SPATIAL RAINFALL DISTRIBUTION IN THE EASTERN PART OF WEST JAVA PROVINCE USING PRINCIPAL COMPONENT ANALYSIS AND CLUSTER ANALYSIS

Distribusi Curah Hujan Spasial di Timur Provinsi Jawa Barat Menggunakan Principal Component Analysis dan Cluster Analysis

Nyayu Fatimah Zahroh[1)]*, Rahmawati Syahdiza[1)], M. Bayu Rizky Prayoga[2)], Erwin Mulyana[1)], Samba Wirahma[2)], Purwadi[1)], Yudha Setiawan Djamil[1)]

[1)]Badan Riset dan Inovasi Nasional, Jl. M.H. Thamrin No. 8, Jakarta Pusat 10340
[2)]Badan Meteorologi Klimatologi dan Geofisika, Jl. Angkasa I No.2 Kemayoran, Jakarta Pusat, 10610
**E-mail: nyay002@brin.go.id*

Abstract

Over a 10-year period, daily precipitation totals at 74 sites were used to explore spatial rainfall distribution in eastern West Java Province, Indonesia, which includes Indramayu, Cirebon, Majalengka, Kuningan, Garut, Ciamis, Tasikmalaya, and Banjar. The study area also includes the border areas of West Java and Central Java provinces, namely Brebes, Cilacap, and Banyumas. S-mode Principal component analysis (PCA) is used to identify the main rainfall patterns, with stations as component loadings and days as component scores. The first three principal components (PCs) were retained, accounting for nearly 43% of the cumulative variance. PC 1, as the most important pattern, produces a north-to-south decline in rainfall in the study area and is associated with wind direction. Then, PC 2 shows the most rainfall in Cirebon, and the variability reduces in the south (Tasikmalaya). The last, PC 3, produces wetter conditions in Majalengka than in its southern and northern areas (which are drier). Then, the result of PCA was analysed using wind data from ERA5 and topography information, and the westerly wind shows its impact on rainfall in windward areas. Finally, cluster analysis is used to determine which regions have consistent rainfall, and the results were compared with the Koppen-Geiger classification. The application of this study can be helpful for activities related to rainfall distribution, such as agriculture and water resource management.

Keywords: *Rainfall Distribution, Principal Component Analysis, Cluster Analysis, West Java*

Intisari

Selama periode 10 tahun, total curah hujan harian di 74 lokasi digunakan untuk mengeksplorasi distribusi curah hujan spasial di Provinsi Jawa Barat bagian timur, Indonesia, yang meliputi Indramayu, Cirebon, Majalengka, Kuningan, Garut, Ciamis, Tasikmalaya, Banjar. Wilayah studi juga memasuki daerah perbatasan provinsi Jawa Barat dan Jawa Tengah, yaitu Brebes, Cilacap, dan Banyumas. Analisis komponen utama (PCA) S-mode digunakan untuk mengidentifikasi pola curah hujan utama, dengan stasiun sebagai pemuatan komponen dan hari sebagai skor komponen. Tiga komponen utama (PC) pertama mencakup hampir 43% dari varians kumulatif. PC 1, sebagai pola yang paling penting, menghasilkan penurunan curah hujan dari utara ke Selatan dari wilayah studi dan berasosiasi dengan arah angin. Kemudian, PC 2 menunjukkan curah hujan paling banyak di Cirebon, dan variabilitas berkurang di selatan (Tasikmalaya). Yang terakhir, PC 3, menghasilkan kondisi yang lebih basah di Majalengka daripada di daerah selatan dan utara (lebih kering). Kemudian, hasil PCA dianalisis menggunakan informasi angin dari ERA5 dan topografi, dan angin barat menunjukkan dampaknya terhadap curah hujan di daerah yang menghadap ke arah angin. Terakhir, analisis klaster digunakan untuk menentukan daerah mana yang memiliki curah hujan yang konsisten, dan hasilnya dibandingkan dengan klasifikasi Koppen-Geiger. Penerapan studi ini dapat bermanfaat untuk kegiatan yang terkait dengan distribusi curah hujan, seperti pertanian dan pengelolaan sumber daya air.

Kata kunci: *Distribusi Curah Hujan, Principal Component Analysis, Cluster Analysis, Jawa* Barat

1. INTRODUCTION

West Java is the second-largest province, covering the western and central parts of Java Island, with an area of more than 35,000 km², including 900,000 hectares of rice fields. Additionally, the population in West Java reached 49 million people, making it the most densely populated region in Indonesia, according to the 2020 population census (BPS, 2023). With the vastness of the agricultural land and the large number of people living in the area, a hydrometeorological disaster will have a significant impact on both the population and agricultural land.

The eastern region of West Java Province is known for its abundant agricultural production. Indramayu was recorded as the largest rice producer in West Java, with a production of 1.4 million tons in 2022 (BPS, 2024). Other areas east of West Java Province also excel in producing a variety of vegetables and fruits. Therefore, this area needs information on precipitation variability to support agriculture, horticulture, and water management. For example, there are several dams and reservoirs, such as Ciwaringin Dam in Majalengka Regency, Cibatukuda Dam in Tasikmalaya Regency, and Darma Reservoir in Kuningan Regency, which serve as means of managing water resources and mitigating hydrometeorological disasters. Hence, studies on the characteristics of rainfall variability in this region are crucial for effective water management and disaster mitigation.

Few studies have focused on rainfall variability in some provinces on Java Island. Perdinan et al. (2017) investigated rainfall variability using monthly data for West Java Province. Other research also used monthly totals to determine spatial variation in rainfall in southern Central Java Province (Nugroho, 2015) and Yogyakarta Province (Nugroho et al., 2020). This study utilised daily rainfall data because one drawback of using monthly rainfall data is that it is challenging to attribute the rainfall total to a single cause, as monthly totals combine multiple rainfall events (Neil & Phillips, 2011).

Wind speed and direction are essential in understanding cloud formation. For example, the stronger the wind speed in the 700–600 hPa layer, the more complex the process of cloud formation (Mulyana, 2014). Besides, the wind direction can determine the source of the moist air mass that forms rain clouds and where the air mass will be taken. For example, air masses from the ocean will be more humid than those from the continent, which has the potential for cloud formation. Also, air masses forced to rise due to orographic factors will form rain clouds on the windward side and dry air masses on the leeward side (Ahrens, 2009). In some cases, at higher mountain heights, the lower slopes on the windward side are wetter than the upper slopes (Houze Jr, 2012).

This paper aimed to identify the geographic rainfall variability pattern over the eastern part of West Java at the daily timescale by investigating how topography and wind direction influence rainfall totals, producing a spatial rainfall distribution over eastern West Java using PCA, and determining rainfall regionalisation over eastern West Java using cluster analysis.

2. METHODS

This research utilised rainfall data from 74 stations spread across the eastern part of West Java Province, with some sites also located in Central Java Province, as shown in Figure 1. Indramayu Regency, Cirebon Regency, and Cirebon City are on the north coast. The Majalengka Regency, Kuningan Regency, Garut Regency, Ciamis Regency, Tasikmalaya Regency, Tasikmalaya City, and Banjar City are located in the central and southern parts of the area. The study areas also encompass the Central Java Province region, specifically Brebes Regency, Cilacap Regency, and Banyumas Regency. This study area has a monsoon rainfall climate type, and based on the data, the annual rainfall totals in the eastern West Java Province range from approximately 1,700 to 3,800 mm. According to Aldrian and Susanto's (2003) research, two monsoons have a significant impact on the study area: the northwest monsoon, which occurs during the wet season (November-March), and the southeast monsoon, which occurs during the dry season (May-September).

2.1. Principal Component Analysis

This research used daily rainfall data from 74 stations in the eastern West Java Province to perform principal component analysis (PCA) and cluster analysis. This research used S-mode in PCA to identify groups of stations that covary together. From 1992 to 2011, Perusahaan Umum Jasa Tirta II offered seventy-four stations over various periods. Perusahaan Umum Jasa Tirta II is an Indonesian state-owned enterprise engaged in water resources management. Furthermore, 347 days of rainfall data are available for running PCA, which is a relatively large sample size. Based on Comrey and Lee's (1992) classification, a sample size of 100 is considered weak, 300 is considered good, and 1000 is considered excellent. Then, this study utilised SPSS statistical software, developed by IBM, to perform PCA and cluster analysis.

Before conducting PCA, we checked station distributions to ensure sites were evenly spread. Some stations that are close together must be discarded to ensure every site has the same probability of being selected onto PC 1. This research excluded three sites from three groups of nearby stations. Stations Ciamis (89), Losari (226), and Pajajar (243) were discarded due to their proximity and high correlation with the other stations.

After determining the station distribution, the next step is to assess whether the rainfall data are normally distributed. All data must be normally distributed when performing PCA. Since daily rainfall data are typically positively skewed, normal distribution conversion must be done to the rainfall data. Therefore, a base-ten logarithm can transform skewed data into a normal, bell-shaped curve by reducing the high-value effect. A zero value was substituted with 0.01 because it cannot be used in calculating logarithms. Since the data were normally distributed after transformation, PCA was performed on the correlation matrix, as it assigns equal weight to variables based on their variability, thereby solving the measurement scale problem.

Following the assessment of data skewness, it was necessary to evaluate its applicability using the Kaiser-Meyer-Olkin (KMO) measure of sample adequacy and Bartlett's test of sphericity. The result from SPSS indicates that the KMO is close to one (0.934), which means that the correlations are compact and distinct, making them suitable for performing PCA. Besides, the significant value of Bartlett's Test of Sphericity offers information on the association between variables, and the result suggests that variables are significant at the 99.9 % confidence level. Therefore, both KMO and Bartlett's Test results indicate that the variables are adequate for PCA.

Following KMO and Bartlett's Test, we perform statistical rotation to get a modest structure that does not burden the PC1 and is easier to explain. To ensure all PCs independence, this research used the Varimax rotation (an orthogonal rotation) approach, the most commonly utilized approach and usually set as the default in many statistical software packages (Jolliffe, 2002), and applied the eigenvalue greater than one rule by Kaiser (1960) and Catell's scree plot (Catell, 1966) to determine the number of PCs to retain.

After determining the number of PCs to retain, we employ an interpolation method to estimate the barrier between the PC loadings of the stations, which is useful for mapping the PC loadings. The inverse distance weighting (IDW) interpolation method in ArcMap 10.8 was used to map the PC loadings. The maps have a 0.1-value interval for PC loadings to facilitate easier reading.

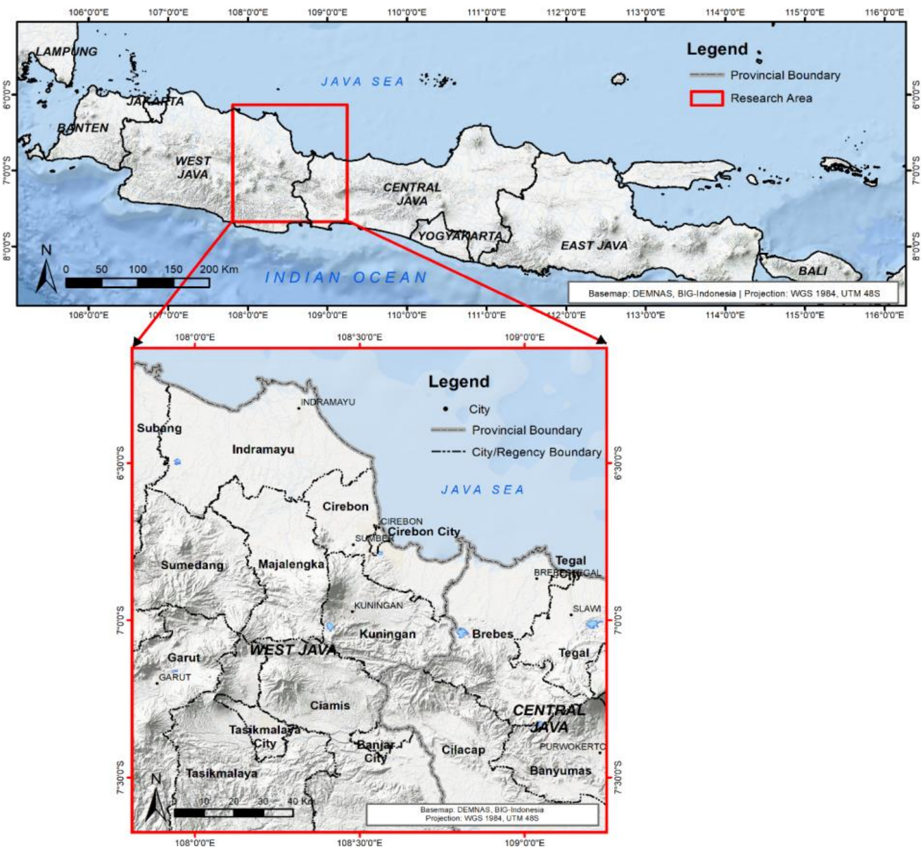

Figure 1. Java Island map, with the red box indicating as the study area.

2.2. Cluster Analysis

Cluster analysis distinguished rainfall regions in the eastern West Java Province, and this research used the dendrogram of Ward's method (a clustering algorithm) to identify the number of clusters. A dendrogram is a tree diagram that represents the relationship between sites; the more correlated they are, the closer they are to each other in the diagram, and vice versa. We chose the agglomerative hierarchical clustering method developed by Ward (1963), which utilises squared Euclidean distances (SED) as an indicator for dissimilarity because of its ability to generate clusters with a similar number of sites compared to alternative methods. Ward's method is quite popular in rainfall regionalisation studies, such as Raja & Aydin (2019), Shirin & Thomas (2016), and Teodoro et al. (2016). In addition, this study employed Clark and Hosking's empirical approach (Eq. 1) to determine the maximum cluster to be utilised, where K denotes the total number of sites and N is the total number of sites (Clark & Hosking, 1986). These methods were compared to decide the number of clusters to analyse.

$$K = 1 + (3.3\log10(N)) \quad (1)$$

3. Result and Discussion

3.1. Principal Component Analysis

According to Kaiser's rule, 14 PCs must be retained with eigenvalues greater than one, which is too many PCs to describe the main rainfall variability. Due to the result not being satisfactory, this study applied Catell's scree plot to determine the number of PCs by breaking the line graph between the steep slope and the tail, as seen in Figure 2. The y-axis describes the eigenvalue, and the x-axis shows the PC number. It can be seen from the figure that the line starts to level off after PC 3. Therefore, this research separated the steep slope and the tail, and decided that 3 PCs had to be retained.

Here is the explanation for the three retained PCs. Table 1 presents the results from the SPSS analysis, indicating that PCs 1 to 3 collectively explain approximately 43% of the variance, with PC 1 accounting for nearly 30%. PC 2 explained 8.5% of the rainfall variability, and PC 3 explained about 4.5% of the variation in precipitation in eastern West Java Province.

Table 1. PCA variance explained.

Total Variance Explained						
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	21.146	29.784	29.784	21.146	29.784	29.784
2	6.055	8.528	38.311	6.055	8.528	38.311
3	3.149	4.435	42.746	3.149	4.435	42.746

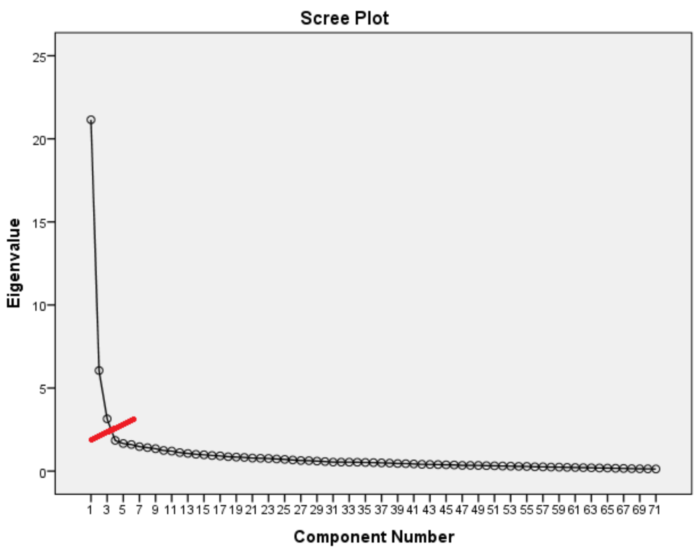

Figure 2. Catell's scree plot with the red line as the break to decide the number of PC to be retained.

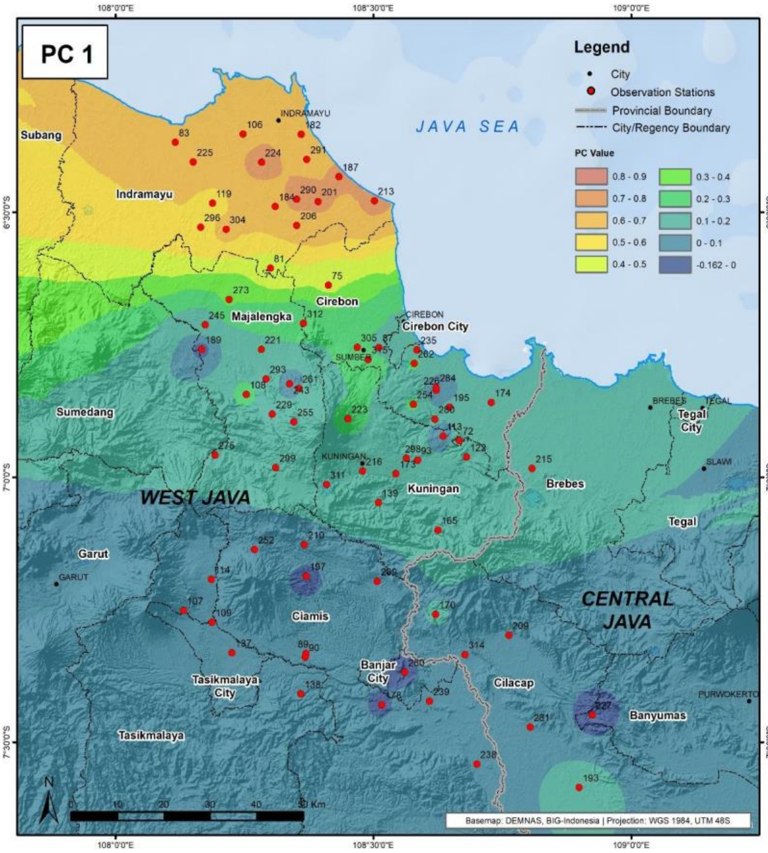

Figure 3. PC 1 loadings, the red dots indicate station locations with the number.

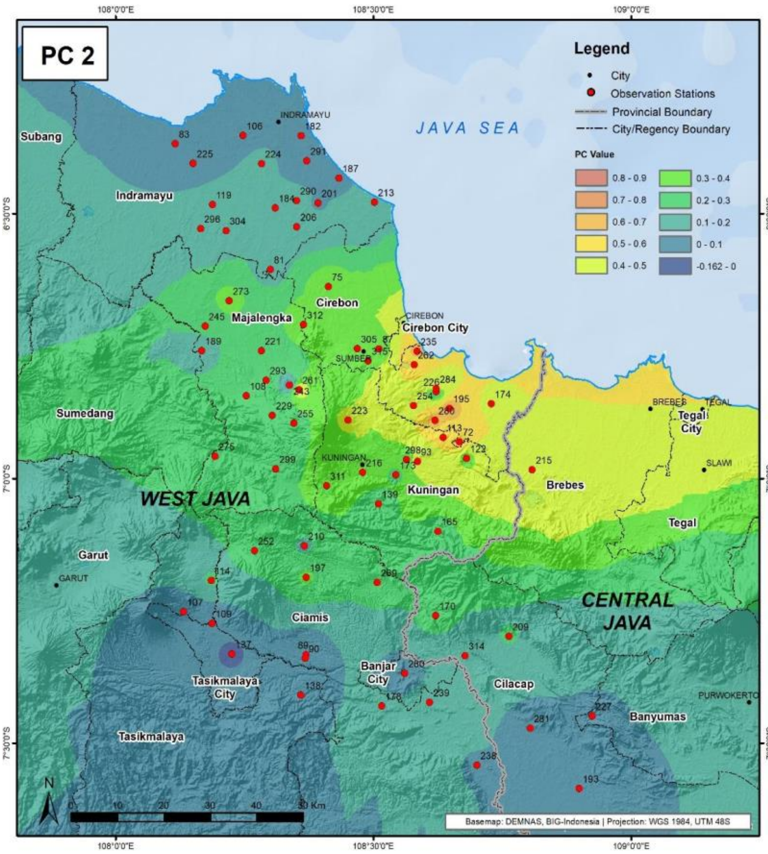

Figure 4. PC 2 loadings, the red dots indicate station locations with the number.

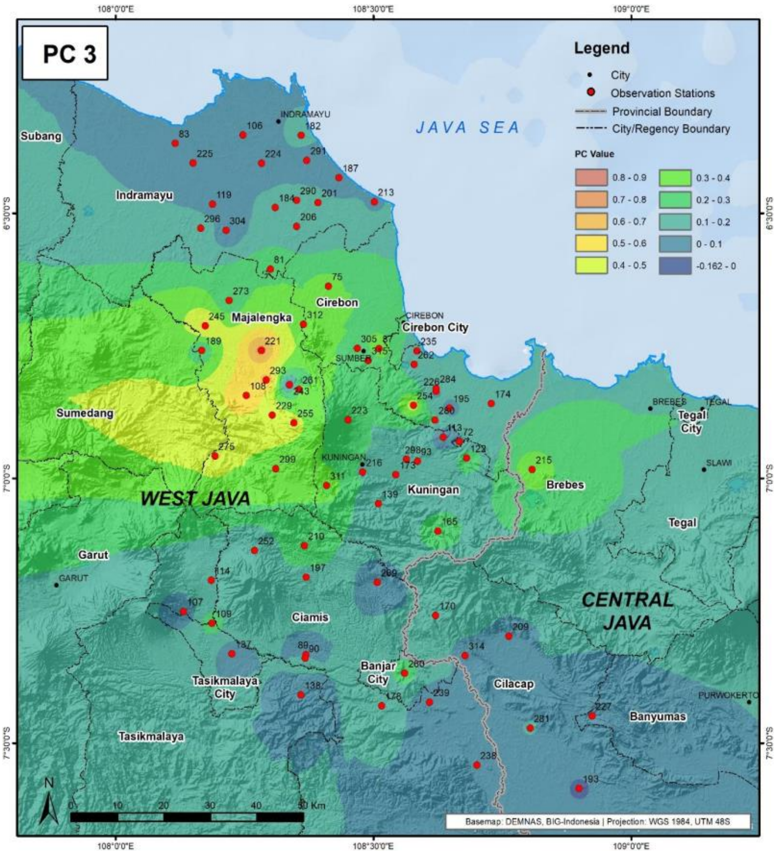

Figure 5. PC 2 loadings, the red dots indicate station locations with the number.

The mapping of PC loadings for estimating rainfall variance can be seen in Figures 3–5. The component loading indicates groups of sites that co-vary together. It is clear that PC1 (Figure 3) exhibits a north-to-south rainfall gradient, accounting for approximately 30% of the variability in the rainfall dataset. The greatest loading of 0.78 focused on Sudikampiran (290) in Indramayu and decreased to the south, with the lowest loading (-0.04) located in Lumbir (227), in Banyumas, East Java. Twenty-one days were picked from PC1 scores above 2.1 (Table 2). The component score indicates the rainfall pattern on a specific day. A strong positive component score suggests that the pattern is apparent on that day. Conversely, days with a high negative score generally experience drier regional conditions. Due to those days having high positive PC scores, rainfall in high-positive loading stations tends to be wetter than in low-loading stations. Table 2 shows that, on those days, Indramayu had higher precipitation than Banyumas.

The PC 2 pattern (Figure 4; explained variance = 8.5%) saw higher PC loadings in Cirebon and decreased to the north (Indramayu) and the south (Tasikmalaya). The highest PC loading was observed at Karangwareng (195), a site in Cirebon, with a value of 0.731, and at Cimulu (137), a station in Tasikmalaya, with the lowest loading at -0.023. Seventeen days with high PC 2 scores were selected, as shown in Table 2. On days with high positive PC 2 scores (>2), rainfall totals in Cirebon tend to be higher than in Tasikmalaya.

PC 3 explained about 4.5% of the variability in the rainfall dataset. The greatest PC loading of 0.733 was observed at Leuweung Gede (221), Majalengka, and decreased to the north and south, with the lowest loading at Karang Bawang (193), Cilacap, at -0.003 (Figure 5). Table 2 shows days with high PC 3 scores (>2) that describe high rainfall in the highest loading and low precipitation intensity in the lowest loading. There were 17 days with high positive PC3 scores, which indicate that Majalengka was wetter than Cilacap.

Table 2. Days with high positive PCs scores, and the rainfall totals at the stations with the highest and lowest rotated PCs component loadings

PC 1				PC 2				PC 3			
	Rainfall (mm)				Rainfall (mm)				Rainfall (mm)		
Date	lumbi227	sudik290	PC scores	Date	cmulu137	kware195	PC scores	Date	kbawa193	leuge221	PC scores
04/03/92	0	23	2.23	01/03/92	0	0	2.03	02/03/92	0	34	2.10
26/03/92	0	17	2.30	09/03/92	6	24	2.04	03/03/92	2	37	2.70
12/04/92	0	23	2.21	09/05/92	0	18	2.40	05/04/92	15	18	2.31
14/04/92	0	29	2.57	29/05/92	6	16	2.10	11/04/92	0	12	3.03
26/04/92	0	58	2.45	20/01/11	4	5	2.33	20/04/92	0	12	2.25
09/05/92	20	29	2.39	27/01/11	11	51	2.19	21/04/92	0	14	2.18
11/05/92	0	4	2.84	29/01/11	0	33	2.10	08/06/92	6	74	2.09
03/06/92	0	8	2.22	12/02/11	0	12	2.04	23/08/92	0	68	2.56
07/01/11	0	15	2.21	24/02/11	0	40	2.30	29/08/92	52	21	2.24
11/01/11	0	17	2.17	12/03/11	6	41	2.52	06/01/11	0	12	2.88
12/01/11	0	23	3.20	04/04/11	51	49	2.05	14/02/11	17	15	2.12
13/01/11	0	21	2.36	07/04/11	2	44	2.26	11/03/11	12	57	2.51
18/01/11	0	2	2.83	08/04/11	0	50	2.06	16/03/11	0	32	2.70
30/01/11	0	18	2.25	10/04/11	10	24	2.38	17/03/11	0	19	2.20
15/02/11	0	5	2.35	05/05/11	18	7	2.09	21/04/11	0	36	2.79
13/03/11	0	70	2.22	17/05/11	0	24	2.72	22/04/11	0	26	2.20
14/03/11	0	48	2.33	29/06/11	0	34	2.41	01/05/11	8	30	3.04
11/04/11	0	9	2.69								
15/04/11	0	49	2.78								
26/04/11	11	10	2.24								
03/05/11	10	27	2.18								

3.2. Cluster Analysis

Figure 6 depicts the dendrogram of Ward's method, split into eight clusters, to achieve optimal rainfall regionalisation in the eastern West Java Province. Clark and Hosking's equation (Eq. 1) results show that the K number for 74 stations was 7.17, which recommends 7-8 clusters as the maximum number. These two results are aligned, and this research applied eight clusters for analysis.

Figure 7 represents the clustering result using Ward's method. Cluster 1 covers the southern research area, namely Ciamis Regency, Tasikmalaya Regency, Tasikmalaya City, Banjar City, and parts of Cilacap Regency. While Cluster 2 is spread out by flanking Cluster 1, two sites are in Garut, and the others are in the western Central Java Province (Cilacap and Banyumas). Between the two clusters lies Cluster 3, which consists of a single site located in Cilacap.

Furthermore, Cluster 4 encompasses Majalengka, Kuningan, the eastern part of Cirebon, and the southern part of Indramayu. Cluster 5 is located in the middle of Cluster 4 in eastern Cirebon. Cluster 6 covers the northern area of Indramayu, while Cluster 7 covers the south of Indramayu and the west of Cirebon. The last, Cluster 8, is south of the study area in Cilacap Regency.

3.3. Discussion

3.3.1. PC1

Zonal and meridional winds from ERA5 (Fifth generation ECMWF reanalysis) are used to determine the dominant winds on high-positive PC1 score days. The predominant wind direction on high PC 1 score days tends to be westerly and from the north, with an average speed of 4-6 m/s. One of the days when the PC1 score was high was March 14, 2011 (Figure 8), when the air mass originated from the Indian Ocean or west of Java Island. On that day, the wind speed in Indramayu was 5 m/s, while in Banyumas, it was 10 m/s. Based on ERA5 data for relative humidity (RH) and convective available potential energy (CAPE) that day, humidity conditions were quite wet in Indramayu (80–90%), and CAPE was high, reaching more than 1000 kg/J. Humidity and CAPE support the totals of 48 mm in Indramayu, but dryness is observed in the southern region.

3.3.2. PC 2

The analysis using the average zonal and meridional winds on PC 2 scores days above 2 indicates that westerly and northern winds are more dominant, with wind speeds of 3–4 m/s. One of the days when the PC 2 score was high was 8 April 2011 (Figure 9), when winds at the 850 hPa level blew from the west, bringing air masses from the Indian Ocean and producing 50 mm of rain in Cirebon and 0 mm in Tasikmalaya. The information about RH and CAPE from ERA5 also supports the wet condition in Cirebon, with values of approximately 80% and 1000 kg/J, respectively. Besides the wind influence, the topography can contribute to the PC2 pattern, as Tasikmalaya has a higher elevation than Cirebon, which is located in a coastal area. Because Tasikmalaya is located south of Mount Ciremai, the city is on the leeward side when northerly winds flow. The air moves downhill, bringing warm and dry air since the air moisture (as cloud and rainfall) was removed on the windward side (Ahrens, 2009).

3.3.3. PC3.

Based on zonal and meridional wind data at the 850 hPa layer, days with a high positive PC 3 score (above 2) tend to have westerly winds and strong southerly flows. The average wind speed is around 3 m/s. On 2 March 1992 (Figure 10), the wind blew from the south. However, based on point data in Leuweung Gede, Majalengka, the wind tends to be neutral with low wind speeds (average less than 1 m/s). In contrast, in Karang Bawang, Cilacap, the wind tends to be stronger at around 4.8 m/s with strong westerly and southerly winds. The high wind velocity in Cilacap impacted the dry weather (0 mm). However, lower wind speeds caused wetter conditions in Majalengka, with rainfall of 34 mm. According to Mulyana (2014), high wind speeds below the 600 hPa layer can disrupt the cloud growth process.

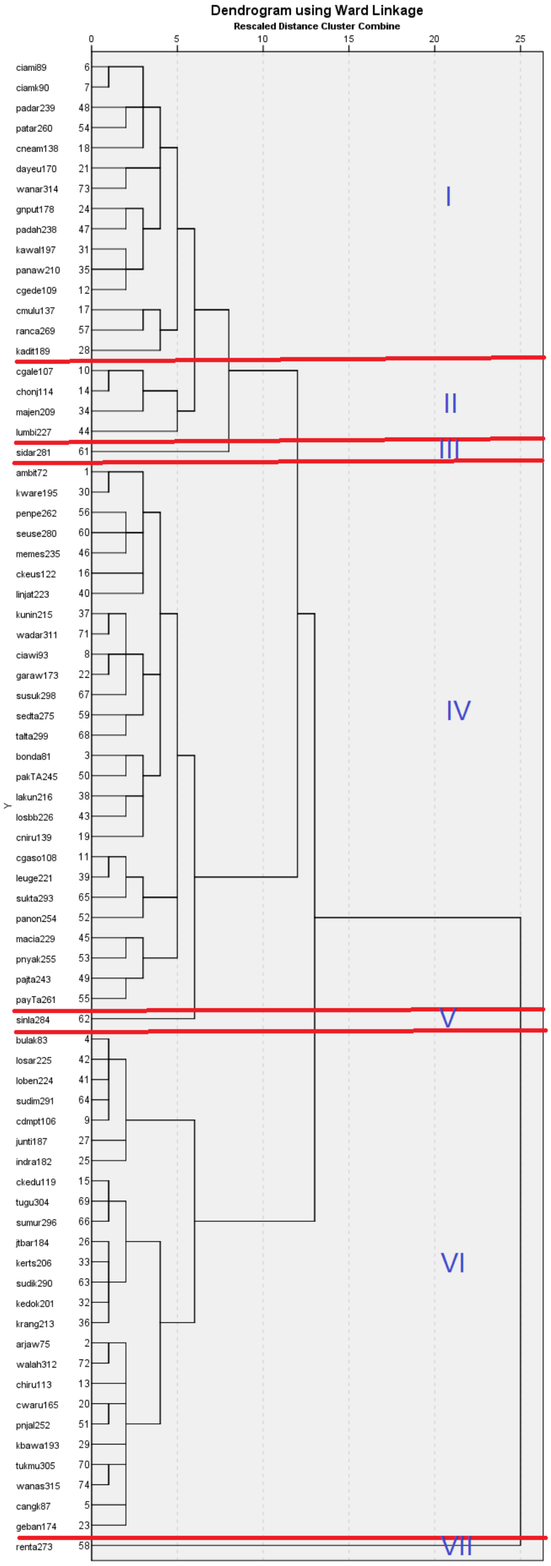

Figure 6. Dendrogram of Ward's method, the red lines separate each cluster.

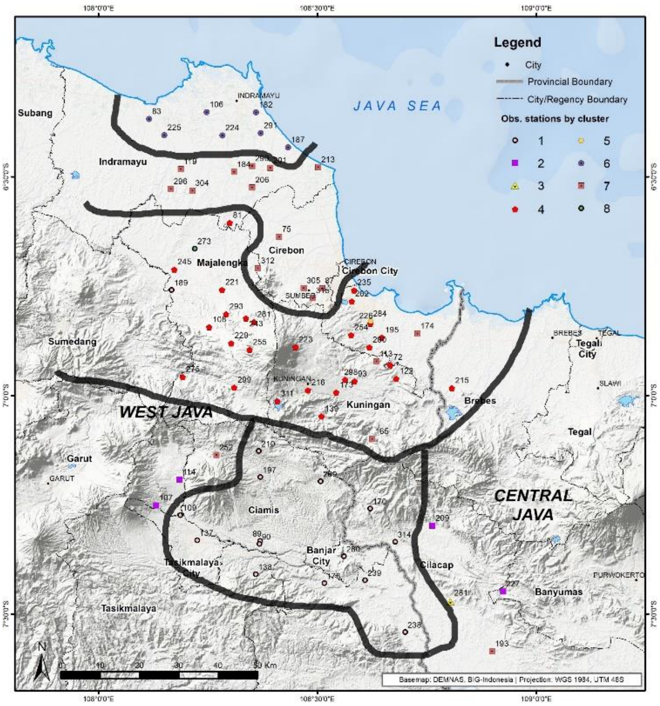

Figure. 7. Rainfall regionalisation in eastern of West Java Province, the black curves separate each cluster region.

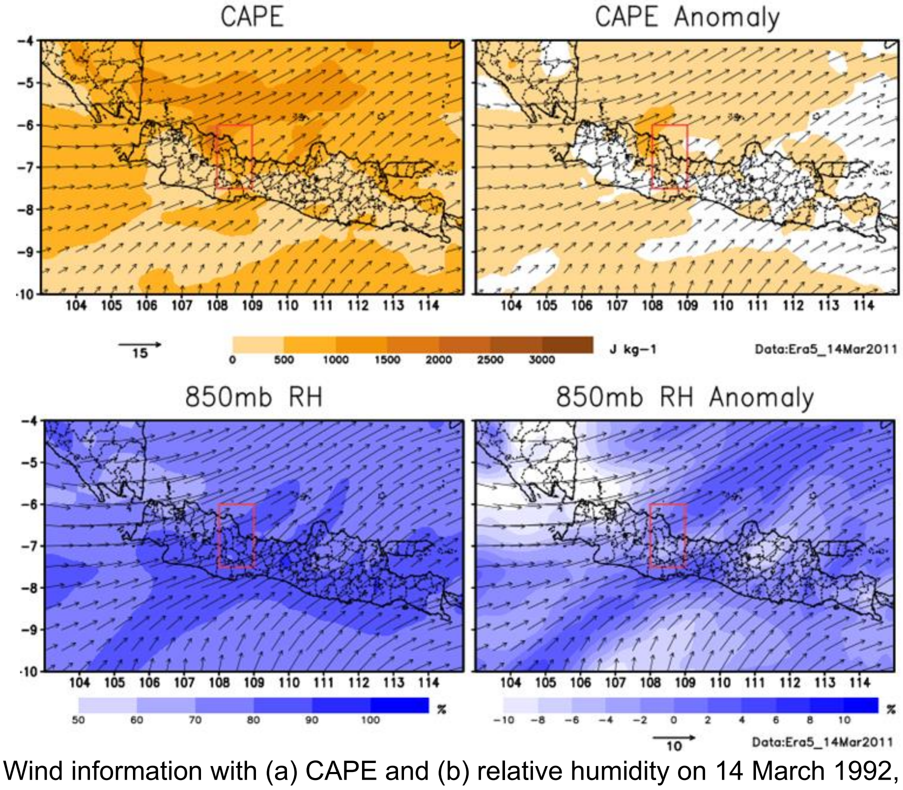

Figure 8. Wind information with (a) CAPE and (b) relative humidity on 14 March 1992, black arrows show the wind direction, and the length arrow indicates the wind speed (the longer, the faster), the red box shows the study area.

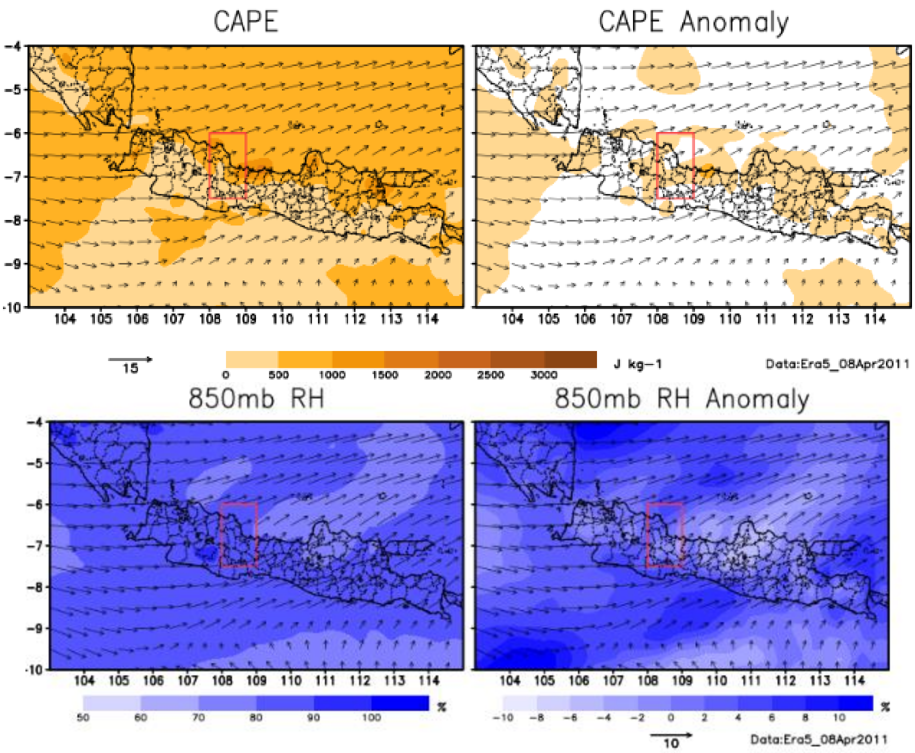

Figure 9. Wind information with (a) CAPE and (b) relative humidity on 8 April 2011, black arrows show the wind direction, and the length arrow indicates the wind speed (the longer, the faster), the red box shows the study area.

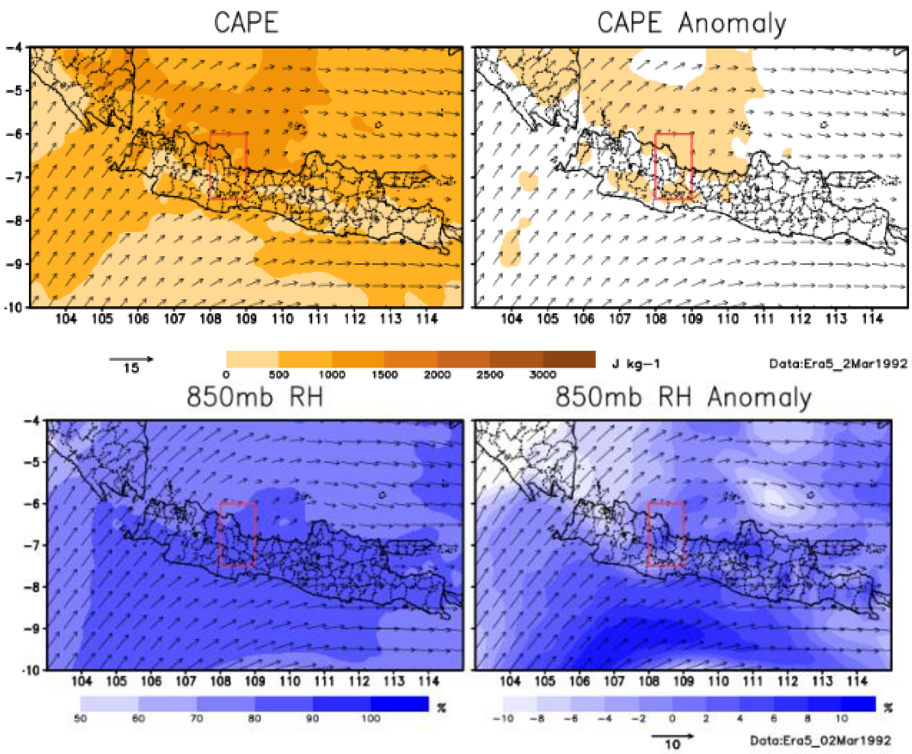

Figure 10. Wind information with (a) CAPE and (b) relative humidity on 2 March 1992, black arrows show the wind direction, and the length arrow indicates the wind speed (the longer, the faster), the red box shows the study area.

According to Table 2, the days included in PC 1, PC 2, and PC 3 are primarily during the wet seasons (November-March), as classified by Aldrian and Susanto (2003). However, some other days fall into the transition month (April) and the dry seasons (May-September). A longer data period is required to observe the association between rainfall and climate phenomena, such as the MJO and ENSO. Additionally, due to the similarity in composition days among each PC, it can be assumed that local factors, such as topography and local winds, more dominantly influence rainfall distribution in eastern West Java.

3.3.4. Cluster Analysis

Several stations are far from their cluster or enter other cluster areas. Cluster 7 has five stations that appear randomly in the different clusters: three stations enter Cluster 4 (site numbers 174, 113, and 165), and two stations enter Cluster 2 (site numbers 252 and 193). Additionally, there are stations 189 (Cluster 1), 273 (Cluster 8), and 284 (Cluster 5), which are located in the Cluster 4 area.

Based on the clustering outcome, there is a justifiable basis for questioning the integrity of the data or the accuracy of the site's placement on the map. The use of Pearson correlation coefficients (r) in computing the correlation between the anomalous site and other sites may serve as a means to validate the accuracy of the data. When two stations are close, the correlation will be close to +1. Otherwise, if the distance is farther, the correlation will decrease. To begin with, it is necessary to compute the Pearson correlation coefficient between the anomalous site and the adjacent site using SPSS. Subsequently, the correlation between the anomalous site and the site belonging to the same cluster should be determined.

Tables 3–5 show the Pearson correlation between the anomalous and other sites from the SPSS run. In Table 3, sites 174, 113, and 165 correlate more closely with site 195 in Cluster 4 than with site 213 (which is in the same cluster with three stations). It can also be seen from Table 4 that site numbers 252 and 193 (Cluster 7) are closer to site number 114 (Cluster 2) than site 213, which is in the same cluster as both sites. Additionally, site 189 correlates more closely with the site in Cluster 4 than with the site in Cluster 1 (Table 5). However, because Cluster 5 and Cluster 8 only have one station, it can be suggested that both sites are part of Cluster 4.

Table 3. Pearson Correlation between the Cluster 4 site and Cluster 7 sites

Pearson Correlation	lggeban174(Cluster 7)	lgchiru113 (Cluster 7)	lgcwaru165 (Cluster 7)
lgkware195(Cluster 4)	0.484	0.426	0.365
lgkrang213 (Cluster 7)	0.266	0.231	0.230

Table 4. Pearson Correlation between the Cluster 2 site and Cluster 7 sites

Pearson Correlation	lgpnjal252 (Cluster 2)	lgkbawa193 (Cluster 2)
lgchonj114 (Cluster 2)	0.383	0.268
lgkrang213 (Cluster 7)	0.173	0.142

Table 5. Pearson Correlation between the Cluster 4 site and Cluster 1 sites

Pearson Correlation	lgkadit189 (Cluster 1)
lgpakta245 (Cluster 4)	0.433
lgpanaw210 (Cluster 1)	0.273

The map from cluster analysis can be compared with the Köppen-Geiger climate classification map, as depicted in Beck *et al*. (2018). The classification is defined by threshold values, as well as the seasonality of monthly air temperature and precipitation totals (Beck *et al*., 2018). If we focus on the study area, there are four climate classifications: tropical rainforest, tropical monsoon, tropical savannah, and temperate. Compared with cluster analysis results, Cluster 6 in the coastal area of Indramayu has a tropical savannah climate. Additionally, the areas of Cluster 7 and 4 (in southern Indramayu, Cirebon, Majalengka, Kuningan, and Brebes) fall under the tropical monsoon climate type. Additionally, Clusters 1 and 2 encompassed Garut, Ciamis, Tasikmalaya, Banjar, Cilacap, and Banyumas, which are characterised by tropical rainforest climates. The temperate climate type describes the Mount Ciremai region considered in Cluster 4.

4. Conclusion

The pattern of PCA results shows that wind direction and velocity influence rainfall variability in eastern West Java Province. Lower wind speeds support cloud formation more effectively than areas with higher wind speeds. Additionally, information related to relative humidity and CAPE at the 850 hPa level suggests wetter tropospheric conditions and potential for cloud formation. Apart from wind information, the PC 2 pattern is also influenced by topographic conditions, where windward areas are wetter than leeward ones. From the results of the cluster analysis, it can be seen that the pattern can be compared to the Köppen-Geiger climate classification.

Suggestion

Suggestions for future research include expanding the area of coverage and extending the data period to provide a more comprehensive analysis of spatial rainfall variability. Climate variability modes, such as ENSO and IOD, can also be utilised in analyses of studies with longer data periods.

References

Ahrens, C. D. (2009). *Meteorology today: an introduction to weather, climate, and the environment*, *Ninth Edition*. California: Brooks/Cole

Aldrian, E., & Susanto, R. D. (2003). Identification of three dominant rainfall regions within Indonesia and their relationship to sea surface temperature. *International Journal of Climatology*, *23*, 1435-1452.

Badan Pusat Statistik Provinsi Jawa Barat. (2024, November 19). Produksi Tanaman Padi (Ton), 2023-2024. https://jabar.bps.go.id/id/statistics-table/2/MzAxIzI=/produksi-tanaman-padi.html

Badan Pusat Statistik Provinsi Jawa Barat (2024, November 19). Jumlah Penduduk Menurut Kabupaten/Kota (Jiwa), 2018-2020. https://jabar.bps.go.id/indicator/12/133/1/jumlah-penduduk-menurut-kabupaten-kota.html

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific data*, *5*(1), 1-12.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate behavioral research*, *1*(2), 245-276.

Clark, W. A., & Hosking, P. L. (1986). Statistical methods for geographers.

Comrey, A. L., & Lee, H. B. (2013). *A first course in factor analysis*. Psychology press.

Houze Jr, R. A. (2012). Orographic effects on precipitating clouds. *Reviews of Geophysics*, *50*(1).

Jolliffe, I.T. (2002) *Principal Component Analysis.* Second Edition. New York: Springer-Verlag.

Kaiser, H. F. (1960). The Application of Electronic Computers to Factor Analysis. *Educational and Psychological Measurement, 20*(1), 141-151.

Mulyana, E. (2014). Pengaruh Angin Terhadap Pertumbuhan Awan Hujan di DAS Waduk PLTA Kota Panjang. *Jurnal Sains & Teknologi Modifikasi Cuaca, 15*(2), 81-88.

Neal, R. A., Phillips, I. D. (2011). Winter daily precipitation variability over Cumbria, Northwest England. *Theoretical and applied climatology, 106*, 245-262.

Nugroho, B.D.A. (2015). Relationships between Sea Surface Temperature (SST) and rainfall distribution pattern in South-Central Java, Indonesia. *The Indonesian Journal of Geography, 47*(1), 20.

Nugroho, B. D. A., Nurrohman, R. K., Sudira, P. (2020). Relationships among global climate indices and rainfall pattern to detect impact of climate change in yogyakarta special region, Indonesia. *IOP Conference Series: Earth and Environmental Science, 515*(1).

Perdinan, Adi, R. F., Sugiarto, Y., Arifah, A., Arini, E. Y., Atmaja, T. (2017). Climate regionalization for main production areas of Indonesia: Case study of West Java. *IOP Conference Series: Earth and Environmental Science 54*(1).

Raja, N. B., & Aydin, O. (2019). Regionalization of precipitation in Mauritius: a statistical approach. *Meteorological Applications*, *26*(4), 711-719.

Shirin, A. S., & Thomas, R. (2016). Regionalization of rainfall in Kerala state. *Procedia Technology*, *24*, 15-22.

Teodoro, P. E., de Oliveira-Júnior, J. F., Da Cunha, E. R., Correa, C. C. G., Torres, F. E., Bacani, V. M., ... & Ribeiro, L. P. (2016). Cluster analysis applied to the spatial and temporal variability of monthly rainfall in Mato Grosso do Sul State, Brazil. *Meteorology and Atmospheric Physics*, *128*, 197-209.

Ward Jr, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, *58*(301), 236-244.