ability to rapidly calculate the discrete Fourier transform, revealing the frequency components that make up a given signal. By decomposing a signal into its constituent frequencies, FFT enables the identification of periodic patterns, harmonics, and hidden structures within the data. At different frequencies it decomposes a function (such as forecast errors) into a sum of sinusoidal components. The Fast Fourier Transform (FFT) is an efficient algorithm to compute the discrete Fourier transform (DFT) of a sequence. For a sequence x(t) of length N, the DFT is given by:

$$\text{X(f)} = \sum_{t=0}^{N-1} x(t) e^{i\frac{2\pi}{N}f t} \tag{5}$$

In which, x(t) is the time-domain signal (e.g., forecast errors). X(f) is the transformed signal in the frequency domain. F represents the frequency component. N is the number of time points in the data.t is the time index, and i is the imaginary unit.

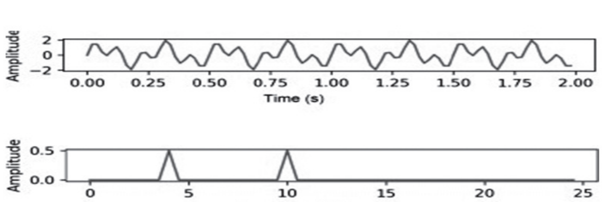

Fig 2. Signal representation in Time-domain (Top) and Frequency-domain (Bottom)

D. Error dynamics in statistical forecasting:

In statistical forecasting in Indian rain fall errors often arise s from inadequate handling of seasonality, model misspecification, stochastic nature of weather patterns, we can analyse frequency domain characteristic of this errors using FFT if the forecasting model produces errors e(t) over time Intra-seasonal oscillations[17] (like the Madden-Julian Oscillation, which affects monsoon activity).Seasonal cycles (such as the annual monsoon onset and withdrawal).Long-term trends (e.g., multi-decadal oscillations like the Indian Ocean Dipole).

The output of the FFT, Xk, is complex and can be represented in terms of magnitude and phase
Magnitude | Xk | : Indicates the strength (or power) of the error signal at frequency k

$$|X\_K| = \sqrt{(X_K)^2 + Im(X_K)^2)} \tag{6}$$

Where Re(X_k)and Im(X_K) are the real and imaginary parts of X_K, respectively.

Phase θk: Indicates the phase shift of the error signal at frequency k

$$\theta k = \text{atan2}(\text{Im}(X\_k),\text{Re}(X\_k)) \tag{7}$$

E. Steps for analysing error dynamics with FFT

1. Obtain the forecast errors : the difference between the observed an forecast values that the errors e(T) at time t
2. Preprocess Data: long term trends will be remove from the error data that might mask cyclical patterns. Use linear regression to detrend the data. The statistical properties of the error series are constant over time like mean variance. Otherwise difference the data to achieve stationarity.
3. Perform FFT: by applying FFT the series of error converts from time series to frequency component. These frequency represents cyclical error patterns.
4. Analyse the frequency spectrum: the power spectrum is computed to analyse how much variance is related with each frequency.
5. Interpret the result: Frequencies are related to physical phenomena like ENSO, MJO, Monsoon onset , retreat.
6. Improve the forecast model: Based on the dominant frequency model will be refined.
7. Visualize the result: To illustrate the error dynamics plot time series and power spectrum.
8. Cross validate: Test findings across different periods and models for validations.

III. CONCLUSION

In this Articlewe discuss about the error dynamics of statistical method of Indian monsoon and various types of methodology like SVM, Random forest [20], linear regression and FFT are used for predicting error dynamics. After reviewing the various method the result shows that the statistical method for predicting error gives the better accuracy than the dynamical method and in the high dynamical error growth rate region. SVR is a powerful approach for modelling error dynamics in Indian monsoon. Through the correction mechanism forecast accuracy can be improved by training SVR on historical errors. SVR is highly adaptable to the complex, multivariate nature of the monsoon system due to its linear and non-linear relationship. In our proposed work we proposed a model for predicting or forecasting error dynamics for modes of oscillation Indian monsoon with statistical forecasting and our proposed model had a good accuracy we will used the linear regression and fast Fourier transform for predicting error dynamics.

REFERENCES

[1] A. K. Mittal, U. P. Singh, A. Tiwari, S. Dwivedi, M. K. Joshi, K. C. Tripathi, “Short-term predictions by statistical methods in regions of varying dynamical error growth in a chaotic system” Springer-Verlag Wien 2015

[2] J. D. Farmer, J.J. Sidorowich, “Predicting chaotic time series,” Phys Rev Lett 59:845–848,1987.

[3] J. Murphy, “An evaluation of statistical and dynamical techniques for downscaling local climate,” J Clim 12:2256–2284,1999.

[4] R. Bhatla, M. Pant, S. Ghosh, S. Verma, N. Pandey, S. BIST, “Variations in Indian summer monsoon rainfall patterns in changing climate,” MAUSAM, 74, 639-650, July 2023.

[5] I. Gouirand, R. Mark, Jury, B. Sing, “An Analysis of Low- and High-Frequency Summer Climate Variability around the Caribbean Antilles,” 3942 Journal of Climate Volume 25

[6] S. Raj, S. Tripathi, K. C. Tripathi, “ArDHO-deep RNN: autoregressive deer hunting optimization based deep recurrent neural network in investigating atmospheric and oceanic parameters,” Dec 2021

[7] B. Singh, K.C. Tripathi, Y.D.S. Arya, “Deep learning analysis on impact of polar ice effects on Indian climate,” 2023

[8] B. Singh, Y.D.S. Arya, K.C. Tripathi “Deep Learning Analysis of Impact of Arctic Sea Extent Over Indian Ocean Sea Surface Temperature,” International Journal of Intelligent Engineering & System 2023.

[9] S. Garai, R. Kumar Paul, M. D. Yeasin, H. S. Roy, A. K. Paul, “Machine learning algorithms for predicting rainfall in India,”, in Current Science · February 2024.

[10] E. Sreehari, G. S. Pradeep, “Climate Changes Prediction Using Simple Linear Regression” Ghantasala Journal of Computational and Theoretical Nanoscience Vol. 16, 1–4, 2019.

[11] J. Refonaa, M. Lakshmi, Raza Abbas, Mohammad Raziullha, “Rainfall Prediction using Regression Model,” International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8, Issue-2S3. 2019.

[12] Ismail, Z., et al, “Forecasting gold pieces using multiple linear regression method,” American Journal of Applied Sciences, 2009.

[13] S. Behera, and T. Yamagata, “Imprint of the El Nin˜o Modoki on decadal sea level changes,” Geophys. Res. Lett., 37, L23702, doi:10.1029/2010GL045936, 2010.

[14] J. Refonaa, M. Lakshmi, Raza Abbas, Mohammad Raziullha, “Rainfall Prediction using Regression Model,” International Journal of Recent Technology and Engineering (IJRTE)ISSN: 2277-3878, Volume-8, Issue-2S3, July 2019.

[15] Q. Liu, Y. Zou, X. Liu, N. Linge, “A survey on rainfall forecasting using artificial neural network,” International Journal of Embedded Systems 11 (2), 240–248, 2019.

[16] K.C. Miao, T. Han, Y. Q. Yao, H. Lu, P. Chen, B. Wang, & J. Zhang, “Application of LSTM for short term fog forecasting based on,” 2020.

[17] M. Qiu, P. Zhao, K. Zhang, J. Huang, X. Shi, X. Wang, W. Chu, “A short-term rainfall prediction model using multi-task convolutional neural networks,” In: 2017 IEEE International Conference on Data Mining (ICDM), IEEE, Piscataway, NJ, USA, pp. 395–404, 2017.

ABOUT THE AUTHORS

Ms. Anuradha M Dhumale is Assistant professor at Ajay Kumar Garg Engineering College Ghaziabad. And perusing Ph.D in Computer Science and Engineering from Noida International University, Greater Noida. She received M.Tech degree from Dr, APJ Abdul kalam Technical university and bachelor’s degree in information and Technology from Rashtrasant Tukdoji Maharaj Nagpur university in 2013. Her Research Interest is in Weather Forecasting with Machine Learning.

Dr. Sneha Mishra, Assistant Professor, AKGEC, Ghaziabad and Ph.D from Galgotias University, has done her Bachelor’s degree in Computer Science and Engineering from Rishi Institute Of Engineering And Technology affiliated from U.P.T.U, in 2012 and her Master’s degree from Amity University, Noida, U.P. in 2015. Her current research interests include image processing, neural networks and Computer Vision. She has various research papers in national and international journals.