Enhancing Extreme Rainfall Forecasting Through Machine Learning Techniques

Ashish Alone[1,*], Ankur Srivastava[1], Maheswar Pradhan[1]

[1]Indian Institute of Tropical Meteorology, Ministry of Earth Sciences, Dr. Homi Bhabha Road, Pashan, Pune, Maharashtra-411008, India

Email[*]: ashish.alone@tropmet.res.in

Introduction

Extreme rainfall events during the Indian Summer Monsoon (ISM) are critical phenomena that significantly influence the socio-economic and environmental landscape of Central India. The ISM occurring primarily between June and September is characterized by heavy and widespread rainfall that sustains agriculture, replenishes water resources, and supports biodiversity. However, extreme rainfall events, defined as unusually intense and short-duration rainstorms, have become more frequent and unpredictable in recent decades. Central India, a region highly vulnerable to such events, often experiences devastating impacts, including urban flooding, loss of agricultural productivity, infrastructure damage, and displacement of communities. These extreme events are driven by complex interactions between atmospheric dynamics, regional climate patterns and global climate change, making their prediction and mitigation particularly challenging. Understanding and forecasting these events are vital for disaster preparedness, resource management, and the development of climate-resilient infrastructure in this densely populated and agriculturally reliant region.

The Arabian Sea plays a pivotal role in modulating the Indian Summer Monsoon, as it serves as a key source of moisture for rainfall over the Indian subcontinent. Study shows a synergistic relationship between SST anomalies and wind speed over the Arabian Sea. Warmer SST often coincide with strengthened westerly winds, amplifying the potential for extreme rainfall events. On the other hand, disruptions in this relationship, such as weakened winds despite warmer SSTs, can lead to anomalous rainfall patterns. This complex interaction underscores the importance of including SST and wind speed as key parameters in monsoon forecasting models to improve the prediction of extreme rainfall events over Central India.

Model and Data

The frequency of widespread extreme rainfall events (Rainfall > 150mm in at least 10 grid points) over Central India (76°E-86°E and 19°N-26°N) during the Indian Summer Monsoon has been analyzed for the period 1951-2022. To understand the atmospheric and oceanic drivers of these events, a composite evolution of sea surface temperature (SST) anomalies and wind data was carried out for the period 1982-2020, spanning -20 lag to +20 lead days time, with day 0 corresponding to the extreme rainfall event. Based on the composite evolution, the Arabian Sea was divided into three distinct domains: Domain-1 (05°S-10°N, 40°E-60°E), Domain-2 (10°N-20°N, 50°E-75°E), and Domain-3 (20°N-26°N, 60°E-70°E). Comparative plots between extreme rainfall event means and SST anomalies for these regions reveal that Domain-1 and Domain-2 exhibit negative SST anomaly correlations, while Domain-3 shows a positive correlation. Similarly, for wind data, Domain-2 exhibits a strong positive correlation, whereas Domains-1 and 3 show moderate correlations (As shown in Fig. 1). These analyses highlight the spatial heterogeneity of SST and wind influences on extreme rainfall events over Central India.

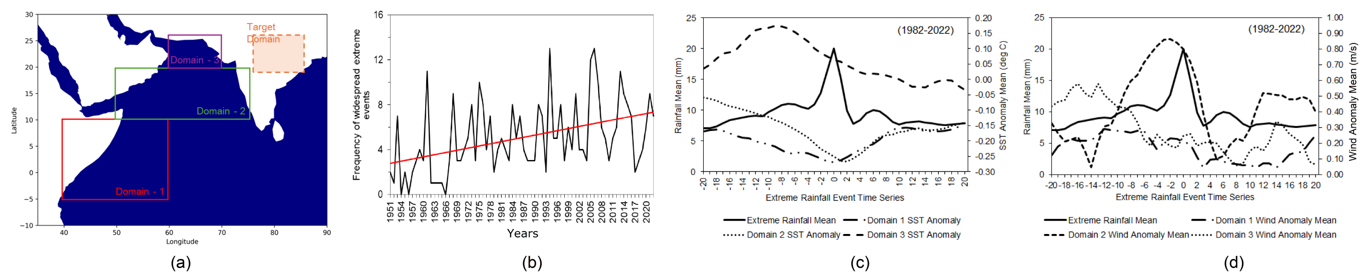

Figure 1: (a) Target domain and variable domains selected (b) Frequency of extreme rainfall events (1951-2022) (c) Extreme events rainfall mean and corresponding SST anomaly mean for domains (d) Extreme events rainfall mean and corresponding Wind anomaly mean for domains

The study utilizes a modified Random Forest Regression (RFR) model to predict extreme rainfall events over Central India. The model is trained and tested on data spanning from 2000 to 2020, with a 75:25 split ratio for