Analysis of Homogeneity of Rainfall Data Between Stations in the Mapat River Catchment Area

Dani Helmi[1], Nurhayati[2], Danang Gunarto[3]
[1,2,3]Department of Civil Engineering, Universitas Tanjungpura
Jl. Prof. Dr. H. Hadari Nawawi, Pontianak, Kalimantan Barat
[1]E-mail: danihelmi@student.untan.ac.id

Abstract — Homogeneous rainfall data is essential in hydrological analysis because it affects the accuracy of calculations and modelling. Inhomogeneity can be caused by environmental changes, topographical differences, or recording errors; therefore, it is necessary to evaluate it before using the data in follow-up analysis. This study aims to test the homogeneity of rainfall data between stations in the Mapat River catchment area, Bengkayang Regency. The data used is the maximum daily rainfall for the period 1993–2022 from six stations: Dawar, Bengkayang, Sanggau Ledo, Karangan, Tebas, and Serukam. The data were directly tested for homogeneity using a two-sided t-test to compare the averages between station pairs at a significance level of 1%. The results of 15 combinations of station pairs showed that the total calculated t-value was smaller than the critical t (2.663), indicating that there was no significant difference in the average maximum daily rainfall between stations. The highest t-value was recorded in the Dawar–Bengkayang pair (2.19), which, although close to the critical limit, remained in the homogeneous category. These findings suggest that the variation in data is due to local climatic factors, rather than differences in instruments or recording methods. Complete homogeneity ensures the feasibility of data for a wide range of hydrological analyses, including river discharge modeling, flood analysis, and water resource management planning. The results of this study also demonstrate the consistency in the management of rain stations in the Bengkayang area over the past three decades.
Keywords: rainfall data; homogeneity, double-sided t-test, Rainfall Catchment Area; Mapat River.

I. INTRODUCTION

Rainfall data plays a significant role in hydrological analysis and water resource planning. The use of accurate and relevant rainfall data is crucial for understanding the hydrological dynamics of an area and managing water resources effectively, particularly in terms of the effect of spatial and temporal rainfall distribution on the accuracy of rainfall models. The variability of rainfall data is a primary source of error in hydrological analysis, making it crucial for researchers to consider this aspect when investigating the impact of rainfall on soil moisture and river flows (Anggraheni et al., 2022). The completeness of rainfall data poses a challenge in hydrological modelling, where incompleteness can occur due to equipment malfunctions or other factors (Atica et al., 2022). The accuracy of discharge calculations, flood analysis, and water balance calculations is highly dependent on the quality of the data used. One of the primary indicators of data quality is homogeneity, which refers to the uniformity of statistical characteristics of the data over time or across observation locations (Jung & Kim, 2023; Makanda et al., 2022). The homogeneous use of data facilitates more reliable analysis, allowing researchers and water resource managers to make more informed decisions in downstream water management and disaster response (Harris et al., 2023; Kumar et al., 2023). Inhomogeneous data can lead to significant analytical errors and inaccurate evaluations of water resource conditions (Makanda et al., 2023).

The homogeneity of rainfall data in meteorological and hydrological research cannot be relied upon solely because of the inhomogeneity caused by various environmental factors, including changes in the surrounding environment that can alter the characteristics of the received rain data. For example, a rain measuring device that serves as a data point may undergo modification or change, which may ultimately affect both the accuracy and consistency of the measurement (Iryani et al., 2022). The quality and accuracy of rainfall data are crucial, particularly in the context of hydrological modeling and water resource management (Iqbal et al., 2022). Topographic differences are key variables that can contribute to the inhomogeneity of rainfall data. Areas with diverse topography have the potential to have very different rainfall patterns, resulting in non-uniform temporal and spatial distributions (Mahmoud et al., 2020). Rainfall data collected from locations with varying meteorological conditions can yield inconsistent results when applied to large hydrological models, particularly in regions with distinct climates, such as West Africa (Dembélé et al., 2020; Ekwezuo et al., 2024). Factors, both human and mechanical, can also cause inhomogeneity. Errors in recording rainfall measurement results can arise from how the data is transmitted or stored (Johnson et al., 2021). Data collection techniques from satellites and ground-based measuring instruments may, in many cases, provide different data, which requires researchers to exercise caution when interpreting the results (Iryani et al., 2022; Mahmoud et al., 2020). The challenges in maintaining the homogeneity of rainfall data are complex and involve a variety of environmental and technical changes. Understanding these inhomogeneities is crucial for accurately analyzing climate impacts and rainfall behaviour that changes over time, particularly in the context of water resources management and climate adaptation planning (Al-Muhyi et al., 2024; Howard et al., 2021). Although many studies have examined the quality of rainfall data, there are still areas that lack a specific homogeneity study, including the Mapat River catchment area in Bengkayang Regency.

A crucial question: Does the rainfall data from the stations in the Mapat River catchment area have an adequate level of homogeneity to be used in advanced hydrological analysis? The answer to this question is crucial, considering that the data serve as a basis for various planning and management of water resources in the region.

A statistical homogeneity test is needed that can provide certainty about the uniformity of data between stations. The double-sided t-test method was chosen because it is simple, effective, and widely used in hydrological analysis to compare averages between two data groups. This study aims to test the homogeneity of rainfall data between stations in the Mapat River catchment area using the two-sided t-test method, in order to ensure the data's suitability for use in subsequent hydrological analyses.

II. LITERATURE REVIEW

Rainfall data homogeneity refers to the uniformity of the statistical distribution of data over time or across interrelated locations. Homogeneous data reflect the absence of systematic changes due to external factors other than natural climate variability. Inhomogeneity may arise due to instrument changes, changes in station location, modification of the surrounding environment, or recording errors (Soewarno, 1995).

Rainfall homogeneity data is a crucial element in climate analysis, which focuses on the uniformity of the statistical distribution of rainfall data over time or between interrelated locations. Homogeneous data are needed to ensure that the variations detected in rainfall data are primarily due to natural climate change, in the absence of systematic interference from external factors such as instrument changes or environmental modifications. Inhomogeneity can arise due to several factors, including instrument changes, changes in station locations, or recording errors, all of which can impact the accuracy and reliability of the data (Ibebuchi & Abu, 2023; Paramasivam et al., 2023).

Reliable statistical methods are essential for assessing the homogeneity of the data. Rainfall variability analysis using data from measurement stations reveals inconsistencies in the distribution of rainfall across different regions, which can be attributed to several factors mentioned above. The presence of influences such as environmental modification and instrument changes also contributes to this inhomogeneity (Ibebuchi & Abu, 2023), indicating that the representation of rainfall patterns in the tropics is a challenge without adequate improvement of the observational data (Ibebuchi & Abu, 2023).

Rainfall data in different regions often necessitate the use of recommended statistical approaches for analyzing rainfall distribution and testing for homogeneity. Homogeneous regional boundaries are important for estimating the maximum possible rainfall (Paramasivam et al., 2023). They utilized more than 25 years of data from 161 stations in Malaysia to estimate the maximum daily rainfall, highlighting the importance of homogeneous data in these estimates (Paramasivam et al., 2023).

Inhomogeneity in rainfall data can have serious consequences in decision-making related to water resource management and infrastructure planning. In-depth understanding and careful evaluation of statistical features and uncertainties in rainfall data should be conducted throughout the research process to

ensure that the proposed policy is based on an accurate understanding of prevailing rainfall patterns (Santos & Nascimento Duarte, 2023).
The double-sided t-test is one of the statistical methods used to determine whether two data groups have significantly different averages. In the context of rainfall data, this method is used to compare the average annual data from two different stations. If the value $|t| < t$ critically, the data of the two stations is considered homogeneous. This method is simple, easy to apply, and has been widely used in hydrological studies to verify data quality. It is important because data homogeneity is a significant prerequisite in many further statistical analyses (Francsdito et al., 2023). The use of the t-test method in hydrological studies has been demonstrated through several studies, which show the application of statistics in validating rainfall data. Francsdito et al. pointed out that, before conducting further analysis, it is necessary to conduct a data quality test at two rain stations, namely Cawang and Kemayoran, which confirms the importance of data validation before comparing averages (Francsdito et al., 2023). In addition, research by Alnino et al. also emphasizes the importance of validating data from various sources, including satellite data and observation stations, to ensure that the results obtained can be accurately accounted for (Alnino et al., 2022). This validation aligns with the research conducted by Cipto et al., who stated that the use of satellite data should be accompanied by statistical testing to determine its reliability (Hastina et al., 2023).

III. METHOD

1. Research Location

The research was conducted in the Mapat River catchment area, Bengkayang Regency, West Kalimantan. This area has six rain stations managed by the Balai Wilayah Sungai Kalimantan (BWSK) I in Pontianak, namely: Dawar Station, Bengkayang Station, Sanggau Ledo Station, Karangan Station, Tebas Station, and Serukam Station.

2. Data and Observation Period

The data used are in the form of daily rainfall for 30 years (1993-2022) from each station. The data is processed into annual rainfall for homogeneity testing purposes.

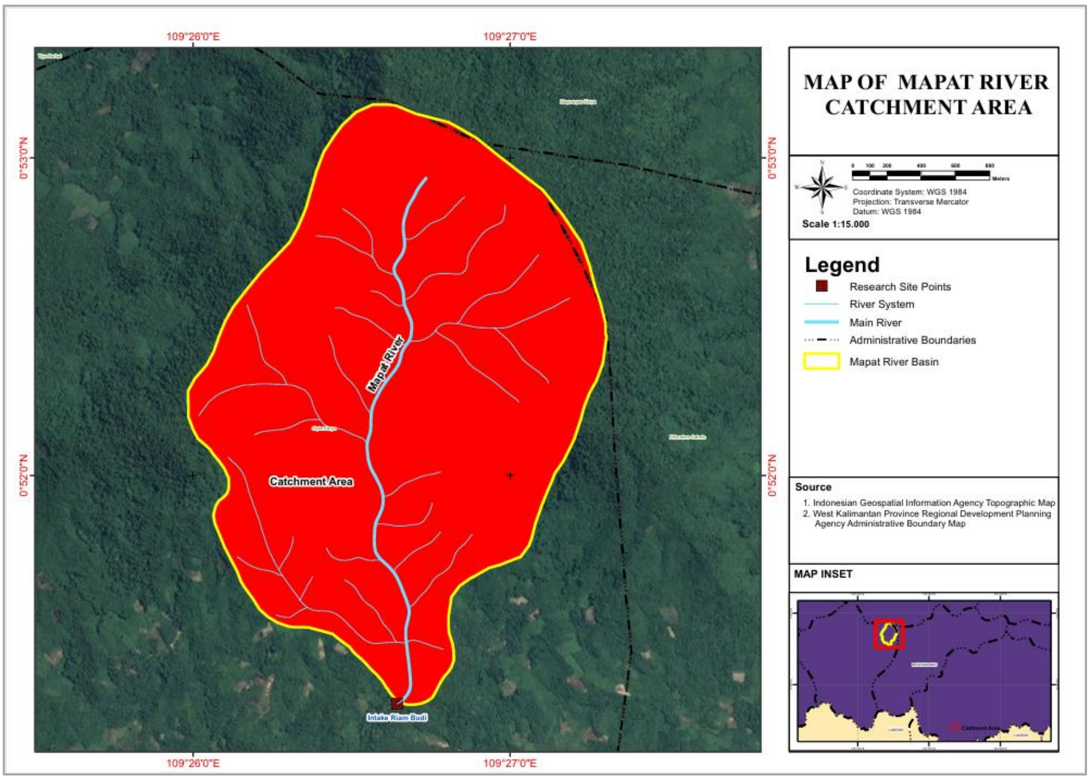

Figure 1. Map of the Mapat River catchment area

Figure 1 shows a map of the research location, specifically the Mapat River catchment area. The geographical location of the study is at 0º8'54.6" N and 109°44'42.93" E.

3. Research Stages
   a. Data Collection – Collect rainfall data from six stations in the Mapat River catchment area.
   b. Data Processing – Converting daily data into annual data and checking the completeness of the data.
   c. Homogeneity Testing – Using a double-sided t-test method to compare average data between station pairs. The t-test formula used is:

$$t = \frac{\overline{X}_1 - \overline{X}_2}{S\sqrt{\frac{1}{N_1} + \frac{1}{N_2}}} \quad (1)$$

with

$$S = \sqrt{\frac{(N_1 - 1)S_1^2 + (N_2 - 1)S_2^2}{N_1 + N_2 - 2}} \quad (2)$$

   d. Interpretation of Results – Comparing the value of $|t|$ with the t critical at a significance level of 5% to determine the homogeneity of the data.

Decision criteria:
Homogeneous: $|t| < t$ critical
Not Homogeneous: $|t| \geq t$ critical

Table 1. *t* Value Two-sided (DATAtab, 2025)

df	0.000	0.500	0.600	0.700	0.800	0.900	0.950	0.980	0.990	0.998	0.999
1	0.000	1.000	1.376	1.963	3.078	6.314	12.710	31.820	63.660	318.310	636.620
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300

IV. RESULTS AND DISCUSSION

Homogeneity testing was conducted on rainfall data in the Mapat River catchment area to ensure statistical uniformity between stations. This study utilizes annual rainfall data from 1993 to 2022 for six stations: Dawar, Bengkayang, Sanggau Ledo, Karangan, Tebas, and Serukam. The data provides an overview of long-term rainfall conditions in the study area.

The double-sided t-test method is used on all station pairs to assess the uniformity of the average annual rainfall. The test process yielded

15 combinations of station pairs, each with a calculated t-value compared to the critical t-value at a 5% significance level. The results of the comparison indicate whether there is a significant difference between stations during the analysis period.

The maximum daily rainfall data is displayed from each station before the results of the homogeneity test are discussed further. The data graphs are presented in Figures 2, 3, 4, 5, 6, and 7 to show the variation in extreme rainfall at the six observation stations.

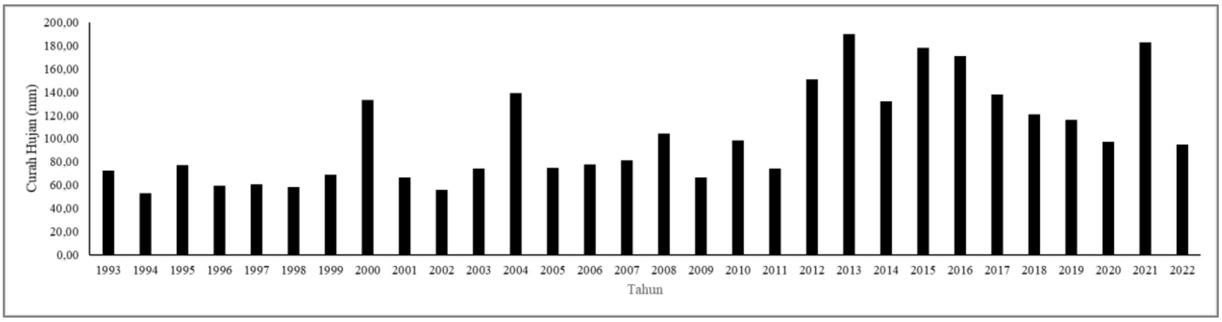

Figure 2. The daily rainfall at Dawar station (SBS-17) (Pemerintah Republik Indonesia, 2023).

Based on Figure 2, the rainfall patterns at Dawar station during the period 1993-2022 show temporally increased variability with three distinct identifiable phases. The initial period (1993-2002) exhibited a relatively stable and low rainfall pattern, with values ranging from 50 to 80 mm per year, except for the anomaly in 1999, which reached 140 mm. The transition period (2003-2011) began to exhibit increased variability, with several peaks in 2004 and 2008, reaching approximately 140 mm and 110 mm, respectively. The highest period occurred from 2012 to 2019, with a maximum peak in 2013 at 190 mm, followed by consistently high years (150-180 mm). The final period (2020-2022) exhibits a downward trend, with the value returning to around 90-180 mm.

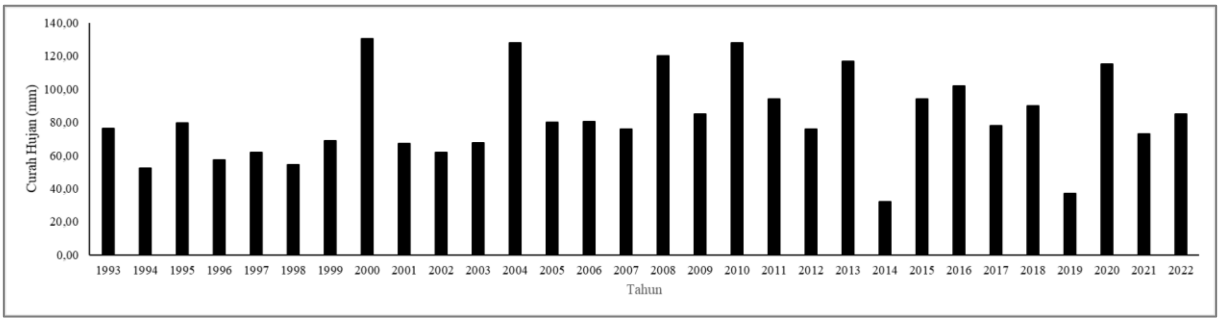

Figure 3. The daily rainfall of Bengkayang station (SBS-07) (Pemerintah Republik Indonesia, 2023).

Based on Figure 3, the rainfall pattern of Bengkayang station from 1993 to 2022 exhibits different characteristics, with more moderate fluctuations compared to previous data. The data show three major peaks reaching about 130 mm in 1999, 2004, and 2020, with a relatively uniform distribution pattern throughout the observation period. The 1993-1998 and 2014-2019 periods exhibit low variability, with rainfall ranging from 55 to 80 mm. In contrast, the 1999-2013 and 2020-2022 periods display higher variability, characterized by alternating between high (100-130 mm) and low (30-80 mm) values.

Figure 4 shows the rainfall pattern of Sanggau Ledo Station, which is almost identical to the previous data. There is a consistent baseline of around 50-80 mm throughout the observation period, with five extreme peaks reaching 140-210 mm. Significant peaks occurred in 2000 (133.14 mm), 2004 (139.05 mm), 2016 (164 mm), 2021 (156 mm), and the highest peak in 2022 (205 mm).

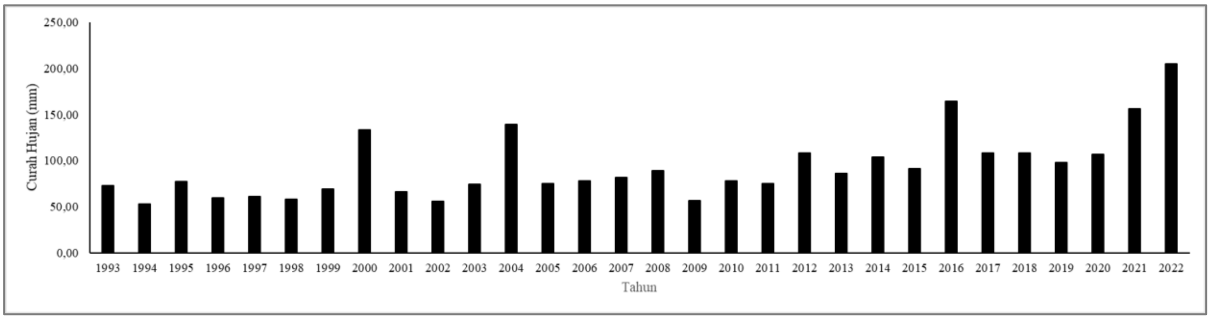

Figure 4. The daily rainfall of Sanggau Ledo station (SBS-03) (Pemerintah Republik Indonesia, 2023).

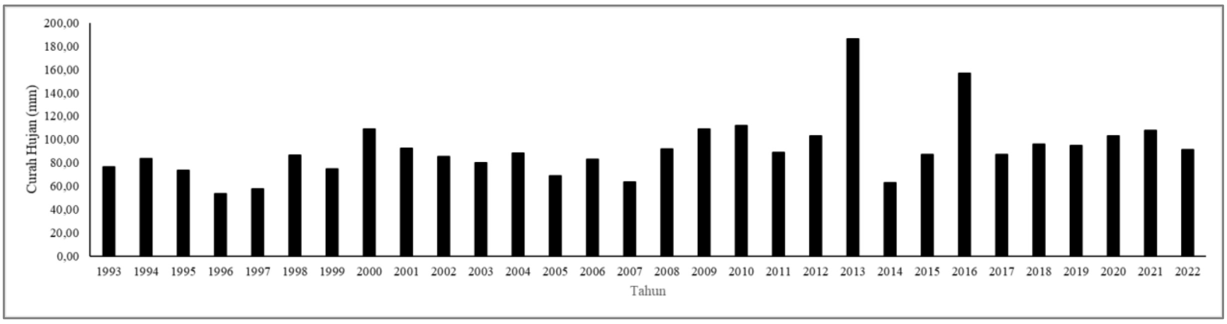

Figure 5. The daily rainfall of Karangan station (PTK-05) (Pemerintah Republik Indonesia, 2023)

Figure 5 shows the rainfall pattern at Karangan Station, which gradually increases, consistent with the previous analysis, and is characterized by three distinct temporal phases. The initial period (1993-1999) showed low rainfall with a baseline of 50-80 mm and minimal variability, followed by a relatively stable transition period (2000-2007) in the range of 80-100 mm.

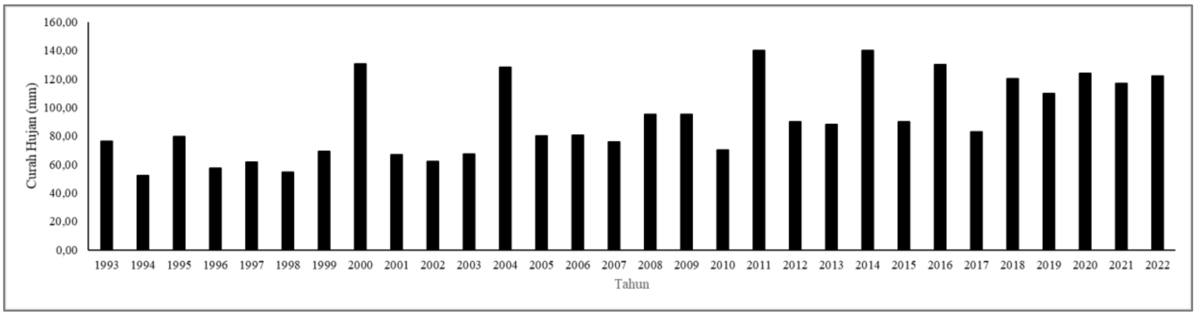

Figure 6. The daily rainfall of Tebas station (SBS-08) (Pemerintah Republik Indonesia, 2023).

Rainfall data at Tebas Station for the period 1993-2022 exhibit a distinct pattern from other stations, with a significant increase in baseline rainfall from the initial to the final period, characterized by relatively high variability throughout the observation period. The early period (1993-2003) was dominated by low to moderate rainfall (50-80 mm) with some peaks reaching 135 mm in 2000 and 2003, while the later period (2010-2022) showed a baseline increase to 90-120 mm with consistently high peaks reaching 140 mm in 2010, 2013, and 2014.

Rainfall data at Serukam Station from 1993 to 2022 exhibit a pattern of high-low fluctuations, with a predominance of low to moderate rainfall and an average annual rainfall of around 85 mm. High rainfall patterns occurred in 2000, 2004 (128.10 mm each), and clustering in the 2019-2022 period, with the highest peaks in 2019 (210 mm) and 2020 (169 mm). Rainfall patterns dominated most of the period (approximately 87%), with a range of 50-80 mm, evenly distributed from 1993 to 1999, 2001 to 2003, and 2005 to 2018.

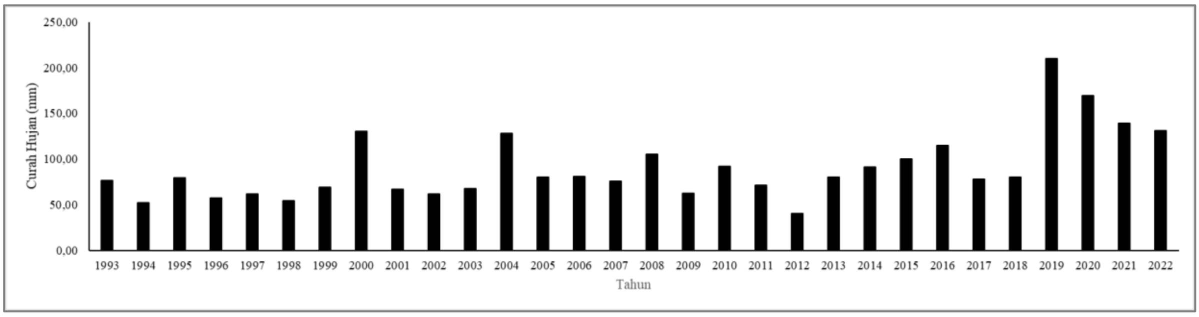

Figure 7. The daily rainfall of Serukam station (SBS-11) (Pemerintah Republik Indonesia, 2023).

The results of the homogeneity test of rain data on 15 combinations of rain station pairs from 5 Dawar rain stations, Sanggau Ledo rain station, Serukam rain station, Tebas rain station, Karangan rain station, and Bengkayang rain station are presented in Table 2.

Table 2. Rainfall data homogeneity test results for 15 rain station pairs combinations

No.	Station Pairs	*t* count	critical *t* (α=1%)	Information
1	Dawar – Bengkayang	2.19	2.663	Homogeneous
2	Dawar – Sanggau Ledo	0.91	2.663	Homogeneous
3	Dawar – Karangan	1.14	2.663	Homogeneous
4	Dawar – Tebas	1.12	2.663	Homogeneous
5	Dawar – Serukam	1.16	2.663	Homogeneous
6	Bengkayang – Sanggau Ledo	1.10	2.663	Homogeneous
7	Bengkayang – Karangan	0.92	2.663	Homogeneous
8	Bengkayang – Tebas	0.83	2.663	Homogeneous
9	Bengkayang – Serukam	1.01	2.663	Homogeneous
10	Sanggau Ledo – Karangan	0.47	2.663	Homogeneous
11	Sanggau Ledo – Tebas	0.37	2.663	Homogeneous
12	Sanggau Ledo – Serukam	0.56	2.663	Homogeneous
13	Karangan – Tebas	0.08	2.663	Homogeneous
14	Karangan – Serukam	0.23	2.663	Homogeneous
15	Tebas – Serukam	0.16	2.663	Homogeneous

Table 2 presents the results of the rainfall data homogeneity test between six observation stations (Dawar, Bengkayang, Sanggau Ledo, Karangan, Tebas, and Serukam) using the *t-test with a significance level of 1% (α =* 0.01). The homogeneity test applied in this study uses the *t-distribution with a critical value of 2.663 at a 1% significance level*. The selection of a 1% significance level indicates a conservative approach to statistical decision-making, where only statistically significant differences that are highly unlikely to occur by chance would be considered an indication of inhomogeneity. The natural variability of rainfall tends to be high, so strict criteria are needed to distinguish between natural variation and true inhomogeneity (Domonkos, 2022). The critical value *t* critical 2,663 obtained from the distribution *t* with a certain degree of freedom and α=0.01 for the two-way test. The interpretation of the test results is as follows: if the t-value of t-is < the critical value, then the null hypothesis (H0: homogeneous data) is accepted, which means that there is no strong enough statistical evidence to state that the two datasets differ significantly (Izsák et al., 2022). Conversely, if t is calculated to be greater than or equal to critical t, then H0 is rejected and the data are declared to be inhomogeneous.

The test results showed that all 15 station pairs were declared homogeneous, with no single value *t* calculation exceeding the critical value of 2,663. This finding is significant because it indicates that the six observation stations are within one relatively uniform climatological region, or at least have statistical characteristics of rainfall that do not differ significantly. This indicates regional consistency in rainfall patterns within the study area, which is likely due to the similarity of climatic factors, including the influence of atmospheric pressure systems, monsoon wind patterns, regional topography, and proximity to water vapour sources (Hewer et al., 2021; Malhi et al., 2020). A deeper analysis of the values reveals an interesting stratification, where the values range from 0.08 (Karangan-Tebas) to 2.19 (Dawar-Bengkayang). All values are still within homogeneous limits; this variation provides important information about the degree of relative similarity between stations. Studies suggest that local climate variations can provide insights into the adaptations of different species in response to environmental factors influenced by local geographical and climatic characteristics (Fan et al., 2023; Saalu et al., 2020).

The stratification of t-values can be categorized into several groups based on their degree of homogeneity. The very high homogeneity category (t < 0.5) includes six pairs of stations that show almost identical rainfall characteristics, namely Sanggau Ledo-Karangan (0.47), Sanggau Ledo-Tebas (0.37), Sanggau Ledo-Serukam (0.56), Karangan-Tebas (0.08), Karangan-Serukam (0.23), and Tebas-Serukam (0.16). This pattern suggests that the four stations (Sanggau Ledo, Karangan, Tebas, and Serukam) exhibit almost statistically identical rainfall characteristics, indicating they are likely to be in the same microclimate zone. The high homogeneity category ($0.5 \leq t < 1.0$) included three pairs, namely Dawar-Sanggau Ledo (0.91), Bengkayang-Karangan (0.92), and Bengkayang-Tebas (0.83), which showed an excellent but slightly lower level of similarity. The category of moderate homogeneity ($1.0 \leq t < 2.0$) included five pairs with relatively higher t-values but still within the homogeneous limits, namely Dawar-Karangan (1.14), Dawar-Tebas (1.12), Dawar-Serukam (1.16), Bengkayang-Sanggau Ledo (1.10), and Bengkayang-Serukam (1.01). Meanwhile, only one pair was in the low homogeneity category ($t \geq 2.0$), namely Dawar-Bengkayang (2.19), with *a calculated t-value* that was close to the critical limit but still within the homogeneity tolerance.

The observed homogeneity pattern provides a strong indication of the influence of geographical and climatological factors on regional rainfall characteristics. Apparent spatial clustering was observed at the stations of Sanggau Ledo, Karangan, Tebas, and Serukam, which showed high homogeneity with each other, indicating that the four stations were located in the same climatological zone or had similar geographical characteristics, such as altitude, topographic orientation, and the influence of regional weather systems. Dawar and Bengkayang stations showed relatively higher t-values when paired with other stations, especially the Dawar-Bengkayang pair, which had the highest value (2.19), which may indicate the influence of local factors that distinguish their rainfall characteristics, such as orographic effects, proximity to water bodies, or differences in exposure to the dominant weather system. Topographic factors, geographical distance, and the influence of water bodies can explain the difference in homogeneity values between station pairs. Stations that exhibit high homogeneity are likely to be located in similar geographical and topographic conditions.

The consistent homogeneity results across the station pairs indicated that the rainfall data used were of good quality and consistent throughout the observation period. Inhomogeneous data is often caused by factors such as changes in measurement instruments, station relocations, environmental changes around the station, or errors in measurement and recording procedures. The consistency of homogeneity suggests that these factors are unlikely to have a significant effect on data quality or have been well controlled over the observation period. The high homogeneity between the stations also confirms that the six stations can be considered as representative of regional climatological conditions, which has important implications for practical applications such as the validity of the use of combined data from multiple stations for regional analysis, the possibility of interpolation or extrapolation of data from one station to another location within the same region, reliability in the development of regional rainfall prediction models, and the validity of regional climate trend and variability analysis.

The homogeneity of rainfall data between stations has direct and significant implications for water resource infrastructure planning and regional hydrological management. When data from multiple stations exhibit homogeneous characteristics, this allows for the use of a regional approach in planning, where infrastructure design can utilise statistical parameters derived from the combined data of all stations. This approach enhances statistical robustness in the calculation of the design flood discharge, the analysis of extreme rainfall frequencies, and the estimation of regional water availability, as it utilises a larger, more representative sample. Homogeneity between stations also supports the development of an integrated early warning system for regional areas, where data from one or more stations can be used as indicators of rainfall conditions for a wider area, allowing for a more efficient and cost-effective implementation of warning systems. In addition, the homogeneity of the data supports the validity of regional analyses for climate change studies, where observed trends and variability can be considered representative of regional climatological conditions rather than local anomalies.

The homogeneity analysis of rainfall data from the six observation stations yielded very positive results, with all 15 station pairs being declared homogeneous at a fairly strict significance level of 1%. These findings confirm the validity and high reliability of the rainfall data used in the study, indicating that the six stations are located within a single, relatively uniform climatological region with consistent rainfall characteristics. Variations in *t*-values, ranging from 0.08 to 2.19, provide valuable additional information about the degree of relative similarity between stations. There is apparent clustering between stations with high characteristic similarities, especially the Sanggau Ledo-Karangan-Tebas-Serukam group, which shows very high homogeneity. These results have very significant positive implications for a wide range of practical applications in water resource management, hydrological infrastructure planning, early warning system development, and regional climatological analysis, while providing a solid and reliable basis for further hydrological research and analysis in the study area.

V. CONCLUSION

The results of the homogeneity testing of rainfall data in the Mapat River catchment area, Bengkayang Regency, showed extraordinarily significant findings. All 15 combinations of station pairs showed perfect homogeneity, with the highest calculated *t*-value of 2.19 (Dawar-Bengkayang) still far below the critical *t*-value of 2.663 at a significance level of 1%. The consistency of this statistic proves that the six rain stations (Dawar, Bengkayang, Sanggau Ledo, Karangan, Tebas, and Serukam) have uniform and reliable rainfall characteristics during the 1993-2022 period.

Theoretically, this study makes a fundamental contribution to the development of a methodology for validating tropical hydrometeorological data. Two-sided *t*-tests with a significance level of 1% have proven to be robust and effective in identifying long-term data consistency. The findings of 30 years of perfect homogeneity break the assumption that rainfall data in the tropics are inherently unstable, while reinforcing the theory of the stability of equatorial regional hydroclimatological systems when external factors can be well controlled.

In practical terms, this homogeneity of data is revolutionising the approach to regional water resource infrastructure planning. The high statistical validity enables the use of combined parameters from all stations for hydrological analysis, thereby improving the accuracy of design flood discharge calculations, water availability estimation, and extreme rainfall frequency analysis. Practical implementation encompasses infrastructure design optimisation with appropriate safety factors, the development of integrated early warning systems, and water management planning grounded in solid scientific data.

REFERENCES

Al-Muhyi, A. H., Aleedani, F. Y., Albattat, M. Q., & Badr, J. M. (2024). Rainfall repercussions: Assessing climate change influence on Iraq precipitation patterns. *Al-Kitab Journal for Pure Sciences, 8*(1), 92–103. https://doi.org/10.32441/kjps.08.01.p9

Alnino, N. F., Suhartanto, E., & Fidari, J. S. (2022). Analisis hujan TRMM (Tropical Rainfall Measuring Mission) menjadi debit dengan metode NRECA pada DAS Bango. *Jurnal Teknologi dan Rekayasa Sumber Daya Air, 2*(1), 561–569. https://doi.org/10.21776/ub.jtresda.2022.002.01.45

Anggraheni, E., Sutjiningsih, D., Mulyono, B. H., Guswanto, Ningrum, I. A., & Yahya, D. M. (2022). Pengaruh sebaran spasial hujan terhadap pemilihan metode hujan wilayah berbasis analisis geospasial. *Jurnal Teknik Sumber Daya Air, 2*(2), 81–92. https://doi.org/10.56860/jtsda.v2i2.41

Atica, A. N. T. L., Alik, G., & Saifurridzal. (2022). Prediksi curah hujan menggunakan data hujan satelit CHIRPS dan PERSIANN-CDR di DAS Bedadung Kabupaten Jember. *Jurnal Teknik Sumber Daya Air, 2*(2), 69–80. https://doi.org/10.56860/jtsda.v2i2.36

DATAtab. (2025). *t-distribution table*. DATAtab. https://datatab.net/tutorial/t-distribution

Dembélé, M., Schaefli, B., van de Giesen, N., & Mariéthoz, G. (2020). Suitability of 17 gridded rainfall and temperature datasets for large-scale hydrological modelling in West Africa. *Hydrology*

and Earth System Sciences, 24(11), 5379–5406. https://doi.org/10.5194/hess-24-5379-2020
Domonkos, P. (2022). Automatic homogenization of time series: How to use metadata? *Atmosphere, 13*(9), 1379. https://doi.org/10.3390/atmos13091379
Ekwezuo, C. S., Ezeh, C. U., Sogbedji, J. M., & Phil-Eze, P. O. (2024). Analysing spatial variability of monsoon rainfall over West Africa. *International Journal of Environment and Climate Change, 14*(9), 77–91. https://doi.org/10.9734/ijecc/2024/v14i94394
Fan, Y., Dai, J., Wei, Y., & Liu, J. (2023). Local adaptation in natural populations of *Toona ciliata* var. *pubescens* is driven by precipitation and temperature: Evidence from microsatellite markers. *Forests, 14*(10), 1998. https://doi.org/10.3390/f14101998
Francsdito, M., Juwono, P. T., & Suhartanto, E. (2023). Mitigasi dampak hidrologi dan hidrolika akibat pelaksanaan pembangunan rumah pompa Ancol Sentiong. *Jurnal Teknologi dan Rekayasa Sumber Daya Air, 3*(2), 617–625. https://doi.org/10.21776/ub.jtresda.2023.003.02.052
Harris, A. R., Daly, S. W., Pickering, A. J., Mrisho, M., Harris, M., & Davis, J. (2023). Safe today, unsafe tomorrow: Tanzanian households experience variability in drinking water quality. *Environmental Science & Technology, 57*(45), 17481–17489. https://doi.org/10.1021/acs.est.3c05275
Hastina, Harisuseso, D., & Fidari, J. S. (2023). Studi pemanfaatan data satelit CHIRPS untuk estimasi curah hujan di Sub DAS Abab. *Jurnal Teknologi dan Rekayasa Sumber Daya Air, 3*(2), 540–549. https://doi.org/10.21776/ub.jtresda.2023.003.02.046
Hewer, M. J., Beech, N., & Gough, W. A. (2021). Development and validation of the Climate Model Confidence Index (CMCI): Measuring ability to reproduce historical climate conditions. *Theoretical and Applied Climatology, 144*(3–4), 1059–1075. https://doi.org/10.1007/s00704-021-03581-5
Howard, G., Nijhawan, A., Flint, A., Baidya, M., Pregnolato, M., Ghimire, A., Poudel, M., Lo, E., Sharma, S., Mengustu, B., Ayele, D. M., Geremew, A., & Wondim, T. (2021). The how tough is WASH framework for assessing the climate resilience of water and sanitation. *NPJ Clean Water, 4*(1), 39. https://doi.org/10.1038/s41545-021-00130-5
Ibebuchi, C. C., & Abu, I.-O. (2023). Rainfall variability patterns in Nigeria during the rainy season. *Scientific Reports, 13*(1), 7888. https://doi.org/10.1038/s41598-023-34970-7
Iqbal, Z., Shahid, S., Ismail, T., Sa'adi, Z., Farooque, A., & Yaseen, Z. M. (2022). Distributed hydrological model based on machine learning algorithm: Assessment of climate change impact on floods. *Sustainability, 14*(11), 6620. https://doi.org/10.3390/su14116620
Iryani, S. Y., Alia, F., Tauhid, M. A., Muhtarom, A., & Usman, A. P. (2022). Utilization of GPM satellite and PERSIANN satellite data for estimated monthly rainfall in South Sumatera. *UKaRsT, 6*(2), 174. https://doi.org/10.30737/ukarst.v6i2.3482
Izsák, B., Szentimrey, T., Lakatos, M., Pongrácz, R., & Szentes, O. (2022). Creation of a representative climatological database for Hungary from 1870 to 2020. *Időjárás, 126*(1), 1–26. https://doi.org/10.28974/idojaras.2022.1.1
Johnson, K. A., Smithers, J. C., & Schulze, R. E. (2021). A review of methods to account for impacts of non-stationary climate data on extreme rainfalls for design rainfall estimation in South Africa. *Journal of the South African Institution of Civil Engineering, 63*(3), 1–7. https://doi.org/10.17159/2309-8775/2021/v63n3a5
Jung, W. S., & Kim, Y. D. (2023). Evaluation of watershed water quality management according to flow conditions through factor analysis and naïve Bayes classifier. *Sustainability, 15*(13), 10038. https://doi.org/10.3390/su151310038
Kumar, L., Kumari, R., Kumar, A., Tunio, I. A., & Sassanelli, C. (2023). Water quality assessment and monitoring in Pakistan: A comprehensive review. *Sustainability, 15*(7), 6246. https://doi.org/10.3390/su15076246
Mahmoud, M. T., Mohammed, S. A., Hamouda, M. A., & Mohamed, M. M. (2020). Impact of topography and rainfall intensity on the accuracy of IMERG precipitation estimates in an arid region. *Remote Sensing, 13*(1), 13. https://doi.org/10.3390/rs13010013
Makanda, K., Nzama, S., & Kanyerere, T. (2022). Assessing the role of water resources protection practice for sustainable water resources management: A review. *Water, 14*(19), 3153. https://doi.org/10.3390/w14193153
Makanda, K., Nzama, S., & Kanyerere, T. (2023). Assessing feasibility of water resource protection practice at catchment level: A case of the Blesbokspruit River catchment, South Africa. *Water, 15*(13), 2394. https://doi.org/10.3390/w15132394
Malhi, Y., Franklin, J., Seddon, N., Solan, M., Turner, M. G., Field, C. B., & Knowlton, N. (2020). Climate change and ecosystems: Threats, opportunities and solutions. *Philosophical Transactions of the Royal Society B: Biological Sciences, 375*(1794), 20190104. https://doi.org/10.1098/rstb.2019.0104
Paramasivam, R., Alias, N. E., & Lokoman, R. M. (2023). Validation of Hershfield probable maximum precipitation estimation using homogeneous region in Malaysia. *IOP Conference Series: Earth and Environmental Science, 1143*(1), 012009. https://doi.org/10.1088/1755-1315/1143/1/012009
Pemerintah Republik Indonesia. (2023). *Data hujan harian*. Balai Wilayah Sungai Kalimantan I, Departemen Pekerjaan Umum dan Perumahan Rakyat.
Saalu, F. N., Oriaso, S., & Gyampoh, B. (2020). Effects of a changing climate on livelihoods of forest dependent communities. *International Journal of Climate Change Strategies and Management, 12*(1), 1–21. https://doi.org/10.1108/IJCCSM-01-2018-0002
Santos, H. T. dos, & Nascimento Duarte, S. (2023). Rainfall data adjustment to Volta Redonda macro-region. *Revista Brasileira de Engenharia de*

Biossistemas, 16. https://doi.org/10.18011/bioeng.2022.v16.1177

Soewarno. (1995). *Hidrologi: Aplikasi metode statistik untuk analisa data*. Nova.