

Decadal Rainfall Distribution Shift in Northern Coastal Java: Histogram-Based Clustering Using Observations and CHIRPS

Iis Widya Harmoko^{1,3*}, Muhammad Zainuri², Anindya Wirasatriya², Supari³, and Aviv Maghridlo⁴

¹Faculty of Fisheries and Marine Sciences, Diponegoro University, Indonesia

²Department of Oceanography, Faculty of Fisheries and Marine Sciences, Diponegoro University, Indonesia

³Meteorology, Climatology, and Geophysics Agency (BMKG), Jakarta Pusat, Indonesia

⁴State College of Meteorology Climatology and Geophysics (STMKG), Tangerang, Indonesia

ARTICLE INFO

Received: 8 Nov 2025

Received in revised: 19 Mar 2026

Accepted: 23 Mar 2026

Published online: 4 Jun 2026

DOI: 10.32526/ennrj/24/20250315

Keywords:

Coastal rainfall/ Distributional change/ Decadal shift/ Density estimation/ Composite index/ Unsupervised clustering

* Corresponding author:

E-mail: iis.harmoko@bmkg.go.id

ABSTRACT

Coastal lowlands are highly vulnerable to seasonal changes and rainfall variability, which affect the risk of floods and droughts. Therefore, this study aimed to analyze decadal changes in the monthly rainfall distribution along the northern coast of Central Java (Pantura), using data from 51 stations of the Meteorology, Climatology, and Geophysics Agency (BMKG) and the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) v2.0. A smoothed histogram was used to calculate the probability density of rainfall for each decade. Subsequently, the distribution between decades was compared using 6 metrics to assess distribution displacement, divergence, overlap, and shift in peak, centroid, and width. The level of uncertainty was measured using the bootstrap method at confidence levels of 5%, 10%, and 25%. The six metrics were normalized and combined into a single composite index called the Total Normalized Index (TNI). K-means clustering was then applied to map changes according to the same rainfall patterns. Analysis of both sources showed a gradient pattern from west to east. Rainfall was concentrated in the west, with moderate changes in structure in the central area, and a more dispersed rainfall pattern with high variability in the east, particularly from 2011 to 2020. CHIRPS data recorded the regional trends but tended to cloud out local fluctuations. This study contributed to new investigations since the implementation of histogram-based rainfall distribution analysis into a composite index. The application of this method as part of a spatial clustering model was proposed to map alterations in rainfall distribution each decade. The method also helped in evaluating coastal climate risk and adaptation planning.

HIGHLIGHTS

- Six distribution metrics show an apparent decadal shift in rainfall structure across Pantura.
- Observations and CHIRPS show consistent west-east gradients in rainfall change.
- TNI confirms significant long-term rainfall transformation in both datasets.
- K-means clustering delineates spatially coherent clusters of rainfall change.
- Recent decades show an eastward shift in rainfall variability signals.

1. INTRODUCTION

The hydrological cycle is enhanced by global warming, which causes more moisture in the atmosphere and risks of causing heavy rainfall (Xiong and Yang, 2024). Consistent with the Clausius-Clapeyron relationship, which describes an approximate 7 percent increase in atmospheric water holding capacity per one degree Celsius warming, rainfall extremes are expected to intensify.

Meanwhile, the variations will be spatially heterogeneous across different areas. This tendency is supported by the Intergovernmental Panel on Climate Change report, which documents rainfall increases in humid tropical areas and high latitudes, while rainfall decreases occur in dry subtropical areas (Li et al., 2025; Parmesan et al., 2023). The spatial-temporal changes also increase the risk of hydrometeorological events, such as floods,

Citation: Harmoko IW, Zainuri M, Wirasatriya A, Supari, Maghridlo A. Decadal rainfall distribution shift in Northern Coastal Java: Histogram-based clustering using observations and CHIRPS. Environ. Nat. Resour. J. 2026;24(4):482-495. (<https://doi.org/10.32526/ennrj/24/20250315>)

landslides, and droughts, particularly in susceptible areas, including Central Java (Putri et al., 2022).

The Asian and Australian responses to climate in the area were complex. Monsoon rainfall has been reported to increase significantly in intensity in South and Southeast Asia (Varghese et al., 2025). During the last 70 years, the rainfall in India has tripled (Roxy et al., 2017). However, southern Australia has experienced a significant decline in seasonal rainfall due to anthropogenic factors (McKay et al., 2023). Climate variability, such as the El Niño-Southern Oscillation (ENSO) and the Indian Ocean Dipole (IOD), influences decadal anomalies, as observed in Indonesia (An et al., 2023; Pillai et al., 2021; Supari et al., 2018).

A clear example of high climate vulnerability is the northern coastal area of Central Java (Pantura). Due to its low topography, the Pantura area is vulnerable to potential hydrometeorological disasters, which are worsened by rapid urbanization and its sensitivity to monsoon variability. Several studies show that the rainy season is characterized by high rainfall intensity, and the dry season is prolonged, with flooding occurring more than 70% of the time between January and March. There was also an accumulation of monthly rainfall exceeding 1,000 mm (Agusta and Putri, 2023; Handayani et al., 2020). These hydrometeorological dynamics show the importance of analyzing decadal-scale spatial and temporal distributions.

Histogram-based distribution diagnostics and coastal-scale clustering have rarely been applied in Pantura, Central Java, despite related studies in other areas (Australia, Panama, Turkey) (Handayani et al., 2020; Hatvani et al., 2025; Kömüschü et al., 2022). In Indonesia, recent studies documented changes in rainfall frequency and intensity, but without integrating histogram metrics with local coastal clustering, as well as jointly analyzing station data and CHIRPS in this framework (Pattipeilohy et al., 2024; Pertiwi et al., 2024; Kurnianingsih et al., 2025; Funk et al., 2015). This study assesses the decadal (1991-2020) shift in monthly rainfall distribution using 51 BMKG stations and CHIRPS, combining six distribution metrics into a Total Normalized Index (TNI) and applying K-means to map spatial transformation patterns relevant to climate adaptation.

Operational in detecting decadal changes in rainfall distribution, this study evaluates three testable hypotheses that can be associated with each distribution metric:

- H1: Decadal changes are evident from the shift in the peak of rainfall intensity and the centroid of the distribution, as well as the decrease in distribution overlap.

- H2: Rainfall variability tends to widen in recent decades, as reflected in the increased distribution width and the distance between decades.

- H3: The transformation of rainfall distribution shows a coherent west-central-east gradient pattern, which can be identified through the TNI and K-means clustering.

2. METHODOLOGY

2.1 Data source

This study analyzed monthly rainfall data from 51 rain gauge stations owned by the Meteorology, Climatology, and Geophysics Agency (BMKG) along the North Coast of Central Java (Pantura) during the period 1991-2020 (Figure 1). As grid-based comparative data, this study used monthly rainfall from the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) version 2.0 with a spatial resolution of 0.05° (UCSB, 2025; Funk et al., 2015). CHIRPS data from 1991-2020 were downloaded from the Climate Hazards Center portal (<https://data.chc.ucsb.edu/products/CHIRPS-2.0/>) and processed in Python, then bilinearly interpolated at BMKG station areas from the four nearest grids.

The handling of missing data has been predetermined, and months with anomalous data due to station damage are treated as missing. Only the decade-station data series with $\geq 90\%$ monthly values were used in the analysis, while the rest were excluded from the distribution estimation. No temporal imputation was applied, and CHIRPS pixels with NoData status were also excluded (Afrifa-Yamoah et al., 2020).

2.2 Decadal segmentation

The 30-year dataset was segmented into three decades, namely (1) 1991-2000, (2) 2001-2010, and (3) 2011-2020. This division followed a climatological framework that used 30 years as the basis for climate conditions, with updates every 10 years (WMO, 2017). Each ten-year period contained 120 months, which enabled analysis of rainfall patterns using an equal sample size while focusing on long-term changes and minimizing the influence of individual events. Two comparative periods were constructed, namely Decade-a: 2001-2010 vs 1991-2000; and Decade-b: 2011-2020 vs 2001-2010. The

decade-to-decade assessment was also made available by the use of the adjacent decades, with a provision for a stable evaluation of progressive changes in the

rainfall distribution patterns and variability in both observations and CHIRPS.

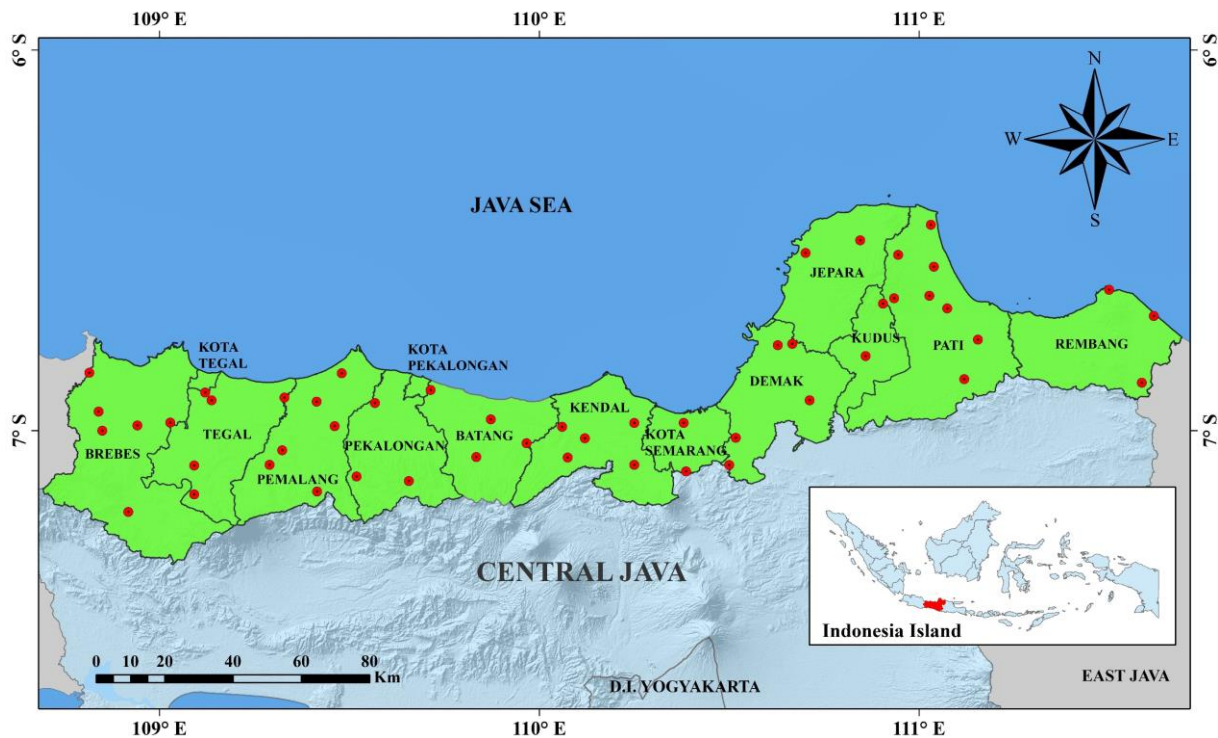


Figure 1. Area of BMKG observation stations across the North Coast of Central Java (Pantura)

2.3 Probability distribution estimation

The probability distribution of each decade and station estimated with a smoothed histogram were used to evaluate decadal changes in rainfall patterns (Figure 2). This measure provided a strong representation of distributional shape and showed lower sensitivity to extreme values. The resulting density curves would serve as a source to assess changes in the intensity of the peaks in terms of the width distribution and general displacement across decades.

The use of histograms in the current study, rather than a single parametric fit and kernel density estimation (KDE), was due to the fact that aggregates of precipitation were constrained to 0 and often skewed to the right. However, the adequacy of a given parametric distribution varied across datasets and aggregation periods. KDE was highly sensitive to bandwidth choice, as insufficient smoothing introduced local spurious features, while excessive smoothing obscured significant distributional structure. In comparison, a histogram built with an objective data-based rule of bin-width gives a clear and repeatable foundation on which to base inter-

decadal comparison of distributional morphology (Freedman and Diaconis, 1981; Jones et al., 1996; Svensson et al., 2017).

A histogram was plotted using the same bin edges for a given station, with two decades of monthly data combined to determine the bin edges, facilitating direct results comparison. A bin width was objectively defined with the help of the Freedman-Diaconis rule, $h=2IQRn^{-1/3}$, taking into consideration the interquartile range and the size of the sample. The counts in each of the bins were then translated into a probability density with a cumulative total of 1, and the outcome was smoothed using a Gaussian filter of width 1 bin. The x-axis of both density plots was the amount of rainfall in mm/month in a month, with the y-axis showing the probability density in units of 1/mm.

Statistical significance testing was carried out using bootstrap confidence intervals for each metric. Global curve difference tests, such as the Kolmogorov-Smirnov, were unnecessary because these methods showed general differences without providing details on specific changes in distributional features.

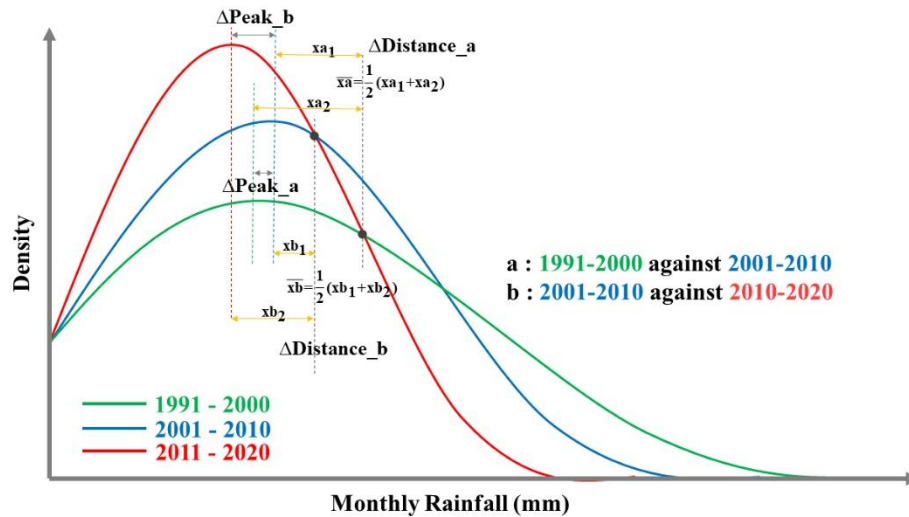


Figure 2. Illustration of histogram-based probability density estimation (x-axis: monthly rainfall; y-axis: probability density)

2.4 Distribution metrics

Six metrics (Table 1) were used to characterize inter-decadal variations. For brevity, these metrics were referred to as WD, KL, OVL, ΔPeak , $\Delta\text{Distance}$, and ΔWidth (Supplementary data).

- Wasserstein distance (WD) for overall displacement between distributions.
- Kullback-Leibler divergence (KL) for changes in distributional shape.
- Overlap coefficient (OVL) for similarity between decades.

- Peak shift (ΔPeak) for changes in predominant rainfall intensity.
- Centroid shift ($\Delta\text{Distance}$) for changes in the distribution mean (centroid).
- Width shift (ΔWidth) for changes in dispersion (standard deviation).

The metrics collectively captured changes in rainfall intensity, timing, and distribution morphology (Tolhurst et al., 2023; Magyar and Sambridge, 2023; Shen et al., 2025). Full mathematical formulations were provided in Supplementary data.

Table 1. Variables used to construct the TNI for two decadal comparisons: (a) 2001-2010 vs 1991-2000 and (b) 2011-2020 vs 2001-2010

Index component	Statistical description	Symbol	Unit	Climatological meaning
Wasserstein distance (WD)	Distribution displacement between decades	WD-a; WD-b	mm/month	General displacement of the monthly rainfall distribution
Kullback-Leibler divergence (KL)	Distributional divergence (shape difference)	KL-a; KL-b	unitless	Change in distribution shape between decades
Overlap coefficient (OVL)	Distributional similarity (overlap area)	OVL-a; OVL-b	Unitless (0-1)	Higher overlap shows a more stable distribution
Peak shift (ΔPeak)	Change in modal (peak) rainfall intensity	ΔPeak -a; ΔPeak -b	mm/month	Shift in the most frequent monthly rainfall level
Centroid shift ($\Delta\text{Distance}$)	Change in distribution centroid (mean)	$\Delta\text{Distance}$ -a; $\Delta\text{Distance}$ -b	mm/month	Shift in the center of mass of the rainfall distribution
Width shift (ΔWidth)	Change in dispersion (standard deviation)	ΔWidth -a; ΔWidth -b	mm/month	Broadening or narrowing of monthly rainfall variability
Total Normalized Index (TNI)	Composite index (equal-weight integration of six normalized metrics)	TNI-a; TNI-b	unitless	Integrated magnitude of decadal distributional change

All signed shift metrics are computed as the value in the later decade minus the value in the earlier decade.

2.5 Index Normalization and TNI Construction

The six-distribution metrics were calculated at each station to compare two pairs of decades, namely (a) 2001-2010 vs 1991-2000, and (b) 2011-2020 vs

2001-2010. Each metric was then normalized to a unitless scale of 0-1 using the min-max method. This normalization was carried out separately for the BMKG and CHIRPS datasets in each decade comparison.

Let $M_{i,c}$ denote the value of a given metric at station i for comparison $c \in \{a,b\}$. The normalized metric $M_{i,c}^*$ is defined as:

$$M_{i,c}^* = \frac{M_{i,c} - M_{\min,c}}{M_{\max,c} - M_{\min,c}} \quad (1)$$

Where; $M_{\min,c}$ and $M_{\max,c}$ are the minimum and maximum values across all stations within the same dataset and comparison period c .

The OVL decreases as distributional change strengthens, and reverse normalization ensures that larger values correspond to greater change:

$$OVL_{i,c}^* = 1 - \frac{OVL_{i,c} - OVL_{\min,c}}{OVL_{\max,c} - OVL_{\min,c}} \quad (2)$$

For signed shift metrics (ΔPeak , $\Delta\text{Distance}$, ΔWidth), the magnitude of change was used for index construction:

$$\Delta\text{Peak}_{i,c}^{\text{abs}} = |\Delta\text{Peak}_{i,c}|, \Delta\text{Distance}_{i,c}^{\text{abs}} = |\Delta\text{Distance}_{i,c}|, \Delta\text{Width}_{i,c}^{\text{abs}} = |\Delta\text{Width}_{i,c}| \quad (3)$$

Each magnitude term was then normalized using the same min-max scaling:

$$(\Delta^{\text{abs}})_{i,c}^* = \frac{\Delta_{i,c}^{\text{abs}} - \Delta_{\min,c}^{\text{abs}}}{\Delta_{\max,c}^{\text{abs}} - \Delta_{\min,c}^{\text{abs}}} \quad (4)$$

The TNI was computed as the aggregation (sum) of the six normalized metrics:

$$TNI_{i,c} = WD_{i,c}^* + KL_{i,c}^* + OVL_{i,c}^* + (\Delta\text{Peak}^{\text{abs}})_{i,c}^* + (\Delta\text{Distance}^{\text{abs}})_{i,c}^* + (\Delta\text{Width}^{\text{abs}})_{i,c}^* \quad (5)$$

Based on this definition, the TNI value for each station and period ranged from 0 to 6. The average calculation only changed the scale without affecting the ranking results, spatial patterns, or clustering. Therefore, TNI interpretation remained comparative within each dataset and period, which was supported by all metrics.

Min-max normalization had limitations due to its sensitivity to extreme values, necessitating comparative TNI interpretation within a given dataset and decade. Z-scores or other scaling methods were not applied in this study and were recommended for future analysis (Mazziotta and Pareto, 2022).

2.6 Significance assessment

Statistical significance was assessed by calculating nonparametric bootstrap confidence intervals (CIs) for the metric at each station and decade (Efron and Tibshirani, 1994). Only station-decade data with at least 90% of the months (≥ 108 months) were used for resampling, and all metrics were recalculated at each bootstrap replicate.

Percentile CIs were calculated at the 95%, 90%, and 75% confidence levels ($\alpha=5\%$, 10% , and 25%). A change was considered significant when the confidence interval did not include zero, namely for WD and KL. The lower bound of the CI must be greater than 0. For OVL, the upper bound of the CI must be less than 1. In the case of ΔPeak , $\Delta\text{Distance}$, and ΔWidth , the CI must not contain 0. The final inference method used this bootstrap-based criterion without the additional use of the Wilcoxon signed-rank test, as the analysis focused on the distribution function, and uncertainty was already captured by the bootstrap method. In the visualization, the significance levels $\alpha = 5\%$, 10% , and 25% were marked with the symbols X, Y, and Z, respectively.

2.7 Spatial clustering

K-means clustering was used on station TNI values to establish spatially coherent groups that share common characteristics of change. The elbow, silhouette, and Davies-Bouldin criteria methods were used to pick the number of clusters (Rahman et al., 2024; Pratama and Salam, 2025). K-means was selected because it was a centroid-based partitioning algorithm that was relatively easy to implement and computationally inexpensive. The method could also be readily understood through centroid representation and minimization of within-cluster sum of squares (Ikotun et al., 2023; Wani, 2024). An example of a very brief sensitivity analysis using close-to-neighbour K values showed the same most pronounced west-central-east pattern. However, alternative clustering algorithms were not evaluated (Roux et al., 2021).

3. RESULTS AND DISCUSSION

Understanding the evolution of decadal rainfall requires not only determining changes in totals but also examining how the shape, intensity, timing, and variability of rainfall distributions vary over time. In this section, six distribution-based measures and a bootstrap significance test ($\alpha=5\%$, 10% , 25%) are used to analyze the spatial distribution of rainfall transformation in the North Coast of Central Java, using both the Observation and CHIRPS data. The results of the study were categorized into two major groups. The first group consisted of shape-based metrics, including displacement, divergence, and overlap of distribution (WD, KL, and OVL). Meanwhile, the second group comprised morphology-based metrics, which compare changes in the peak, centroid, and width of the distribution (Δ Peak, Δ Distance, Δ Width). These six measures were

integrated into a TNI and then spatially analyzed to extract major areas where a change in the distribution of rainfall is taking place.

In order to promote comprehension, every metric was only described once and applied throughout the results. These included WD, KL, OVL, and Δ Peak, Δ Distance, and Δ Width (morphology measures). Changes between decades will be denoted as change a between 2001-2010 and 1991-2000, as well as change b between 2011-2020 and 2001-2010.

3.1 Shape-based distribution metrics

Figure 3 shows the spatial distribution of three shape-based measures of observations (A1-C1) and CHIRPS (A2-C2). All these measures together show the changes in the distribution and structure of rainfall over decades.

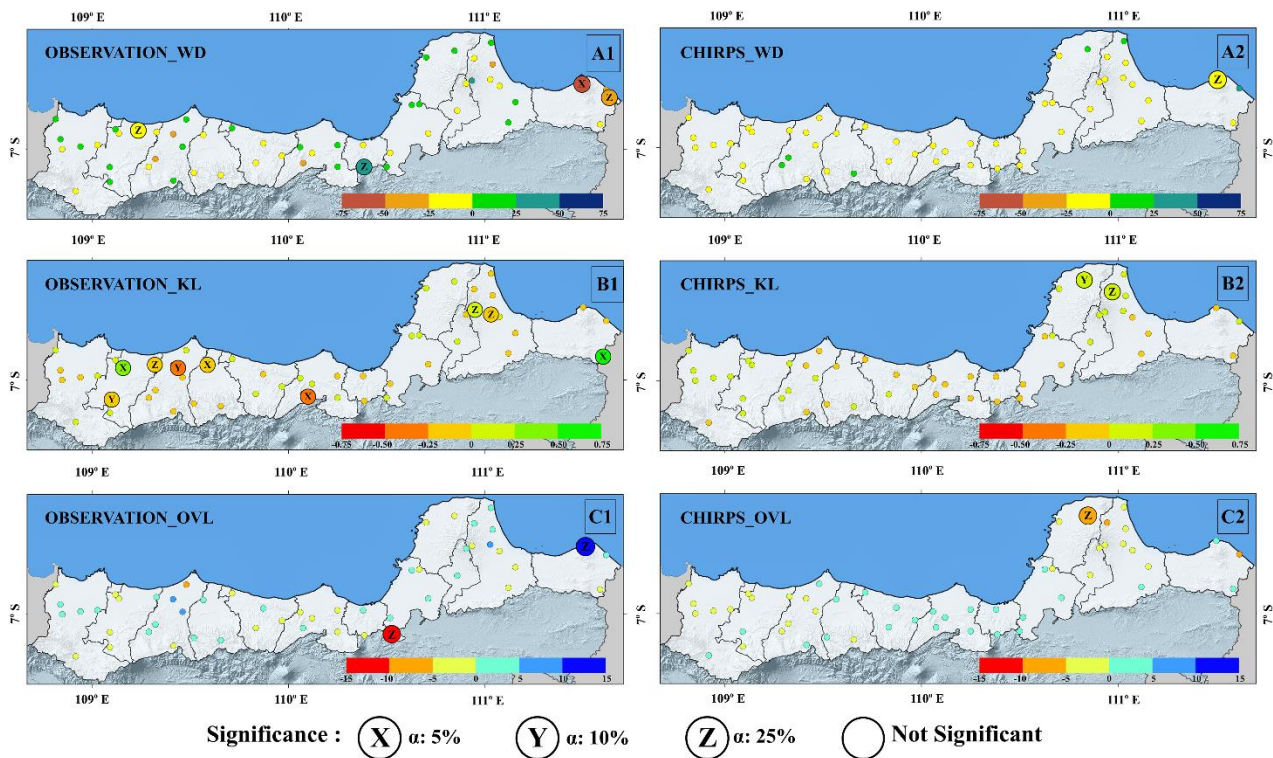


Figure 3. Shape-based rainfall measures (WD, KL, OVL) of observations (A1-C1) and CHIRPS (A2-C2) with significant levels, X (5%), Y (10%), and Z (25%).

The WD showed that the western and eastern areas exhibited a significant difference in terms of movement of rainfall, with the western corridor registering a greater change. These results suggested that there were alterations in the intensity of rainfall and dramatic structural changes in the western area over the past decade. CHIRPS data exhibited the same

trend, but the shift appeared smoother, reflecting the spatial smoothing inherent in gridded data.

KL showed a nonlinear change in the shape of the distribution, specifically in the central and eastern areas. Higher values of KL in observations suggested that there was greater variability in rainfall. CHIRPS captured the broader gradient while underestimating localized differences. The OVL was built upon the last

result by identifying the areas where distribution has grown dissimilar over decades. The western and northeast areas of the coastlines had a lower value of OVL, which showed a high level of structural variation. The central area had a greater value, suggesting relative stability. Evaluation of all three measures showed that significant markers of change (X, Y, Z) established the presence of distributional changes at several stations. Observational data showed

more localized changes, while CHIRPS data had greater areal consistency.

3.2 Morphology-based metrics

The morphology-based statistics of Observations (A1-C1) and CHIRPS (A2-C2) (Figure 4) showed the changes in the rainfall intensity, central tendency, and seasonal variability during the decades.

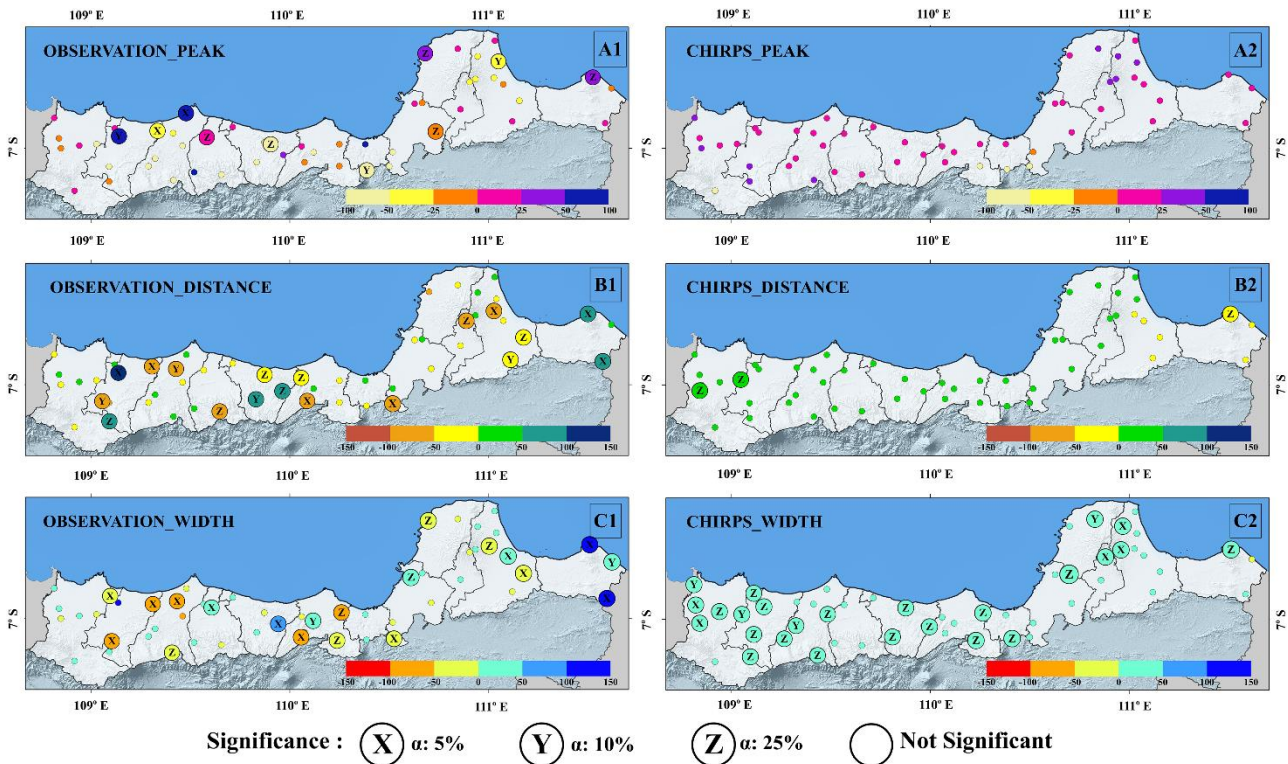


Figure 4. Morphology-based rainfall metrics (Δ Peak, Δ Distance, Δ Width) for observations (A1-C1) and CHIRPS (A2-C2), with significance markers X (5%), Y (10%), and Z (25%).

The Δ Peak represented the growth in rainfall intensity in the western and northeastern range, with positive values denoting heavier rainfall in the last decades. A significant portion of the eastern areas had smaller peak intensities, showing varied rainfall-intensification patterns throughout the area. CHIRPS could capture the general trends, but only a few important stations.

Δ Distance represented the separation between the centroid of the distribution, showing significant movement in the western and eastern areas. This observation showed the tendency of rainfall rearrangement, and could be moving either way in terms of intensities, to the heavy or light end, based on the area. Observation data contained more detailed spatial features, while CHIRPS data showed more gradual decadal variations.

Δ Width showed the opposite rainfall variability regimes. Western Pantura exhibited negative Δ Width predominantly, process distribution was narrow, and rainfall events were concentrated. Eastern Pantura exhibited a positive change in Δ Width, suggesting that intra-annual variability was greater and the rainfall pattern was more dispersed. These identified spatial contrasts were also consistent across datasets, but CHIRPS contained more generalized gradients.

The Δ Peak- Δ Distance- Δ Width measured all points to a consistent morphological change. An intensification and westward concentration tendency occurred, accompanied by expansion in the eastern region. Statistical significance further confirmed the strength of these changes.

3.3 Transformation of the TNI

The 6 measures were summed to obtain the TNI, a unified, holistic measure of the changes in the rainfall distribution. Figure 5 shows the patterns of TNI of observations and CHIRPS, and Table 2 contains the bootstrap value of significance.

Figure 6 provided a station-wise comparison between observation-based and CHIRPS-based TNI for each decadal transition, using an identity line to

visualize agreement and systematic bias. The western and eastern areas exhibited high TNI values (0-6), representing more substantial structural alterations. Meanwhile, the central area showed moderate TNI values, consistent with transitional rainfall patterns. All transitions in Figure 5 (a)-(d) showed a statistically significant shift (lower CI>0), suggesting that the transformation of rainfall in Pantura was both systematic and meaningful.

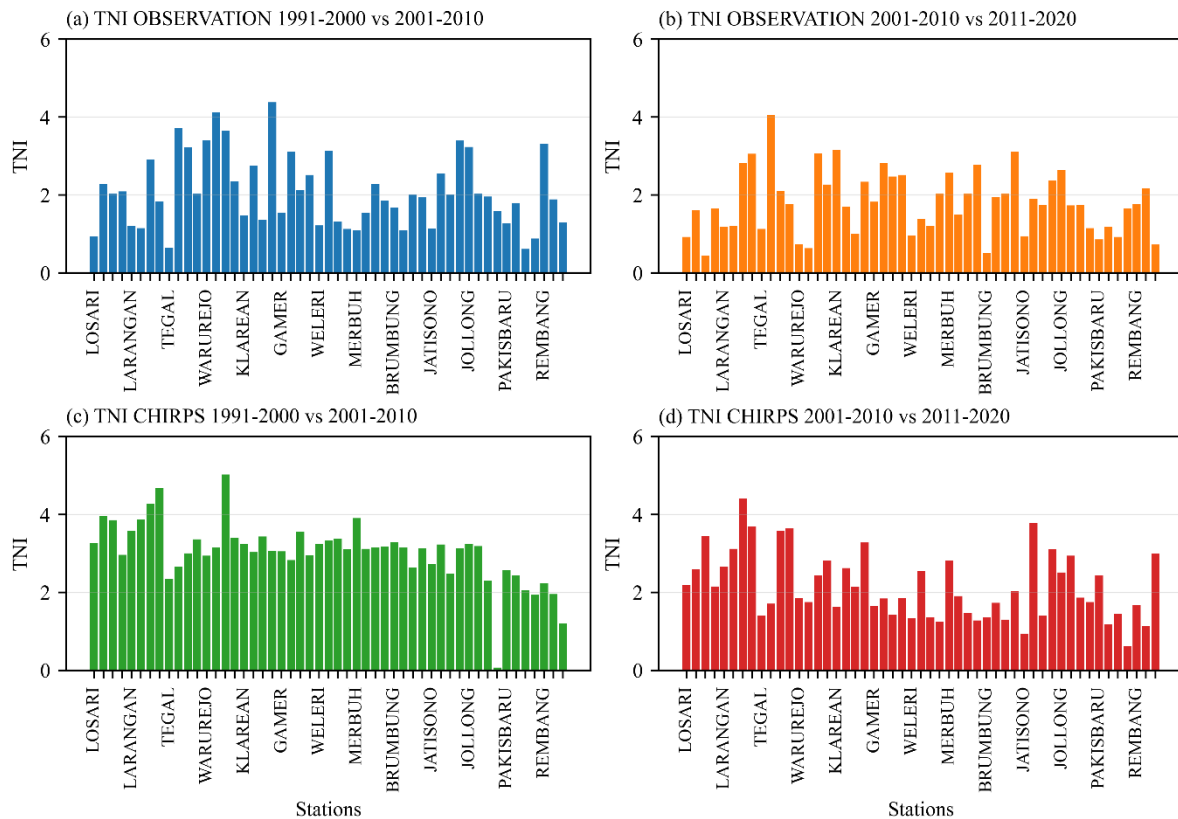


Figure 5. Spatial pattern of TNI in stations ranked by longitude (west to east) in two decadal comparisons (a-b) observations and (c-d) CHIRPS.

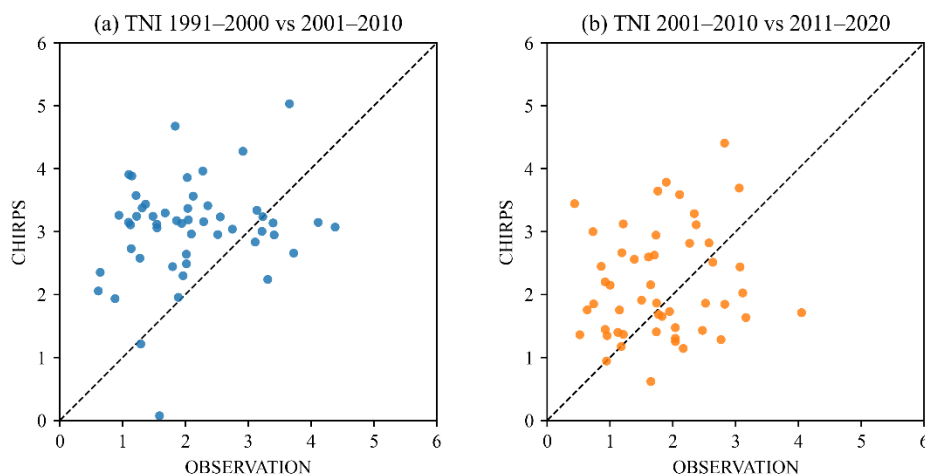


Figure 6. Station-wise relationship between observation-based TNI and CHIRPS-based TNI for the two decadal transitions. The dashed line denotes the identity line. (a) 2001-2010 vs 1991-2000; (b) 2011-2020 vs 2001-2010.

Table 2. Mean total normalized index (TNI; unitless, range 0-6) and bootstrap confidence intervals (BCIs) for two adjacent-decade comparisons. BCIs are reported as lower-upper bounds; significance flags show whether the CI excludes the null.

Dataset and decadal comparison	Mean TNI	BCI 95% (lower-upper)	Sig. ($\alpha=0.05$)	BCI 90% (lower-upper)	Sig. ($\alpha=0.10$)	BCI 75% (lower-upper)	Sig. ($\alpha=0.25$)
Observation (2001-2010 vs 1991-2000)	2.081	1.823-2.339	Yes	1.874-2.288	Yes	1.933-2.229	Yes
Observation (2011-2020 vs 2001-2010)	1.804	1.588-2.020	Yes	1.618-1.990	Yes	1.674-1.934	Yes
CHIRPS (2001-2010 vs 1991-2000)	3.038	2.829-3.247	Yes	2.859-3.217	Yes	2.914-3.162	Yes
CHIRPS (2011-2020 vs 2001-2010)	2.160	1.921-2.399	Yes	1.968-2.352	Yes	2.023-2.297	Yes

CHIRPS tended to produce larger TNI values because of its less abrupt transitions over decades, while the spatial gradients of CHIRPS were similar to those of observations. As shown in [Figure 6](#), the X-Y relationship plot suggested that CHIRPS generally followed the transformation signal obtained at the station during both transitions, although the points tended to be above the identity line. This result suggested that the CHIRPS-based TNI was higher systematically but was otherwise consistent with the general trend of the gradient.

3.4 Spatial clustering of rainfall distribution shift

K-means clustering integrated the six distributional measures into consistent spatial clusters, which showed major trends of decadal rainfall change. Combining validation indices (Silhouette, Davies-Bouldin and Calinski-Harabasz), the number of clusters to be used during each transition period was selected automatically ([Tables 3-4](#)). The results showed that there were significant differences between Observations and CHIRPS regarding spatial heterogeneity.

Table 3. Observation K-means clustering 2001-2010 vs 1991-2000 and 2011-2020 vs 2001-2010

2001-2010 vs 1991-2000		2011-2020 vs 2001-2010	
Cluster	Regency	Cluster	Regency
0	Brebes, Pekalongan, Pemalang, Tegal	0	Demak, Kendal, Semarang
1	Demak, Kudus, Pati, Rembang	1	Jepara, Kudus, Pati
2	Batang, Kendal, Pekalongan, Pemalang, Tegal	2	Batang, Demak, Kendal, Pekalongan, Semarang
3	Kudus, Pati, Rembang	3	Brebes, Pekalongan, Pemalang
4	Batang, Demak, Kendal, Semarang	4	Pemalang, Tegal
5	Brebes, Pekalongan, Pemalang, Tegal	5	Jepara, Pati, Rembang
6	Jepara, Pati	6	Brebes, Pekalongan, Pemalang, Tegal
		7	Pati, Rembang
		8	Demak, Kudus, Pati, Rembang
		9	Pemalang, Tegal

Table 4. CHIRPS K-means clustering 2001-2010 vs 1991-2000 and 2011-2020 vs 2001-2010

2001-2010 vs 1991-2000		2011-2020 vs 2001-2010	
Cluster	Regency	Cluster	Regency
0	Demak, Kudus	0	Brebes, Kendal, Pekalongan, Pemalang
1	Brebes, Pekalongan, Pemalang	1	Jepara, Pati, Rembang
2	Pati	2	Demak, Kudus, Pati, Rembang
3	Pekalongan, Pemalang, Tegal	3	Brebes, Pemalang
4	Batang, Demak, Kendal, Semarang	4	Batang, Demak, Kendal, Semarang
5	Pati, Rembang	5	Pekalongan, Pemalang, Tegal
6	Pemalang, Tegal	6	Brebes, Pekalongan, Pemalang, Tegal
7	Pati, Rembang	7	Jepara, Kudus, Pati, Rembang
8	Brebes		
9	Jepara, Kudus, Pati		

3.4.1 Based on observational data

Clustering of observations (Figure 7) showed a high level of spatial fragmentation, suggesting that microclimatic and topographic factors exerted a significant impact on the variation of rainfall in the North Coast. Within the first transition period (1991-2000 to 2001-2010), clusters were mostly found along the western central corridor. Consequently, early decadal variability was localized in areas where there was a strong convective velocity and strong land-sea interaction.

In the second transition (2001-2010 to 2011-2020), the cluster arrangement was reorganized seriously to the north-east direction with the precipitation transformation shift to the east. This change was consistent with the increasing influence of sea surface temperature anomalies in the Java Sea and enhanced convective activity in the eastern area during the past decade. The observation clusters were also finer compared to CHIRPS, capturing the fine-scale variations caused by local rainfall characteristics in certain stations.

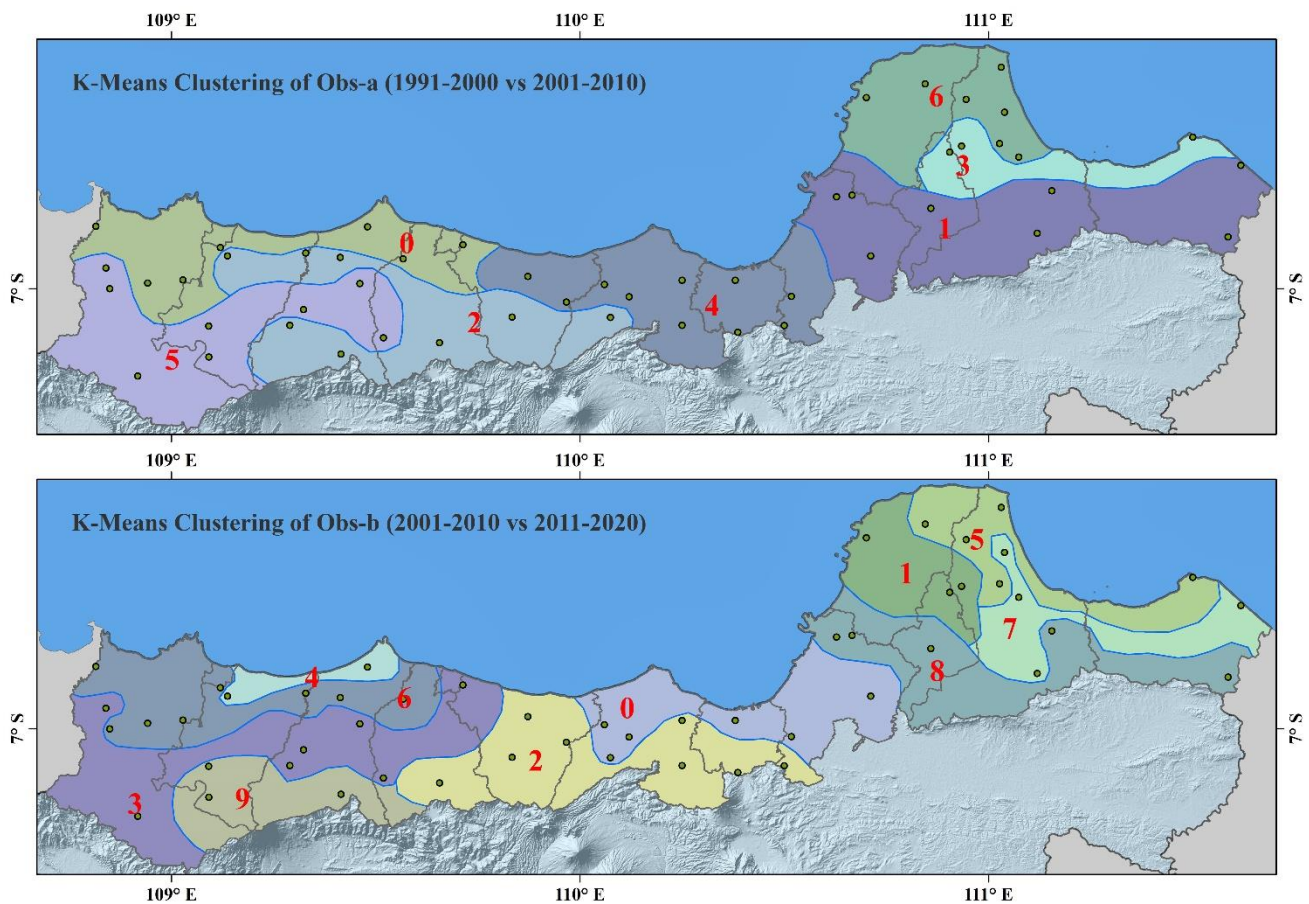


Figure 7. Two periods of transition (Obs-a and Obs-b) between observational data and the clustering of rainfall distribution using K-means. The clusters are the spatial areas that have a similar decadal variation of rainfall based on the six distribution measures.

3.4.2 Based on CHIRPS data

Figure 8 provides an even more coherent space zone in the regional clustering of CHIRPS data. The western and central areas were variable, accumulating in the early period of transition, and the eastern area was rather steady. The second transition period saw a change in the pattern of clustering to the east and showed a high variation in the area of Demak-Kudus-Jepara-Pati-Rembang. This easterly trend was consistent with the tendencies of observation, but

CHIRPS had fewer definite frontiers, focusing more on regional influences rather than the local ones.

In both transitions, CHIRPS provided 3 wide regional shift, namely (1) Western intensification, (2) Central transitional corridor, and (3) Eastern variability expansion. These areas were consistent with the spatial changes conveyed by the six distributional metrics, particularly ΔPeak , $\Delta\text{Distance}$, and ΔWidth .

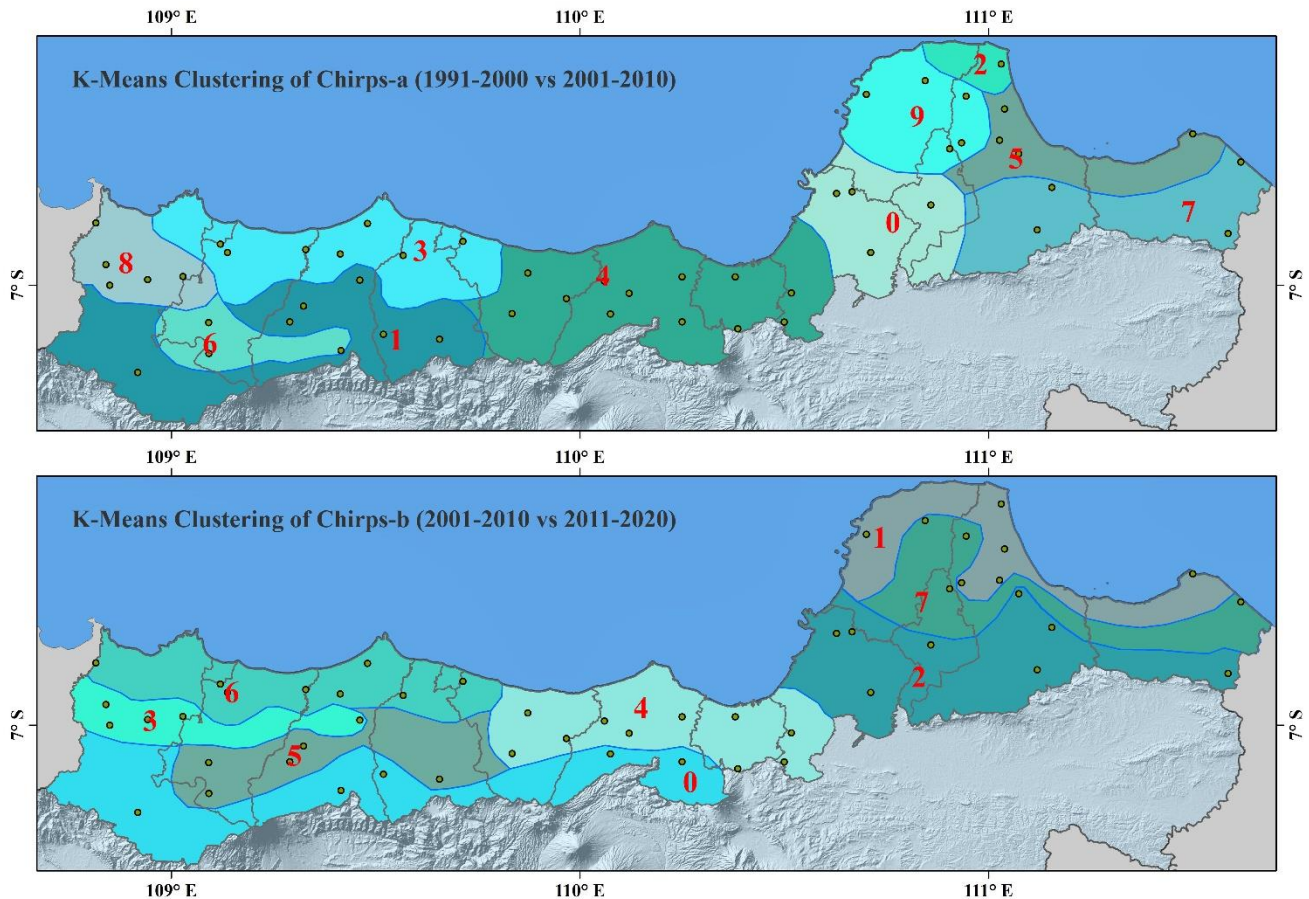


Figure 8. K-means clustering of the rainfall distribution between CHIRPS data between two transition periods (Chirps-a and Chirps-b). The clusters show regionally consistent patterns of rainfall change and emphasize the occurrence of consistent trends of west-east gradients over decades.

3.4.3 Consistency between observations and CHIRPS

Figure 6 assessed the consistency by comparing the TNI values of observational (TNI_Obs) and CHIRPS (TNI_CHIRPS) data on each of the transition periods with the identity line as the reference. In general, the transformation pattern at the station was reproduced by CHIRPS. It provided a greater range of values, showing that there were differences in the representation of the observational and grid data, and the fluctuations around the identity line.

The similarity in the eastward movement of the variability signal in both datasets and the identified transition period showed that the past ten years were characterized by active modification of precipitation patterns in the northeastern coastal area. These results showed the joint influence of monsoonal circulation, trends of changing conditions of the surface in the Java Sea and the intricate coastline topography in the area.

The clustering results collectively provided a comprehensive view of how rainfall transformation occurs in the North Coast. Furthermore, the

complementary results from Observations and CHIRPS, local sensitivity and regional coherence, provided a strong foundation for interpreting long-term climate trends and supporting planning to anticipate the impacts of hydrometeorological disasters on the coast of Central Java.

3.5 Mechanistic interpretation and literature comparison

The focus of this section was on the interpretation of changes in rainfall distribution on a decadal scale by comparing ten-year periods (1991-2000, 2001-2010, 2011-2020) to show interannual variability and also long-term climate evolution. Therefore, ENSO and IOD were viewed only as interannual modulators that alter annual variations or seasonal extreme events, not as causes of decadal rainfall changes as shown in this study. To date, there have been limited studies on histogram-based distribution diagnosis methods and composite indices. However, the decadal pattern changes observed were consistent with other studies that used trend analysis,

rainfall indices, or extreme indices, as well as with the presence of multidecadal changes in tropical coastal areas affected by the monsoon. The result of this present study showed a transformation of moderate to strong rainfall at the station level, as evidenced by the TNI values. These values were derived from observational data (average TNI around 1.8-2.1) and CHIRPS data (around 2.2-3.0), on a scale range of 0-6 over the two decades compared (Table 2).

Similar patterns of changes in rainfall distribution have been reported in several other areas. These included patterns found in a study on seasonal changes in the watershed of the Panama Canal (Hatvani et al., 2025) and shift in distribution in the coastal area of Victoria, Australia (Tolhurst et al., 2023). Although the metric and scopes used vary, these changes appeared to form a distinct regional gradient rather than an average change. These results showed the need to know more about the distribution patterns to determine the real processes of climate change taking place in the area.

In Indonesia, rainfall was very diverse, with the monsoon system influencing the southern area of Java. Similarly, the monsoon influenced the northern region of Central Java (Pantura), with its effects shaped by changes in circulation and sea surface temperature both regionally and globally. This result initiated alterations in the distribution of rainfall that were sensitive to changing regional circulation patterns and sea surface temperatures (Aldrian and Susanto, 2003).

In most instances, the coastal areas, such as the Pantura, enjoyed heavy rainfall along the coastline, and this happened daily due to the interplay of sea and land breezes. This natural process established a space gradient, and the rain pattern in Pantura was different and distinct compared to other areas (Yamanaka, 2016).

ENSO and IOD had a very strong effect on year-to-year changes in rainfall in Indonesia (particularly during the JJA-SON), which grew stronger during SON (Nur'utami and Hidayat, 2016; Supari et al., 2018). The west-central-east pattern uncovered in the rainfall distribution of Pantura showed that day-to-day coastal processes and the monsoon background were the primary forces that influenced the formation of a structured spatial gradient, such as the peak, centroid and width measures. ENSO and IOD continued to regulate annual rainfall variation during certain seasons, making the annual patterns in Indonesia highly dynamic (Kurniadi et al., 2021).

ENSO and IOD provided a reasonable foundation for understanding rainfall variations in Indonesia, and the monsoon circulation was suspected to influence how these two phenomena affect coastal rainfall. This study did not conduct direct attribution analysis using statistical methods, such as correlation or composite indices (Niño 3.4 and DMI), or diagnostics using monsoon indices. Decadal variations and interannual variability were not also divided effectively. In future studies, Additional studies were needed using more detailed methodologies, such as seasonal correlation, composite analysis, or stratification, to better connect rainfall distribution measurements with ENSO, IOD, or monsoon dynamics.

4. CONCLUSION

In conclusion, this study calculates decadal shift in the monthly rainfall distribution using smoothed histograms, six measures of distribution, and the aggregate TNI. The aim is to compare the changes in monthly rainfall distribution in the Pantura (North Coast) of Central Java in terms of 51 BMKG stations and CHIRPS v2.0 (1991-2020). The major results show a distinct west-central-east gradient, with concentration increasing in the west, moderate restructuring in the central area, and a wider distribution with higher variability in the eastern area over the past decade. The CHIRPS data summarizes regional trends but has a tendency to blur out local trends as opposed to observed data.

Changes in the distribution of rainfall should be taken into account when dealing with coastal risk management. The areas where the rainfall concentration is rising should enhance drainage and flood prevention. Furthermore, the mitigation planning should consider peak intensity and variability in order to cope with short and wet months with a lot of rainfall. Increased variability also demands greater flexibility in preparedness plans, including adjustments to planting schedules and drought management. The constant control with the help of rain gauges and satellite data is essential to keep risk maps up to date and prioritize interventions.

A key limitation of this study is that it did not formally attribute change to particular drivers of climate. There is also no explicit distinction between low-frequency (decadal) change and interannual variability, such as ENSO or IOD. Future studies need to measure the connection with Niño 3.4 or DMI

correlation-composites and add monsoon indices as an important background circulation comparator.

ACKNOWLEDGEMENTS

The authors are grateful to the Meteorology, Climatology and Geophysics Agency, especially the Center for Human Resource Development, Central Java Climatology Station, and the Faculty of Fisheries and Marine Sciences, Diponegoro University, for permission to use the data and analysis as part of the study for the Doctoral Program in Marine Science Dissertation, Faculty of Fisheries and Marine Sciences, Diponegoro University.

AUTHOR CONTRIBUTIONS

Conceptualization, Methodology, Validation, Formal analysis, Resources, Data Curation, Visualization, Funding acquisition, and Writing-Original Draft Preparation, Iis Widya Harmoko; Supervision, Writing-Review and Editing, Muhammad Zainuri, Anindya Wirasatriya, and Supari; Software, Formal analysis, Investigation, Aviv Maghridlo.

DECLARATION OF CONFLICT OF INTEREST

The authors declare no conflicts of interest.

REFERENCES

- Afrifa-Yamoah E, Mueller UA, Taylor SM, Fisher AJ. Missing data imputation of high-resolution temporal climate time series data. *Meteorological Applications* 2020;27(1):e1873.
- Agusta VC, Putri MR. Identification of rainfall variability in the northern coast of Java. *AIP Conference Proceedings* 2023;2719(1):Article No. 020015.
- Aldrian E, Susanto RD. Identification of three dominant rainfall regions within Indonesia and their relationship to sea surface temperature. *International Journal of Climatology* 2003; 23(12):1435-52.
- An D, Eggeling J, Zhang L, He H, Sapkota A, Wang YC, et al. Extreme precipitation patterns in the Asia-Pacific regions and its correlation with El Niño-Southern Oscillation (ENSO). *Scientific Reports* 2023;13(1):Article No. 11068.
- Efron B, Tibshirani RJ. *An Introduction to the Bootstrap*. Boca Raton, FL, USA: Chapman and Hall/CRC; 1994.
- Freedman D, Diaconis P. On the histogram as a density estimator: L_2 theory. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 1981;57(4):453-76.
- Funk C, Peterson P, Landsfeld M, Pedreros D, Verdin J, Shukla S, et al. The climate hazards infrared precipitation with stations: A new environmental record for monitoring extremes. *Scientific Data* 2015;2(1):1-21.
- Handayani W, Chigbu UE, Rudiarto I, Putri IHS. Urbanization and increasing flood risk in the northern coast of Central Java—Indonesia: An assessment towards better land-use policy and flood management. *Land* 2020;9(10):Article No. 343.
- Hatvani IG, de Castro Souza NA, Harmon RS, Espinosa JA, Kern Z. Assessment of multidecadal precipitation seasonality in the Panama Canal Watershed. *Journal of Hydrology: Regional Studies* 2025;58:Article No. 102213.
- Ikotun AM, Ezugwu AE, Abualigah L, Abuhaija B, Heming J. K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. *Information Sciences* 2023;622:178-210.
- Jones MC, Marron JS, Sheather SJ. A brief survey of bandwidth selection for density estimation. *Journal of the American Statistical Association* 1996;91(433):401-7.
- Kömüşcü AÜ, Turgu E, DeLiberty T. Dynamics of precipitation regions of Turkey: A clustering approach by K-means methodology in respect of climate variability. *Journal of Water and Climate Change* 2022;13(10):3578-606.
- Kurniadi A, Weller E, Min S-K, Seong M-G. Independent ENSO and IOD impacts on rainfall extremes over Indonesia. *International Journal of Climatology* 2021;41(6):3640-56.
- Kurnianingsih O, Hadiani RR, Setiawan B, Sobriyah. Evaluation and reliability of CHIRPS satellite rainfall products against station rainfall data in the Siwaluh Basin. *Engineering Proceedings* 2025;84(1):Article No. 94.
- Li C, Liu J, Du F, Zwiers FW, Feng G. Increasing certainty in projected local extreme precipitation change. *Nature Communications* 2025;16(1):Article No. 850.
- Magyar JC, Sambridge M. Hydrological objective functions and ensemble averaging with the Wasserstein distance. *Hydrology and Earth System Sciences* 2023;27(5):991-1010.
- Mazziotta M, Pareto A. Normalization methods for spatio-temporal analysis of environmental performance: Revisiting the Min-Max method. *Environmetrics* 2022;33(5):e2730.
- McKay RC, Boschat G, Rudeva I, Pepler A, Purich A, Dowdy A, et al. Can southern Australian rainfall decline be explained? A review of possible drivers. *Wiley Interdisciplinary Reviews: Climate Change* 2023;14(2):e820.
- Nur'utami MN, Hidayat R. Influences of IOD and ENSO to Indonesian rainfall variability: Role of atmosphere-ocean interaction in the Indo-Pacific Sector. *Procedia Environmental Sciences* 2016;33:196-203.
- Parmesan C, Morecroft MD, Trisurat Y. Terrestrial and freshwater ecosystems and their services. In: *Climate Change: Impacts, Adaptation and Vulnerability (Working Group II contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change)*. Cambridge (UK): Cambridge University Press; 2023. p. 197-378.
- Pattipeilohy WJ, Barung FM, Ronsumbre EY. Analysis of changes in seasonal patterns and distribution of rainfall frequency in Sentani. *Buletin GAW Bariri* 2024;5(1):1-9 (in Indonesian).
- Pertiwi DAS, Sutisna S, Supriyatno M, Yosik N, Paski JAI. Study of changes in the distribution of rainfall frequency in Jakarta for the period 1991-2020. *Jurnal Geografi Edukasi dan Lingkungan* 2024;8(2):183-91 (in Indonesian).
- Pillai PA, Ramu DA, Nair RC. Recent changes in the major modes of Asian summer monsoon rainfall: Influence of ENSO-IOD relationship. *Theoretical and Applied Climatology* 2021; 143(3):869-81.
- Pratama YD, Salam A. Comparison of data normalization techniques on KNN classification performance for Pima Indians Diabetes Dataset. *Journal of Applied Informatics and Computing* 2025;9(3):693-706.
- Putri IHS, Buchori I, Handayani W. Hydrometeorological disaster assessment: Study of risk and loss assessment of disaster events in Central Java. *Sustainability and Climate Change* 2022;15(6):446-60.
- Rahman FD, Mulki MIZ, Taryana A. Clustering and classification of Cilacap weather data using the K-means and random forest

- methods. *Jurnal SINTA: Sistem Informasi dan Teknologi Komputasi* 2024;1(2):90-7 (in Indonesian).
- Roux S, Buis S, Lafolie F, Lamboni M. Cluster-based GSA: Global sensitivity analysis of models with temporal or spatial outputs using clustering. *Environmental Modelling and Software* 2021;140:Article No. 105046.
- Roxy MK, Ghosh S, Pathak A, Athulya R, Mujumdar M, Murtugudde R, et al. A threefold rise in widespread extreme rain events over central India. *Nature Communications* 2017;8(1):Article No. 708.
- Shen C, Liu H, Wang J, Yang Z, Hai C. Kullback-Leibler divergence-based distributionally robust chance-constrained programming for PV hosting capacity assessment in distribution networks. *Sustainability* 2025;17(5):Article No. 2022.
- Supari, Tangang F, Salimun E, Aldrian E, Sopaheluwakan A, Juneng L. ENSO modulation of seasonal rainfall and extremes in Indonesia. *Climate Dynamics* 2018;51(7):2559-80.
- Svensson C, Hannaford J, Prosdociimi I. Statistical distributions for monthly aggregations of precipitation and streamflow in drought indicator applications. *Water Resources Research* 2017;53(2):999-1018.
- Tolhurst G, Hope P, Osburn L, Rauniyar S. Approaches to understanding decadal and long-term shift in observed precipitation distributions in Victoria, Australia. *Journal of Applied Meteorology and Climatology* 2023;62(1):13-29.
- University of California, Santa Barbara (UCSB) Climate Hazards Center. CHIRPS: Rainfall Estimates from Rain Gauge and Satellite Observations [Internet]. 2025 [cited 2025 Jun 14]. Available from: <https://data.chc.ucsb.edu/products/CHIRPS-2.0/>.
- Varghese SJ, Pentakota S, Thadivalasa P, Podapati G, Ashok K. Changes in physical characteristics of extreme rainfall events during the Indian summer monsoon based on downscaled and bias-corrected CMIP6 models. *Scientific Reports* 2025; 15(1):Article No. 3679.
- Wani AA. Comprehensive analysis of clustering algorithms: Exploring limitations and innovative solutions. *PeerJ Computer Science* 2024;10:e2286.
- World Meteorological Organization (WMO). WMO Guidelines on the Calculation of Climate Normals [Internet]. 2017 [cited 2026 Mar 6]. Available from: <https://library.wmo.int/records/item/55797-wmo-guidelines-on-the-calculation-of-climate-normals>.
- Xiong J, Yang Y. Climate change and hydrological extremes. *Current Climate Change Reports* 2024;11(1):Article No. 1.
- Yamanaka MD. Physical climatology of the Indonesian maritime continent: An outline to comprehend observational studies. *Atmospheric Research* 2016;178:231-59.