Evaluation of remote sensing and reanalysis based precipitation products for agro-hydrological studies in semi-arid tropics of Tamil Nadu

Aatralarasi Saravanan[1,2], Daniel Karthe[1,2,3], Selvaprakash Ramalingam[4] and Niels Schütze[2]

[1]United Nations University – Institute for Integrated Management of Material Fluxes and of Resources (UNU-FLORES), Dresden, 01067, Germany
[2]Faculty of Environmental Sciences, Technische Universität Dresden, Dresden, 01069, Germany
[3]Faculty of Engineering, German-Mongolian Institute for Resources and Technology (GMIT), Nalaikh, Mongolia
[4]Division of Agricultural Physics, Indian Agricultural Research Institute, New Delhi, 110012, India

Correspondence to: Niels Schütze (niels.schuetze@tu-dresden.de)

Abstract. This study provides a comprehensive evaluation of eight high spatial resolution gridded precipitation products in semi-arid regions of Tamil Nadu, India, focusing specifically on Coimbatore, Madurai, Tiruchirappalli and Tuticorin, where both irrigated and rainfed agriculture is prevalent. The study regions lack sufficiently long-term and spatially representative observed precipitation data, which is essential for agro-hydrological studies and better understanding and managing the nexus between food production and water and soil management. Hence, the present study evaluates the accuracy of five remote sensing-based precipitation products viz. Tropical Rainfall Measuring Mission (TRMM), Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks – Climate Data Records (PERSIANN CDR), CPC MORPHing technique (CMORPH), Integrated Multi-Satellite Retrievals for Global Precipitation Measurement (GPM-IMERG) and Multi-Source Weighted-Ensemble Precipitation (MSWEP) and three reanalysis-based precipitation products viz. National Center for Environmental Prediction Reanalysis 2 (NCEP2), and European Centre for Medium-Range Weather Forecast (ECMWF) Reanalysis version 5 Land (ERA5-Land), Modern-Era Retrospective analysis for Research and Application version 2 (MERRA 2) against the station data. Linearly interpolated precipitation products were statistically evaluated at two spatial (grid and district-wise) and three temporal (daily, monthly and yearly) resolutions for the period 2003-2014. Based on overall statistical metrics, ERA5-Land was the best-performing precipitation product in Coimbatore, Madurai, and Tiruchirappalli, with MSWEP following closely behind. In Tuticorin, however, MSWEP outperformed the others. On the other hand, MERRA2 and NCEP2 performed the worst in all the study regions, as indicated by their higher Root Mean Square Error (RMSE) and lower correlation values. Except in Coimbatore, most of the precipitation products underestimated the monthly monsoon precipitation, which highlights the need for a better algorithm for capturing the convective precipitation events. Also, the Percent Mean Absolute Error (%MAE) was higher in non-monsoon months, indicating that these product-based agro-hydrological modelling, like irrigation scheduling for water-scarce periods, may be less reliable. The ability of precipitation products to capture the extreme precipitation intensity differed from the overall statistical metrics, where MSWEP performed the best in Coimbatore and Madurai, PERSIANN CDR in Tiruchirappalli and ERA5-Land in Tuticorin. This study offers

crucial guidance for managing water resources in agricultural areas, especially in precipitation data-scarce regions, by helping to select suitable precipitation products and bias correction methods for agro-hydrological research.

1 Introduction

Tamil Nadu, an agriculturally important state in India, is endowed with only 3 percent of the country's water resources. Of this limited supply, more than 95 percent of surface water and 80 percent of groundwater are already allocated for various uses, including human and animal consumption, irrigation and industrial activities (TN-ENVIS, 2023). The state, which is in the rain shadow region of the Western Ghats, receives an average annual precipitation of about 925 mm (IMD). Agriculture relies heavily on precipitation and provides livelihood to nearly 40 percent of the people (Government of Tamil Nadu, 2022). The gross sown area is 5.94 million ha, while the net sown area is 4.73 million ha. The total irrigated area in the state is 3 million hectares; hence, around 50% of the gross sown area depends on precipitation as its primary water resource. In recent years, the state's agriculture sector has grown around 3% annually while its overall economy has risen by 6-9% (Statistical Handbook, 2020-21). The decreased performance of the agricultural sector is collectively attributed to multiple factors like water shortages during critical stages of crop irrigation, low level of crop diversification, increasing climate change threats, the rigid mindset of the farmers in refusing to accept the improved technologies, weak market development, and high rates of post-harvest losses (Arjune and Kumar (2023), Balaganesh et al. (2020), Gardas et al. (2018), Malaiarasan et al. (2021)). Compared to other factors, climate change has differential impacts on the region's monsoons (Lal, 2003). Climate change has led to a noticeable rise in the intensity of rainfall in the months of August, October, and November, along with greater variation. For the agricultural sector, the most important implication of climate warming is an increase in irrigation water requirement, which is estimated to range between 0.6 to 3.7% for temperature increases between 0.5 to 3.0°C (Surendran et al., 2021). In addition to irrigated crops, climate-resilient rainfed crops, which include cotton, pearl millet, and sorghum, are expected to produce less by 2030 in all of Tamil Nadu's agroclimatic zones. Significant reductions are expected in the Northeast zone (6.07%), the Cauvery delta zone (3.55%), and the South zone (3.54%) (Arumugam et al., 2013). Furthermore, there has been an increase in the frequency of extreme hydrological events in the semi-arid region of the state, including the floods in Chennai in 2015 and 2022 and the hydrological droughts in 2018 caused by the failure of the northeast monsoon. Besides that, the floods that occurred in the Tuticorin district during December 2023 show that the state's extreme hydrological events are spreading to newer regions (Radhakrishnan et al. (2024), Lalmuanzuala et al. (2023), Paramasivam (2023)). IPCC's Sixth Assessment Report projects an increase in heavy precipitation events linked to tropical cyclones and worsening drought in certain regions, hence hydrological modelling will play a critical role in early warning systems. This modelling lays the foundation for the development of suitable environmental policies, which are essential given the state's current inconsistencies in warming, climate change, and rainfall impacts.

Establishing dense meteorological observatories and disseminating the state of weather variables at frequent intervals is essential for adapting to the regional and local consequences of climate change (Wilby and Yu, 2013). Developing countries

like India are more vulnerable in this regard due to inadequate ground weather stations and, thus, a lower adaptation capacity. The distribution of rain gauges in one of the largest countries, India, is inadequate to accurately predict and assess various rain-induced events/processes, like droughts, flash floods and their implications, such as dam failures (Singh et al., 2018). Despite the accurate measurement of rainfall by rain gauges, their heterogeneous distribution over land and relatively sparse measurements over large oceans limit their use. Also, gauge data are available as point observations with potentially limited spatial representativeness; inconsistencies exist between other data products, which is largely due to a limited number of ground stations, merging and interpolation methods, limited time resolution and limited documentation quality (Huffman et al., 2009; Nikulin et al., 2012., Sylla et al., 2013). The gaps and inhomogeneities in precipitation time series collected by the rain gauges at the time of severe weather conditions also make it difficult to force hydrological models that require a continuous time series of precipitation data.

The alternate solution is multi-source merged precipitation in general, including ground-based weather radars, meteorological satellites, and reanalysis products. These provide a comprehensive map of rainfall at homogeneous spatial and temporal resolutions (Tapiador et al., 2012; Kucera et al., 2013). Several satellite-based rainfall estimates have been developed over the last decade (Sapiano and Arkin, 2009; Zambrano-Bigiarini et al., 2017). For India, a list of precipitation products is available that can be used for climate change impact studies, such as the Global Precipitation Climatology Project (GPCP) from the National Centers for Environmental Information of the National Oceanic and Atmospheric Administration (NOAA) (Adler et al., 2003), CPC Merged Analysis of Precipitation (CMAP) from the National Center for Atmospheric Research (NCAR) (Xie et al., 2007), Tropical Rainfall Measuring Mission (TRMM) (Huffman et al., 2007). In addition, Global Satellite Mapping of Precipitation (GSMaP) (Ushio et al., 2009), Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks – Climate Data Records (PERSIANN CDR) (Ashouri et al., 2015), Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks – Cloud Classification System (PERSIANN CCS) (Sorooshian et al., 2000), CPC MORPHing technique(CMORPH) (Joyce et al., 2004), Global Precipitation Measurement (GPM-IMERG) (Hou et al., 2014) and Multi-Source Weighted-Ensemble Precipitation (MSWEP) (Beck et al., 2017) are also available with at varying spatial and temporal resolutions for longer periods.

As another source of climate information, climate model-derived reanalysis data are suitable for assessing climate variability and change. Global reanalysis-based climate products include National Center for Environmental Prediction Reanalysis 2 (NCEP2) (Kanamitsu et al., 2002), European Centre for Medium Range Weather Forecast (ECMWF) Reanalysis version 5 Land (ERA5-Land), 20th Century Reanalysis version 2 (20CRv2) (Compo et al., 2011), Climate Forecast Reanalysis System (CFSR) (Saha et al., 2010), Japanese 55-year Reanalysis (JRA-55) (Ebita et al., 2011), and Modern-Era Retrospective analysis for Research and Application version 2 (MERRA 2) (Gelaro et al., 2017).

Precipitation products from non-station sources must be assessed against ground station data before being imported into the Agro-hydrological models. The differential statistical assessment of the product at multiple spatial and temporal scales will highlight the accuracy of the precipitation product, quantify bias and lead to choosing the appropriate bias correction methods before forcing them into a hydrological model. Further, previous studies have shown the differential performance of

precipitation products in capturing extreme precipitation events, as well as their spatial variability. Hence, assessing these unique criteria is important to reduce the input uncertainty of hydrological modelling. Also, the precipitation data must be available for extended periods (at least 10 years) to study the impact of climate change on the hydrological cycle. The best-performing precipitation products when fed into Agro-hydrological models like the AquaCrop, Soil and Water Assessment Tool (SWAT, Neitsch et al., 2002), MODFLOW (McDonald and Harbaugh, 1984), Water evaluation and Planning (WEAP), Revised Universal Soil loss equation (RUSLE, Renard et al., 1997) can provide accurate projections with lower uncertainty. Thorough assessment of various precipitation products requires both dense station data availability and continuous time series of precipitation data for longer periods of time. Such assessment studies are not available for Tamil Nadu, particularly for the semi-arid regions on a daily timescale, which is crucial for comprehending hydrological variability in the context of climate change.

However, a limited number of studies at daily intervals are available based on the river basin spatial scale (Yashwanth et al., 2023; Kolluru et al., 2020). Further, Singh et al., (2018) analyzed only heavy rainfall occurrences in peninsular India during the 2015 winter monsoon, taking grid-level spatial scale into consideration. Other investigations were carried out either at weekly or monthly scales (Dubey et al., 2021; Singh et al., 2019).

Thus, the purpose of this study is to compare and assess, using the most widely used and recognized statistical and graphical evaluation methods, the available precipitation products for Coimbatore, Madurai, Tiruchirappalli, and Tuticorin at the highest possible spatial and temporal (daily, monthly, and yearly) resolution against station data. The results of our study will help to overcome the data scarcity in the study area with regard to the spatial and temporal resolution gaps of daily, monthly, and annual precipitation data products that can be used for hydrologic and environmental change and impact studies at a regional scale.

2 Methodology

2.1 Study regions

The study's main objective is to assess precipitation data sources for semi-arid regions of Tamil Nadu (SAT), specifically Coimbatore, Madurai, Tiruchirappalli and Tuticorin, on a daily, monthly, and yearly basis. The study regions are representative of the following agroclimatic zones and are known to contribute significantly to the state's agricultural output: the Cauvery Delta Zone (Tiruchirappalli), the Western Zone (Coimbatore), and the Southern Zone (Madurai, Tuticorin) (Rajkumar et al., 2020; Chandrasekar et al., 2009). The state of Tamil Nadu receives precipitation during North-East monsoon (44%), South-West monsoon (41.27%), Hot weather season (11.41%) and Winter season (2.7%) (Dhar et al., 1981; Lakshmi et al., 2021). The study region's precipitation pattern had inter-annual variability, and the results of the Mann Test produced non-significant p-values indicating no trend in the long term (Table 1). The land surface heterogeneity, the intricate topography, and their interactions with global climatic forcing systems are some of the local elements contributing to the inter-annual variability in the precipitation pattern in these locations. Similarly, weather and climate variability, both geographical and

temporal, as well as tendencies toward rising temperatures and falling precipitation, are experienced throughout the study's investigated locations. Additionally, research indicates that the rainfall window for the Southwest monsoon is getting shorter, while the northeast monsoon crops are vulnerable to flooding during their early stages (Varadan et al., 2017). Climate change has already started to change rainfall pattern in a way that is unfavorable for most agricultural crops, so measures for sustainable adaptation and mitigation are needed. These strategies rely on comprehensive hydrological and climatic models and the corresponding data inputs.

Table 1. General characteristics of selected study regions

Study Locations	Area (km^2)	Average area elevation (m)	Precipitation (mm)					Mann Kendall score (p-value)
			Northeast Monsoon (October to December)	Southwest Monsoon (June to September)	Winter Season (January to February)	Summer Season (March to May)	Annual	
Coimbatore	4732	427	343.8	686.8	19.6	164.8	1215.0	-5.0 (0.7)
Madurai	3710	101	418.8	325.2	23.3	146.9	914.2	-7.0 (0.5)
Tiruchirappalli	4509	74	394.2	276.6	18.1	96.7	785.6	-11.0 (0.3)
Tuticorin	4745	27 to few metres	427.7	64.7	41.4	113.9	647.7	-7.0 (0.5)

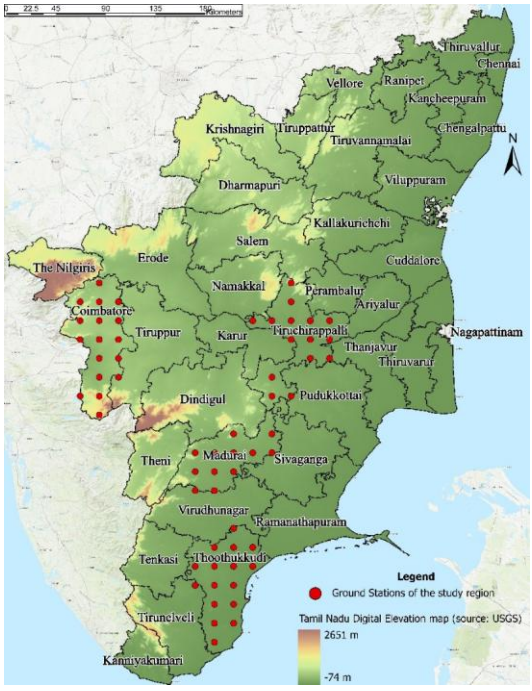


Figure 1. Map of Tamil Nadu and study regions (Coimbatore, Madurai, Tiruchirappalli, Tuticorin) with ground stations. The stations were aligned into (an almost) regular grid by the state authorities/institutions for the study period (2003 – 14), which makes the districts selected "ideal" study areas for this evaluation study.

2.2 Data sets

2.2.1 Ground stations

The points in Fig. 1 are the geographical location of the available ground station data. In most regions of India, the density and availability of field-based meteorological stations are limited, and their accessibility is restricted for many reasons. Hence, the selection of ground station data was based on the availability, quality and density of field-based meteorological stations from 2003 to 2014. The daily data was provided by the Public Works Department (PWD) of Tamil Nadu, which collects and maintains meteorological data for the entire state. Additionally, each selected study region has a state-owned agricultural university with many crop-specific research institutes, including the Wheat Research Station in Coimbatore, located at a higher altitude. These institutes provide both agro-meteorological and hydro-meteorological data, enabling

coverage even in mountainous regions. The stations were aligned into (an almost) regular grid by the state authorities/institutions, which makes the districts selected "ideal" study areas for this evaluation study. The data were carefully and extensively checked for their quality, and data for 2005 and 2010 was eliminated due to the large number of missing inputs. Instead of filling the gaps with uncertain results using imputation methods, we decided to skip the years. Finally, datasets used for validation included 17 stations in Coimbatore, 12 stations in Madurai, 16 in Tuticorin and 15 in Tiruchirappalli (Table S1a).

2.2.2 Satellite and Reanalysis-based Precipitation Products

Among the several global satellite and reanalysis-based precipitation datasets, those with the maximum time period coverage and the highest spatial resolution were chosen. We selected products with spatial resolutions between 0.1° and 0.25°. While reanalysis products typically have coarser resolutions, we included MERRA2 and NCEP2 in our study due to their unique advantages. Both products provide global coverage on a daily timescale and incorporate significant advancements such as the inclusion of cloud-top cooling (NCEP2, Kanamitsu et al., 2002) and atmospheric aerosols (MERRA2, Bosilovich et al., 2015). Daily temporal scale precipitation products were selected because ground station data was only available on a daily basis, even though several precipitation products were available at half-hourly and sub-daily temporal scales. We selected five satellite-based and three reanalysis-based rainfall estimates using the aforementioned standards (Table 2). The five satellite-based precipitation estimates fulfilling are the TRMM 3B42 (Huffman et al., 2007), PERSIANN CDR (Ashouri et al., 2015), CMORPH (Joyce et al., 2004), GPM-IMERG (Hou et al., 2014) and MSWEP (Beck et al., 2017). Also, the reanalysis products considered are MERRA-2 (Reichle et al., 2017), NCEP2 and ERA5-Land. Reanalysis products like 20CRv2, CFSR and JRA-55 are not considered in this study due to their limited spatial and temporal resolution.

Launched in late 1997, the TRMM products are among the satellite precipitation products that have been used extensively because of their acceptable accuracy (Huffman et al., 2007; Sahoo et al., 2015). The mission uses 5 instruments: Precipitation Radar (PR), TRMM Microwave Imager (TMI), Visible Infrared Scanner (VIRS), Clouds and Earths Radiant Energy System (CERES) and Lightning Imaging Sensor (LSI). The product TRMM 3B42, which is available at 0.25° spatial resolution and daily temporal resolutions, is used in this study (Huffman et al., 2023).

PERSIANN-CDR is a satellite-based precipitation estimation product that provides more than three decades (from 1983 to present) of daily precipitation estimates at 0.25°×0.25° spatial resolution for the 60°S–60°N geographical extent. PERSIANN-CDR utilises the archive of infrared brightness temperature from the gridded satellite dataset (GridSat-B1), as the input to the Artificial Neural Network based PERSIANN algorithm that converts the input infrared data in degrees Kelvin into rain rate (RR) data in mm/hour.

CMORPH was generated on an 8kmx8km grid over the global domain (60°S-60°N) in 30-min intervals from January 1, 1998 to the present. In the first step, high-resolution global satellite precipitation estimates from passive microwave (PMW) measurements along the cloud motion vectors (the raw CMORPH) were integrated. In the second step, bias in the raw

CMORPH was removed through Probability Density Function (PDF) matching against the CPC daily gauge analysis over land (Xie et al., 2010) and through adjustment against the pentad GPCP merged analysis over ocean (Xie et al., 2003).

GPM Core Observatory satellite, launched on February 27, 2014, operates in low Earth orbit, carrying passive multi-channel conical-scanning microwave radiometer called the GPM Microwave Imager (GMI) that can measure light to heavy precipitation and an active Dual-frequency Precipitation Radar (DPR) which can measure precipitation characteristics in the atmospheric column in three dimensions (Smith et al., 2007). In this study, the Level 3 IMERG Final Run product, which includes the research-quality gridded global multi-satellite precipitation estimates, was used.

MSWEP Version 2 is the first fully global precipitation dataset with a 0.1° resolution derived by optimally merging a range of gauge, satellite, and reanalysis estimates (Beck et al., 2017). Its previous version was released in May 2016. MSWEP V2 merges the highest quality precipitation information from multiple sources thereby taking advantage of the complementary strengths of gauge-, satellite-, and reanalysis-based data to provide reliable precipitation estimates over the entire globe. It is a global precipitation dataset with 0.1° spatial resolution and has been available from 1979 to the present. In this study, MSWEP version 2 with no gauge data is used.

NCEP 2 was developed from NCEP1, which covered the period from 1948 onwards (Kalnay et al., 2001) and is based on the T62 operational NCEP model in which the physical parameterisation was run on 192×94 Gaussian grid (1.92°×1.875°). NCEP2 reanalysis (Kanamitsu et al., 2002) was run at the same resolution (T62) as NCEP1, covering the period from 1979 onwards. An important improvement was the updated precipitation parameterisations and more realistic cloud-top cooling. The products are available in a 1.875°×1.875° grid with a 6-hour temporal resolution.

MERRA2 is a NASA atmospheric reanalysis that began in 1980. MERRA2 replaced the original MERRA dataset to leverage advances in assimilation systems, allowing integration of modern hyperspectral radiance, microwave observations, and GPS-Radio Occultation data (Bosilovich et al., 2015, 2016). The current version's advantages over the previous one include enhanced use of satellite observations, assimilation of aerosol information (MERRAero), NASA's ozone profile observations and advances in both the GEOS model and the GSI assimilation system (Cullather et al., 2014; Wu et al., 2002).

ERA5-Land is a reanalysis-based product that covers the period from January 1950 to 2-3 months before the present. ERA5 was produced using 4D-Var data assimilation and model forecasts in CY41R2 of the ECMWF Integrated Forecast System (IFS). H-TESSEL is the land surface model that is the basis of ERA5-Land. The data is available on a regular latitude/longitude grid of 0.1° × 0.1°. In particular, ERA5-Land runs at enhanced resolution (9 km vs 31 km in ERA5). The temporal frequency of the output is hourly, and the fields are masked for all oceans, making them lighter to handle (Munoz-Sabater et al., 2021).

Table 2. An overview of precipitation products used in the study

No	Precipitation Product	Concept	Primary band/model	Station Data	Spatial*	Temp#
1	CMORPH	Data-driven	PMW – 10.7 to 190 GHz	CPC Gauge data (land), GPCP (Ocean)	0.25°	Daily
2	GPM-IMERG	Data-driven	Dual Frequency Precipitation Radar - 35.5 GHz & 13.6 GHz; GPM Microwave Imager - 10 GHz to 183 GHz	GPCC gauge analysis	0.1°	Daily
3	MSWEP (Past no gauge)	Data-driven	Weighted mean average merge	-	0.1°	Daily
4	PERSIANN CDR	Data-driven	GridSat-B1 IR window - 11µm	GPCP monthly data	0.25°	Daily
5	TRMM 3B42	Data-driven	Precipitation Radar & TRMM Microwave Imager	GPCC gauge	0.25°	Daily
6	ERA5-Land	Model-based	H-TESSEL	Gridded dataset based on station time series from ECA&D, US Climate Reference Network	0.1°	Hourly
7	MERRA2	Model-based	GEOS-5.12.4 system	Sea surface observations from International Comprehensive Atmosphere-Ocean Data set (ICOADS); Land surface observation from National Centers for Environmental Predictions (NCEP)	0.5°×0.63°	Daily
8	NCEP2	Model-based	T62 operational NCEP model	GPCC gauge	1.875°×1.875°	Daily

*Spatial resolution given in latitude × longitude, #Temporal resolution of the product used in this study

2.3 Comparison of ground data with satellite and observational reanalysis-based data

The most commonly used method to compare ground observations with other data products, such as satellite-based precipitation estimates and reanalysis products, is point (station) to pixel comparison. It can be challenging to acquire reasonable agreement when comparing daily rainfall, particularly in a very complex topography, on a point-to-pixel basis (Gebrechorkos et al., 2018). The study used 10 years of precipitation data for evaluation of the products. We followed the most

common practice of interpolating the point-based rain gauge data to 0.1° spatial resolution using the linear method, as it was relatively easy to implement and a popular application in evaluation studies (Duan et al. 2016, Liu et al. 2015, Shukla et al. 2019). Since the real stations are arranged almost in a regular grid, we can expect low interpolation errors. This is also shown by the results of LOOCV analysis (Table S1). The interpolated 0.1° station dataset was used as ground truth to evaluate all the other precipitation products. Using the linear method, the satellite and reanalysis rainfall products with the coarse spatial resolution were downscaled to 0.1°. The precipitation products were downscaled to higher resolution with an aim to capture spatial rainfall variability, which will be better exhibited at 0.1° as compared to their original coarse resolution.

We evaluated the performance of selected precipitation products at three temporal resolutions (daily, monthly and yearly). For each temporal resolution, evaluations were conducted at two spatial scales: the grid scale and the district scale.

2.3.1 Evaluation at grid scale

The interpolated station dataset was used as ground truth to evaluate all the other downscaled precipitation products at 0.1° spatial scales. The evaluation was considered for grids having atleast one rain gauge in the corresponding grid or in the surrounding grids.

2.3.2 Evaluation at district scale

The gridded precipitation from the grids considered at the grid-scale evaluation were averaged to estimate the district precipitation amount. This way, district-level precipitation was estimated for the station data and the precipitation products. The developed district-level data was then evaluated, keeping the station value as a ground truth for the precipitation products.

The study region has abundant rainfall during the monsoon months, which is significantly higher than the rest of the months, hence we evaluated the performance of the products during the monsoon and non-monsoon seasons. The study region falls under the North East monsoon zone from October to December (Table 1). So, average precipitation values of October, November, and December were taken as monsoon months. The average precipitation for the rest of the year, i.e. from January to September, was taken as Non-monsoon months.

2.4 Evaluation metrics

Evaluations were primarily based on the widely used statistical indices, such as the Pearson correlation coefficient (CC), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Bias (systematic error), Relative Bias (R_{bias}) and Index of Agreement (IA) were used. CC (Equation 1) is applied to evaluate the agreement of individual precipitation products (P) to station data (O). A value of 1 shows a perfect fit between the products and station data.

$$\text{Correlation Coefficient} = \frac{\sum_{i=1}^{N}(P_i-\bar{P}).(O_i-\overline{O})}{\sqrt{\sum_{i=1}^{N}(P_i-\bar{P})^2}\ \sqrt{\sum_{i=1}^{N}(O_i-\bar{O})^2}} \quad \text{(Equation 1)}$$

In Equation 1, P_i is the value of the precipitation product at each time step (i), $\bar{P}$ is the mean value of the precipitation product over the entire time step. Similarly, Oi represents the precipitation value recorded at the ground station at time step (i), $\bar{O}$ is the mean of the ground station recorded precipitation. Each product's average differences and systemic bias are given as Bias (Equation 2) and R_{bias} (Equation 3). Based on the underestimation or overestimation of the product, Bias and R_{bias} can be positive or negative.

$$\text{Bias (Systematic error)} = \frac{\sum(P_i - O_i)}{N} \quad \text{(Equation 2)}$$

$$R_{bias} = \frac{\sum_{i=1}^{N}(P_i - O_i)}{\sum_{i=1}^{N} O_i} \times 100 \quad \text{(Equation 3)}$$

The MAE and RMSE (Equations 4 and 5) are well-known and accepted indicators of goodness of fit, which shows the difference between ground observations and precipitation products under study (Legates and McCabe, 1999).

$$MAE = \frac{\sum_{i=1}^{N}|O_i - P_i|}{N} \quad \text{(Equation 4)}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(O_i - P_i)^2}{N}} \quad \text{(Equation 5)}$$

The IA (Willmott, 1981) is another widely used indicator of goodness of fit between station data and different products. Equation 6 describes the amount of precipitation product that is error-free as compared to the station data.

$$IA = \frac{\sum(P_i - O_i)^2}{\sum(|P - \bar{O}| + |O - \bar{O}|)} \quad \text{(Equation 6)}$$

Additional analyses were conducted at the district spatial scale, which includes Intensity-Duration-Frequency (IDF) curve analysis at a daily time scale, seasonal evaluation at a monthly timescale and Taylor plot-based multimodel comparison at a yearly timescale. IDF curve analysis was done using the Gumbel distribution methodology. It is the most widely used distribution for IDF analysis owing to its suitability for modelling maxima. Frequency precipitation (PT in mm, Equation 7) for each duration with a specific return period T (in year) is given by the following equation:

$$PT = P_{ave} + KS \quad \text{(Equation 7)}$$

Where K is the Gumbel frequency factor (Equation 8), P_{ave} is the average of the maximum precipitation corresponding to a specific duration, and S is the standard deviation of the precipitation data (Equation 9). Then, the Precipitation Intensity, It (mm/hour) for return period T, is obtained using Equation 10.

$$\text{Gumbel frequency factor, } K = \frac{\sqrt{6}}{\Pi}[0.5772 + \ln[\ln[\frac{T}{T-1}]]] \quad \text{(Equation 8)}$$

$$\text{Standard deviation of precipitation data} = \frac{1}{n-1}\sum_{i=1}^{n}(P_i - P_{ave})^2]^{0.5} \quad \text{(Equation 9)}$$

$$I_t = \frac{P_t}{T_d} \quad \text{(Equation 10)}$$

Where T_d is the duration in hours

In the monthly timescale, Percent Mean Absolute Error (%MAE) (Equation 11) was computed as the ratio of MAE to the mean precipitation from corresponding station data. Hence, it enables the precipitation products' intercomparison, considering the precipitation's monthly variation (Duan et al., 2016).

$$\%MAE = \frac{\text{Monthly MAE of the precipitation product}}{\text{Mean Precipitation from the Station data}} \quad \text{(Equation 11)}$$

In addition to the above statistical methods, the Taylor diagram (Taylor, 2001) was used to summarize the statistical relationship between ground station data and precipitation products at yearly timescale evaluations. This diagram explains the relationship between the two fields by correlation coefficient (R), standard deviation (σ) and centred root mean square (RMS) difference (E'). A single point on the diagram displays three statistical values (E',R and σ) and their relationship is given by Equation 12. Percent Bias, either in a positive or negative way, is indicated by the upward or downward notation of the product.

$$E'^2 = \sigma_f^2 + \sigma_r^2 - 2\,\sigma_f \sigma_r R \quad \text{(Equation 12)}$$

Where σ_f^2 and σ_r^2 are the variances of the satellite / reanalysis and station data, and R is the correlation coefficient between the two products. The distance between the reference point (station data) is given in the diagram as the centred RMS difference of the two fields (E'^2, Equation 13). A satellite/reanalysis product with no error would show a perfect correlation to the station data.

$$E'^2 = \frac{1}{N}\sum_{n=1}^{N}[(f_n - \bar{f}) - (r_n - \bar{r})]^2) \quad \text{(Equation 13)}$$

Where f is the model (e.g. satellite or reanalysis) data, and r is the reference (station data) data, whereas σ_f and σ_r are the standard deviations of the model and reference fields (Equation 14 and Equation 15).

$$\sigma_f = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(f_n - \bar{f})^2} \quad \text{(Equation 14)}$$

$$\sigma_r = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(r_n - \bar{r})^2} \quad \text{(Equation 15)}$$

3 Results

3.1 Evaluation of precipitation products at the daily time scale

3.1.1. Grid-Scale

The statistical metrics for daily precipitation are summarized in Tables 3 and S2. For the study region Coimbatore at grid scale, ERA5-Land had the highest correlation value (CC = 0.366), followed by MSWEP. The lowest CC value of 0.052 was observed for the NCEP2 product. All the products overestimated the rainfall with a Bias and R_{bias} value ranging from 0.197 to 18.293 and 0.087 to 8.036, respectively. MSWEP recorded the lowest RMSE, indicating the product's fit with the station data compared to the other products. Interestingly, ERA5-Land was observed to have both the lowest MAE and highest IA. In Madurai, correlation trends were similar to those in Coimbatore. ERA5-Land, followed by MSWEP, had the highest correlation coefficient, while NCEP2 had the lowest correlation. ERA5-Land was observed to have the lowest MAE and RMSE values of 3.859 mm/day and 7.511 mm/day. Five precipitation products underestimated the rainfall with Bias and R_{bias} values ranging from -2.039 to -1.123 and -0.456 to -0.253, respectively. On the other hand, NCEP2 and MERRA2 overestimated

precipitation, with the former having comparatively higher values of Bias and R_{bias} (10.875 and 2.434). Regarding IA, ERA5-Land and NCEP2 recorded the highest and lowest values.

In Tiruchirappalli, correlation and IA had similar trends with respect to Coimbatore and Madurai. Table 3 shows that the correlation value ranged from 0.096 (NCEP2) to 0.371 (ERA5-Land). In terms of error, MERRA2 was observed to have the highest MAE (11.186 mm/day) and RMSE (33.08 mm/day). On the other hand, ERA5-Land had the lowest MAE (3.237 mm/day), and MSWEP had the lowest RMSE (7.229 mm/day). Five products underestimated the precipitation, with Bias and R_{bias} ranging from -1.187 to -0.549 and -0.342 to -0.158, respectively. Furthermore, NCEP2 and MERRA2 overestimated the precipitation with values ranging above zero. ERA5-Land, followed by PERSIANN CDR, had the highest IA, while MERRA2 had the lowest value. In Tuticorin, MSWEP had the highest correlation (CC = 0.456), while NCEP2 had the lowest (CC = 0.117). MAE values ranged from 3.225 mm/day to 11.888 mm/day, with the lowest recorded by MSWEP and the highest recorded by MERRA2. Similar trends were observed for RMSE, with values ranging from 6.623 to 33.268 mm/day. CMORPH, GPM-IMERG, MSWEP, PERSIANN CDR, TRMM and ERA5-Land had underestimated the precipitation with negative Bias and R_{bias} values. MERRA2 and NCEP2 overestimated rainfall with positive Bias and R_{bias} values.

The performance of IDF curves generated using remote sensing and reanalysis-based precipitation products was evaluated by comparing them with the estimated IDF curves using station data. It demonstrates that rainfall intensity decreases with duration. In the Coimbatore district, MSWEP and PERSIANN CDR had the closest precipitation intensity estimation with respect to the station data, as shown in Fig. 2a. The intensity of precipitation produced by station data was 16.056 mm/hour and 1.929 mm/hour for 1 hour and 24 hour duration, respectively. PERSIANN CDR produced 15.952 mm/hour and 1.917 mm/hour for 1 hour and 24 hour duration, respectively. ERA5Land highly underestimated precipitation intensity, while GPM-IMERG, CMORPH, TRMM, NCEP2 and MERRA2 overestimated. For the Madurai district, PERSIANN CDR, ERA5-Land and MSWEP underestimated the precipitation intensity, while GPM-IMERG produced the closest estimate. Station data's precipitation intensity was 26.033 mm/hour and 3.128 mm/hour for 1 hour and 24 hour duration, respectively. GPM-IMERG produced 25.93 mm/hour and 3.117 mm/hour for 1 hour and 24 hour duration, respectively. From Fig. 2b, it is also observed that TRMM, CMORPH, NCEP2 and MERRA2 overestimated the precipitation intensity. In Tiruchirappalli district, all the precipitation products overestimated the precipitation intensity in comparison to the station data (Fig. 2c). PERSIANN CDR produced the closest precipitation intensity estimation with slight overestimation. In the Tuticorin district, PERSIANN CDR underestimated the precipitation intensity, and ERA5-Land produced the closest estimate with respect to the station data. Station data's precipitation intensity was 20.406 mm/hour and 2.452 mm/hour for 1 hour and 24 hour duration, respectively. ERA5-Land produced 20.269 mm/hour and 2.436 mm/hour for 1 hour and 24 hour duration, respectively. The remaining precipitation product overestimated the intensity at different time periods, with the highest being produced by MERRA2 (Fig. 2d).

The range of precipitation frequency at continuous intensity intervals (1-2, 2-5, 5-10, 10-20, 20-50 and >50 mm/day) was evaluated for the eight gridded precipitation products against the station data (Fig. S3). Since the frequency class 0–1 mm/day accounts for more than 60% of the daily grid-scale precipitation, it was investigated separately using Fig. S2. At the

grid scale in Coimbatore, more than 65% of daily precipitation from the interpolated rain gauge falls in the intensity of 0-1 mm/day. At this intensity range, GPM-IMERG (66%) had very close interpolation to the station data, followed by ERA5-Land (61%). From the occurrence frequency of precipitation at a different intensity, as shown in Fig. S3a, CMORPH, GPM-IMERG, and TRMM products followed the station data trend, underestimating the precipitation in the 1 to 20 mm/day range and thereafter overestimating it.

For Madurai district, at grid scale, the frequency plot of 0-1 mm/day is shown in Fig. S2b. PERSIANN CDR product had a close estimate with the station data in a range of 0 to 1 mm/day. By observing the precipitation occurrence frequency at higher ranges (i.e. from 5 to 50 mm/day), CMORPH, GPM-IMERG, TRMM, and ERA5-Land had produced underestimated values consistently, as signified by the negative Bias value in Table 3. ERA5-Land and MSWEP were observed to have a similar trend as the gridded station data from 1-10 mm/day; thereafter, it diverged. Although NCEP2 had overestimated in the 2 to 10 mm/day range, it was found to capture gridded station data frequency at remaining intensity levels (Fig. S3b). Among the products, MERRA2 had the highest overestimation frequency percentage, which is highly substantiated by the positive Bias and R_{bias} values in Table 3.

In Tiruchirappalli, at grid scale, nearly 60 % of the interpolated ground station precipitation falls in the range of 0 to 1 mm/day (Fig. S2c). In frequency range from 1 mm/day to 50 mm/day, MSWEP was observed to have a close relation with the station data from 1 to 5 mm/ day, NCEP2 followed by PERSIANN CDR from 5 to 20 mm/ day, TRMM followed by GPM-IMERG from 20 to 50 mm/day and ERA5-Land at more than 50 mm/ day range, respectively (Fig. S3c). Further, Fig. S2d shows the grid scale precipitation data of Tuticorin at 0 to 1 mm/day intensity range. It is observed that more than 55% of the interpolated ground station precipitation falls in the 0 to 1 mm/day range, which NCEP2 and PERSIANN CDR closely captured. In the precipitation range of 0-1 mm/day, NCEP2, followed by PERSIANN CDR, had estimated closer frequency percentages with respect to the interpolated ground station data. Further, MSWEP is followed by ERA5-Land at 1 to 2 mm/day, NCEP2 followed by MSWEP at 2 to 5 mm/day, NCEP2 at 5 to 20 mm/day, TRMM at 20 to 50 mm/day and PERSIANN CDR at precipitation range above 50 mm/day (Fig. S3d).

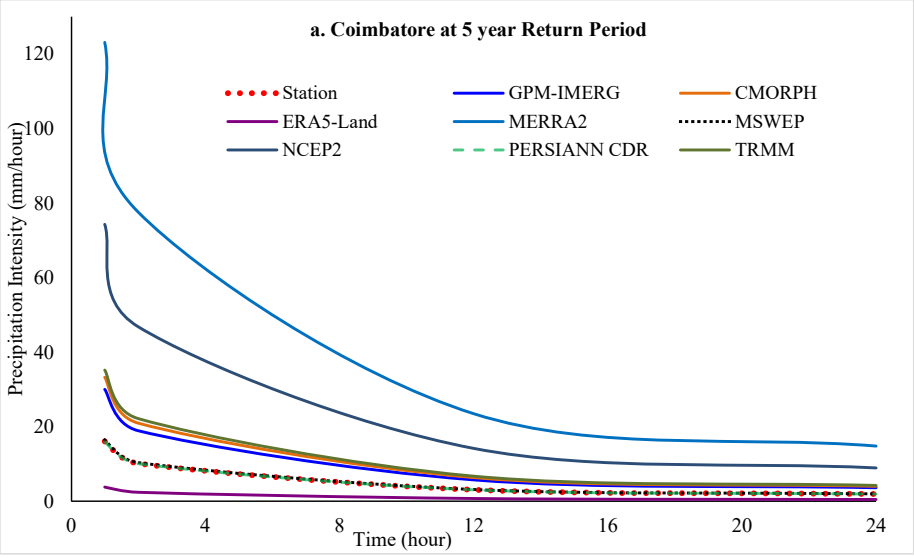
3.1.2 District-Scale

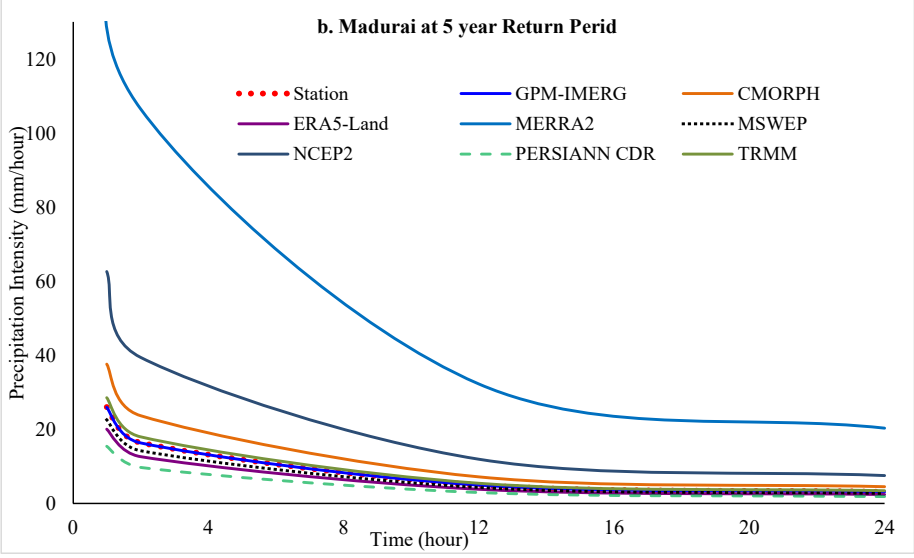
The statistical evaluations perform better at the district scale and typically reflect the grid scale trend. In Coimbatore, ERA5-Land recorded the lowest RMSE with 0.614 mm/day as compared to MSWEP at the grid scale. Also, PERSIANN CDR was observed to have the lowest Bias, R_{bias} and highest IA. In Madurai, the statistical metrics at the district level followed the grid scale trend. In Tiruchirappalli, ERA5-Land recorded the lowest RMSE with 6.663 mm/day, and other metrics followed the grid scale trend. In Tuticorin, PERSIANN CDR was observed to have the highest correlation coefficient of 0.466 and the lowest RMSE of 6.025 mm/day.

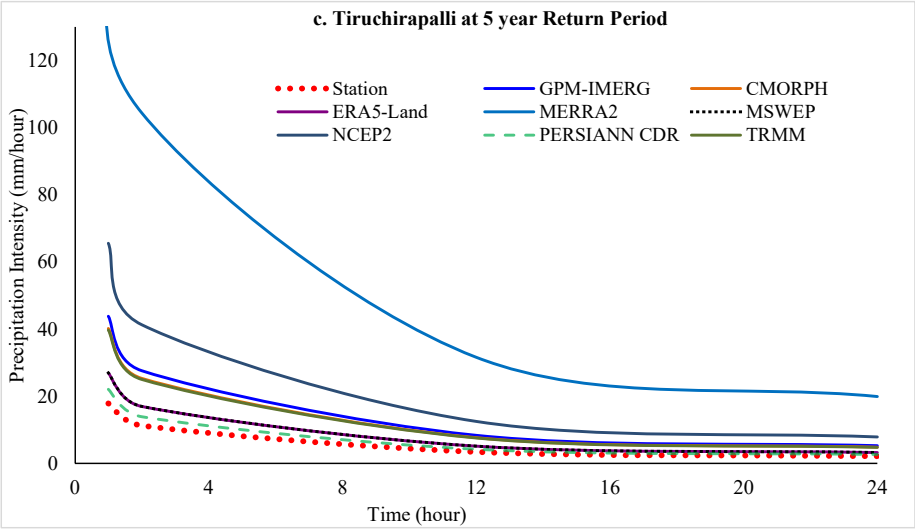
420

Table 3. Daily precipitation characteristics of Coimbatore, Madurai, Tiruchirappalli and Tuticorin

		Coimbatore			Madurai			Tiruchirappalli			Tuticorin		
Scale	**Product**	**CC**	**RMSE**	**RB**	**CC**	**RMSE**	**RB**	**CC**	**RMSE**	**RB**	**CC**	**RMSE**	**RB**
Grid	**CMORPH**	0.19	8.92	0.26	0.24	10.92	-0.25	0.20	9.46	-0.28	0.33	8.50	-0.45
	GPM-IMERG	0.21	9.00	0.44	0.29	9.20	-0.36	0.26	9.53	-0.16	0.38	7.94	-0.34
	MSWEP	0.33	6.07	0.16	0.41	7.87	-0.41	0.33	7.23	-0.33	0.46	6.62	-0.29
	PERSIANN CDR	0.20	7.72	0.16	0.35	8.00	-0.43	0.33	7.32	-0.33	0.43	6.53	-0.34
	TRMM	0.21	8.76	0.09	0.26	10.13	-0.28	0.23	9.65	-0.19	0.36	8.47	-0.41
	ERA5-Land	0.37	6.32	0.15	0.45	7.51	-0.46	0.37	6.92	-0.34	0.39	7.23	-0.39
	MERRA2	0.24	41.36	8.04	0.31	35.74	2.43	0.21	33.08	2.70	0.30	33.27	2.83
	NCEP2	0.05	15.77	1.14	0.07	14.00	0.10	0.10	13.33	0.31	0.12	13.59	0.40
District	**CMORPH**	0.23	8.09	0.26	0.25	10.56	-0.25	0.21	9.16	-0.28	0.37	7.75	-0.45
	GPM-IMERG	0.25	8.03	0.44	0.31	8.81	-0.36	0.27	9.10	-0.16	0.43	7.24	-0.34
	MSWEP	0.38	5.37	0.16	0.43	7.59	-0.41	0.35	6.94	-0.33	0.46	6.53	-0.29
	PERSIANN CDR	0.23	7.15	-0.15	0.36	7.80	-0.43	0.35	7.08	-0.33	0.47	6.03	-0.34
	TRMM	0.23	1.87	0.63	0.27	9.78	-0.28	0.24	9.30	-0.19	0.40	7.73	-0.41
	ERA5-Land	0.39	0.61	-0.45	0.48	7.24	-0.46	0.39	6.66	-0.34	0.46	6.37	-0.39
	MERRA2	0.29	39.9	8.04	0.32	35.41	2.43	0.22	32.84	2.70	0.33	32.24	2.83
	NCEP2	0.06	15.51	1.14	0.07	13.91	0.10	0.10	13.21	0.31	0.12	13.36	0.40


a. Coimbatore at 5 year Return Period
Station
GPM-IMERG
CMORPH
ERA5-Land
MERRA2
MSWEP
NCEP2
PERSIANN CDR
TRMM
Precipitation Intensity (mm/hour)
Time (hour)
0
20
40
60
80
100
120
0
4
8
12
16
20
24


b. Madurai at 5 year Return Perid
Station
GPM-IMERG
CMORPH
ERA5-Land
MERRA2
MSWEP
NCEP2
PERSIANN CDR
TRMM
Precipitation Intensity (mm/hour)
Time (hour)
0
20
40
60
80
100
120
0
4
8
12
16
20
24



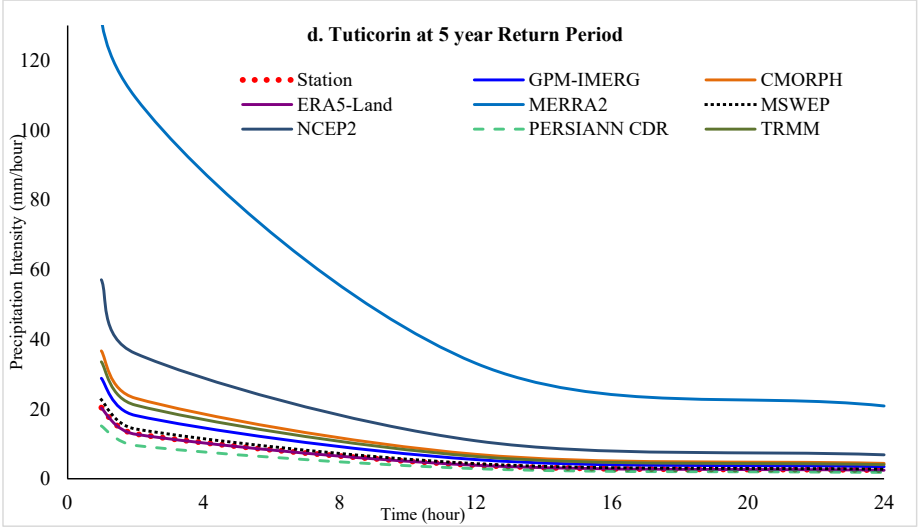
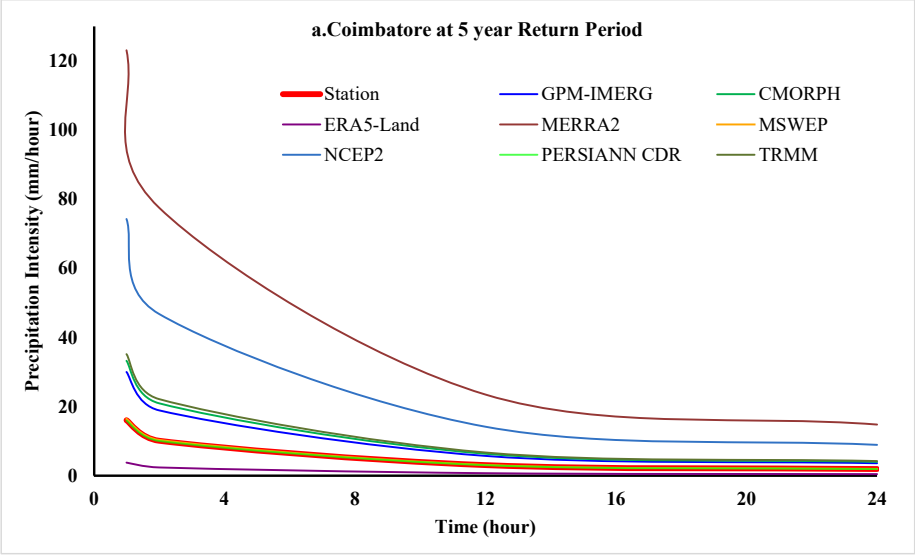
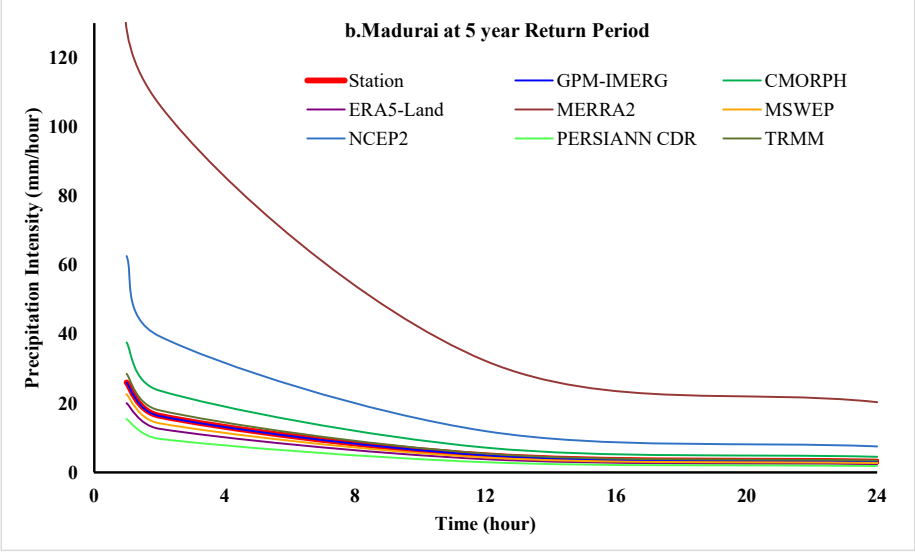
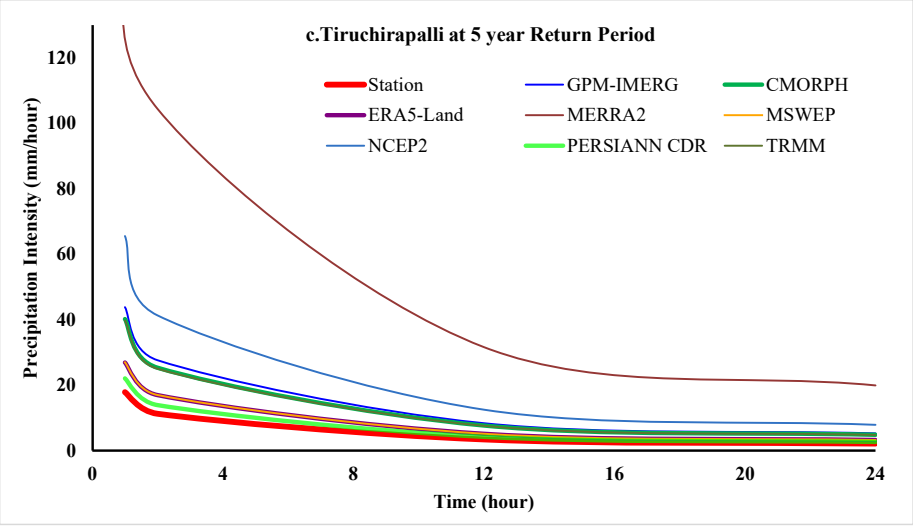


Figure 2. Intensity Duration Frequency (IDF) curves based on daily extreme (maximum) precipitation values for the study regions at 5 years return period

425


a.Coimbatore at 5 year Return Period
Precipitation Intensity (mm/hour)
Time (hour)
Station
GPM-IMERG
CMORPH
ERA5-Land
MERRA2
MSWEP
NCEP2
PERSIANN CDR
TRMM
0
20
40
60
80
100
120
0
4
8
12
16
20
24


b.Madurai at 5 year Return Period
Precipitation Intensity (mm/hour)
Time (hour)
Station
GPM-IMERG
CMORPH
ERA5-Land
MERRA2
MSWEP
NCEP2
PERSIANN CDR
TRMM
0
20
40
60
80
100
120
0
4
8
12
16
20
24



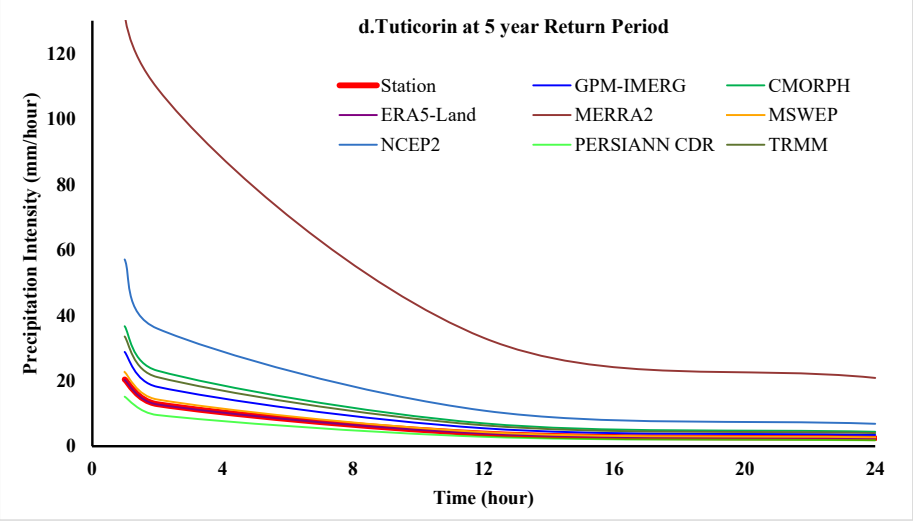


Figure 2. Intensity Duration Frequency (IDF) curves based on daily extreme (maximum) precipitation values for the study regions at 5 years return period

Formatted: Left, Line spacing: single, Tab stops: 10.62 cm, Left

3.2 Evaluation of precipitation products at the monthly time scale

3.2.1 Grid Scale

When comparing the estimations to the daily time scale, the monthly timeframe exhibited a considerably lower error and stronger correlations with the station data. This could be because statistical metrics take less variability into account when analysing data with a lower temporal resolution. At the grid scale in Coimbatore, ERA5-Land had the highest correlation coefficient, followed by MSWEP (Table 4 and S3). The precipitation products NCEP2 and MERR2 were observed to have a correlation with the station data, indicating poor representation at this spatial scale. Further, ERA5-Land was observed to have the lowest MAE and RMSE, with values of 43.632 mm/month and 62.799 mm/month, respectively. All the products were observed to overestimate the grid scale monthly rainfall, with MERRA2 having the highest Bias and R_{bias} values. On the other hand, TRMM and ERA5-Land were observed to have the lowest Bias value and R_{bias} values of 0.197 and 0.079, respectively. Further, MSWEP had the highest IA and NCEP2 had the lowest. In the Madurai region, the performance of the precipitation products at the grid was observed to be highly varied, as shown in Table 4. ERA5-Land recorded the highest correlation coefficient of 0.814, followed by MSWEP (CC = 0.8). TRMM was observed to record the lowest MAE and RMSE. Further, five products underestimated the precipitation with negative Bias and R_{bias} values. NCEP2 and MERRA2 overestimated precipitation at the monthly grid scale with positive Bias and R_{bias} values. NCEP2 was observed to have the lowest Bias, whereas MERRA had the highest. Further, NCEP2 had the lowest IA and TRMM had the highest.

In Tiruchirappalli, ERA5-Land had the highest correlation coefficient (CC = 0.791), followed by PERSIANN CDR (r = 0.770). Also, as observed in Table 4, PERSIANN CDR recorded the lowest MAE, RMSE, Bias, R_{bias}, and IA. Five products underestimated monthly grid-scale precipitation at Tiruchirappalli, while two overestimated it. In Tuticorin, MSWEP was observed to have the highest correlation coefficient (CC = 0.796) and IA (0.862). Further, MSWEP recorded the lowest MAE and RMSE of 47.07 mm/month and 66.35 mm/month. Among the products, PERSIANN CDR was found to have the minimum Bias and R_{bias} values of -29.843 and -0.265, respectively (Table 4).

Table 4. Monthly precipitation characteristics of Coimbatore, Madurai, Tiruchirappalli and Tuticorin

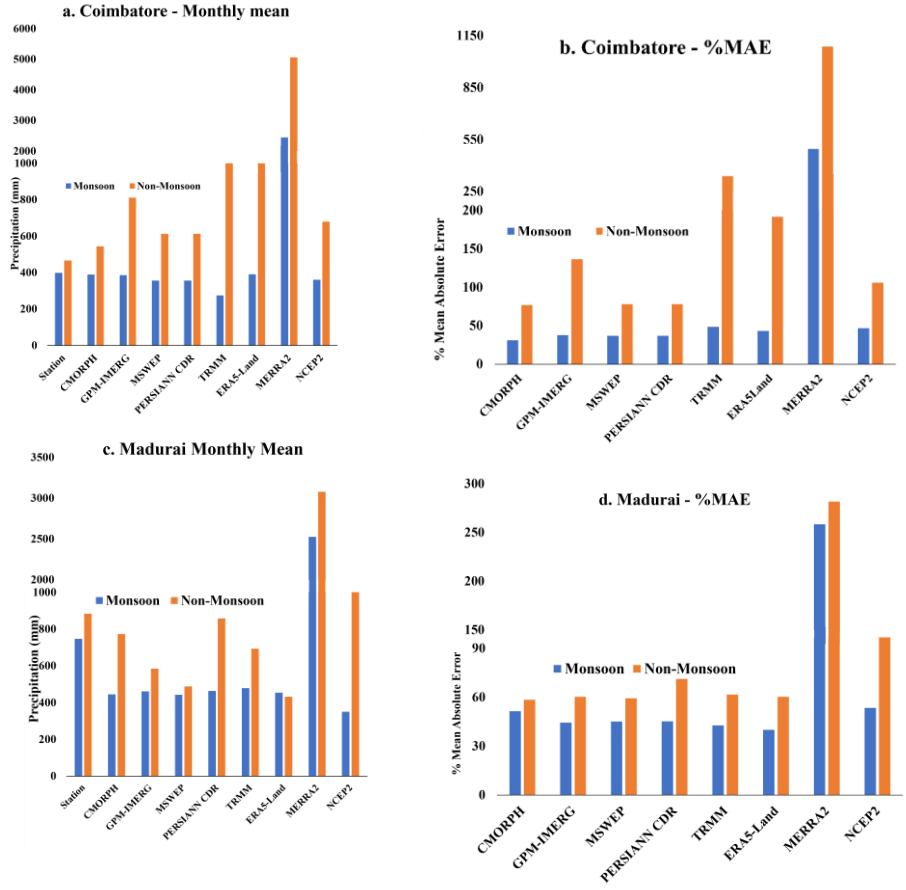
		Coimbatore			Madurai			Tiruchirappalli			Tuticorin		
Scale	**Product**	**CC**	**RMSE**	**RB**	**CC**	**RMSE**	**RB**	**CC**	**RMSE**	**RB**	**CC**	**RMSE**	**RB**
Grid	**CMORPH**	0.54	72.09	0.26	0.16	91.80	-0.25	0.64	79.11	-0.28	0.73	83.80	-0.45
	GPM-IMERG	0.48	83.12	0.44	0.78	87.07	-0.36	0.71	72.25	-0.16	0.77	75.04	-0.34
	MSWEP	0.59	63.01	0.12	0.80	91.04	-0.43	0.75	69.96	-0.35	0.80	66.35	-0.31
	PERSIANN CDR	0.34	113.09	0.79	0.69	87.18	-0.19	0.77	61.65	-0.07	0.73	69.23	-0.26
	TRMM	0.10	109.78	0.09	0.75	84.70	-0.28	0.69	74.32	-0.19	0.77	77.44	-0.41
	ERA5Land	0.62	62.80	0.08	0.81	93.25	-0.46	0.79	66.13	-0.34	0.65	89.45	-0.39
	MERRA2	-0.15	96.82	8.04	0.71	459.54	2.43	0.59	416.93	2.70	0.64	448.33	2.83
	NCEP2	-0.08	225.95	1.14	-0.02	190.56	0.10	0.17	160.10	0.31	0.11	169.54	0.40
District	**CMORPH**	0.63	60.63	0.26	0.69	89.22	-0.25	0.65	76.60	-0.28	0.80	75.62	-0.45
	GPM-IMERG	0.56	73.52	0.44	0.80	85.03	-0.36	0.73	69.47	-0.16	0.80	68.40	-0.34
	MSWEP	0.69	50.66	0.12	0.81	89.25	-0.43	0.77	67.24	-0.35	0.80	65.58	-0.31
	PERSIANN CDR	0.39	107.45	0.79	0.70	85.39	-0.19	0.79	58.07	-0.07	0.82	61.57	-0.26
	TRMM	-0.01	14.08	0.50	0.77	82.40	-0.28	0.71	71.46	-0.19	0.83	70.34	-0.41
	ERA5Land	0.72	50.35	0.08	0.83	91.38	-0.46	0.81	63.27	-0.34	0.76	75.15	-0.39
	MERRA2	0.56	705.76	8.04	0.72	457.92	2.43	0.60	415.75	2.70	0.68	443.23	2.83
	NCEP2	-0.09	223.0	1.14	-0.02	189.86	0.10	0.18	158.98	0.31	0.11	167.21	0.40

3.2.2 District Scale

All the precipitation products overestimated rainfall at the district scale in Coimbatore, similar to their performance at the grid scale. ERA5-Land recorded the highest CC and IA in the Coimbatore district scale. Contrasting to grid-scale performance, TRMM produced the lowest MAE (35.842 mm/month) and RMSE (50.348 mm/month), which is followed by ERA5-Land (Table 4 and S3). District scale performance of precipitation products at Madurai and Tiruchirappalli mostly depicted similar trends as observed in the grid scale (Table 4). MAE trends at the district scale differed from those of grid-scale performance in Madurai. TRMM (62.883 mm/month) and MERRA2 (333.675 mm/month) recorded the lowest and highest MAE in Madurai. In Tuticorin, TRMM recorded the highest CC, followed by GPM-IMERG (Table 4 and S3). In terms of error, PERSIANN CDR was found to have the lowest MAE (44.033 mm/month) and RMSE (61.570 mm/month). The performance of Bias, R_{bias}, and IA followed a similar trend to that of the grid scale.

Further, we also investigated the seasonal trend with %MAE during monsoon (October to December) and non-monsoon months (January to November) of the study period. Figure 3a shows that ERA5-Land, followed by CMORPH and GPM-IMERG, closely depicts the station's monsoon precipitation at Coimbatore. On the other hand, all the products overestimated the precipitation for the non-monsoon month, with CMORPH having the closest fit. From Fig. 3b, monsoon months have a lower %MAE compared to the non-monsoon months. CMORPH, followed by PERSIANN CDR, recorded the lowest %MAE during the monsoon and non-monsoon months at Coimbatore. In Madurai, most of the products underestimated the monsoon month precipitation while MERRA2 overestimated it (Fig. 3c). During non-monsoon months, PERSIANN CDR had the closest depiction with the station precipitation data. Fig. 3d shows that ERA5-Land and MERRA2 had the lowest monsoon month %MAE, respectively. On the other hand, CMORPH, followed by MSWEP, had the minimum %MAE during

non-monsoon months at Madurai. In Tiruchirappalli, except MERRA2, all the products underestimated the monsoon month precipitation. Fig. 3e depicts that GPM-IMERG only produced the closest value with respect to the station data. Further, CMORPH, followed by PERSIANN CDR, had the closest estimation during the non-monsoon months. ERA5-Land recorded the lowest %MAE during both periods, as shown in Fig. 3f. Like Madurai and Tiruchirappalli, monsoon month precipitation was underestimated by most of the products in Tuticorin (Fig. 3g). MSWEP and PERSIANN CDR estimated the closest value with respect to the station data during the monsoon and non-monsoon months, respectively. Further, ERA5-Land estimated the lowest %MAE during the monsoon months and PERSIANN CDR during the non-monsoon months (Fig. 3h).



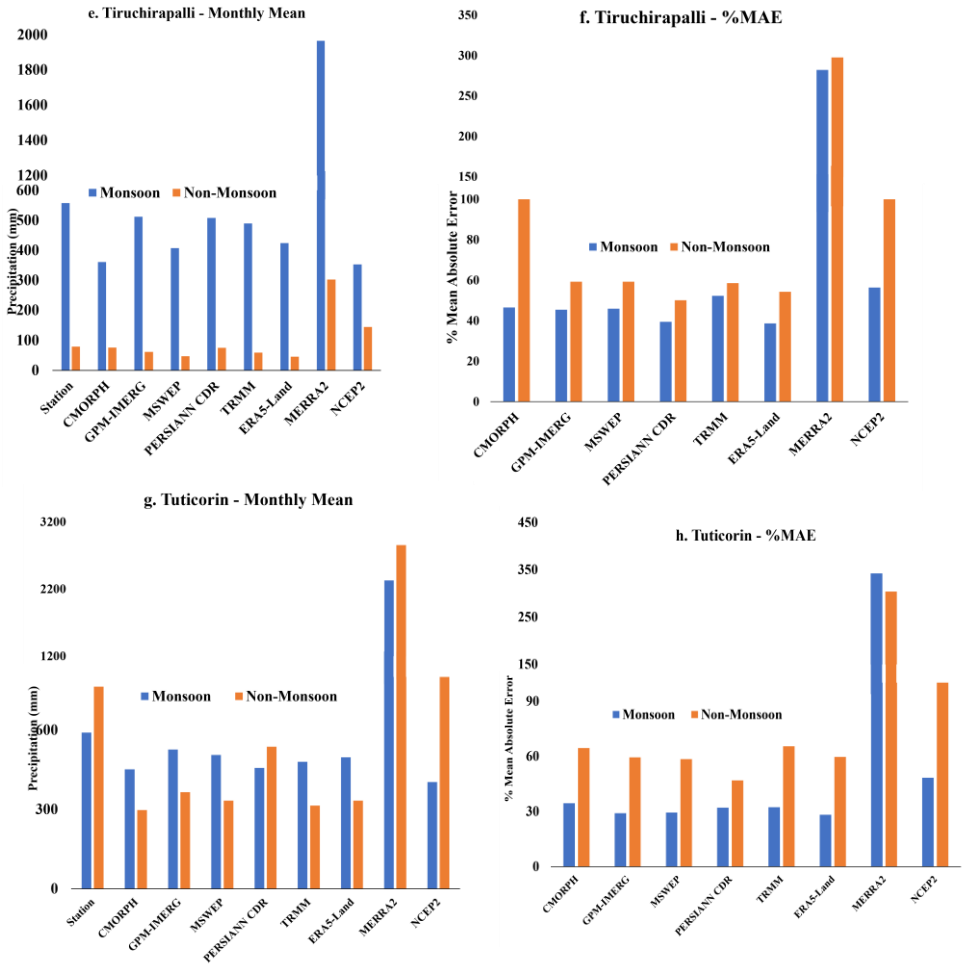
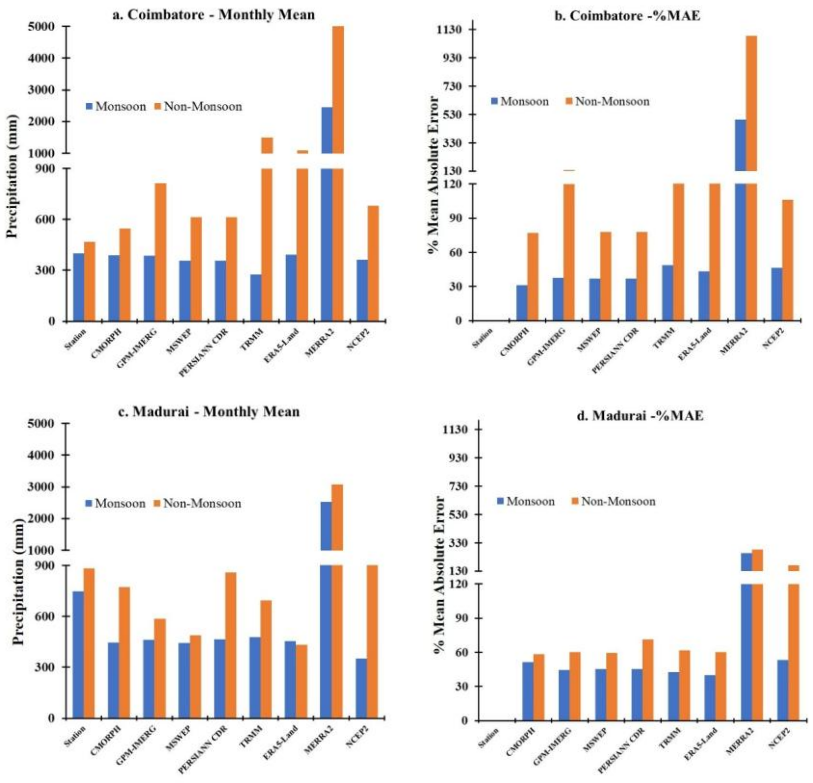


Figure 3. Monthly Mean and %MAE of different precipitation products with respect to Station data in Coimbatore (a,b), Madurai (c,d), Tiruchirappalli (e,f) and Tuticorin (g,h)


a. Coimbatore - Monthly Mean
Precipitation (mm)
Monsoon
Non-Monsoon
b. Coimbatore -%MAE
% Mean Absolute Error
c. Madurai - Monthly Mean
d. Madurai -%MAE
Station
CMORPH
GPM-IMERG
MSWEP
PERSIANN CDR
TRMM
ERA5-Land
MERRA2
NCEP2

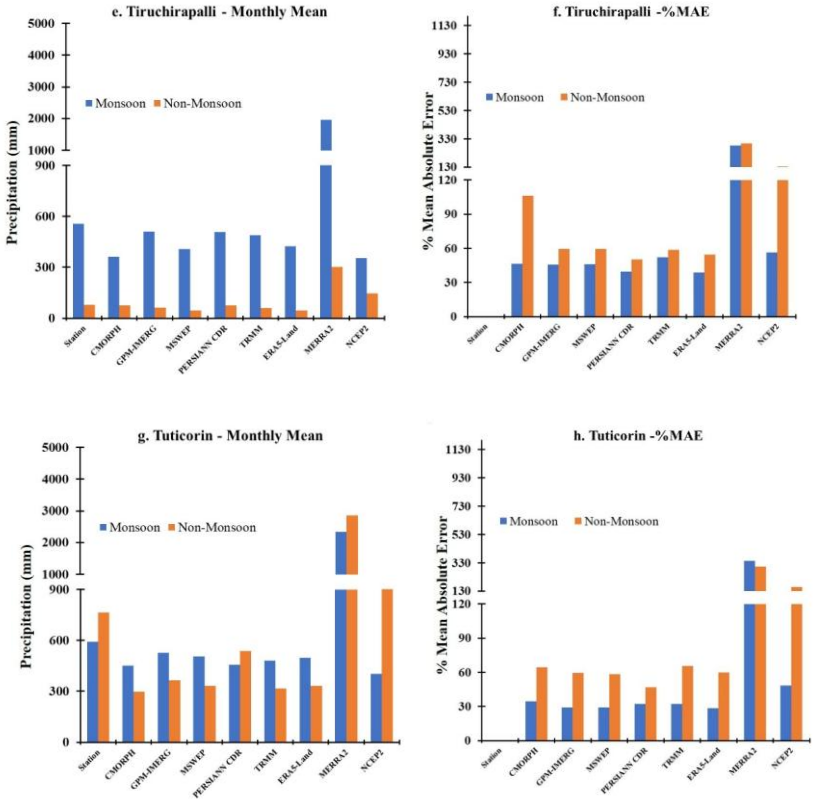


Figure 3. Monthly Mean and %MAE of different precipitation products with respect to Station data in Coimbatore (a,b), Madurai (c,d), Tiruchirappalli (e,f) and Tuticorin (g,h)

4.3 Evaluation of precipitation products at the annual time scale

4.3.1 Grid Scale

The grid-scale daily precipitation was summed as the annual precipitation and was statistically evaluated against the station data. PERSIANN CDR recorded the highest CC in Coimbatore, while ERA5-Land had the best metrics regarding MAE, RMSE and IA. All the products overestimated the precipitation, with the highest overestimation caused by ERA5-Land (Table

5 and S4). In Madurai and Tiruchirappalli (Table 5), PERSIANN CDR was found to have the lowest MAE, RMSE and highest IA. On the other hand, ERA5-Land had the highest CC. Except for NCEP2 and MERRA2, all the products underestimated the precipitation. In Tuticorin, MSWEP was found to have the highest CC, IA and lowest RMSE (Table 5 and S4). PERSIANN CDR recorded the lowest MAE, Bias and R_{bias}.

The yearly precipitation was further evaluated with Taylor Plots. It enables the comparison of multiple models with reference to the station value in a single plot. In general, precipitation products show overestimation in Coimbatore and underestimation in the remaining study regions, as substantiated by their Bias value in Tables 5 and S4. Also, NCEP2 and MERRA2 could not be included in the same plot as the other products due to their higher standard deviation values. By examining the scatter of the product's RMS at the grid spatial scale, PERSIANN CDR had the closest relationship to the station data in Coimbatore, ERA5-Land in Madurai and Tiruchirappalli, MSWEP in Tuticorin (Fig. 4). Similar trends were observed at the district scale, except in Tuticorin, where ERA5-Land lay closer to the station data (Fig. S4). From the combined Taylor plot at the annual scale, it is observed that ERA5-Land performance in Tiruchirappalli is the closest to the origin. The RMS spread of GPM-IMERG, ERA5-Land, and PERSIANN CDR are similar in Madurai. Partial Clustering of PERSIANN CDR with GPM is observed in Tiruchirappalli. Most of the precipitation product's RMS deviates from the observed data for the Coimbatore region (Fig. S5a). For the combined Taylor Plot in monsoon months, GPM-IMERG of Madurai lies the closest to the station RMS, as shown in Figure S5b. Unlike annual and non-monsoon plots, NCEP2 was observed to be included in the monsoon Taylor plot. Performance of precipitation products like TRMM, PERSIANN CDR of Madurai and Tuticorin is clustered around the similar RMS value. During non-monsoon months, ERA5-Land in Tiruchirappalli lies closer to the standard deviation of the station. The deviation of precipitation products from station standard deviation is greater in Coimbatore for non-monsoon months. Clustering of multiple precipitation products for Tuticorin is observed, indicating their similarity in performance (Fig. S5c).

Geospatial maps of the annual rainfall were plotted using Inverse Distance Weighing (IDW) of the grid-scale values using ArcGIS Pro software. The accuracy of different precipitation products in terms of capturing the variability can be assessed through these plots. In the station data spatial map, the distribution of ground station points and their respective linearly interpolated grids are plotted to understand the precipitation variation across the grids. The red circular dots represent the locations of ground stations from which precipitation data was collected for the period 2003–2014. The black stars indicate the linearly interpolated 0.1° grids. For the evaluation, only grids surrounded by at least one rain gauge were considered (Fig. 5a, 6a, 7a, 8a). The station values from Coimbatore showed a large variability from south to north. The annual precipitation varied from 400 mm in the southern region to 1200 mm in the northern region. Also, a few pockets in the northern region received precipitation higher than 1200 mm. CMORPH and partially GPM-IMERG represented the precipitation in the northern region. All the products overestimated the precipitation in the southern region. Based on the correlation values and RMSE given in the Table. 5, along with Figure 8, it is concluded that ERA5-Land produces a close estimate of the station data.

(Fig. 5). In Coimbatore, CMORPH and GPM-IMERG closely depicted the station precipitation pattern in the northern regions. On the other hand, ERA5-Land produced the closest estimate in the southern portion.

Table 5. Yearly precipitation characteristics of Coimbatore, Madurai, Tiruchirappalli and Tuticorin

		Coimbatore			Madurai			Tiruchirappalli			Tuticorin		
Scale	**Product**	**CC**	**RMSE**	**RB**	**CC**	**RMSE**	**RB**	**CC**	**RMSE**	**RB**	**CC**	**RMSE**	**RB**
Grid	**CMORPH**	0.21	423.07	0.24	0.42	539.15	-0.26	0.32	468.57	-0.29	0.46	693.33	-0.46
	GPM-IMERG	0.12	536.78	0.44	0.64	644.06	-0.36	0.68	283.43	-0.16	0.71	517.85	-0.34
	MSWEP	0.05	397.89	0.12	0.45	767.09	-0.43	0.68	481.53	-0.35	0.77	425.33	-0.31
	PERSIANN CDR	0.39	731.66	0.79	0.61	416.82	-0.19	0.68	217.84	-0.07	0.48	463.14	-0.26
	TRMM	-0.61	939.29	0.09	0.48	560.14	-0.28	0.60	333.92	-0.19	0.60	617.91	-0.41
	ERA5-Land	0.33	346.07	0.08	0.66	791.22	-1.00	0.82	461.09	-0.34	0.21	669.71	-0.39
	MERRA2	-0.01	6878.40	8.04	0.35	4148.24	2.43	0.31	3523.97	2.70	0.30	3923.29	2.83
	NCEP2	-0.23	1145.97	1.14	-0.75	609.66	0.10	-0.23	567.99	0.31	-0.54	746.16	0.40
District	**CMORPH**	0.42	56.92	0.24	0.44	200.58	-0.26	0.35	160.87	-0.29	0.73	187.02	-0.46
	GPM-IMERG	0.41	85.22	0.44	0.67	241.07	-0.36	0.74	92.69	-0.16	0.88	138.19	-0.34
	MSWEP	0.43	53.03	0.12	0.47	287.87	-0.43	0.74	165.86	-0.35	0.81	295.53	-0.31
	PERSIANN CDR	0.47	138.93	0.79	0.63	154.01	-0.19	0.79	62.87	-0.07	0.85	113.78	-0.26
	TRMM	0.37	39.02	0.09	0.49	208.98	-0.28	0.67	110.92	-0.19	0.90	164.62	-0.41
	ERA5-Land	0.69	41.64	0.08	0.71	296.84	-0.46	0.92	157.49	-0.34	0.90	154.23	-0.39
	MERRA2	0.30	1394.66	8.04	0.36	1565.42	2.43	0.33	1245.12	2.70	0.44	1127.75	2.83
	NCEP2	-0.48	225.68	1.14	-0.77	228.17	0.10	-0.24	197.06	0.31	-0.71	207.55	0.40

Madurai district mostly had an even distribution of precipitation within 1600-1800 mm (Fig. 6). In contrast to the results of statistical evaluation (Table 5), NCEP2 had the closest estimates. Except for MERRA2, all the products have highly underestimated the precipitation. Annual Precipitation in the Tiruchirappalli district falls in the range of 1100-1300 mm, with the eastern portion having 1300-1500 mm. Although PERSIANN CDR didn't capture the variability in the eastern region, it produced the closest estimate. Most of the precipitation products underestimated the precipitation as expected from the negative RB values in Table 5 and did not capture its regional variability (Fig. 7). In Tuticorin, a high precipitation variability is observed from the station data. The southern and partially eastern portions received precipitation in the 1400-1600 mm range, with some portions even exceeding 1600 mm (Fig. 8). The remaining portions mostly had precipitation in the 1200-1400 mm range, and part of the western portion had 1000-1200 mm. Except for MERRA2 and NCEP2, all the other products underestimated the precipitation. Among the products, MSWEP and GPM produced the closest estimates compared to the station data which is also evident from the higher CC values in Table 5.

4.3.2 District Scale

On a district scale, similar results were observed for Coimbatore, Madurai and Tiruchirappalli. In contrast to the grid scale performance, at the Tuticorin district scale, TRMM and ERA5-Land recorded the highest CC, while PERSIANN CDR was found to have the lowest RMSE (Table 5 & S4).

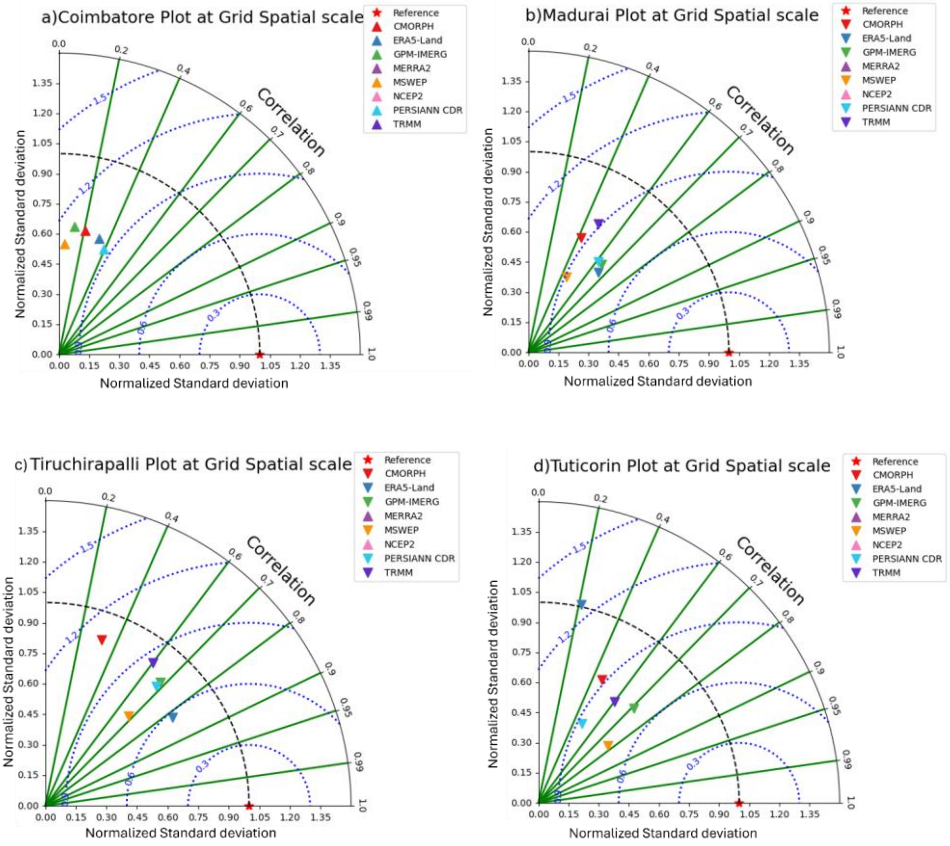


Figure 4. Taylor diagram depicting the agreement between station data and precipitation products at a yearly grid scale for (a) Coimbatore, (b) Madurai, (c) Tiruchirappalli and (d) Tuticorin

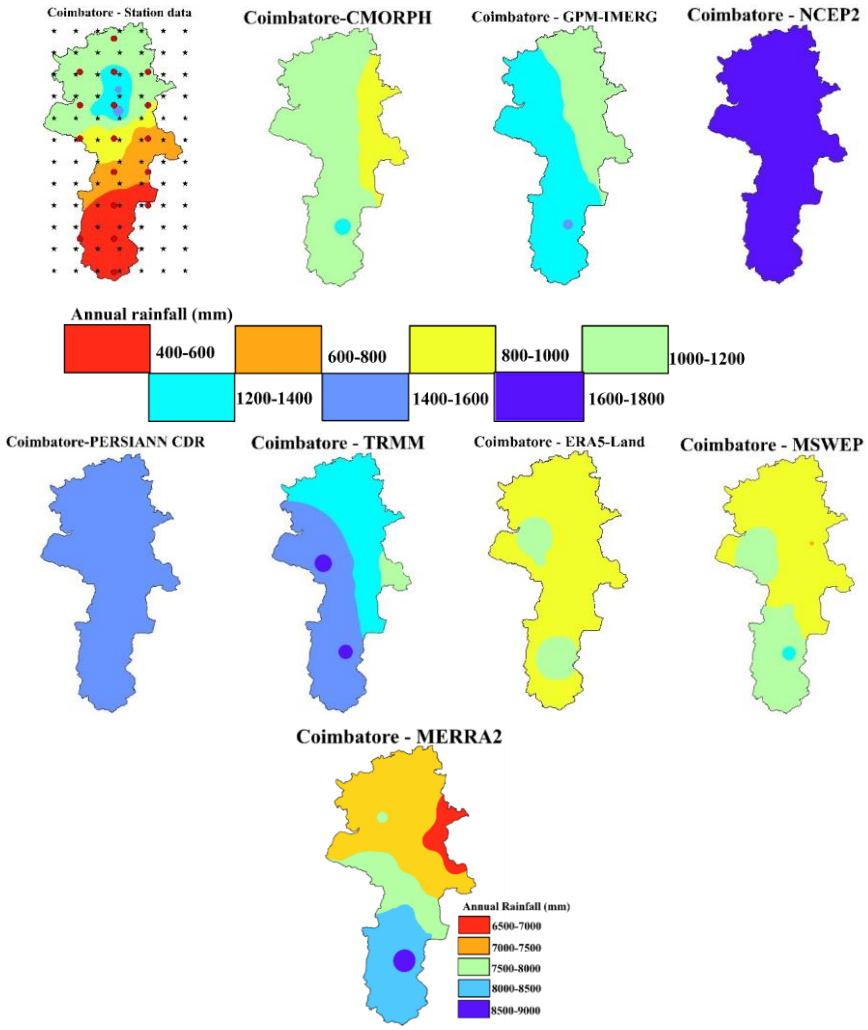
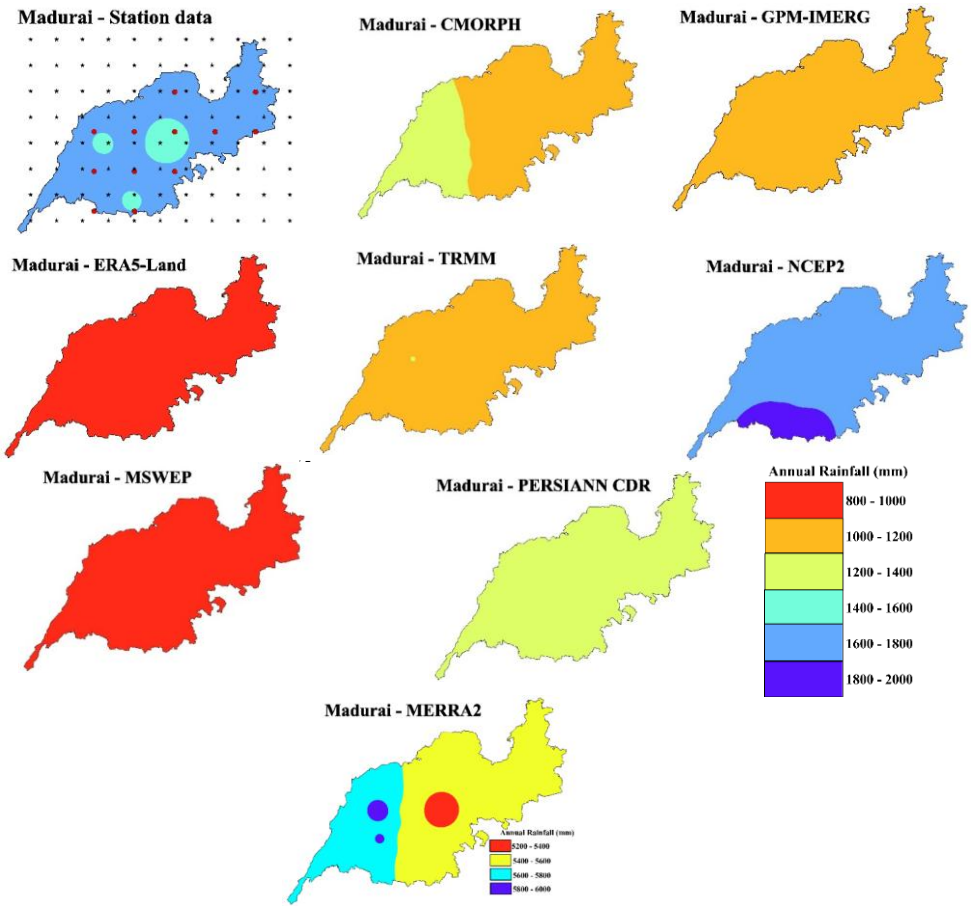
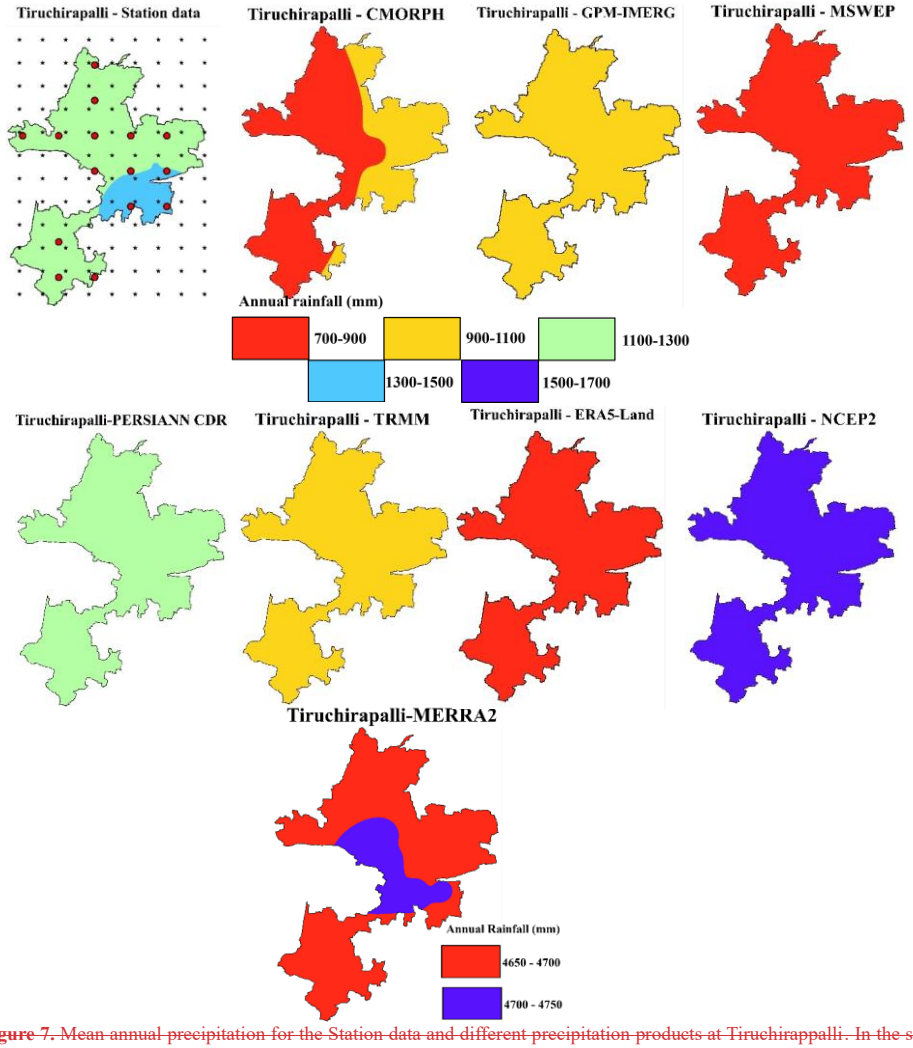


Figure 5. Mean annual precipitation for the Station data and different precipitation products at Coimbatore. In the station data spatial map, red dots represent the station data, and black stars denote the linearly interpolated grids.



675

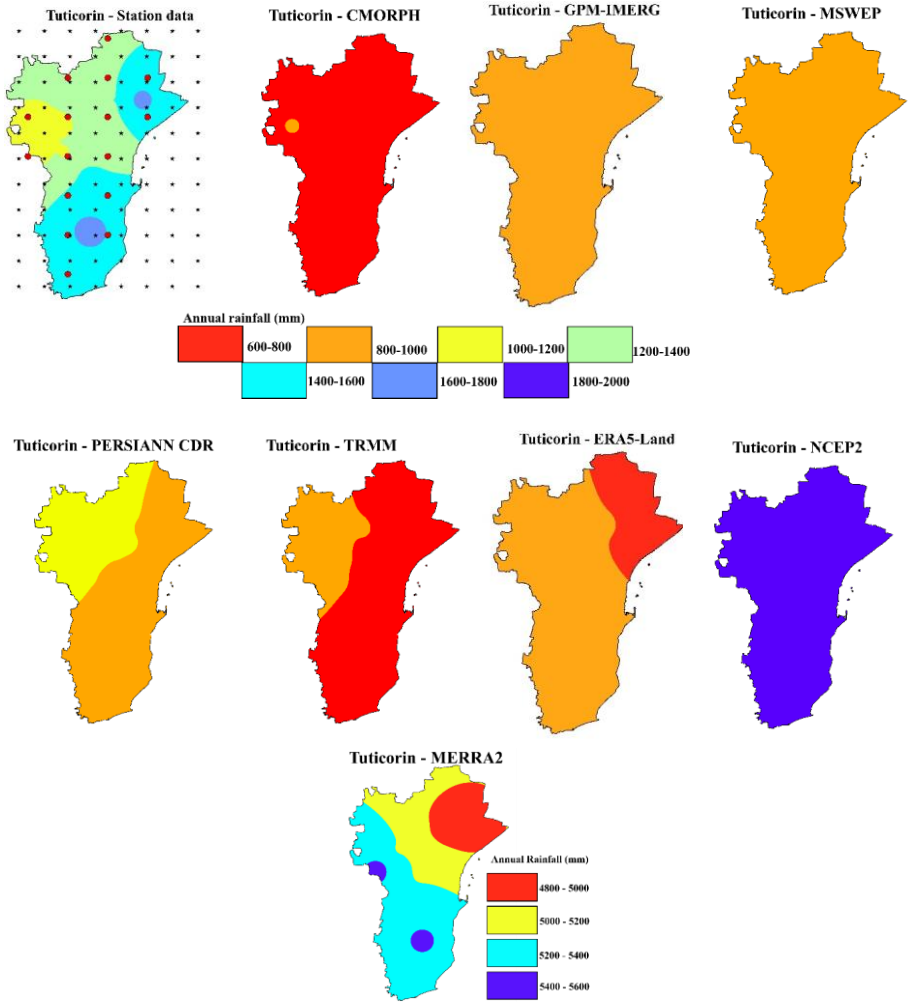
680 **Figure 6.** Mean annual precipitation for the Station data and different precipitation products at Madurai. In the station data spatial map, red dots represent the station data, and black stars denote the linearly interpolated grids.



695

700

Figure 7. Mean annual precipitation for the Station data and different precipitation products at Tiruchirappalli. In the station data spatial map, red dots represent the station data, and black stars denote the linearly interpolated grids.



~~**Figure 8.** Mean annual precipitation for the Station data and different precipitation products at Tuticorin. In the station data spatial map, red dots represent the station data, and black stars denote the linearly interpolated grids.~~

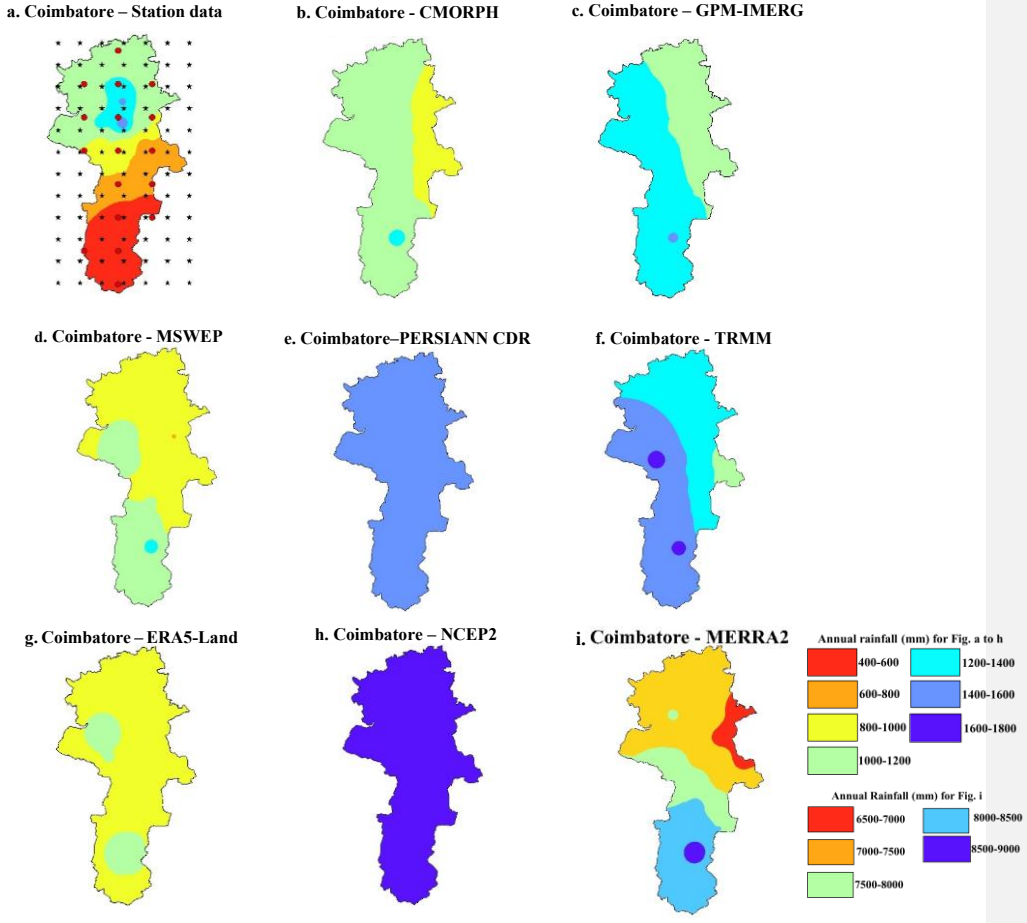


Figure 5. Mean annual precipitation for the Station data and different precipitation products at Coimbatore. In the station data spatial map, the distribution of ground station points and their respective linearly interpolated grids are plotted to understand the precipitation variation across the grids. The red circular dots represent the locations of ground stations from which precipitation data was collected for the period 2003–2014. The black stars indicate the linearly interpolated 0.1° grids. For the evaluation, only grids surrounded by at least one rain gauge were considered.

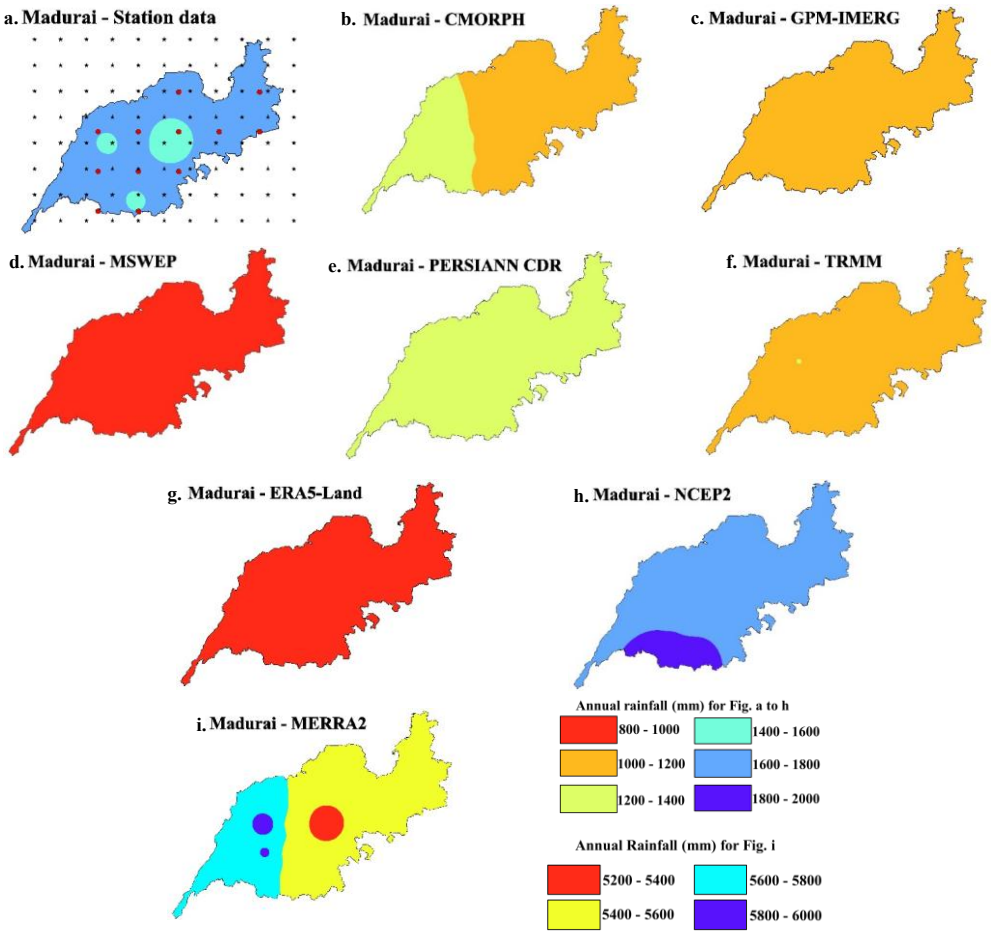


Figure 6. Mean annual precipitation for the Station data and different precipitation products at Madurai. In the station data spatial map, the distribution of ground station points and their respective linearly interpolated grids are plotted to understand the precipitation variation across the grids. The red circular dots represent the locations of ground stations from which precipitation data was collected for the period 2003–2014. The black stars indicate the linearly interpolated 0.1° grids. For the evaluation, only grids surrounded by at least one rain gauge were considered.

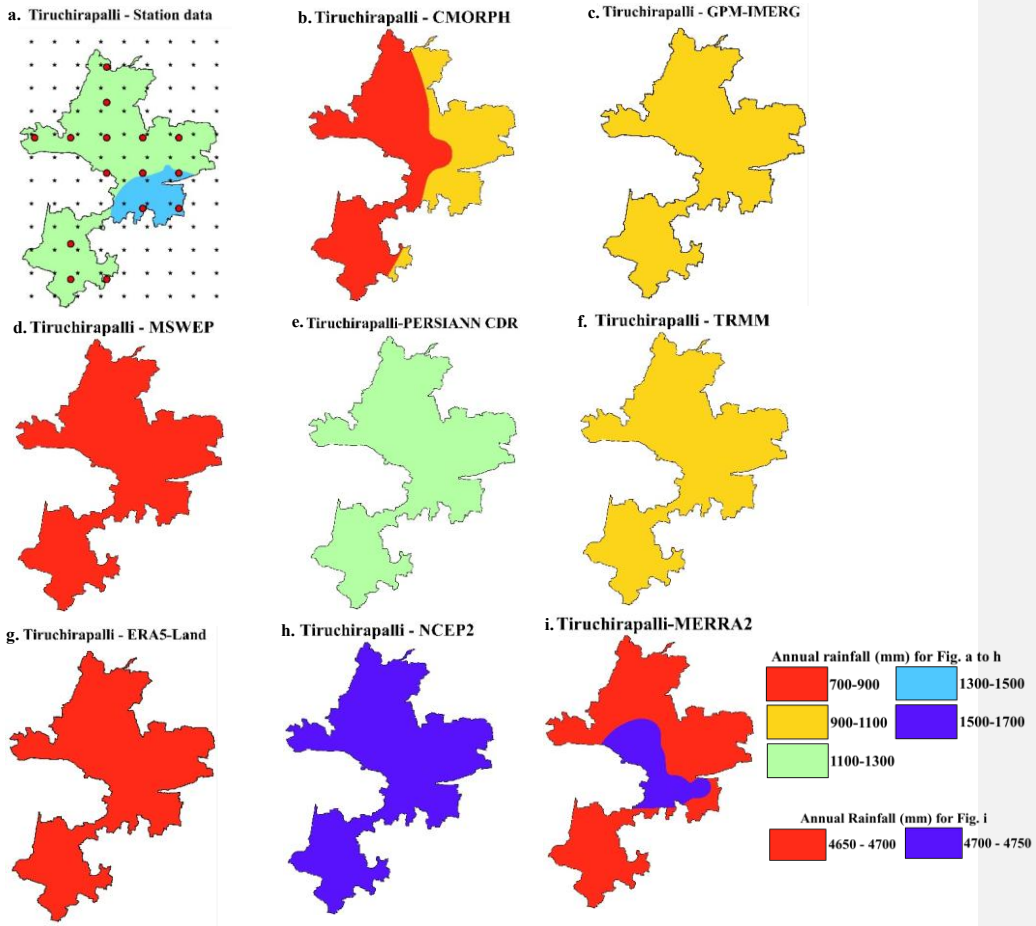


Figure 7. Mean annual precipitation for the Station data and different precipitation products at Tiruchirappalli. In the station data spatial map, the distribution of ground station points and their respective linearly interpolated grids are plotted to understand the precipitation variation across the grids. The red circular dots represent the locations of ground stations from which precipitation data was collected for the period 2003–2014. The black stars indicate the linearly interpolated 0.1° grids. For the evaluation, only grids surrounded by at least one rain gauge were considered.

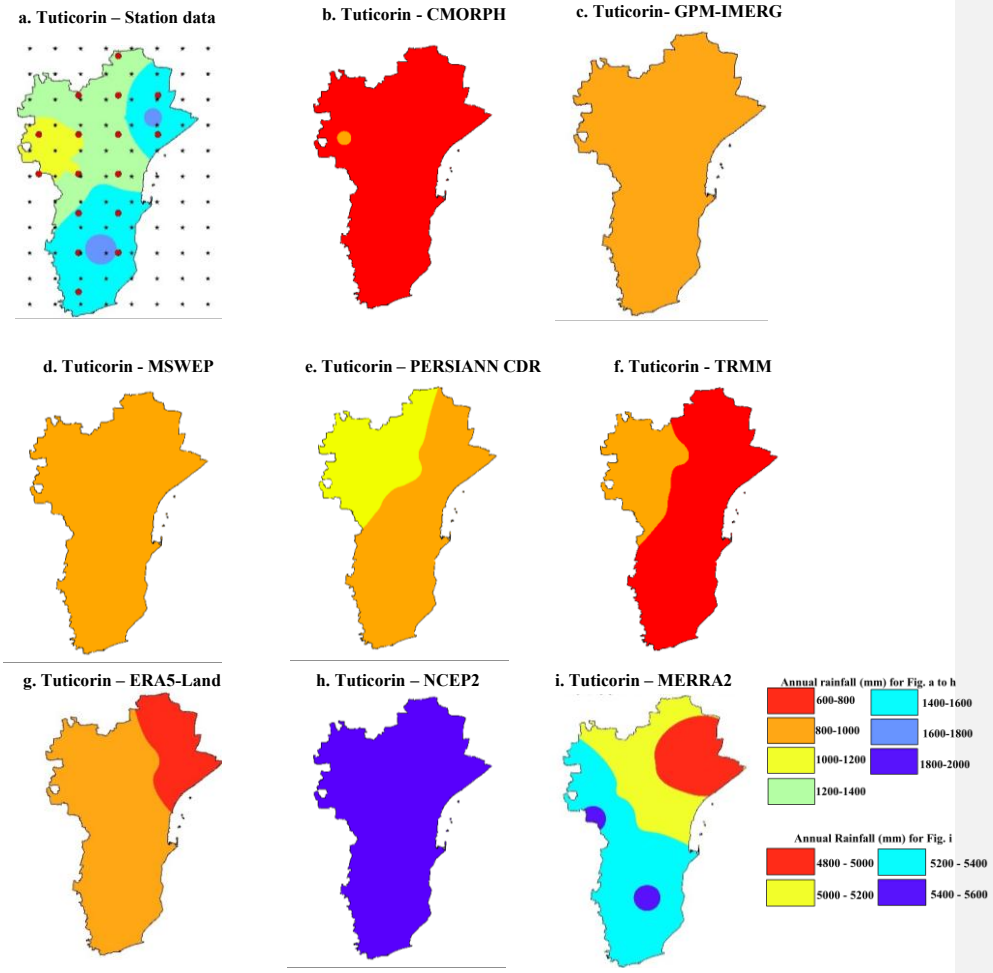


Figure 8. Mean annual precipitation for the Station data and different precipitation products at Tuticorin. In the station data spatial map, the distribution of ground station points and their respective linearly interpolated grids are plotted to understand the precipitation variation across the grids. The red circular dots represent the locations of ground stations from which precipitation data was collected for the period 2003–2014. The black stars indicate the linearly interpolated 0.1° grids. For the evaluation, only grids surrounded by at least one rain gauge were considered.

5 Discussion

Comparing all the statistical metrics, ERA5-Land, followed by MSWEP, is the best precipitation product in Coimbatore, Madurai and Tiruchirappalli. ERA5-Land is attributed as the best-performing product because of its higher CC combined with lower RMSE. In a similar study on the Godavari river basin, which covers the neighbouring states, Reddy and Saravanan (2023) found that MSWEP performed better than other products during the monsoon and pre-monsoon seasons. ERA5-Land performs better than other products as it was developed based on a replay of the land components with the H-TESSEL model. ERA5-Land is the land portion of ERA5 reanalysis data, which provides more accurate land parameters and better characterizes land status with its advanced 0.1° resolution (Muñoz-Sabater et al., 2021). Further, PERSIANN CDR had lower MAE and RMSE in the three regions but substantially lower CC and IA. When the remaining precipitation products, GPM and TRMM—are considered, it is shown that GPM has a stronger correlation with less error than TRMM. GPM-IMERG has the provision of high-frequency channels contributed by GMI (GPM Microwave Imager) – 166 GHz and 183 GHz. Thus, GPM-IMERG provides significantly improved spatial resolution in covering moderate to heavy precipitation over TRMM (Beria et al., 2017; Shen et al., 2022). In accordance with our findings, Venkatesh et al. (2021) also reported that GPM, TRMM and PERSIANN CDR produced a higher monthly correlation coefficient in comparison with the station data in Tamil Nadu. The study also reported that PERSIANN CDR produced the minimum RMSE in December. Of all the precipitation products examined in this study, only CMORPH exhibited good performance at Coimbatore during the monsoon season.

Tuticorin's precipitation product performance differed from the other research regions since it is a low-lying area with less precipitation (Table 1 and 2). MSWEP performed the best in Tuticorin, followed by PERSIANN CDR. The MSWEP products consistently exhibited good performance for daily rainfall over the Indian subcontinent, as shown by a similar analysis conducted by Nair and Indu (2017). This indicates that the product performed best even without station data-based bias correction (Table 2). MSWEP takes advantage of the complementary strengths of gauge, satellite and reanalysis-based data to provide reliable precipitation estimates. Further, using Cumulative distribution functions and precipitation frequency corrections in the present version overcomes the previous limitations.

In all four study locations, MERRA2 and NCEP2 had the highest MAE and RMSE. They performed the worst by overestimating the precipitation in all studied regions, regardless of the spatial and temporal scale of evaluation. This could result from the original product's availability at a coarser spatial resolution than that of other products (Table 2). Similar overestimations were also reported by Gupta et al., (2024) during the summer month precipitation in Tamil Nadu and Andhra Pradesh.

The performance of precipitation products in representing extreme precipitation events varied from their overall performance. In all study regions except Tiruchirappalli, ERA5-Land underestimated the intensity of extreme rainfall. It struggles to detect precipitation intensity of more than 1 mm/hour in Coimbatore and more than 5 mm/hour in Madurai as observed in Fig. 2. This underestimation might probably be due to the inability of the product to capture lightning-associated

heavy rainfall, which was reported in ERA5 by Kumar et al. (2024). Overall, MSWEP provided the most accurate estimates that were closest to the station data in most study regions. The study by Yaswanth et al. (2023) also pointed out that satellite-based precipitation products had a negative bias in capturing the peak discharge, particularly in the Adyar River basin of Tamil Nadu. Furthermore, GPM-IMERG, which closely captured station data in Madurai, was observed to produce Peninsular India's accurate estimates of heavy rainfall at the grid level during the 2015 winter monsoon by Singh et al., (2018).

The monsoon month precipitation in all locations, except Coimbatore, was notably underestimated by the precipitation products. Table 2 indicates that most algorithms used by the products combine high-temporal-frequency Visible/Infra-Red (V/IR) estimations with high-spatial-resolution Passive Microwave (MW) estimations. To address sampling issues caused by low orbits of current microwave sensors, passive microwave observations and infrared geostationary measurements are integrated. The continuous measurement provided by the infrared images effectively tracks system evolution, while the microwave rain retrieval serves as a reference (Sorooshian et al., 2002; Tapiador et al., 2004). The study areas, located in tropical zones, are prone to convection-driven precipitation. In these regions, convection events can result in rain rates exceeding 50 mm/hour, lasting anywhere from less than 30 minutes to up to 6 hours (Prigent, 2010). Despite the V/IR-MW integration being designed to capture light rainfall, it still leads to significant underestimations of precipitation in these areas. Coimbatore, however, being at a relatively higher altitude, is also subject to orographic precipitation (Table 1 and Fig. 1), leading to comparatively better performance of the precipitation products.

Moreover, based on the annual spatial plots, it is evident that the precipitation products could not accurately represent the spatial variability and generally underestimated the amount of precipitation. This aligns with similar observations in eastern China, where satellite-based precipitation products also underestimated precipitation and failed to capture the variability in the complex terrain and rainy areas of the south (Shaowei et al., 2022).

This study offers essential input for agro-hydrological modelling in the study region, addressing gaps left by previous research. Beyond the standard statistical metrics, the study incorporates additional criteria for evaluating precipitation products. In the context of climate change and the associated rise in flood occurrences, Intensity-Duration-Frequency (IDF) analysis is essential for selecting appropriate products for flood modelling. Frequency plot analysis was performed to assess the differential capability of various products in capturing both low and high-intensity precipitation on a daily scale. Comparing the mean monthly values for monsoon and non-monsoon seasons will aid future studies in selecting the most suitable precipitation product for each season. Additionally, district-scale spatial plots on a yearly scale provide insights into the variability of precipitation across different parts of the districts, as recorded by ground station data, and evaluate the products' accuracy in capturing this variability.

In addition to directly integrating precipitation products for agro-hydrological modelling, many studies also make use of pre-compiled datasets. For example, in regions with limited ground stations, numerous agrometeorological studies rely on NASA Power for hydrometeorological modelling, yield predictions, and other analyses. As NASA Power is based on MERRA2 datasets, it is important to note that using it without bias correction can lead to significant errors, as the product tends to overestimate, as evident from the analysis of the study. Therefore, the findings of this study can help in selecting

appropriate precipitation products and pre-compiled precipitation datasets, identifying biases at different time scales, and choosing the right bias correction methods. The main limitations of this study include the lack of access to station-based precipitation datasets after 2014 and the unavailability of hourly precipitation datasets for sub-daily analysis. Future studies could also examine the impact of different bias correction methods on enhancing the performance of various precipitation products.

6 Conclusion

In this study, the performance of multiple satellite-based (CMORPH, GPM-IMERG, MSWEP, PERSIANN CDR and TRMM) and reanalysis-based (ERA5-Land, MERRA2 and NCEP2) precipitation products were statistically evaluated against the station data for the period 2003–2014. The evaluations were conducted at two spatial (Grid and District) and three temporal (Daily, Monthly and Yearly) scales. In addition to the statistical analysis, the frequency of extreme precipitation events on a daily scale, monsoon and non-monsoon season precipitation variability on a monthly scale, and spatial precipitation variability on a yearly scale were also analyzed. ERA5-Land performed the best in Coimbatore, Madurai and Tiruchirappalli, whereas MSWEP performed well at Tuticorin. The accuracy of precipitation products was better at monthly and yearly scales as compared to the daily scale. Further, the increase of spatial scale from grid to district did not strongly affect the performance of precipitation products. The ability of precipitation products to capture extreme rainfall intensity is different from the statistical metrics. ERA5-Land underestimated the extreme rainfall intensity at all the study regions except in Tiruchirappalli. MSWEP captured the closest rainfall intensity in Coimbatore and Madurai, PERSIANN CDR in Tiruchirappalli and ERA5-Land in Tuticorin. Most products underestimated the mean monthly precipitation in monsoon and non-monsoon months. Additionally, the %MAE was higher in non-monsoon months, indicating that these product-based projections for water-scarce periods may be less reliable. The present precipitation algorithm is unable to capture the convective-based low precipitation events in the study regions (Madurai, Tiruchirappalli and Tuticorin), leading to underestimation. Hence, algorithm improvisation or prior bias corrections are needed before considering the evaluated products for future modelling studies.

The evaluation's findings help select the precipitation product for the non-gauged area in agro-hydrological modelling. The study results can be useful for the districts falling within the same agro-climatic regions. The state of Tamil Nadu is divided into 7 agroclimatic zones. The present study includes 3 important agricultural zones. The study results can be transferred to areas in similar agro-climatic zones. This includes other districts within the same agroclimatic zones in Tamil Nadu, for which less ground station data is available. The findings of this study are designed to support field-level experimentation and also provide a proof of concept for modellers developing climate data products, with the potential for extrapolation to other regions with similar agro-climatic conditions. Additional accuracy and the selection of a better product for future analysis can be obtained from the event-level assessment of extreme precipitation events, monsoon, and non-monsoon precipitation. Precipitation is the major source of soil moisture, which in turn drives food production. Identifying the accurate precipitation is one of the underlying factors of the nexus between food production (agriculture), water and soil

management. Hence, the study has significant implications for managing the nexus between food production and water and soil management in agricultural regions.

Data availability. In this study, publicly available daily precipitation data are retrieved from multiple climate data sources, as provided below. Satellite based precipitation product used in this study are the CPC MORPHing technique(CMORPH) available at National Centres for Environmental Information-Climate Data Records (NCEI-CDR) library (Index of /data/cmorph-high-resolution-global-precipitation-estimates/access/daily (noaa.gov)), Global Precipitation Measurement (GPM) and Tropical Rainfall Measuring Mission (TRMM) available at NASA Earth Data search library (Earthdata Search | Earthdata Search (nasa.gov)), Multi-Source Weighted-Ensemble Precipitation (MSWEP) available at GloH2O library (MSWEP - GloH2O) , Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks – Climate Data Records (PERSIANN CDR) available at NCEI-CDR library (Index of /data/precipitation-persiann/access (noaa.gov)), European Centre for Medium Range Weather Forecast (ECMWF) Reanalysis version 5 Land (ERA5-Land) available at Copernicus Climate Change Service (C3S) Climate Data Store (ERA5-Land hourly data from 1950 to present (copernicus.eu)), Modern-Era Retrospective analysis for Research and Application version 2 (MERRA 2) available at Global Modeling and Assimilation Office library(https://disc.gsfc.nasa.gov/datasets?page=1&subject=Precipitation&source=Models%20MERRA-2&project=MERRA-2&temporalResolution=1%20day), National Center for Environmental Prediction Reanalysis 2 (NCEP2) available at NOAA Physical Science Laboratory (NCEP/DOE Reanalysis II: NOAA Physical Sciences Laboratory NCEP/DOE Reanalysis II).

Author contributions. AS, DK, and NS designed the research. AS performed the data analysis and prepared the manuscript with substantial contributions from DK, NS and SR.

Competing interests. The authors declare that they have no conflict of interest.

References

Adler, R.F., Huffman, G.J., Chang, A., Ferraro, R., Xie, P.P., Janowiak, J., Rudolf, B., Schneider, U., Curtis,S., Bolvin,D., Gruber, A., Susskind, J., Arkin, P. and Nelkin, E.: The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979-present). *J. Hydrometeorol*. 4, 1147–1167, 2003.

Arjune, S., and Kumar, V. S.: Precision Agriculture: Influencing factors and challenges faced by farmers in delta districts of Tamil Nadu. In 2022 OPJU International Technology Conference on Emerging Technologies for Sustainable Development (OTCON) (pp. 1-6). IEEE, 2023.

Arumugam, S., Ashok, K. R., Kulshreshtha, S. N., Vellangany, I., and Govindasamy, R.: Yield variability in rainfed crops as influenced by climate variables: A micro level investigation into agro-climatic zones of Tamil Nadu, India. International Journal of Climate Change Strategies and Management 7(4):442-459, 2013.

Ashouri, H., Hsu, K.L., Sorooshian, S., Braithwaite, D.K., Knapp, K.R., Cecil, L.D., Nelson, B. R. and Prat, O. P.: PERSIANN-CDR Daily Precipitation Climate Data Record from Multisatellite Observations for Hydrological and Climate Studies. Bull. Am. Meteorol. Soc. 96 , 69-71, 2015.

Balaganesh, G., Malhotra, R., Sendhil, R., Sirohi, S., Maiti, S., Ponnusamy, K., and Sharma, A. K.: Development of composite vulnerability index and district level mapping of climate change induced drought in Tamil Nadu, India. Ecol. Indic., 113, 106197, 2020.

Beck, H. E., van Dijk, A. I. J. M., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., and de Roo, A.: MSWEP: 3-hourly 0.25◦ global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data, Hydrol. Earth Syst. Sci., 21, 589–615, https://doi.org/10.5194/hess-21-589-2017, 2017.

Beria, H., Nanda, T., Singh Bisht, D., and Chatterjee, C.: Does the GPM mission improve the systematic error component in satellite rainfall estimates over TRMM? An evaluation at a pan-India scale, Hydrol. Earth. Syst. Sci., 21, 6117–6134, https://doi.org/10.5194/hess-21-6117-2017, 2017.

Bosilovich, R., Lucchesi., and Suarez.: MERRA-2: File specification. NASA GMAO Office Note 9 (version 1.1), 75 pp, https://gmao.gsfc.nasa.gov/pubs/docs/ Bosilovich785.pdf., 2016.

Bosilovich, M. G., Lucchesi, R., and Suarez, M.: MERRA-2: File specification, GMAO Office Note No. 9 (Version 1.0), Microsoft Word - GMAO-OfficeNote-09-V1-M2FileSpec-Final.docx, 2015.

Chandrasekar, K., Sesha Sai, M. V., Roy, P. S., Jayaraman, V., and Krishnamoorthy, R.: Identification of Agricultural Drought Vulnerable Areas of Tamil Nadu, India using GIS-based Multi-Criteria Analysis, Asian Journal of Environment and Disaster Management 1(1): 40-61, 2009.

Compa, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., Vose, R. S., Rutledge, G., Bessemoulin, P., Bronnimann, S., Brunet, M., Crouthamel. R. I., Grant, A. N., Groisman, P. Y., Jones, P. D., Kruk, M. C., Kruger, A.C., Marshall, G. J., Maugeri, M., Mok. H. Y., Nordli, O., Ross, T. F., Trigo, R. M., Wang, X. L.,

Woodruff, S. D., and Worley, S. J.: The Twentieth Century Reanalysis Project, Q. J. R. Meteorol. Soc.,,137:1-28, 2011.

Cullather, R. I., Andrews, L. C., Croteau, M. J., Digirolamo, N. E., Hall, D. K., Lim, Y.-K., Loomis, B, D., Shuman, C. A., and Nowicki, S. M. J.: Anomalous circulation in July 2019 resulting in mass loss on the Greenland Ice Sheet. Geophys. Res. Lett., 47, e2020GL087263. https://doi.org/10.1029/2020GL087263, 2020.

Dhar, O. N., Rakhecha, P. R., and Mandal, B. N.: Some facts about Indian rainfall – A brief appraisal from hydrological considerations. Indian. J. Power. River. Val. Dev., 31, Nos. 7,8, 117-125, 1981.

Duan, Z., Liu, J., Tuo, Y., Chiogna, G., and Disse, M.: Evaluation of eight high spatial resolution gridded precipitation products in Adige Basin (Italy) at multiple temporal and spatial scales, Sci. Total Environ., https://doi.org/10.1016/j.scitotenv.2016.08.213, 573, 1536–1553, 2016.

Dubey, S., Gupta, H., Goyal, M. K. and Joshi, N.: Evaluation of precipitation datasets available on Google earth engine over India, Int. J. Climatol., 41(10), 4844-4863, 2021.

Ebita, A., Kobayashi, S., Ota, Y., Moriya, M., Kumabe, R., Onogi, K., Harada, Y., Yasui, S., Miyaoka, K., Takahashi, K., Kamahori, H., Kobayashi, C., Endo, H., Soma, M., Oikawa, Y., and Ishimizu, T.: The Japanese 55-year reanalysis "JRA-55": an interim report. Sola, 7, 149-152, 2011.

Gardas, B. B., Raut, R. D., & Narkhede, B.: Evaluating critical causal factors for post harvest losses (PHL) in the fruit and vegetables supply chain in India using the DEMATEL approach, J. Clean. Prod., 199, 47-61, 2018.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M. and Zhao, B.: The modern-era retrospective analysis for research and applications, version 2 (MERRA-2), J. Clim., 14, 5419-5454, doi:10.1175/JCLI-D-16-0758.1, 2017.

Government of Tamil Nadu: Statistical Handbook 2020-21: Sub chapter: Rainfall, 2022.

Gupta, A., Jain, M. K., Pandey, R. P., Gupta, V., and Saha, A.: Evaluation of global precipitation products for meteorological drought assessment with respect to IMD station datasets over India, Atmos. Res., 297: 107104. DOI: https://doi.org/10.1016/j.atmosres.2023.107104, 2024.

Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow,C. D., Kojima, M., Oki R., Nakamura, K., and Iguchi, T.,: TheGlobal Precipitation Measurement Mission, B. Am. Meteorol.Soc., 95, 701–722, 2014.

Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., Hong, Y., Stocker, E. F., and Wolff, D. B.: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales, J. Hydrometeorol., 8:38-55, 2007.

Huffman, G. J., Adler, R. F., Bolvin, D. T., and Gu, G.: Improving the global precipitation record: GPCP Version 2.1, Geophys. Res. Lett., 36, L17808, https://doi.org/10.1029/2009GL040000, 2009.

Huffman, G. J., and, Pendergrass, A.: The Climate Data Guide: TRMM: Tropical Rainfall Measuring Mission, National Center for Atmospheric Research Staff (Eds), Retrieved from https://climatedataguide.ucar.edu/climate-data/trmm-tropical-rainfall-measuring-mission on 2024-01-01, 2023.

IPCC: Summary for Policymakers, in: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Core Writing Team, Lee, H., and Romero, J., IPCC, Geneva, Switzerland, 1–34, https://doi.org/10.59327/IPCC/AR69789291691647.001, 2023.

Joyce, R. J., Janowiak, J. E., Arkin, P. A., and Xie, P.: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. J. Hydrometeorol., 5(3), 487-503, 2004.

Kanamitsu, M., Ebisuzaki, W., Woollen, J., Potter, J., Fiorion, M.: An overview of NCEP/DOE Reanalysis-2. Proc. Conf. 2ndIntl. Conf. On Reanalyses, Reading, England, 23–27 Aug. 1999,WCRP, vol. 109. WMO, Geneva, pp. 1 – 4. WMO/TD 985, 2000.

Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S.-K., Hnilo, J. J., Fiorino, M., and Potter, G. L.: Ncep–doe amip-ii reanalysis (r-2), Bull. Am. Meteorol. Soc., 83, 1631–1644, 2002.

Kolluru, V., Kolluru, S., and Konkathi, P: Evaluation and integration of reanalysis rainfall products under contrasting climatic conditions in India, Atm. Res., 246, 105121, 2020.

Kucera, P.A., Ebert, E.E., Turk, F.J., Levizzani, V., Kirschbaum. D., Tapiador. F.J., Loew. A. and Borsche, M.: Precipitation from space: advancing Earth system science. Bull. Amer. Meteor. Soc. 94:365-375, https://doi.org/10.1175/BAMS-D-11-00171.1, 2013.

Kumar, P., Srivastava, S. S., Jivani, N., Varma, A. K., Yokoyama, C., and Kubota, T.: Long-term assessment of ERA5 reanalysis rainfall for lightning events over India observed by Tropical Rainfall Measurement Mission Lightning Imaging Sensor. *Quarterly Journal of the Royal Meteorological Society*, 150, 2472 – 2488, 2024.

Lakshmi, S., Nivethaa, E.A.K., Ahamed Ibrahim, S. N., Ramachandran, A., and Palanivelu, K.: Prediction of Future Extremes during Northeast monsoon in the Coastal District of Tamil Nadu state in India based on ENSO. Pure Appl. Geophys 178: 3207-3228, 2021.

Lal, M.: Global Climate Change – India's Monsoon and its variability, Journal of Environmental Studies and Policy 6(1):1-34, 2003.

Lalmuanzuala, B., Sathyamoorthy, N., Kokilavani, S., Jagadeeswaran, R., and Kannan, B.: Drought analysis in southern region of Tamil Nadu using meteorological and remote sensing indices. MAUSAM, 74(4), 973-988, 2023.

Legates, D. R. and McCabe, G. J.: Evaluating the use of "goodness-of-fit" Measures in hydrologic and hydroclimatic model validation, Water Resour. Res., 35, 233–241, https://doi.org/10.1029/1998WR900018, 1999.

Liu, J., Duan, Z., Jiang, J., and Zhu, A.X.: Evaluation of three satellite precipitation products TRMM 3B42, CMORPH, and PERSIANN over a subtropical watershed in China, Adv. Meteorol, 2015 https://doi.org/10.1155/2015/151239, 2015.

Loucks, D. P., and Van Beek, E. Water resource systems planning and management: An introduction to methods, models, and applications. Springer, 2017.

Malaiarasan, U., Paramasivam, R., and Felix, K. T.: Crop diversification: determinants and effects under paddy-dominated cropping system. Paddy and Water Environment, 19, 417-432, 2021.

McDonald, M. G., and Harbaugh, A. W.: A modular three-dimensional finite-difference ground-water flow model. U.S. Geological Survey Open-File Report, 83–875, 528 p, 1984.

MERRA 2: Land Surface Precipitation in MERRA-2 in: Journal of Climate Volume 30 Issue 5 (2017) (ametsoc.org).

Molod, A., Takacs, L., Suárez, M., and Bacmeister J.: Development of the GEOS-5 atmospheric general circulation model: Evolution from MERRA to MERRA2. Geosci. Model. Dev., 8, 1339–1356, doi:10.5194/gmd-8-1339-2015, 2015.

Molod, A., Takacs, L., Suarez, M., Bacmeister, J, Song, I. S. and Eichmann, A.: The GEOS-5 atmospheric general circulation model: Mean climate and development from MERRA to Fortuna. NASA/ TM-2012-104606, 28, 117 pp, https://gmao.gsfc.nasa.gov/pubs/docs/tm28.pdf, 2012.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J.-N.: ERA5-Land: a state-of-the-art global reanalysis dataset for land applications, Earth Syst. Sci. Data, 13, 4349–4383, https://doi.org/10.5194/essd-13-4349-2021, 2021.

Nair, A. S., and Indu, J.: Performance assessment of multi-source weighted-ensemble precipitation (MSWEP) product over India. Climate, 5(1), 2, 2017.

Neitsch, S. L., Arnold, J. G., and Srinivasan, R.: Pesticides fate and transport predicted by the soil and water assessment tool (SWAT). Atrazine, Metolachlor and Trifluralin in the Sugar Creek Watershed: BRC Report, 3, 2002.

Nikulin, G., Jones, C., Giorgi, F., Asrar, G., Buechner, M., Cerezo-Mota, R., Christensen, O. B., Deque, M., Fernandez, J., Haensler, A., van Meijgaard, E., Samuelsson, P., Sylla, M. B.,and Sushama, L.: Precipitation Climatology in an Ensemble of CORDEX-Africa Regional Climate Simulations, J. Climate, 25, 6057–6078, https://doi.org/10.1175/JCLI-D-11-00375.1, 2012.

Pan, H.-L., Wu, W. S.: Implementing a mass flux convection parameterization package for the NMC medium-range forecast model. Tenth Conf. on Numerical Weather Prediction. Amer. Meteor. Soc., Portland, OR, pp. 96 – 98, 1994.

Paramasivam, P.: Hundreds stranded as parts of India's Tamil Nadu flooded after heavy rain, Reuters: Hundreds stranded as parts of India's Tamil Nadu flooded after heavy rain | Reuters, 2023.

Prigent, C.: Precipitation retrieval from space: An overview. C R. Geoscience 342: 380-389, 2010.

Radhakrishnan, S., Duraisamy Rajasekaran, S. K., Sujatha, E. R., and Neelakantan, T. R.: A Comparative Study on 2015 and 2023 Chennai Flooding: A Multifactorial Perspective. Water, 16(17), 2477, 2024.

Rajkumar, R., Kumaran, M. M., Prasath, M. R. K., and Senthilshanmugavelayuthm, S. P.: Agro-Climatic Zonation for Tamil Nadu Using GIS and AHP Techniques. IOP Conf. Ser.: Mater. Sci. Eng 1006: 012015, 2020.

Randles, C. A., da Silva, A.M., Buchard, V., Darmenov, A., Colarco, P.R., Aquila, V.,Bian, H., Nowottnick, E.P., Pan, X., Smirnov, A., Yu, H.: The MERRA-2 aerosol assimilation. Technical Report Series on Global Modeling and Data Assimilation, Vol. 45, NASA Tech. Rep. NASA/TM– 2016–104606, 143 pp, 2016.

Reichle, R.H., Draper, C.S., Liu, Q., Girotto, M., Mahanama, S.P., Koster, R.D., De Lannoy, G.J.: Assessment of MERRA-2 land surface hydrology estimates. J. Clim. 30, 2937–2960, 2017.

Renard, K.G., Foster, G.R., Weesies, G.A.: Predicting soil erosion by water: a guide to conservation planning with the Revised Universal Soil Loss Equatio (RUSLE). USDA agricultural handbook No. 703, 404 pp, 1997.

Renard, B., Kavetski, D., Kuczera, G., Thyer, M., and Franks, S. W.: Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors. Water Resources Research, 46(5). https://doi.org/10.1029/2009WR008328, 2010.

Rienecker, M. M., Suarez, M.J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G. K., Bloom, S., Chen, J., Collins, D., Conaty, A., da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C.R., Reichle, R., Robertson, F. R., Ruddlick, A. G., Seinkiewicz, M., and Woollen, J.: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. J. Climate,24, 3624–3648, doi:10.1175/JCLI-D-11-00015.1, 2011.

Saha, S., Moorthi, S., Pan, H.L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D. and Liu, H.: The NCEP climate forecast system reanalysis, Bull.Am. Meteorol. Soc.,91, 1015–1057, doi:10.1175/2010BAMS3001.1, 2010.

Sapiano, M. R. P. and Arkin, P. A.: An Intercomparison and Validation of High-Resolution Satellite Precipitation Estimates with 3-Hourly Gauge Data, J. Hydrometeorol., 10, 149–166, https://doi.org/10.1175/2008JHM1052.1, 2009.

Shaowei, N., Jie, W., Juliang, J., Xiaoyan, X., Yuliang, Z., Fan, S., and Linlin, Z.: Comprehensive evaluation of satellite-derived precipitation products considering spatial distribution differences of daily precipitation over eastern China, Journal of Hydrology: Regional Studies, 44: 101242: 2002.

Shen, Z., Yong, B., Yi, L., Wu, H., Xu, H.: From TRMM to GPM, how do improvements of post/near-real-time satellite precipitation estimates manifest?. Atmospheric Research 268: 106029, 2022.

Shukla, A.K., Ojha, C.S.P., Singh, R.P., Pal, L., Fu, D.: Evaluation of TRMM precipitation dataset over Himalayan Catchment: the upper Ganga Basin, India. Water (Switzerland), 11 (3). https://doi.org/10.3390/w11030613, 2019.

Singh, A. K., Singh, V., Singh, K. K., Tripathi, J. N., Kumar, A., Soni, A. K., Sateesh, M., and Khadke, C.: A case study: Heavy rainfall event comparison between daily satellite rainfall estimation products with IMD gridded rainfall over peninsular India during 2015 winter monsoon. Journal of the Indian Society of Remote Sensing, 46, 927-935, 2018.

Singh, A. K., Tripathi, J. N., Singh, K. K., Singh, V., and Sateesh, M.: Comparison of different satellite-derived rainfall products with IMD gridded data over Indian meteorological subdivisions during Indian Summer Monsoon (ISM) 2016 at weekly temporal resolution. *Journal of Hydrology*, *575*, 1371-1379, 2019.

Smith, E. A., Asrar, G., Furuhama, Y., Ginati, A., Mugnai, A., Nakamura, K., Adler, R. F., Chou, M.-D., Desbois, M., and Durning, J. F.: International global precipitation measurement (GPM) program and mission: An overview, Measuring 920 precipitation from space: EURAINSAT and the future, 611–653, 2007.

Sorooshian, S., Gao, X., Hsu, K., Maddox, R.A., Hong, Y., Gupta, H.V., Imam, B.: Diurnal variability of tropical rainfall retrieved from combined GOES and TRMM satellite information. J. Clim. 15, 983–1001, 2002.

Sorooshian, S., Hsu, K., Gao, X., Gupta, H. V., Imam, B., and Braithwaite, D.: Evaluation of PERSIANN system satellite-based estimates of tropical rainfall, B. Am. Meteorol. Soc., 81, 2035–2046, https://doi.org/10.1175/1520-0477(2000)081<2035:EOPSSE>2.3.CO;2, 2000.

Stackhouse, P.W.J., Westberg, D., Hoell, J.M., Chandler, W.S., Zhang, T.: Prediction of world-wide energy resource (POWER)-Agroclimatology methodology-(1.0^{O} longitude spatial resolution). Hampton, NASA Langley Research Centre, 2015.

Surendran, U., Raja, P., Jayakumar, M., and Subramoniam, S. R.: Use of efficient water saving techniques for production of rice in India under climate change scenario: A critical review. J. Clean. Prod., 309: 127272, 2021.

Sylla, M. B., Giorgi, F., Coppola, E., and Mariotti, L.: Uncertainties in daily rainfall over Africa: assessment of gridded observation products and evaluation of a regional climate model simulation, Int. J. Climatol., 33, 1805–1817, https://doi.org/10.1002/joc.3551, 2013.

Takacs, L. L., Suárez, M. J., and R. Todling, R.: Maintaining atmospheric mass and water balance in reanalyses. Quart. J. Roy. Meteor. Soc., 142, 1565–1573, doi:10.1002/qj.2763, 2016.

Tapiador, F.J., Kidd, C., Levizzani, V., Marzano, F.S.: A neural networks-based fusion technique to estimate half-hourly rainfall estimates at 0.18 resolution from satellite passive microwave and infrared data. J. Appl. Meteorol. 43, 576–594, 2004.

Tapiador, F.J., Turk, F.J., Petersen, W., Hou, A.Y., Garcia-Ortega, E., Machado, L.A.T.: Global precipitation measurement: methods, datasets and applications. Atmos. Res. 104, 70–97, 2012.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, J. Geophys. Res.-Atmos., 106, 7183–7192, https://doi.org/10.1029/2000JD900719, 2001.

TN-ENVIS.: http://tnenvis.nic.in/Database/TN-ENVIS_791.aspx, 2023.

Varadan, R. J., Kumar, P., Jha, G. K., Pal, S. and Singh, R.: An exploratory study on occurrence and impact of climate change on agriculture in Tamil Nadu, India., Theor. Appl. Climatol .127: 993-1010, 2017.

Samykannu, V., Pazhanivelan S., and Mrunalini, K.: Performance of satellite-based precipitation products using dense network of rain-gauge observations in Tamil Nadu. SKUAST Journal of Research 23(1): 53-59., 2021.

Wagener, T., and Montanari, A.: Convergence of approaches toward reducing uncertainty in predictions in ungauged basins. Water Resour. Res. 47 (6). http:// dx.doi.org/10.1029/2010wr009469, 2011.

Wheater, H., and Gober, P.: Water security in the Canadian Prairies: science and management challenges. Philos. T. R. Soc. A. 371. http://dx.doi.org/10.1098/Rsta. 2012.0409, 2013.

Wilby, R. L. and Yu, D.: Rainfall and temperature estimation for a data sparse region, Hydrol. Earth Syst. Sci., 17, 3937–3955, https://doi.org/10.5194/hess-17-3937-2013, 2013.

Wilby, R. L., & Dawson, C. W.: The statistical downscaling model: insights from one decade of application. International Journal of Climatology, 33(7), 1707-1719, 2013.

Willems, P., and de Lange, W.J.: Concept of technical support to science-policy interfacing with respect to the implementation of the European water framework directive. Environ. Sci. Policy 10, 464–473. http://dx.doi.org/10.1016/j.envsci.2007.03.006, 2007.

Willmott, C. J.: On the Validation of Models, Phys. Geogr., 2, 184–194, 1981.

Wu, W.-S., Purser, R. J., and Parrish, D. F.: Three-dimensional variational analysis with spatially inhomogeneous co-variances. Mon. Wea. Rev., 130, 2905–2916, doi:10.1175/ 1520-0493(2002)130,2905:TDVAWS.2.0.CO;2, 2002.

Xie, P., Arkin, P.A., and Janowiak, J.E.: CMAP: The CPC Merged Analysis of Precipitation. In: Levizzani, V., Bauer, P., Turk, F.J. (eds) Measuring Precipitation From Space. Advances In Global Change Research, vol 28. Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-5835-6_25, 2007.

Yaswanth, P., Kannan, B. A. M., Bindhu, V. M., Balaji, C., and Narasimhan, B.: Evaluation of Remote Sensing Rainfall Products, Bias Correction and Temporal Disaggregation Approaches, for Improved Accuracy in Hydrologic Simulations. Water Resources Management, 1-24, 2023.

Zambrano-Bigiarini, M., Nauditt, A., Birkel, C., Verbist, K., and Ribbe, L.: Temporal and spatial evaluation of satellite-based rainfall estimates across the complex topographical and climatic gradients of Chile, Hydrol. Earth Syst. Sci., 21, 1295–1320, https://doi.org/10.5194/hess-21-1295-2017, 2017.

Supplementary Tables and Figures

Table S1a. Geographic coordinates of the Meteorological ground stations

Geographic coordinates (°N,°E)			
Coimbatore	Madurai	Tiruchirapalli	Tuticorin
11.1869976,76.8228552	9.9585312,77.7565042	11.3384582,78.539969	9.3464424,78.0663815
11.0298278,76.8263478	9.8040348,77.62401	11.1891953,78.5362835	9.1953367,77.9152757
10.8831359,76.8228552	9.6471797,77.7588629	11.0307185,78.226701	9.1934939,78.0719097
10.4226101,76.8243661	9.9597106,77.9168974	11.0436178,78.383335	9.1898084,78.2248582
11.340301,76.9828428	9.8099316,77.9168974	11.019662,78.5391262	9.0387027,77.76417
11.1873525,76.9754718	9.6495384,77.9145386	11.041775,78.6947602	9.0350172,77.9115902
11.0288758,76.9773145	9.7993174,78.072573	11.034404,78.8384949	9.0350172,78.0719097
10.8796128,76.9883711	9.9561725,78.0725731	10.8888266,78.5381262	9.0258034,78.21933
10.7266643,76.9810001	9.8052142,78.0713938	10.8685563,78.6910747	8.880226,77.7604845
10.5774013,76.9791573	9.9585312,78.2306076	10.870399,78.8550797	8.8839115,77.913433
10.4207673,76.9865283	10.1059514,78.3803866	10.7248216,78.6855464	8.880226,78.0755952
10.2586051,76.9791573	9.9585312,78.385104	10.5718731,78.3796495	8.7328057,77.9171185
11.1855098,77.1339485	-	10.4170818,78.3814922	8.723592,78.0719097
11.0288758,77.1431623	-	10.721136,78.8513942	8.5798572,77.9097495
10.8796128,77.1394768	-	10.4189246,78.539969	8.5780145,78.0719097
10.7248216,77.1321058	-	-	8.4140095,77.913433
10.5847723,77.1339485	-	-	-

(Source: Public Works Department, Government of Tamil Nadu)

Leave One Out Cross Validation (LOOCV) for linear interpolation

We performed a LOOCV analysis on the linear interpolated data to check for bias or uncertainty in the data. LOOCV was performed on linear interpolation based on assuming that the value of a station was unknown. The unknown station was estimated from the value of the neighbouring stations based on linear interpolation. The analysis was systematically carried out in Python. The results of the LOOCV analysis represent the mean RMSE at a monthly timescale, which includes both monsoon and non-monsoon months. The RMSE and %MAE of LOOCV analysis produced lower values, indicating minimal uncertainty due to the linear interpolation. Hence, we proceeded with linear interpolation of station data before grid-to-grid comparison.

Table S1. LOOCV performance of Linear interpolation of precipitation at the monthly timescale

Study Region	Mean RMSE (mm)	Percent Mean Absolute Error (%)
Coimbatore	39.22	36
Madurai	52.50	30
Tiruchirappalli	39.93	29

Tuticorin	42.32	25

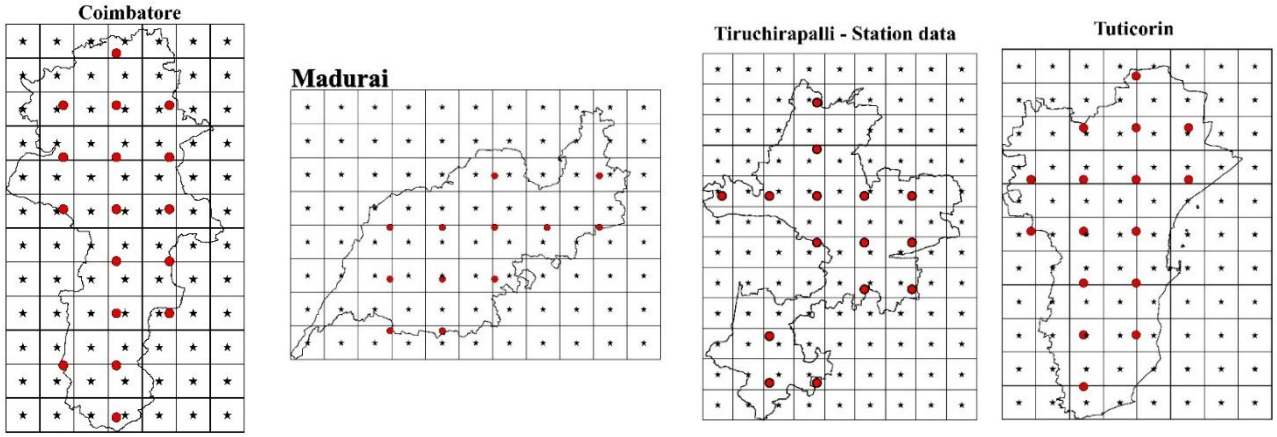
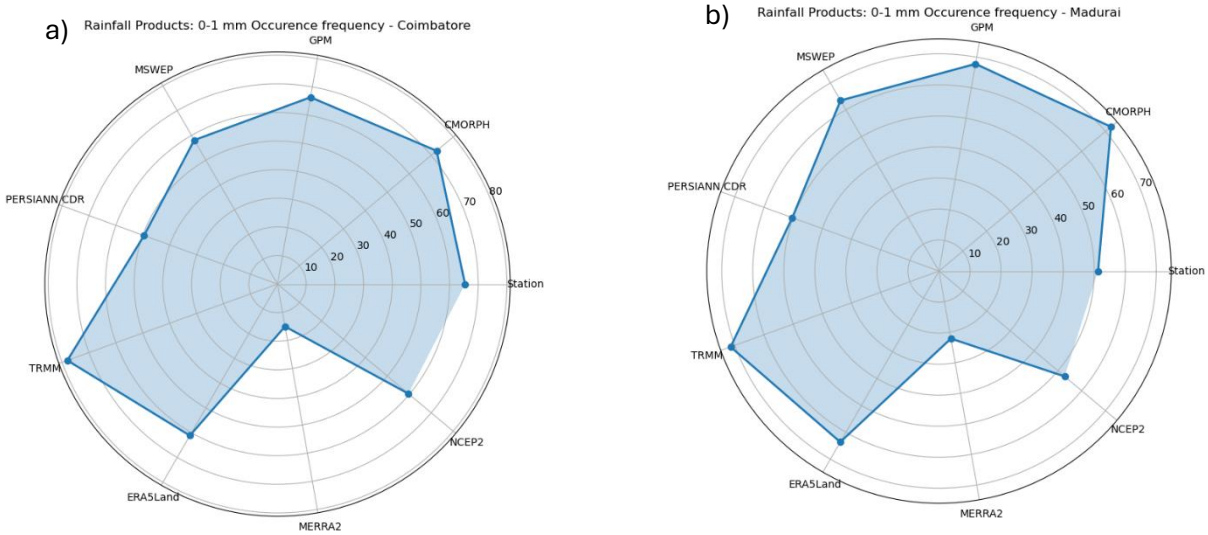


Figure S1: Linearly interpolated grids (0.1°) of the study regions along with the station data. In the figure above, red points denote the distribution of station datasets. The grids and the stars represent the linear interpolation. Only grids that are surrounded by at least one rain gauge station were included in the analysis to avoid uncertainties in the analysis.



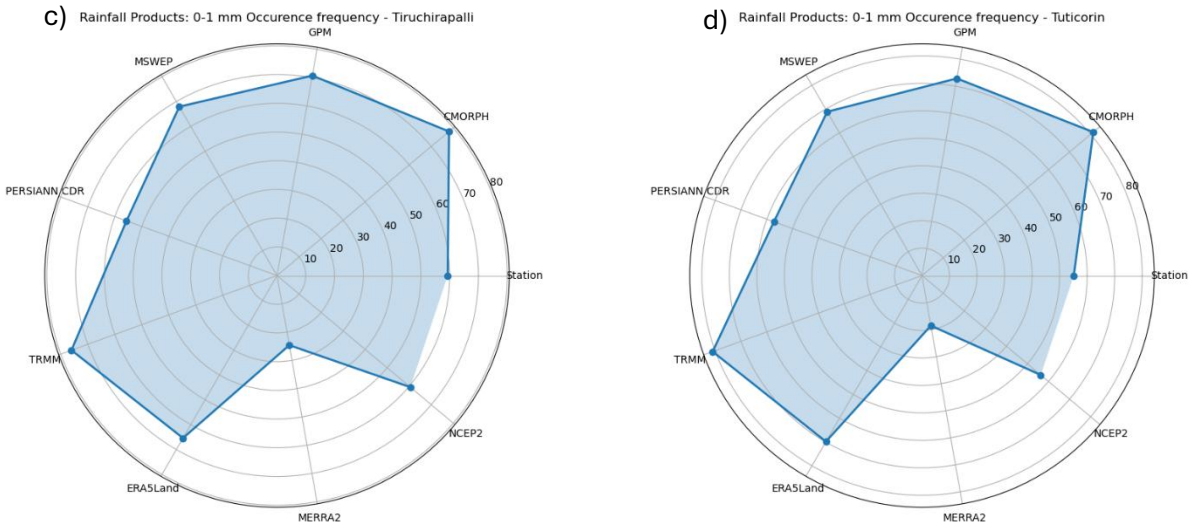
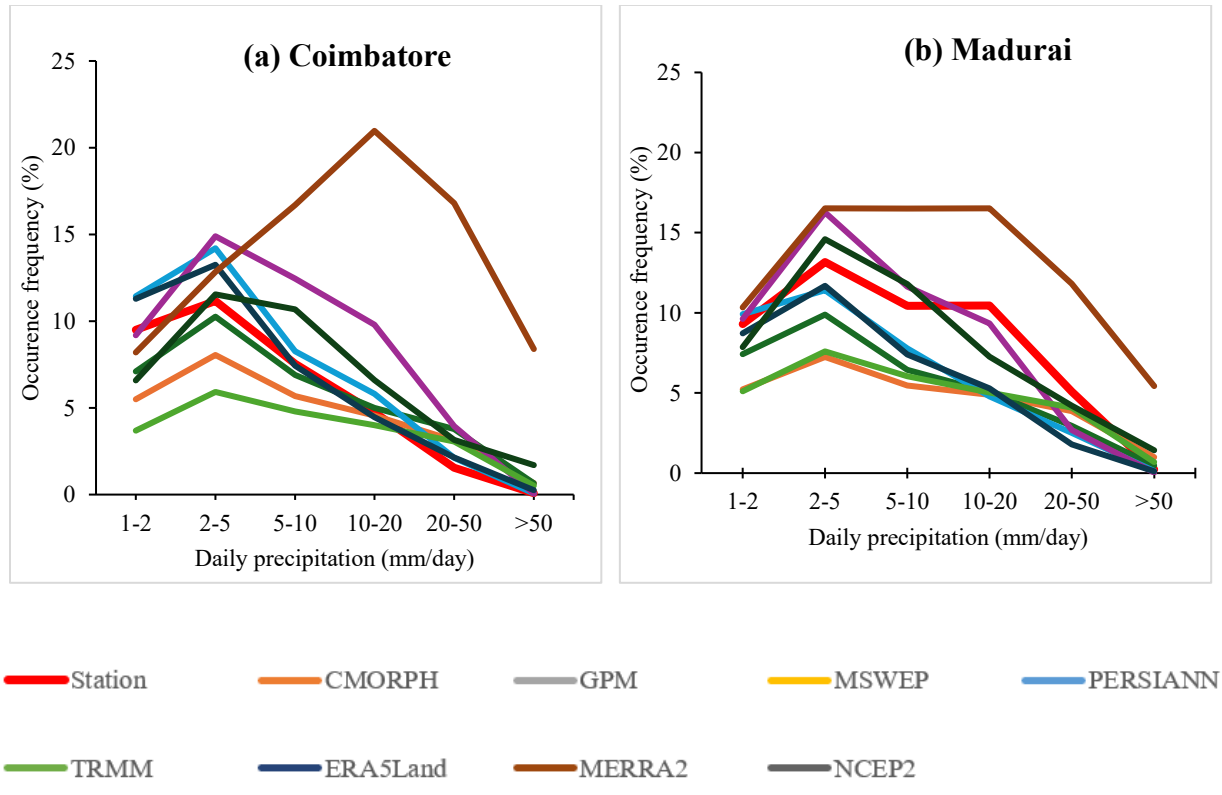


Figure S2 Occurrence frequency of precipitation (0-1 mm) for (a) Coimbatore, (b) Madurai, (c) Tiruchirapalli and (d) Tuticorin



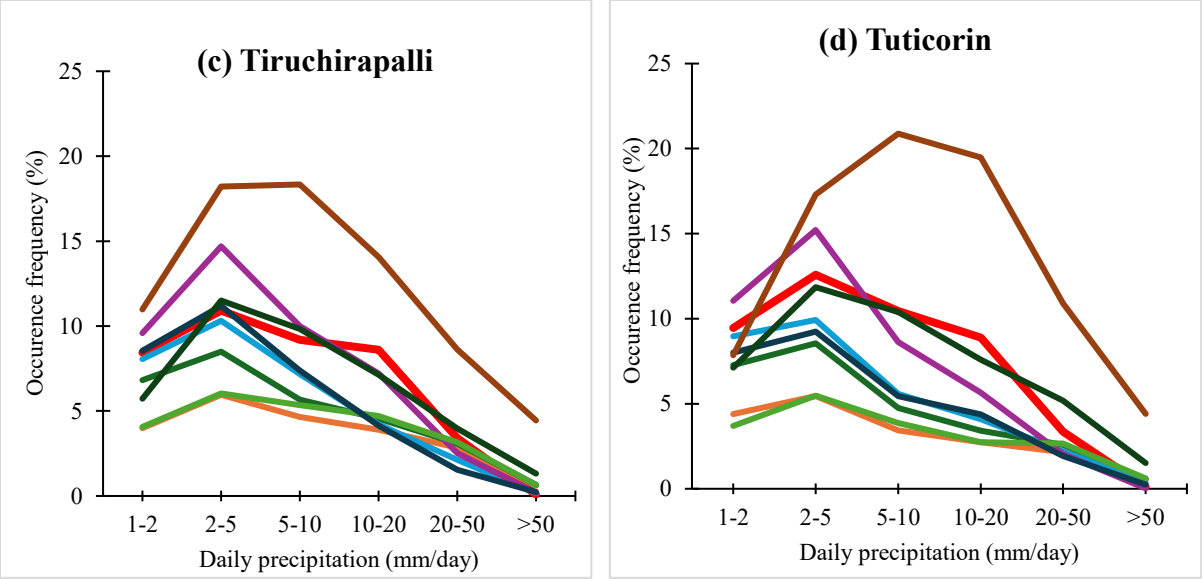
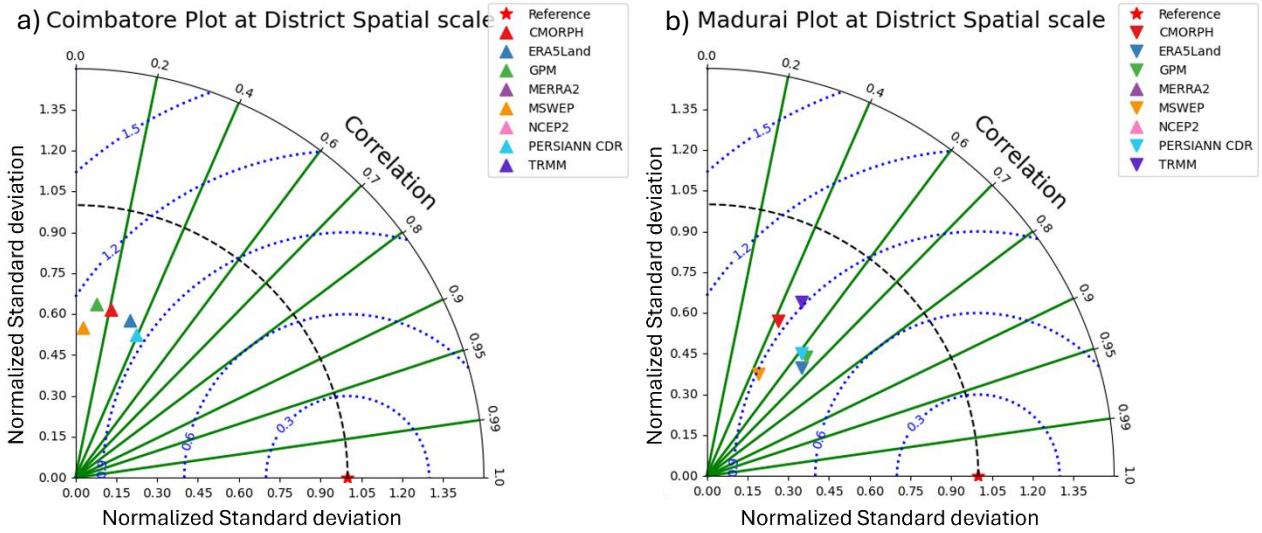


Figure S3 Occurrence frequency of precipitation (1-> 50 mm) for (a) Coimbatore, (b) Madurai, (c) Tiruchirapalli and (d) Tuticorin



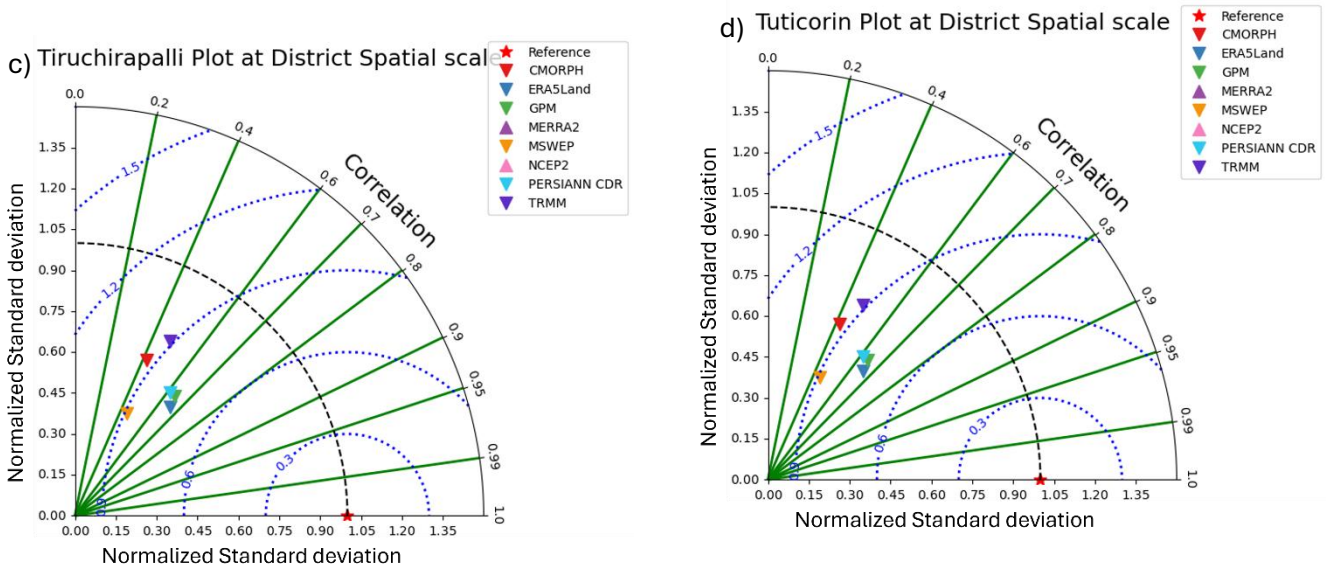
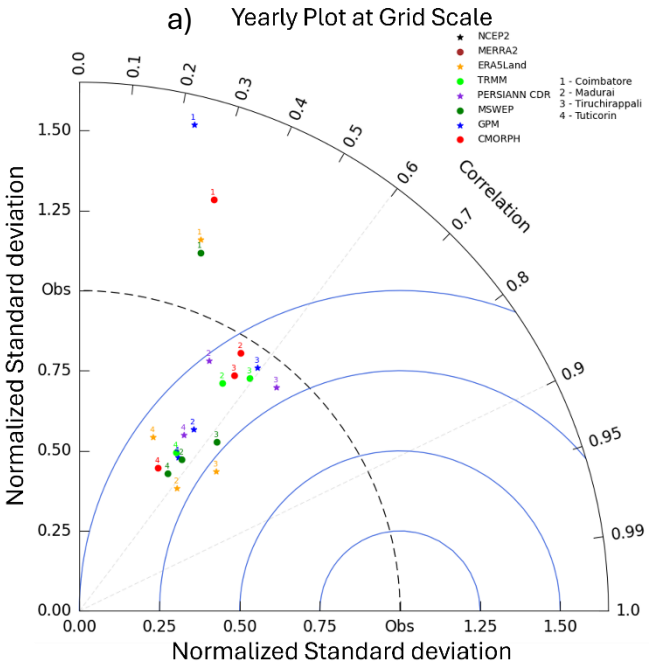


Figure S4 Taylor diagram depicting the agreement between station data and precipitation products at a yearly district scale



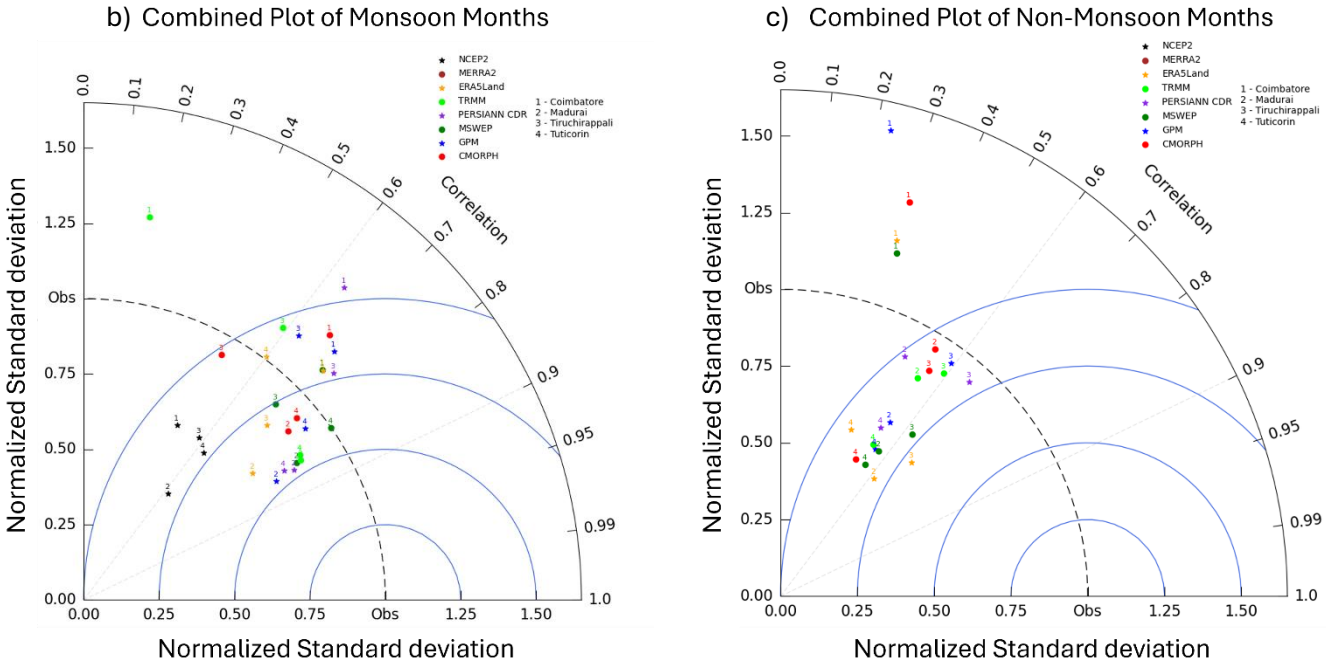


Figure S5 Combined Taylor diagram depicting the agreement between station data and precipitation products at a) yearly grid scale b) monsoon months c) non-monsoon months

Table S2. Daily precipitation characteristics of Coimbatore, Madurai, Tiruchirapalli and Tuticorin

Scale	Product	Coimbatore				Madurai				Tiruchirapalli				Tuticorin			
		MAE	IOA	BIAS	%MAE	MAE	IOA	BIAS	%MAE	MAE	IOA	BIAS	%MAE	MAE	IOA	BIAS	%MAE
Grid	**CMORPH**	3.62	0.36	0.58	1.59	5.21	0.47	-1.12	1.17	4.19	0.42	-0.97	1.21	3.88	0.55	-1.66	1.05
	GPM-IMERG	3.87	0.38	1.00	1.70	4.67	0.53	-1.60	1.05	4.19	0.47	-0.55	1.21	3.99	0.60	-1.26	1.08
	MSWEP	2.96	0.55	0.39	1.30	4.09	0.61	-1.85	0.91	3.45	0.56	-1.15	0.99	3.23	0.66	-0.99	0.87
	PERSIANN CDR	4.25	0.56	0.37	1.87	4.52	0.62	-1.95	1.01	3.91	0.57	-1.17	1.13	3.62	0.63	-1.21	0.98
	TRMM	3.49	0.30	0.20	1.53	5.09	0.50	-1.26	1.14	4.27	0.44	-0.67	1.23	3.96	0.58	-1.53	1.07
	ERA5-Land	2.86	0.57	0.35	1.26	3.86	0.63	-2.04	0.86	3.24	0.59	-1.19	0.93	3.57	0.61	-1.43	0.96
	MERRA2	18.95	0.13	18.29	8.32	13.02	0.25	10.87	2.91	11.19	0.18	9.38	3.22	11.89	0.23	10.49	3.21
	NCEP2	5.40	0.13	2.59	2.37	6.44	0.28	0.45	1.44	5.50	0.26	1.06	1.59	6.16	0.30	1.48	1.66
District	**CMORPH**	3.39	0.39	0.58	1.49	5.12	0.48	-1.12	1.15	4.11	0.43	-0.97	1.18	3.69	0.59	-1.66	1.00
	GPM-IMERG	3.55	0.41	1.00	1.56	4.56	0.55	-1.60	1.02	4.08	0.48	-0.55	1.18	3.62	0.64	-1.26	0.98
	MSWEP	2.70	0.59	0.39	1.18	4.00	0.63	-1.85	0.89	3.36	0.57	-1.15	0.97	3.19	0.66	-0.99	0.86
	PERSIANN CDR	4.04	0.59	-0.17	1.77	4.44	0.64	-1.95	0.99	3.81	0.58	-1.17	1.10	3.38	0.66	-1.21	0.91
	TRMM	0.17	0.37	1.43	0.07	4.99	0.52	-1.26	1.12	4.17	0.45	-0.67	1.20	3.74	0.61	-1.53	1.01
	ERA5-Land	0.06	0.59	-0.73	0.03	3.79	0.65	-2.04	0.85	3.15	0.60	-1.19	0.91	3.28	0.66	-1.43	0.89
	MERRA2	18.70	0.14	18.29	8.21	12.93	0.25	10.87	2.90	11.09	0.17	9.38	3.20	11.62	0.23	10.49	3.14
	NCEP2	5.22	0.13	2.59	2.29	6.40	0.29	0.45	1.43	5.43	0.26	1.06	1.57	6.00	0.29	1.48	1.62

Table S3. Monthly precipitation characteristics of Coimbatore, Madurai, Tiruchirapalli and Tuticorin

		Coimbatore				Madurai				Tiruchirapalli				Tuticorin			
Scale	**Product**	**MAE**	**IOA**	**BIAS**	**%MAE**	**MAE**	**IOA**	**BIAS**	**%MAE**	**MAE**	**IOA**	**BIAS**	**%MAE**	**MAE**	**IOA**	**BIAS**	**%MAE**
Grid	**CMORPH**	53.79	0.72	17.60	0.78	66.47	0.79	-34.17	0.49	55.51	0.77	-29.57	0.527	60.42	0.79	-50.57	0.54
	GPM-IMERG	61.88	0.67	30.57	0.90	65.64	0.80	-48.76	0.48	51.08	0.83	-16.67	0.485	55.31	0.83	-38.41	0.49
	MSWEP	45.99	0.83	8.71	0.67	67.68	0.79	-58.34	0.50	50.99	0.82	-36.54	0.484	47.07	0.86	-32.02	0.42
	PERSIANN CDR	84.50	0.54	54.51	1.23	68.06	0.79	-25.84	0.50	43.51	0.87	-7.32	0.413	50.59	0.83	-29.84	0.45
	TRMM	80.84	0.30	0.20	1.18	63.63	0.82	-38.29	0.47	51.99	0.82	-20.43	0.494	57.74	0.82	-46.41	0.51
	ERA5-Land	43.63	0.78	5.70	0.63	69.98	0.78	-62.09	0.52	49.01	0.83	-36.16	0.465	62.03	0.77	-43.45	0.55
	MERRA2	68.78	0.16	556.73	1.00	333.84	0.42	330.96	2.47	17383.95	0.34	285.37	165	319.79	0.37	319.31	2.85
	NCEP2	121.22	0.13	78.79	1.76	131.16	0.30	13.65	0.97	103.07	0.40	32.27	0.979	114.70	0.38	45.12	1.02
District	**CMORPH**	47.69	0.77	17.60	0.69	65.44	0.79	-34.17	0.48	53.66	0.78	-29.57	0.509	57.18	0.82	-50.57	0.51
	GPM-IMERG	54.81	0.71	30.57	0.80	64.83	0.81	-48.76	0.48	48.52	0.84	-16.67	0.461	51.76	0.85	-38.41	0.46
	MSWEP	38.43	0.83	8.71	0.56	67.00	0.80	-58.34	0.49	49.13	0.83	-36.54	0.466	46.72	0.86	-32.02	0.42
	PERSIANN CDR	78.23	0.54	54.51	1.14	67.76	0.80	-25.84	0.50	41.04	0.89	-7.32	0.39	44.03	0.86	-29.84	0.39
	TRMM	2.28	0.30	25.22	0.03	62.88	0.83	-38.29	0.46	50.09	0.83	-20.43	0.476	4947.30	0.84	-46.41	44.10
	ERA5-Land	35.84	0.84	5.70	0.52	69.45	0.78	-62.09	0.51	47.53	0.84	-36.16	0.451	55.56	0.82	-43.45	0.50
	MERRA2	556.66	0.15	556.73	8.09	333.68	0.42	330.96	2.46	287.05	0.34	285.37	2.725	319.43	0.37	319.31	2.85
	NCEP2	117.64	0.12	78.79	1.71	130.84	0.30	13.67	0.97	102.69	0.40	32.27	0.975	113.03	0.37	45.12	1.01

Table S4. Yearly precipitation characteristics of Coimbatore, Madurai, Tiruchirapalli and Tuticorin

		Coimbatore			**Madurai**			**Tiruchirapalli**			**Tuticorin**		
Scale	**Product**	**MAE**	**IOA**	**%MAE**	**MAE**	**IOA**	**%MAE**	**MAE**	**IOA**	**%MAE**	**MAE**	**IOA**	**%MAE**
Grid	**CMORPH**	344.76	0.47	0.41	433.69	0.52	0.27	382.18	0.49	0.30	625.21	0.44	0.46
	GPM-IMERG	454.90	0.44	0.55	585.09	0.50	0.36	230.68	0.72	0.18	461.92	0.55	0.34
	MSWEP	317.51	0.41	0.38	700.10	0.44	0.43	438.44	0.51	0.35	384.15	0.56	0.28
	PERSIANN CDR	670.97	0.44	0.81	330.26	0.60	0.20	174.31	0.80	0.14	379.93	0.50	0.28
	TRMM	860.22	0.07	1.04	462.26	0.54	0.28	266.01	0.65	0.21	558.37	0.48	0.41
	ERA5-Land	278.73	0.57	0.34	745.05	0.25	0.46	434.31	0.57	0.34	521.35	0.41	0.39
	MERRA2	6680.71	0.07	8.04	3971.54	0.12	2.44	3424.42	0.12	2.70	3831.64	0.13	2.83
	NCEP2	976.94	0.26	1.18	471.97	0.07	0.29	446.51	0.32	0.35	640.77	0.25	0.47
District	**CMORPH**	9.69	0.53	0.28	61.43	0.52	0.26	45.75	0.49	0.29	51.83	0.44	0.46
	GPM-IMERG	15.95	0.40	0.46	83.60	0.51	0.36	25.64	0.74	0.16	38.42	0.54	0.34
	MSWEP	12.50	0.58	0.36	100.03	0.45	0.43	54.80	0.51	0.35	192.15	0.57	1.71
	PERSIANN CDR	27.25	0.31	0.79	45.61	0.61	0.20	17.66	0.85	0.11	29.85	0.56	0.26
	TRMM	6.53	0.57	0.19	65.67	0.54	0.28	30.63	0.67	0.19	46.43	0.49	0.41
	ERA5-Land	9.72	0.80	0.28	106.46	0.46	0.46	54.21	0.58	0.34	43.44	0.51	0.39
	MERRA2	278.36	0.04	8.04	567.34	0.12	2.43	428.06	0.11	2.70	319.30	0.11	2.83
	NCEP2	39.40	0.12	1.14	66.67	0.07	0.29	54.81	0.31	0.35	53.25	0.20	0.47