Synopsis of Proposed Study Submitted for the Degree of

Doctor of Philosophy in Geography

Title of the study

ARIMA BASED TREND PREDICTION AND MODELLING OF RAINFALL AND TEMPERATURE- A SPATIO-TEMPORAL ANALYSIS OF GUJARAT, INDIA

Submitted By
Lakhan Jain
Ph.D. Registration No. FOS/2191
Ph.D. Registration Date: 19/09/2020

Under the Supervision of
Prof. Bindu Bhatt

सत्यं शिवं सुन्दरम्

Department of Geography,
Faculty of Science,
The Maharaja Sayajirao University of Baroda,
Vadodara-390002, Gujarat
March, 2024

1. Background

The Global climate has constantly been changing, but its speed has never been as intense as today. Climate change is widely recognized as the foremost challenge of the twenty-first century, with ongoing occurrences leading to fluctuations in hydrological, climatological, and meteorological conditions worldwide and is governed predominantly by changes in global temperature and rainfall. The IPCC's AR4 (IPCC, 2007), revealed that the global surface temperature increased by approximately 0.74 ± 0.18 °C over the 20th century. This warming trend was not consistent throughout the century, with a notable acceleration in the latter half. Furthermore, the AR5, 2014 confirmed a continuation of this trend, indicating a global average surface temperature increase of about 0.85°C from 1880 to 2012. The worldwide mean temperature in 2022 is estimated to be 1.15 ± 0.13 °C warmer than the 1850-1900 average (WMO, 2022). The eight years from 2015 to 2022 are likely to be the eight warmest years on record. Projections outlined by the IPCC suggest a further increase of 1.8°C by the year 2100.

As global temperatures increase, the frequency and severity of extreme weather events increases. This direct correlation between increased temperatures and intensified extreme events underscores the urgent need for comprehensive strategies to mitigate further temperature increase and its impacts. Excessive temperature, rainfall fluctuations, and an increase in extreme climate events can lead to severe consequences for society. International agreements and conferences have been established over the years to raise awareness about climate change and implement strategies to mitigate its effects (Connors et al., 2022). Rainfall and air surface temperature are commonly used indicators of global climate change due to the availability of long-term data worldwide. Even small-scale variations in rainfall, including duration, distribution, and volume, can significantly impact society (IPCC, 2007; Pai et al., 2014). Therefore, understanding rainfall variability, identifying its underlying mechanisms, and predicting it have been key focuses of climate research. The predicting changes in rainfall and temperature patterns and the frequency of extreme events, attributed to global warming, underscore the importance of this research.

2. Rationale

This study's rationale is grounded in the urgent need to address the escalating impacts of climate change, underscored by alarming trends observed globally and particularly in South Asia. The IPCC's 2021 report reveals a significant increase in global surface temperatures, with an increase of 1.09°C between 2011 and 2020 compared to pre-industrial levels. Each incremental

0.5°C of warming amplifies the frequency and intensity of extreme weather events such as hot extremes, extreme rainfall, and droughts. These changes have profound implications for various sectors, including water resources, the environment, terrestrial ecosystems, oceans, biodiversity, agriculture, and food security, as highlighted by (Praveen et al., 2020). South Asian countries are particularly vulnerable to these impacts, with extreme events like heatwaves, floods, and droughts affecting over half of the region's population annually, imposing substantial economic burdens. Climate change has also accelerated the water cycle, resulting in erratic rainfall patterns and intensified monsoon seasons, with up to 80% of annual rainfall occurring during the monsoon period. However, projections by (Loo et al., 2015; Schewe & Levermann, 2012) suggest that rising temperatures will lead to further variations and shifts in monsoon rainfall, potentially deviating by up to 70% from the average. Moreover, research by (Ashfaq et al., 2009) indicates possible delays of up to 15 days in the onset of the monsoon across Southeast Asia. These temporal and spatial variations in the SWM have critical implications for water resources, power production, agriculture, economy, and ecosystems, emphasizing the pressing need for comprehensive research and sustainable planning (Attri & Tyagi, 2010).

In the specific context of Gujarat, erratic rainfall patterns, particularly attributed to the uneven nature of the SWM, pose significant challenges, leading to recurrent droughts and water scarcity in various regions. These fluctuations in rainfall directly impact river surface flows, thereby affecting domestic and agricultural water availability and, consequently, adversely affecting crop productivity. To address these complex challenges, this study utilizes long-term, high-resolution data and applies ARIMA models for forecasting, leveraging their ability to transform non-stationary time series into stationary ones and extrapolate future values.

Research Area Overview

The study area of Gujarat, situated along the western coast of India, encompasses diverse geographical and climatic features. Spanning between 20°6' and 24°42'N latitude and 68°10' and 74°28'E longitude, Gujarat boasts the longest coastline in India, stretching 1,663 kilometers from Kutch in the north to Valsad in the south. Despite its coastal location adjacent to the Arabian Sea, Gujarat experiences predominantly dry arid to semi-arid climates across most regions, influenced by its proximity to the Thar Desert in the north. The state is geographically partitioned into five distinct physiographic regions: Saurashtra along the southwest coast bordering the Arabian Sea, North Gujarat, Central Gujarat, South Gujarat, and Kutch. Gujarat's climatic conditions vary across its different regions. The southern part, situated below the river Narmada, exhibits subtropical conditions with sub-humid climates. Central Gujarat, between the Narmada and Sabarmati rivers, experiences moderately humid conditions. Along the coastal areas of southern Saurashtra, the climate tends to be humid and hot. In contrast, parts of central Gujarat, north of Ahmedabad, and central Saurashtra, witness dry conditions. The northern regions of Gujarat and Kutch predominantly face arid and semi-arid climates. This diverse climatic and geographical landscape offers a unique opportunity for interdisciplinary research. The distribution of SWM rainfall exhibits significant spatial heterogeneity, with the highest rainfall occurring in the south, gradually decreasing towards the northwest, resulting in varying levels of agricultural productivity and water availability across the state.

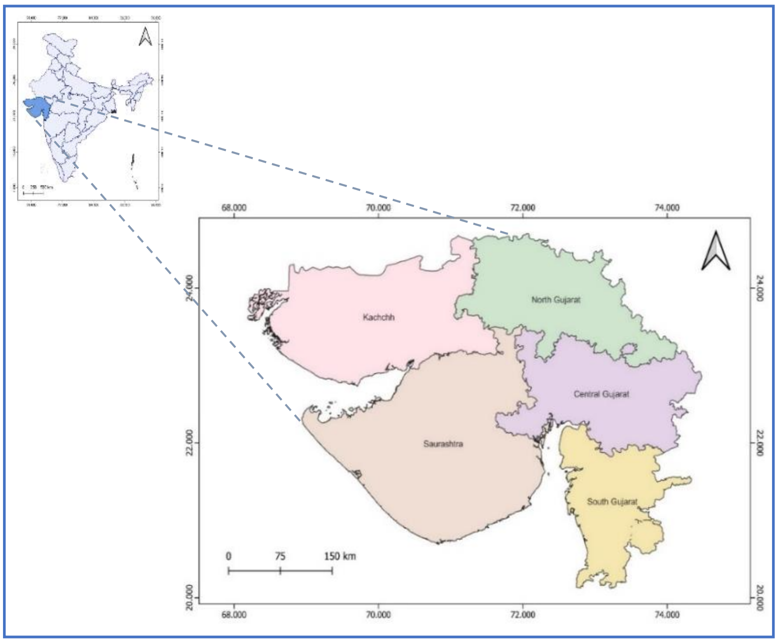

Figure 1. Location of Gujarat and its Physiographic Divisions.

3. Objectives of the Study

- To analyze the spatial and temporal patterns of rainfall and temperature for the state.
- To Identify spatial and temporal trends of temperature extremes in maximum and minimum temperatures and rainfall.
- To predict the future spatial and temporal rainfall and temperature using (ARIMA) Autoregressive integrated moving average model, the SARIMA.
- To propose water deficit and surplus regions on the basis of the variability model.

4. Database and Methodology

4.1 Data Source

This study relied on both primary and secondary data. Primary data was collected to validate forecast results through ground truth checking of variables. While, secondary data were collected from IMDLIB (India Meteorological Department Library), an open-source library developed for accessing, processing, and conducting spatiotemporal analysis of gridded meteorological observation datasets within the geographical scope of India (Nandi et al., 2024).

a) Rainfall

Long-term daily gridded rainfall data from 1961 to 2020 has been collected from the IMD Library. The study is based on a newly developed high spatial resolution IMD4 dataset that was interpolated at fixed spatial grid points of 0.25° x 0.25° resolution India-wide (Pai et al., 2014).

b) Temperature

Long-term daily gridded high-resolution temperature (minimum and maximum) data at 1° x 1° (Srivastava et al., 2009) from 1961 to 2020 has been collected from the IMD Library. The data have been organized based on seasonal criteria outlined by the IMD, including;

Winter (JF), Pre-Monsoon (MAM), Southwest Monsoon (JJAS), Post Monsoon (OND)

The research also utilizes secondary data obtained from the Census of India, the Geological Survey of India, the Library of Maharaja Sayajirao University of Baroda, Academic Journals, and Bhukosh.

Software used: QGIS, IMD Toolbox, MS Office, SPSS, XLSTAT, and RStudio.

4.2 Methodological Approach

The following methodologies were to be adopted and carried out in three phases:

Phase I- In the initial phase, data collection, conversion, and extraction are undertaken. This includes employing an intersection algorithm to identify overlapping features within the input data and overlay layers. The data is initially gathered in a gridded file format (.grd), which must then be converted to a shapefile format (.shp) for the extraction process. Following this, the extracted data is transformed into a Comma-Separated Values (.csv) format for further segregation based on seasonal criteria outlined by IMD.

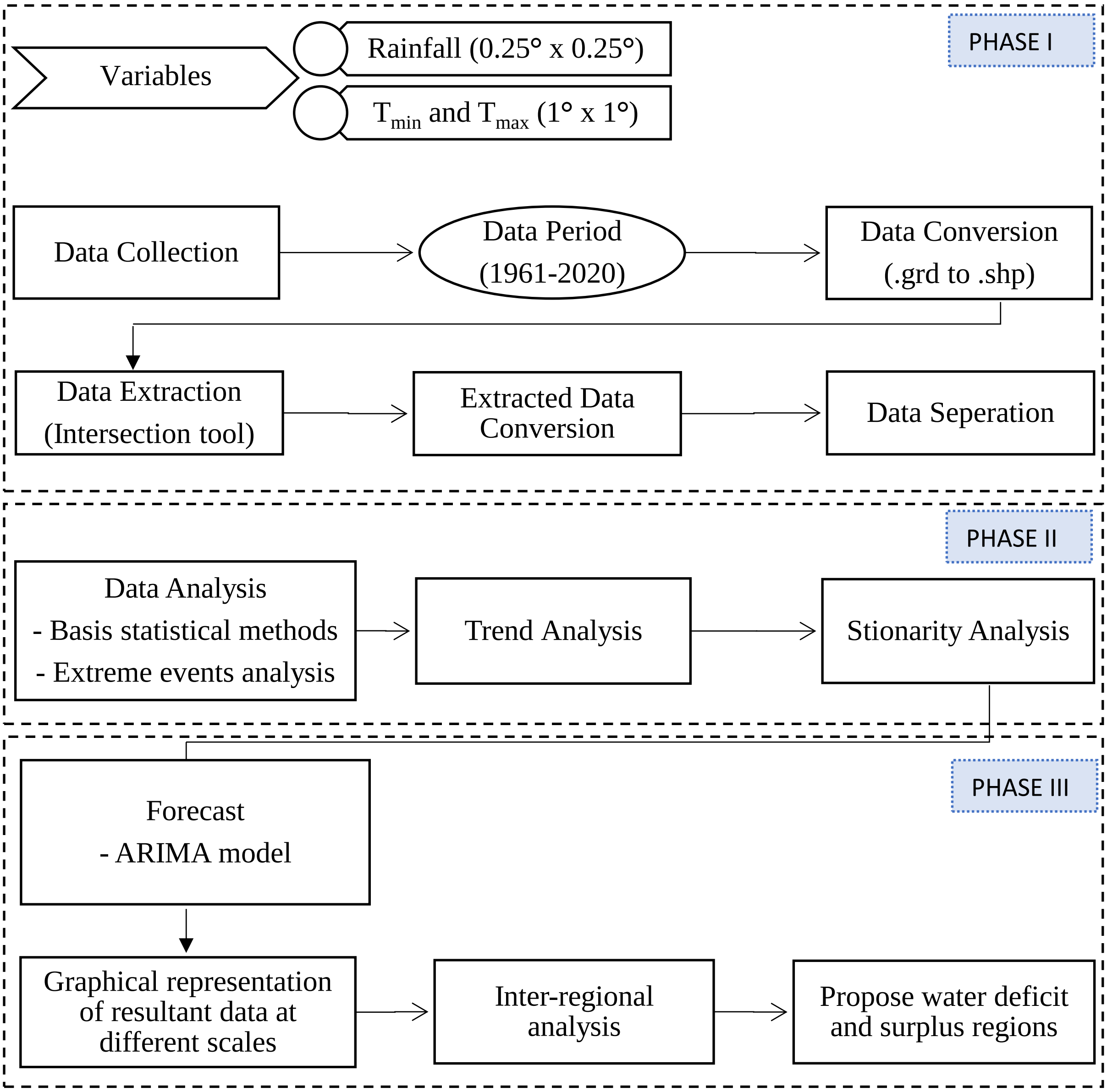

Figure 2. Flowchart of the Methodology adopted.

Phase II- This phase encompasses data filtering, sorting, visualization, executing necessary functions and formulas, and generating charts, tables, and pivot tables to enhance analysis, as well as error checking. Additionally, it involves examining trends and extreme events. Furthermore, both the Mann-Kendall (MK) (Mann, 1945; Kendall, 1957) test and Sen's slope (Sen, 1968; Theil, 1950) estimator (SSE) has been utilized to identify the trend and its magnitude. The MK (Z_{mk}, if n > 10) statistics reveal a trend (either upward or downward) present in the time series. The Z_{mk} statistic's value is determined by computing the variance of S, denoted as Var(S), and S Statistics. The contemporary approach, ITA (Şen, 2012), has been employed to yield effective outcomes and to compare them with the outcomes of the traditional method. Additionally, the time series underwent the Augmented Dickey-Fuller (ADF) (Dickey & Fuller, 1979) test to assess stationarity, a pre-requisite to the ARIMA model.

***Phase III* -** This involves forecasting, analyzing resultant values alongside historical data, and suggesting regions experiencing water deficits or surpluses. The autoregressive integrated moving average (ARIMA) model (Box, G. E. P., & Jenkins, 1976; Geurts et al., 1977) has been used to forecast future values, while Akaike's information criterion (AIC) (Akaike, 1974) has been employed to select the best-fit model. The best model is determined by the one with the lowest AIC value (Eni & Adeyeye, 2015). The forecast results facilitate a comprehensive understanding of the spatial distribution of water deficits and surpluses, thereby informing strategic planning across multiple sectors.

5. Future Scope of Study

In the future, this study could broaden its scope by improving its forecasting methods beyond ARIMA. This might involve using machine learning or combining different approaches to make predictions more accurate. Additionally, including more weather and water-related factors, could offer a more holistic understanding of water dynamics. This study presents an opportunity to conduct micro-level research at the district level. The micro-level approach enables a more nuanced understanding of the complexities involved and can inform targeted interventions and policies tailored to the unique characteristics of each district. Additionally, Integrating the study findings into decision support systems or geospatial platforms can facilitate their use by policymakers, water resource managers, and stakeholders for informed decision-making and planning. Making these tools easy to use and understand can help people from different backgrounds access and use the information.

6. Thesis organization

The study has been organized into seven chapters. All chapters highlight relevant literature.

Chapter 1: Introduction

The study comprises seven chapters, with the introductory chapter covering the background, problem statement, objectives, methodology, and significance. This chapter offers a concise overview of the study's purpose, set against the backdrop of relevant literature. It also introduces key concepts, such as changing trends in climate variables, extreme events, and the significance of future data forecasting.

Chapter 2: Profile of the Study Area

This chapter provides an in-depth exploration of Gujarat's geographical delineation, emphasizing its Physiography, Climate, vegetation, demography, and Socio-economic conditions.

Chapter 3: Literature Review

The chapter focuses on identifying research gaps and areas necessitating further investigation to understand the intricate connection between climate variables. This involves assessing the theoretical models used in the study and uncovering significant trends, drivers, and impacts associated with these changes. The chapter is organized into five subsections: "Climate Change: Global Temperature and Rainfall," "Climate Profile of India," "Extreme Events," "Trend Analysis," and "Forecast Models."

Chapter 4: Temperature, Rainfall, and PCI Trends of Gujarat: A Spatio-Temporal Analysis

The chapter involves a thorough examination and analysis of the spatiotemporal distribution of climate variables in different physiographic regions within Gujarat. It employs the Precipitation Concentration Index (PCI) and basic statistical methods. This chapter aims to unravel the intricate shifts and variations in rainfall and temperature trends that occurred in different physiographic regions over the designated timescales. The resulting SPCI indicates that overall, Gujarat experiences a moderate SPCI on a decadal scale, suggesting a fairly uniform distribution of monsoon rainfall across most regions. However, Kutch stands out as an exception, exhibiting significant irregularities in seasonal rainfall distribution during the SWM months. Over the recent decade, there has been a noticeable increase in irregularity in the seasonal rainfall distribution in different regions of Gujarat. These findings are crucial for understanding changing patterns of variables and developing adaptation strategies, serving as a foundation for subsequent chapters.

Chapter 5: Assessing Spatio-Temporal Trends of Rainfall and Temperature in Gujarat's Physiographic Regions.

The chapter focuses on assessing the trends of climate variables across distinct physiographic regions in Gujarat. Its significance lies in comparing outcomes between traditional and contemporary methods. This study aims to assess and analyze the results on both annual and seasonal scales, contributing to a more comprehensive understanding of climate variables. The findings indicate that the contemporary approach is more effective in monitoring rainfall trends, especially when dealing with highly variable data. Notably, the assessment of rainfall trends reveals significant observations, with the Kutch region experiencing an upward trend in rainfall across seasons, except in winter. Meanwhile, the South Gujarat region shows insignificant variations at different scales. Additionally, the findings highlight an increasing trend in T_{min} and T_{max} across the regions, noticeable at both monthly and seasonal scales. With a few exceptions, both traditional and contemporary approaches yield similar results for Temperature.

Chapter 6: Forecasting Variables using ARIMA: Comparative Analysis of Inter-Regional Changes in Gujarat

The chapter encompasses the forecasting of climate variables, which involves a crucial prerequisite of normalizing the data before conducting the forecast. This normalization step is essential for ensuring the accuracy and reliability of the forecasted outcomes. The ARIMA model (p, d, q) with the lowest AIC value has been chosen as the best fit. The study aids in forecasting climate variables and investigates the comparison between the long-term average of input data and the mean forecasted data. This approach facilitates a comprehensive understanding of the spatial distribution of water deficits and surpluses, thereby informing strategic planning across multiple sectors. An increase in average T_{min} has been forecasted for the next 30 years based on 60 years of observed data across various regions of Gujarat, at different scales. Similarly, with a few exceptions, average T_{max} has also increased across these regions, at varying scales. Meanwhile, the rainfall forecast exhibits high variations across regions and scales.

Chapter 7: Conclusion

Upon examining the spatiotemporal trends in climate variables, notably T_{max} and T_{min}, a notable increase is observed across various scales with a few exceptions, employing both contemporary and traditional analytical methodologies. However, the analysis of rainfall data yields insignificant results due to the high variability within the dataset at different scales, particularly evident when employing traditional trend analysis methods at different confidence intervals. Subsequently, forecasting techniques are applied to assess changes in both long-term average data and forecasted values. The resultant trend over the next three decades (till 2050) indicates an increase in average minimum temperatures, drawing from six decades of observed data spanning different physiographic regions of Gujarat. Likewise, average maximum temperatures have shown an upward trend across distinct physiographic regions, with some exceptions. However, the forecast for rainfall exhibits significant fluctuations across different regions and scales, highlighting the complexity and variability of climate patterns in the area. In light of the projected climate trends and the variability in rainfall patterns across Gujarat, effective water resource management is paramount to ensuring sustainability across various sectors such as agriculture, industry, and domestic consumption. Such findings underscore the importance of targeted interventions and adaptive strategies to mitigate the impacts of water scarcity and ensure sustainable water management in Gujarat.

Suggestions

The study's implications are far-reaching, extending beyond its immediate scope. Its findings not only contribute to the existing body of knowledge but also provide a solid foundation for further research. Policymakers must heed the call to action, prioritizing regions experiencing shifts in rainfall patterns and rising temperatures. Public awareness campaigns play a pivotal role in fostering understanding and driving sustainable practices. Additionally, a robust monitoring system is imperative to track SWM accurately, enabling proactive measures to mitigate risks effectively. Collaboration among stakeholders is essential to address the multifaceted challenges posed by climate variability and ensure a resilient and sustainable future.

References

Akaike, H. (1974). A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, *19*(6). https://doi.org/10.1109/TAC.1974.1100705

Ashfaq, M., Shi, Y., Tung, W. W., Trapp, R. J., Gao, X., Pal, J. S., & Diffenbaugh, N. S. (2009). Suppression of south Asian summer monsoon rainfall in the 21st century. *Geophysical Research Letters*, *36*(1). https://doi.org/10.1029/2008GL036500

Attri, S. D., & Tyagi, A. (2010). Climate profile of India. *Environment Meteorology, India Meteorological Department*.

Box, G. E. P., & Jenkins, G. M. (1976). Time series analysis: forecasting and control. San Francisco, CA: Holden-Day. *[University of Wisconsut. Madison. WI and University OfLancaster, England], 1970*.

Connors, S. L., Nicolai, M., Berger, S., Pidcock, R., Walsh, M., & Hawtin, N. (2022). Co-developing the IPCC frequently asked questions as an effective science communication tool. *Climatic Change*, *171*(1–2). https://doi.org/10.1007/s10584-021-03248-0

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *Journal of the American Statistical Association*, *74*(366). https://doi.org/10.2307/2286348

Eni, D., & Adeyeye, F. J. (2015). Seasonal ARIMA Modeling and Forecasting of Rainfall in Warri Town, Nigeria. *Journal of Geoscience and Environment Protection*, *03*(06). https://doi.org/10.4236/gep.2015.36015

Geurts, M., Box, G. E. P., & Jenkins, G. M. (1977). Time Series Analysis: Forecasting and Control. *Journal of Marketing Research*, *14*(2). https://doi.org/10.2307/3150485

IPCC, I. P. O. C. C. (2007). Climate Change 2007 - The Physical Science Basis: Working Group I Contribution to the Fourth Assessment Report of the IPCC. *Science*, *October 2009*. https://doi.org/volume

Kendall, M. G. (1957). Rank Correlation Methods. *Biometrika*, *44*(1/2). https://doi.org/10.2307/2333282

Loo, Y. Y., Billa, L., & Singh, A. (2015). Effect of climate change on seasonal monsoon in Asia and its impact on the variability of monsoon rainfall in Southeast Asia. *Geoscience Frontiers*, *6*(6). https://doi.org/10.1016/j.gsf.2014.02.009

Mann, H. B. (1945). Nonparametric Tests Against Trend. *Econometrica*, *13*(3). https://doi.org/10.2307/1907187

Nandi, S., Patel, P., & Swain, S. (2024). IMDLIB: An open-source library for retrieval, processing and spatiotemporal exploratory assessments of gridded meteorological observation datasets over India. *Environmental Modelling & Software*, *171*, 105869. https://doi.org/https://doi.org/10.1016/j.envsoft.2023.105869

Pai, D. S., Sridhar, L., Rajeevan, M., Sreejith, O. P., Satbhai, N. S., & Mukhopadhyay, B. (2014). Development of a new high spatial resolution (0.25° × 0.25°) long period (1901-2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region. *Mausam*, *65*(1). https://doi.org/10.54302/mausam.v65i1.851

Praveen, B., Talukdar, S., Shahfahad, Mahato, S., Mondal, J., Sharma, P., Islam, A. R. M. T., & Rahman, A. (2020). Analyzing trend and forecasting of rainfall changes in India using non-parametrical and machine learning approaches. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-67228-7

Schewe, J., & Levermann, A. (2012). A statistically predictive model for future monsoon failure in India. *Environmental Research Letters*, *7*(4). https://doi.org/10.1088/1748-9326/7/4/044023

Sen, P. K. (1968). Estimates of the Regression Coefficient Based on Kendall's Tau. *Journal of the American Statistical Association*, *63*(324). https://doi.org/10.1080/01621459.1968.10480934

Şen, Z. (2012). Innovative Trend Analysis Methodology. *Journal of Hydrologic Engineering*, *17*(9). https://doi.org/10.1061/(asce)he.1943-5584.0000556

Srivastava, A. K., Rajeevan, M., & Kshirsagar, S. R. (2009). Development of a high resolution daily gridded temperature data set (1969-2005) for the Indian region. *Atmospheric Science Letters*, *10*(4). https://doi.org/10.1002/asl.232

Theil, H. (1950). A rank-invariant method of linear and polynomial regression analysis, 3; confidence regions for the parameters of polynomial regression equations. *Indagationes Mathematicae*, *1*(2).

WMO. (2022). Provisional State of the Global Climate in 2022. *Wmo*.