 doi:10.20944/preprints202207.0325.v1

Article

Spatial Frequency Analysis by Adopting Regional Analysis with Radar Rainfall in Taiwan

Che-Hao Chang [1], Riki Rahmad [1,2,*], Shiang-Jen Wu [3], and Chih-Tsung Hsu [4]

[1] Department of Civil Engineering, National Taipei University of Technology, No. 1, Sec. 3, Zhongxiao E. Rd., Da'an Dist., Taipei City, 10608, Taiwan; chchang@ntut.edu.tw
[2] Department of Geography Education, Faculty of Social Sciences, Universitas Negeri Medan, Jl. Willem Iskandar Pasar V, Medan, 20221, Indonesia
[3] Department of Civil and Disaster Prevention Engineering, National United University, No. 2, Lienda, Miaoli, 360302, Taiwan; sjwu@nuu.edu.tw
[4] National Center for High-performance Computing, National Applied Research Laboratories, No. 7, Yanfa 6th Rd., East Dist., Hsinchu City, 30076, Taiwan; hsu_nelson@nchc.org.tw

* Correspondence: rikirahmad@unimed.ac.id

Abstract: This study proposed a spatially and temporally improving methodology adopting the Regional Frequency Analysis with L-moments approach to estimate rainfall quantiles from 22787 grids of radar rainfall in Taiwan for a 24-hour duration. Due to limited radar coverage in the eastern region, significant discordant grids were found in the coastal area of the eastern region. A K-means cluster analysis using scaled at-site characteristics was used to group the QPESUMS grids in Taiwan into 22 clusters/sub-regions based on their characteristics. Spatially, homogeneous subregions with QPESUMS data produce more detailed homogeneous subregions with clear and continuous boundaries, especially in the mountain range area where the number of rain stations is still very limited. According to the results of z-values and L-moment ratio diagrams, the Wakeby (WAK), Generalized Extreme Value (GEV), and Generalized Pareto (GPA) distributions of rainfall extremes fitted well for the majority of subregions. The Wakeby distribution was the dominant best-fitted distribution, especially in the central and eastern regions. The east of the northern part and southern part of Taiwan had the highest extreme rainfall for each return period. Both areas were frequently struck by typhoons. By using grid-based (at-site) as the basis for assessing regional frequency analysis, the results show that the regional approach in determining extreme rainfall is very suitable for large-scale applications and even better for smaller scales such as watershed areas. The spatial investigation was performed by establishing regions of interest in small subregions across the northern part. It showed that regionalization was correct and consistent.

Keywords: Rainfall, QPESUMS, Regional Frequency Analysis, L-moments, Taiwan

1. Introduction

Extreme rainfall events have a significant influence on society and may result in death and property destruction, for example, by creating landslides or floods as a result of dike breaches or dam collapses [1]. Extreme rainfall and record-breaking floods are becoming more frequent and intense in many parts of the world. Floods are growing increasingly prevalent when design limitations for flood protection infrastructure (storm drainage, embankments, and dams) are surpassed and places with a large population suddenly face unexpected rainfall amounts [2]. Numerous recent research have been conducted to determine the patterns and trends of extreme rainfall globally, for example in Jakarta [3], in US [4], in Pakistan [5], in Ethiopia [6], in Korea [7], and in Taiwan [8].

Design rainfall is a stochastic representation of rainfall intensity or depth at a given location for a specific duration and return period [9]. Extreme rainfall modeling is critical for the design of flood preparation systems that rely primarily on rainfall frequency

 doi:10.20944/preprints202207.0325.v1

estimations [10]. A rainfall frequency analysis is used to estimate the frequency with which certain rainfall amounts or depths are predicted to occur. Additionally, the rainfall frequency analysis data may be utilized to quantify the rainfall depth associated with a certain probability of occurrence. Thus, information on the design values (quantiles) of extreme one-day and multi-day rainfall quantities is critical in a variety of sectors of water resources engineering, including dam and sewage system design, flood mitigation, and soil and vegetation loss protection [11].

The majority of hydrologic processes, including rainfall, are stochastic in nature. Because there are no pure deterministic hydrologic processes, considerable use of probability theory and frequency analysis is required to adequately comprehend and characterize the phenomena. The objective of frequency analysis is to determine the quantiles of the distribution of the random variable of interest [12]. [13] proposed probability weighted moments (PWMs) as a substitute for ordinary moments for estimating parameters for distributions whose inverse form is clearly established. However, PWMs can only be read inferentially as scale and shape measurements of a probability distribution. [14] circumvented this by defining L-moments as linear combinations of the PWMs.

There are two approaches to rainfall frequency analysis: one is at-site estimation, which simply uses the data from each station for statistical analysis, and the other is regional estimation, which uses observations from gauges located in a homogeneous region with similar climatological and physical characteristics [15]. The use of at-site frequency analysis with yearly maximum data is a common method for determining extreme quantiles. In this method, an appropriate distribution is chosen to characterize the data on extreme rainfall. Regardless of the parameter estimating technique used (e.g. maximum likelihood or L-moments), the at-site approach is linked with rather high levels of uncertainty, which vary according to the available data. It takes longer sequences to get more reliable design values from traditional at-site hydrological frequency analysis. In general, the fewer data available, the greater the amount of uncertainty associated with both parameters and quantile estimations [16]. Hence, increasing the amount of rainfall data available for fitting the parameters of such parametric distributions is a critical step in improving the probability of extreme rainfall occurrences.

To compensate for the scarcity of at-site observations, in the 1960s, the regional method of frequency analysis was created, which "traded space for time". Since the 1980s, this method, based on the index-flood method, has grown in popularity [17]. The regional approach's basic concept is the substitution of time for space: a multi-site study yields more accurate quantile estimates than an at-site analysis. However, under the regional method, sites cannot be aggregated arbitrarily; the resultant group of sites must satisfy the criteria of homogeneity, which requires that sites pooled together have comparable probability distribution curves of extremes. As a result, one of the most contentious topics in regional frequency analysis is the mechanism for grouping locations.

The index-flood method is one way of estimating rainfall extremes. The basic premise of this approach is that the rainfall distributions at each site within a homogeneous area have the same coefficient of variation and skewness, allowing for the quantile estimation of any return period at any location within the region [18]. By multiplying the yearly maximum rainfall data by a scale or index, the index flood technique begins by normalizing extreme rainfall data for each location within the homogeneous area. The standardized extreme rainfall estimate for the selected return period is generated by applying a regional distribution to the pooled standardized extreme rainfall data from the region's stations. Finally, for the location of interest, regional extreme rainfall estimates are derived by multiplying its index by the standardized extreme rainfall estimates [19]. [12] utilized the L-moments approach to enhance the index flooding method and established the regional frequency analysis method based on L-moments. This technique has been extensively used in the study of regional floods, precipitation, and drought.

Recently, several researchers have shifted their focus to regional frequency studies in order to compensate for data scarcity. [20] employed rainfall regionalization to expand rainfall data to areas where rainfall data is not accessible. [21] used a principal component

 doi:10.20944/preprints202207.0325.v1

analysis in T-mode to determine the regionalization of extreme rainfall in Portugal. The findings reveal three spatial regions derived using PCA. [22] categorized 4 regions as homogeneous regions through cluster analysis in the Yangtze River Delta, China. The spatial pattern of extreme rainfall with a 100-year return period shows two regions with the largest extreme rainfall and a large area with the lowest extreme rainfall. The areas with the lowest extreme rainfall are the most developed and densely populated areas. The aforementioned studies demonstrate that regionalization is a valuable tool for analyzing the frequency of rainfall extremes across large areas, and the results will aid in flood prevention and management in the area of research. However, to predict rainfall extremes, L-moments-based regional frequency analysis needs more representative, consistent, accurate, and long-sequence rainfall data. In locations with complicated terrains, such as Taiwan, rainfall vertical variation is obvious, and the data representative of rain gauge stations is often sparse, complicating infrastructure development and the research of extreme rainfall.

Taiwan is situated in the East Asian monsoon zone, receiving 2,500 mm of rainfall annually [23]. Taiwan is characterized by several adverse natural characteristics, including complicated terrain, short and swift rivers, and a position within the East Asian monsoon region and northwestern Pacific typhoon track. Additionally, during the Mei-Yu Rain Season in May and June, Taiwan is impacted by frontal systems, and the island is pounded by three to four typhoons each year [24]. This extreme meteorological phenomenon may occur during short periods of intense rainfall, resulting in slope collapse, landslides, and floods. On the other hand, during the dry season, from September to April, the southwest section of Taiwan often experiences several days without rainfall, which may result in droughts. A study by [25] found that extreme rainfall has been getting worse and more common for a long time, but the number of consecutive dry days has been going down in southwest Taiwan in recent years. Due to extreme weather and climate events in Taiwan, it is important to look at how the hydrological environment has changed.

Monitoring extreme rainfall is critical for flood predictions, warnings, and mitigation. Precipitation monitoring accuracy is critical for estimating the volume of floods and the spatial and temporal distribution of flooded regions. Estimates of extreme rainfall are often derived from analyses of long-term rain gauge data. Due to the scarcity of rain gauges in many parts of the world, an interpolation approach is often used to provide a spatial map for the predicted frequency [26]. However, the sparse spatial density of rain gauges has a substantial impact on the dependability of estimations of extreme rainfall. For design applications involving large regions, such as river catchments or drainage basins, these point estimates become insufficient.

In comparison to rain gauges, radar-based Quantitative Precipitation Estimates (QPE) offer a significant advantage over sparse rain gauges in that they offer high spatial and temporal resolution with extensive spatial coverage. Typically, a single weather radar station samples an area with a radius of 100–200 kilometers with a spatial resolution of around 1-4 kilometers and a temporal resolution of up to 5 minutes. Despite the inherent errors associated with weather radars as a remote sensing and indirect method, a typical radar station's coverage may theoretically give the equivalent of a dense network of rain gauges. Additionally, this kind of quantitative precipitation estimate product addresses the drawbacks associated with gauge station-based data, such as limited coverage, uneven distribution, and inconsistent results [15].

There have been a few studies examining the use of radar and satellite-based rainfall data. [27] used the extreme value theory to estimate depth-duration-frequency curves for the period 1998–2008 using weather radar data spanning the whole land area of the Netherlands. They calculated the data for extreme rainfall and their associated uncertainty using a regional frequency analysis technique. They determined that radar data, when subjected to stringent quality control, are acceptable for estimating the amounts of extreme area rainfall. [28] used multivariate analysis and L-moment techniques to regionalize rainfall in Iran using GPCC gridded data. They classified Iran into eight homogenous zones and found that for half of the regions, the Wakeby distribution was the best-fitting

 doi:10.20944/preprints202207.0325.v1

distribution. [29] assessed the potential of using a 12-year quantitative precipitation estimate (QPE) radar-based precipitation dataset to examine extreme rainfall in Belgium. The radar dataset spans 2005–2016, has a resolution of 1 km, and is updated every 5 minutes. They discovered that the extreme statistics vary little according to topography.

Despite recent developments in the realm of radar-rainfall monitoring and estimation, the use of radar QPE products for extreme rainfall analysis and the derivation of design frequency estimates has not yet been fully exploited. The primary constraint on using radar datasets for rainfall frequency estimates is their low historical data collection. However, as more radar datasets with lengthy archival records (>10 years) become available, the use of radar-based information for rainfall frequency analysis becomes increasingly viable. In this sense, a combination of radar-based rainfall and the L-moments approach provides a superior tool for examining the properties of extreme rainfall. The objective of this study is to develop a methodology to incorporate the benefit of high spatial resolution radar rainfall data as well as temporal enhancement against the limited historical records by regional frequency analysis . Since the homogeneous regions are identified in regional frequency analysis based on the L-moments approach, the best-fit distributions are therefore selected in order to estimate the rainfall quantiles in each homogeneous region. The index flood method and regional approach to each grid across Taiwan could be applied to explore the spatial variations of rainfall extremes.

2. Materials and Methods

2.1. Study Area

Taiwan, with a total area of 36,000 km^2, is located in Southeast Asia's Western Pacific (between Japan and the Philippines). Taiwan has a long and narrow shape, with the Central Mountain Range extending across the middle of the country. The mountainous area with elevations greater than 1,000 meters comprises 32% of the island, while hills and plateaus with elevations between 100 and 1,000 meters encompass 31% of the island, and the rest of the island is flat with elevations of less than 100 meters.

Typhoons are the primary rainfall sources in Taiwan, followed by meiyu fronts, tropical low-pressure systems, and southwesterly flow occurrences [30]. These extreme meteorological occurrences can bring tremendous rainfall in a short period of time, resulting in flooding, landslides, slope collapse, and other problems. Between 1897 and 2018, the CWB reported that Taiwan issued an average of 6.8 typhoon warnings per year [8]. Taiwan receives an abundance of rainfall each year, ranging from 1,600 to 3,100 mm with an average of approximately 2,500 mm. Additionally, extremely heavy rainfall with a maximum rate of more than 100 mm/hour can occur during the summer season (May–October).

2.2. Data

The radar-derived rainfall data used in this study is QPESUMS (Quantitative Precipitation Estimation and Segregation Using Multiple Sensors). The QPESUMS system was created by Taiwan's Central Weather Bureau (CWB) and the National Severe Storms Laboratory (NSSL) of the United States' National Oceanic and Atmospheric Administration (NOAA). The QPESUMS system is primarily constructed of four weather Doppler radars that cover the entirety of Taiwan and the surrounding ocean. It measures base reflectivity with a high spatial resolution (0.0125° x 0.0125°) and a temporal resolution of 10 minutes [31, 32].

At the end of 2001, Taiwan's Central Weather Bureau completed the Doppler radar observation network and implemented the QPESUMS system, which enhanced the monitoring, analysis, and early warning capability for violent or mutational weather. Quantitative radar rainfall estimation technology has advanced to the point that it can deliver more uniform and high-resolution spatial rainfall distribution information on a larger scale, hence establishing itself as the most effective method for increasing rainfall estimation accuracy [33].

In this study, the records of 22787 grid pixels of QPESUMS are collected to cover the whole of Taiwan. The 24-hour duration of the annual maximum series for the period 2006-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2022 doi:10.20944/preprints202207.0325.v1

2020 data was applied to conduct Regional Frequency Analysis (RFA) in Taiwan. The gridded rainfall dataset is the most recent data that hasn't been widely used in studies.

2.3. L-Moment approach

L-moments are linear combinations of order statistics that provide the expectation of particular linear combinations. The letter L denotes linear, indicating that L-moments are a linear function of the order statistics. L-moments are more robust to outliers in the data than ordinary moments. As a result, L-moments yield more reliable parameter estimations than conventional moments. L-moments was proposed as a modification of probability-weighted moments (PWMs) by [13] and [14]. For a distribution function, there are two cases of PWMs of the *r*-th order, α_r and β_r:

$$\alpha_r = \int_0^1 x(F)(1-F(x))^r\, dF,\ \beta_r = \int_0^1 x(F)(F(x))^r\, dF,\ r = 0, 1, 2, \ldots \quad (1)$$

[14] defined L-moments as linear functions of PWMs for the ordered sample $x1 \le x2 \le \ldots \le xm$ of size m:

$$\lambda_1 = \beta_0,\ \lambda_2 = 2\beta_1 - \beta_0,\ \lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0,\ \lambda_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0 \quad (2)$$

The parameter k_1 is known as the L-location or distribution mean, while the parameter k_2 is known as the L-scale. The L-moment ratios are expressed as follows:

$$\tau_2 = {}^{\lambda_2}/_{\lambda_1},\ \tau_3 = {}^{\lambda_3}/_{\lambda_2},\ \tau_4 = {}^{\lambda_4}/_{\lambda_2} \quad (3)$$

where τ_2 denotes the coefficient of L-variation (L-CV), τ_3 is L-skewness (L-CS), and τ_4 is L-kurtosis (L-CK).

The L-moments and L-moments ratios are analogous to their ordinary moment counterparts, but they are more unbiased. Thus, these values can be used to describe the major characteristics of the distributions that are frequently utilized in extreme value analysis [34]. The L-moments procedures are available as an R package from CRAN (https://cran.r-project.org/package=lmomRFA) [35]. Figure 1 depicts a summary of the Regional Frequency Analysis procedures using L-Moments, and the section that follows explains the regionalization approaches in further depth.

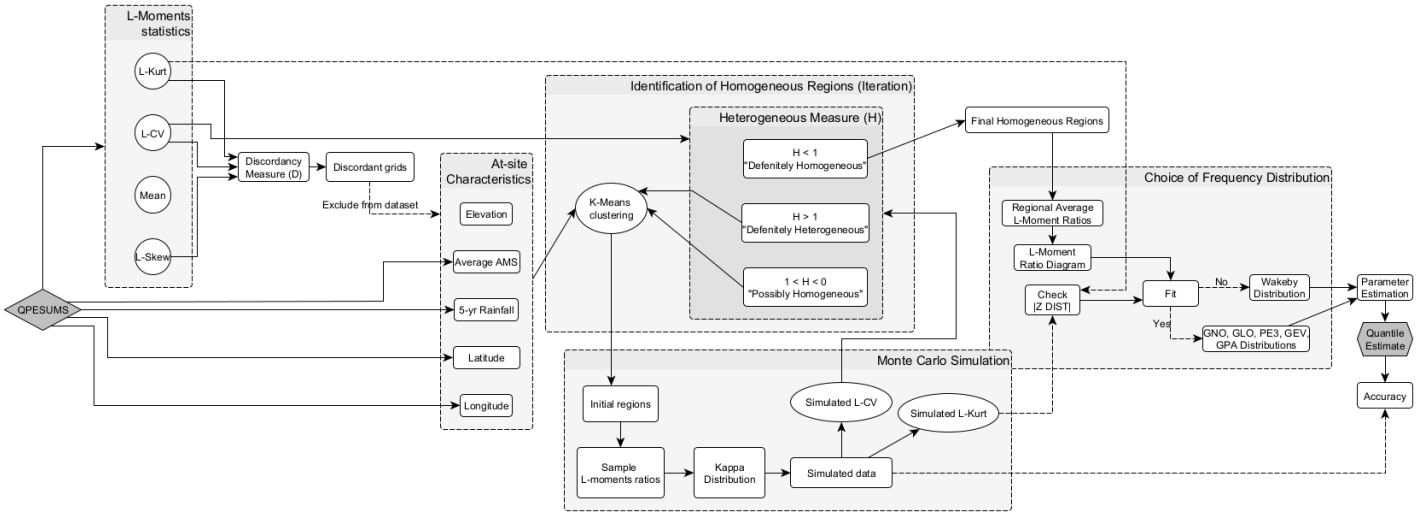


Figure 1. Flowchart of Regional Frequency Analysis using L-Moments

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2022 doi:10.20944/preprints202207.0325.v1

2.4. Discordancy Measure

Hosking & Wallis (*H*) [12] discordancy measure is widely recommended for screening outliers' sites in regional frequency analysis (RFA). Discordancy measures provide a discordancy metric for detecting anomalous sites in a region by utilizing the sample mean and sample covariance matrix of the site's sample L-moment ratios including L-CV, L-skewness, and L-kurtosis. The performance of conventional and robust techniques for identifying discordant sites is evaluated using Monte Carlo simulations.

Assuming N stations are available in the area, the following three-dimensional vector of sample L–moment ratios of the yearly maximum daily rainfall sequence may be constructed and recorded:

$$u_i = \left[t^i, t_3^i, t_4^i\right]^T \tag{4}$$

The following are the discordant detection indicators:

$$D_i = \frac{1}{3}(u_i - \bar{u})^T A^{-1}(u_i - \bar{u}), \tag{5}$$

$$\bar{u} = N^{-1}\sum_{i=1}^{N} u_i, \tag{6}$$

$$A = \sum_{i=1}^{N}(u_i - \bar{u})(u_i - \bar{u})^T. \tag{7}$$

A is a sample covariance matrix, N is the number of property vectors, u_i is the vector containing L-moment ratios for the station i, and $\bar{u}$ is the unweighted regional average value, D_i is the measure of discordancy. Sites with $D_i > 3$ should be considered discordant, according to [12]. To some extent, the criterion for discordance should increase as the number of sites in the region increases. This is due to the fact that sites with high D_i values are more likely to be found in larger regions. In any case, regardless of the magnitude of these values, it is recommended to search the data for the sites with the highest D_i values.

2.5. K-Means Clustering

Cluster analysis is a conventional approach of statistical multivariate analysis that can be used to break down huge and complicated data sets into a small number of data groups where members of a group have characteristics in common [36]. The supervised cluster algorithm approach, K-means, is applied to identify homogeneous regions across Taiwan. The K-mean clustering technique is chosen owing to its successful application in other rainfall regionalization studies [22,36-37]. In the K-means method, each cluster is represented by its centroid, which is the mean (weighted or unweighted average) of feature vectors within the cluster. This method is known for its efficiency in clustering large data sets with numerical attributes [38].

In this study, five dimensions/attributes known as site-characteristics are used in conjunction with the K-means algorithm to regionalize the rainfall extreme. The site-characteristics including latitude, longitude, elevation, the mean annual maximum series of each grid, and 5-year rainfall (Figure 2). Because variables with different units might impact the outcomes of the K-means clustering technique, the data should be standardized using suitable transformation functions before performing the algorithm [39]. As a result, the following variables/attributes of each site should be rescaled as follows:

$$Y = \frac{(X - X_{min})}{(X_{max} - X_{min})} \tag{8}$$

where Y is the normalized value, and X_{min} and X_{max} are the minimum and maximum values for each variable in the data set (X), respectively. Cluster analysis does not always provide the desired results. Subjective enhancements are often required to promote physical continuity and reduce regional heterogeneity.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2022 doi:10.20944/preprints202207.0325.v1

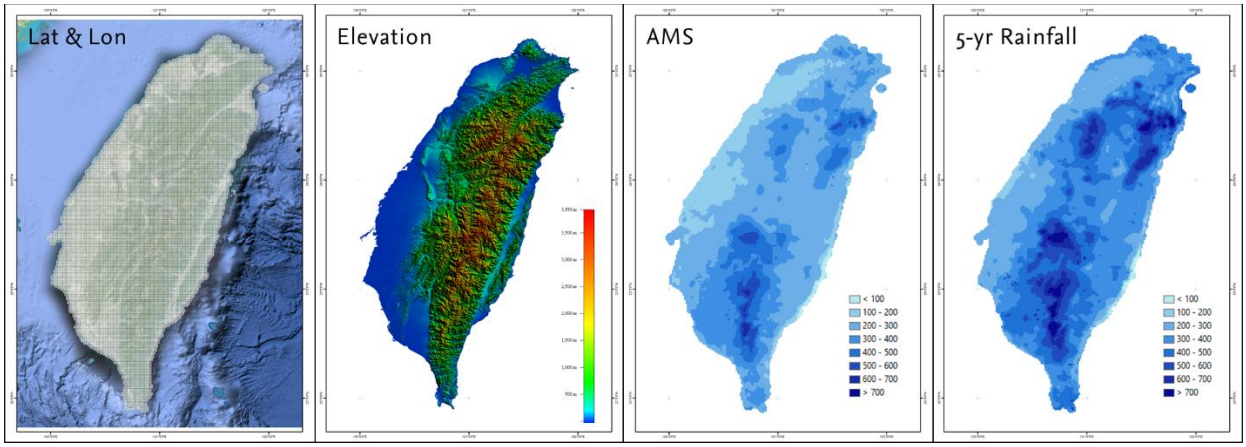


Figure 2. Site-characteristics used in K-means clustering

2.6. Heterogeneity Measure

The heterogeneity measure is used to determine if proposed regions generated by cluster analysis of site attributes could be considered as homogenous. Hosking and Wallis's heterogeneity measure (H) is often used in regional frequency analysis to assess regional homogeneity. This statistic is calculated using dispersion measures, which are defined as the total of the errors between the at-site sample L-moment ratios and the regional average L-moment ratios. Using Monte Carlo simulation, a measure of heterogeneity generates several simulated regions from a flexible probability distribution and compares observed and generated data [40]. [12] estimate heterogeneity by parameterizing the more flexible four-parameter Kappa distribution using data from the candidate region, from which simulated regions are produced using Monte Carlo sampling. A statistic V can be computed for the actual region and all simulated regions as the total of the squared deviations from the mean across all sites for a given L-moment ratios. To estimate H, the mean and standard deviation of V over all simulated regions are calculated and compared to V for the actual data. The heterogeneity measure is calculated as follows:

$$H_1 = \frac{(V_1 - \mu_v)}{\sigma_v}, \tag{9}$$

$$V_1 = \{\sum_{i=1}^{N} n_i (t^i - t^R)^2 / \sum_{i=1}^{N} n_i\}^{\frac{1}{2}}, \tag{10}$$

$$t^R = \sum_{i=1}^{N} n_i t^i / \sum_{i=1}^{N} n_i \tag{11}$$

where V_1 is the standard deviation of at-site sample L-CV and μ_v and σ_v mean and the standard deviation of the simulated values of V, respectively.

[12] classify an area as acceptably homogeneous when $H < 1$, possibly heterogeneous when $1 \leq H < 2$ and definitely heterogeneous when $H \geq 2$. If the value of H_1 did not fulfill the criterion, the site inside the region needed to be modified to meet the requirements for becoming an approved homogenous region. [12] are then proposed several region changes, including the following: (1) relocating a site or a few sites within a region; (2) subdividing the region; (3) merging two or more regions and redefining groups; (4) breaking up the region by reassigning its sites to other regions; and (5) removing a site or sites from the data.

2.7. Selection of Candidate Distributions

[12] suggested two methods for determining the best-fitting distribution: The L-Moment ratios diagram and the Z-test. The L-moment ratio diagram is a visualization of the distribution function's calculated L-Skewness and observed L-Kurtosis values. The curves represent potential links between the proposed distribution's L-Cs and L-Ck [41]. L-moment ratio diagrams have been proposed as a handy tool for differentiating between alternative distributions that might be used to represent regional data. The sample average

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2022 doi:10.20944/preprints202207.0325.v1

and a line of best-fit through the sample L-moment ratios are two graphical tools used to aid in distribution selection. The closeness of the sample average or the record length weighted average to the theoretical curve or point in L-skewness and L-kurtosis space of a given candidate distribution has been regarded as an indicator of the distribution's suitability to represent the regional data.

The Z-test is a statistic that reflects how well the theoretical L-kurtosis of the fitted distribution matches the regional average L-kurtosis of the observed data. The quality of fit is determined for each of the candidate distributions by:

$$|Z^{DIST}| = \frac{\tau_4^{DIST} - \bar{t}_4 + \beta_4}{\sigma_4} \quad (12)$$

where Z^{DIST} denotes the candidate distribution and τ_4^{DIST} denotes the fitted (candidate) distribution's L-kurtosis obtained via simulation for the purpose of estimating the fitted distribution as follows:

$$\tau_4^{DIST} = \sum_{k=0}^{8} A_k \, \tau_3^k \,, \quad (13)$$

where A_k are the coefficients described in [12], β_4 denotes the bias of the regional average sample L-kurtosis ($\bar{t}_4 = \bar{t}_4^{(m)}/m$) determined using simulation approach with $\bar{t}_4^{(m)}$ denoting the sample L-kurtosis of the mth simulation:

$$\beta_4 = \frac{\sum_{m=1}^{N_{sim}} (\bar{t}_4^{(m)} - \bar{t}_4)}{N_{sim}} \quad (14)$$

and σ_4 is the standard deviation of regional average sample L-kurtosis estimated as

$$\sigma_4 = \sqrt{(N_{sim} - 1)^{-1} \left(\sum_{m=1}^{N_{sim}} (\bar{t}_4^{(m)} - \bar{t}_4)^2 - N_{sim} \beta_4^2 \right)} \quad (15)$$

A fit is regarded satisfactory if $|Z^{DIST}|$ is sufficiently close to zero, with $|Z^{DIST}|$ 1.64 being a fair criterion. If there are more than one acceptable distribution, the one with the lowest |ZDIST| is considered to be the best distribution.

Generalized Normal (GNO), Pearson Type III (PE3), Generalized Extreme Value (GEV), Generalized Logistic (GLO), and Generalized Pareto (GPA) Distributions are the candidate distributions considered in this study. If the three-parameter frequency distributions do not match the rainfall data well, we may try fitting using a five-parameter Wakeby distribution. Wakeby distributions have more parameters than other system distributions, allowing for a greater variety of shapes and a relatively good fit to any sample [42].

2.8. Estimation of Rainfall Quantiles

The index-flood approach is used in regional frequency analysis to determine quantiles for each station's distinct return periods. The index flood method assumes that rainfall from multiple locations within a region is normalized by their mean annual rainfall from a single distribution. The index flood technique works on the premise of defining rainfall at each site as two components, one representing the rainfall specific to the location and the other reflecting the rainfall features common to the homogenous region. Equations (16) and (17) determine the parameters of the fitted frequency distribution from the regional average L-moment ratios, which are weighted proportionately to the site's record length, respectively.

$$t^R = \sum_{i=1}^{N} n_i t^{(i)} / \sum_{i=1}^{N} n_i \,, \quad (16)$$

$$t_r{}^R = \sum_{i=1}^{N} n_i t_r{}^{(i)} / \sum_{i=1}^{N} n_i \,, \quad r = 3,4, \ldots \quad (17)$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2022 doi:10.20944/preprints202207.0325.v1

Finally, the quantile estimates at site *i* are obtained from the index flood μ_i and regional quantile function q(F)F given by:

$$Q_i(F) = \mu_i q(F), \quad i = 1, 2, \cdots, N \tag{18}$$

2.9. Accuracy Assessment

Uncertainty is a necessary component of every statistical estimate. The related uncertainty of an estimated quantile indicates how dependable the value is for confident usage. [12] suggested an evaluation approach that includes utilizing a Monte Carlo simulation to generate regional average L-moments. Quantile estimations for different return durations are computed in the simulation. The computed quantiles for non-exceedance probability F is $\hat{Q}_i^{[m]}(F) - Q_i(F)/Q_i(F)$ at the *m*th repetition. The relative error of the estimate at station i for non-exceedance probability F is $\hat{Q}_i^{[m]}(F) - Q_i(F)/Q_i(F)$. This amount can be averaged across all *M* repeats to approximate the estimators' relative RMSE. The relative RMSE for a large M is approximated by:

$$R_i(F) = \left\{ \frac{1}{M} \sum_{m=1}^{M} \left[\frac{\hat{Q}_i^{[m]}(F) - Q_i(F)}{Q_i(F)} \right]^2 \right\}^{1/2} \tag{17}$$

The regional average relative RMSE of quantile estimates provides an overview of the accuracy of quantile estimations across all grids in the region:

$$R^R(F) = \frac{1}{N} \sum_{i=1}^{N} R_i(F) \tag{18}$$

Analogous values can be calculated for growth curve estimations, but with $\hat{Q}_i(F)$ and $\hat{Q}_i^{[the\ m]}$ substituted by $\hat{q}_i(F)$ and $\hat{q}_i^{[m]}$, respectively. The 90% error bounds for $\hat{q}(F)$ are:

$$\frac{\hat{q}(F)}{U_{0.05}(F)} \leq q(F) \leq \frac{\hat{q}(F)}{L_{0.05}(F)} \tag{19}$$

where $U_{0.05}(F)$ and $L_{0.05}(F)$ are the values where approximately 90% of simulated values to true value ratio, i.e., $\hat{q}(F)/q(F)$, lies. We simply multiply the relative error measurements by the calculated quantiles to get the absolute error measures. The flowchart of Regional Frequency Analysis is summarized in Figure 2 and the following section describes the detailed procedure of regionalization techniques.

2.10. Regional and Gird-based Estimates for Watershed Scale

Extreme rainfall patterns are commonly characterized using at-site and regional frequency analysis. While at-site frequency analysis uses only records from the target location (grid/station) to estimate the quantiles, regional frequency analysis makes use of data from stations with comparable rainfall patterns to produce more precise estimates of extreme rainfall intensities. In this part, we were analyzed the difference and similarity between regional and grid-based (at-site) estimate at several return periods at one particular watershed. In this study, we select Touqian watershed. The reason for this comparison is that regional estimates, when applied at the watershed-scale, sometimes have high uncertainties and cannot be applied, especially in regions that have complex topography and rainfall patterns. This is due to the different approaches where at-site analysis considers at-site characteristics & at-site statistics only for the entire watershed area, while regional analysis, when applied to watersheds also considers at-site characteristics & at-site statistics for all grids/stations in each homogeneous region where part of the area is in the watershed.

To analyze the differences between the two approaches, we use a spatial approach with visual and statistical comparisons at 50-year return period. Furthermore, the similarity analysis uses an image matching approach. Normalized Cross-Correlation (NCC) and

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2022 doi:10.20944/preprints202207.0325.v1

Jaccard Similarity Index (JSI) were used to measure the degree of similarity between the two approaches. NCC is one of the most effective and commonly used similarity metrics in matching tasks. The Normalized Cross-Correlation (NCC) of two time series can be defined as

$$NCC = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sigma_X \sigma_Y} \quad (20)$$

where n is sample size, x_i and y_i are the individual grids indexed with i, X and Y are the regional and at-site(grid-based) datasets.

Jaccard Similarity Index (JSI) is a statistical value used to compare the similarity and diversity between two different sample sets. Jaccard similarity is defined as the magnitude of the intersection of the two sets divided by the magnitude of the union of the two sets (Eqs. 21). The larger the value, the more similar condition for the two groups [43].

$$Jaccard\ (A,B) = \frac{|A \cap B|}{|A \cup B|} \quad (21)$$

In this study we used weighted Jaccard Similarity Index. In this measure, the high similarity between a pair of points indicates that the points have high resemblance, meanwhile low similarity indicates that the points are distant. Weighted Jaccard Similarity computed between two vectors or data points x and y where each of vectors has a length of n is the summation of the minimum of two vectors in each dimension divided by the summation of the maximum of those two vectors in each dimension. The calculation of weighted JSI can be seen in the following equation.

$$JSI\ (X,Y) = \frac{\sum_{K=1}^{n} \min\ (X_K, Y_K)}{\sum_{K=1}^{n} \max\ (X_K, Y_K)} \quad (22)$$

Practically, the numerator is the intersection of two vectors and denominator is the union of two vectors. The Jaccard Similarity Index between two vectors is 1 when those two vectors are the same and is 0 when those two vectors do not have anything in common (that is the numerator is zero).

3. Results

3.1. Screening the Data

Validating the acceptability of the data for the analysis by screening for unusual data items is a crucial part of any statistical data analysis. The discordancy measure, $D_i = D(u_i)$, should tell us how distant u_i is from the center of the area in respect to its size. Large D_i values suggest that the ith site should be examined further for the potential of data errors, or that the prospect of relocating the site should be considered.

Validating the data for analytical acceptability by screening for outliers is a critical component of any statistical data analysis. The discordancy measure, $D_i = D(u_i)$, should indicate the distance between u_i and the area's center in proportion to its size. Large D_i values indicate that the ith site should be investigated further for any data inaccuracies or that the site should be relocated.

When assessing the discordancy test results generated for the stations, grids with a D value of 3 and higher are recommended to be eliminated from future regional studies since they could not pass the discordancy test when the cluster size is more than 15. To some degree, the discordancy criteria should be a function of the number of sites in the area. This is because sites with high D_i values are more likely to be found in large regions. In any case, regardless of the size of these values, it is recommended to study the data for the locations with the highest D_i values [12].

Based on the above theories, we conducted two discordance test approaches, first by using a critical D_i of more than 3 and the second using a critical value of > 6. Figures 3a and 3b show the discordant spatial grid distribution (DI >3) and a graph of the relationship between L-skewness and L-CV. A total number of 801 of 22787 grids is categorized as

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2022 doi:10.20944/preprints202207.0325.v1

discordant (3.52% of total grids). The spatial distribution of discordant grids is almost evenly distributed throughout the region, but a continuous significant discordant grid is seen in the southeastern area.

After further testing, we found that, to some extent, the grid with DI values >3 (discordant) did not affect the homogeneity of the region after clustering and heterogeneity tests. therefore, we continued the second analysis by increasing the critical D_i limit by > 6. Based on this critical value, 171 of 22787 grids are categorized as discordant (0.75%) (see Figure 3c and 3d). Grids with discordant value greater than 6 greatly affects the homogeneity of the region. The region that was originally homogeneous shifted to heterogeneous, especially in the southeastern region.

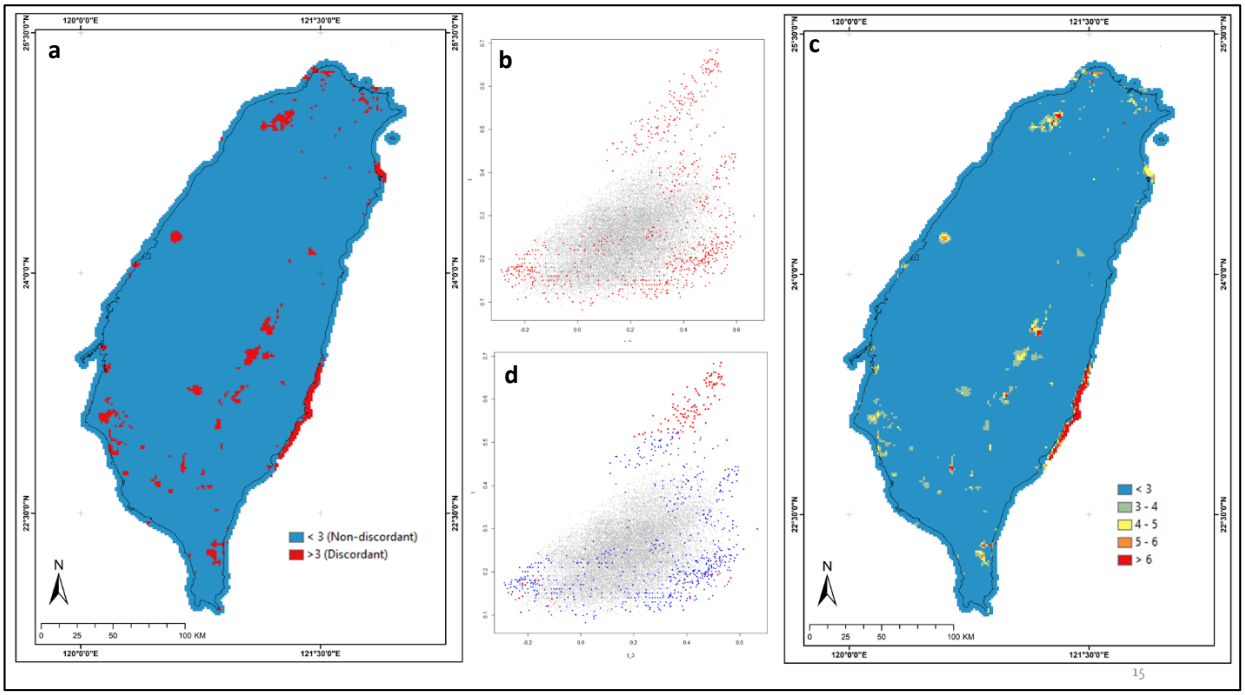


Figure 3. Spatial distribution of QPESUMS discordant grids in Taiwan; a). discordant grids using Di > 3 as critical

After conducting further investigations in the southeast (discordant grid > 6), by comparing the radar estimation value (QPESUMS) with the three closest rainfall stations, we found that the QPESUMS rainfall values were extremely underestimated rain gauge values from 2006-2015 (Figure 4). but after 2016, the value tends to be similar. According to [44], due to the island's two major mountain ranges (the Snow Mountain Range (SMR) and the Central Mountain Range (CMR)) and the significant blockages caused, the S-band radar network's vertical coverage in the inner island was severely limited. For large swaths of the SMR and CMR, the lowest 1 km of the atmosphere is not covered by radar. These vertical coverage gaps adversely impacted the quality of precipitation forcing used in hydrological forecasting, and additional radars are required to address these gaps. Since 2016, four additional C-pol radars have been placed island-wide to enhance radar coverage at low altitudes. As a result, the C-pol radars considerably improved low-level coverage, particularly along the east coast and throughout a major portion of northern and southern Taiwan. Based on that consideration, the discordant grids with DI >6 were excluded for further analysis.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2022 doi:10.20944/preprints202207.0325.v1

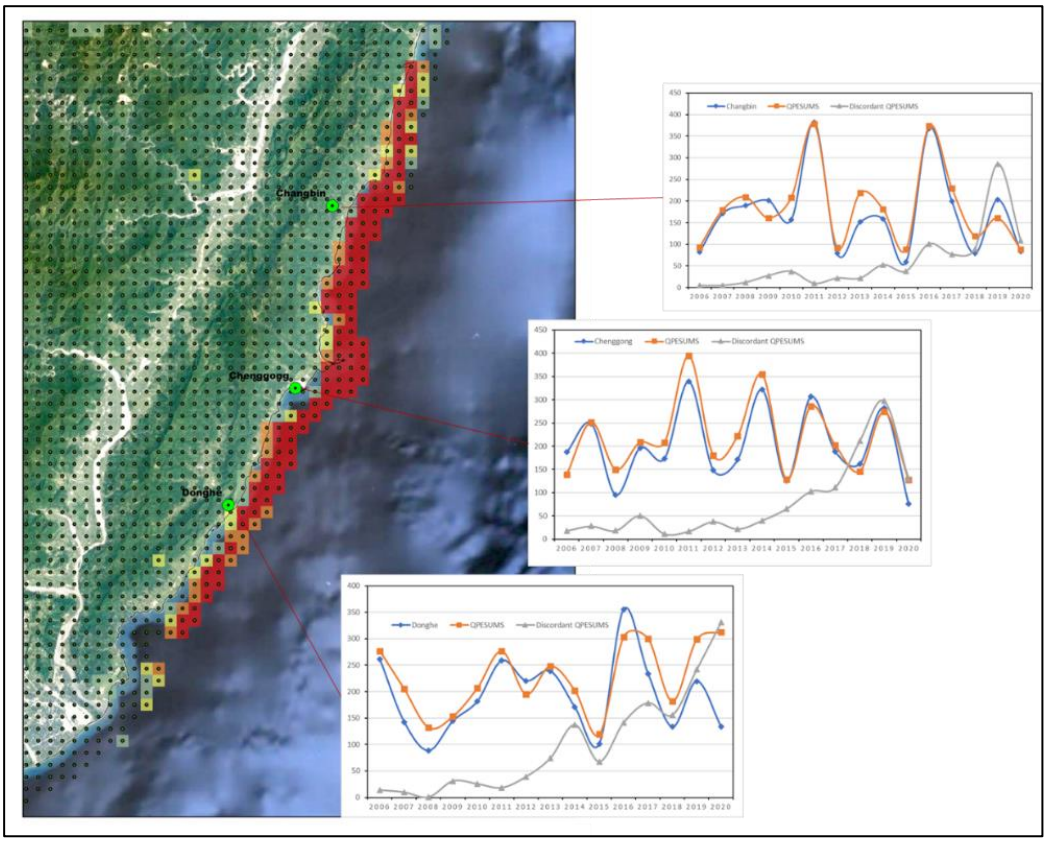


Figure 4. Discordant grids along the coastal in comparison with 3 closest rain gauges

3.2. Identification of Homogeneous Regions

Determining the number of subregions is critical in the work of L-moment because an insufficient number of subregions will result in an inappropriate division pattern that may not meet the RFA's homogeneity requirement [45]. The number of subregions in this research was decided by first establishing an initial number based on regional geographical factors and then shifting the number back and forth until all split subregions passed the homogeneity test. This determination employs the K-means clustering method and subjective adjustment. in K-means clustering method, five factors called site-characteristics including latitude, longitude, elevation, mean annual maxima, and 5-year rainfall are employed to subdivide the whole Taiwan island into subregions, while the subjective adjustment is used to ensure that all regions are homogeneous. Since the QPESUMS data has a very large grid (22787 grids), the use of the k-means clustering method alone to regionalize the frequency of analysis will be very difficult, therefore it needs to be refined by subjective adjustment [36]. The region should be sufficiently large to have a large sample size (many extremes) and small enough to neglect extreme statistic variability [29].

For each region, discordancy and heterogeneity measures (D_i and H_1) defined in Eqs. (5) and (9) were calculated to see if they are spatially continuous and physically reasonable. When the estimated heterogeneity measure (H_1) surpasses the critical value, which indicates "potentially" and "definitely" heterogeneous, K-Means then will be used to regroup the grids in the region into smaller regions. This process was repeated until no further separation of heterogeneous regions was possible. Occasionally, the discordance measure suggested that several surrounding grids within a region are discordant with the remainder of the region. In this case, the homogenous regions were manually refined by using subjective adjustments as described in the methodology section. Examining the discordant grids showed several natural and physically defensible modifications to the clusters, leading to more homogeneous clusters after accounting for the topographical and geographic patterns of the average annual maximum series.

15 initial clusters/sub-regions were determined by using the K-means clustering method. Furthermore, a total of 22 sub-regions are discovered to apply to the study area

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2022 doi:10.20944/preprints202207.0325.v1

after subjective adjustments. The 22 subregions were defined as geographical areas with similar topography and climatological characteristics that experienced similar meteorological conditions during storm events. We depict the boundaries of the 22 subregions in Figure 5. This regional classification is similar to the findings of [46] which divides Taiwan into 17 homogeneous sub-regions. They divided sub-regions using the Principal component analysis method, self-organizing maps, and L moment using 127 rain gauges data. Limited spatial density using uneven rain gauges makes interpolation difficult and makes it difficult to produce maps with clear homogeneous boundaries. [47] in their research that compares mapping approaches of design annual maximum daily precipitation also concludes that one of the problems or difficulties of the regional approach is the discontinuity of boundaries between sub-regions. They suggest that in the future, the use of radar-based rainfall could be a solution. Our findings further confirm that, using QPESUMS which is based on high resolution radar-based rainfall, we can produce homogeneous regions that are not only similar, but also more detailed with clear boundaries between regions, especially in the center/mountain range area where the number of rain stations is still very limited. In other words, we have better results in terms of detail.

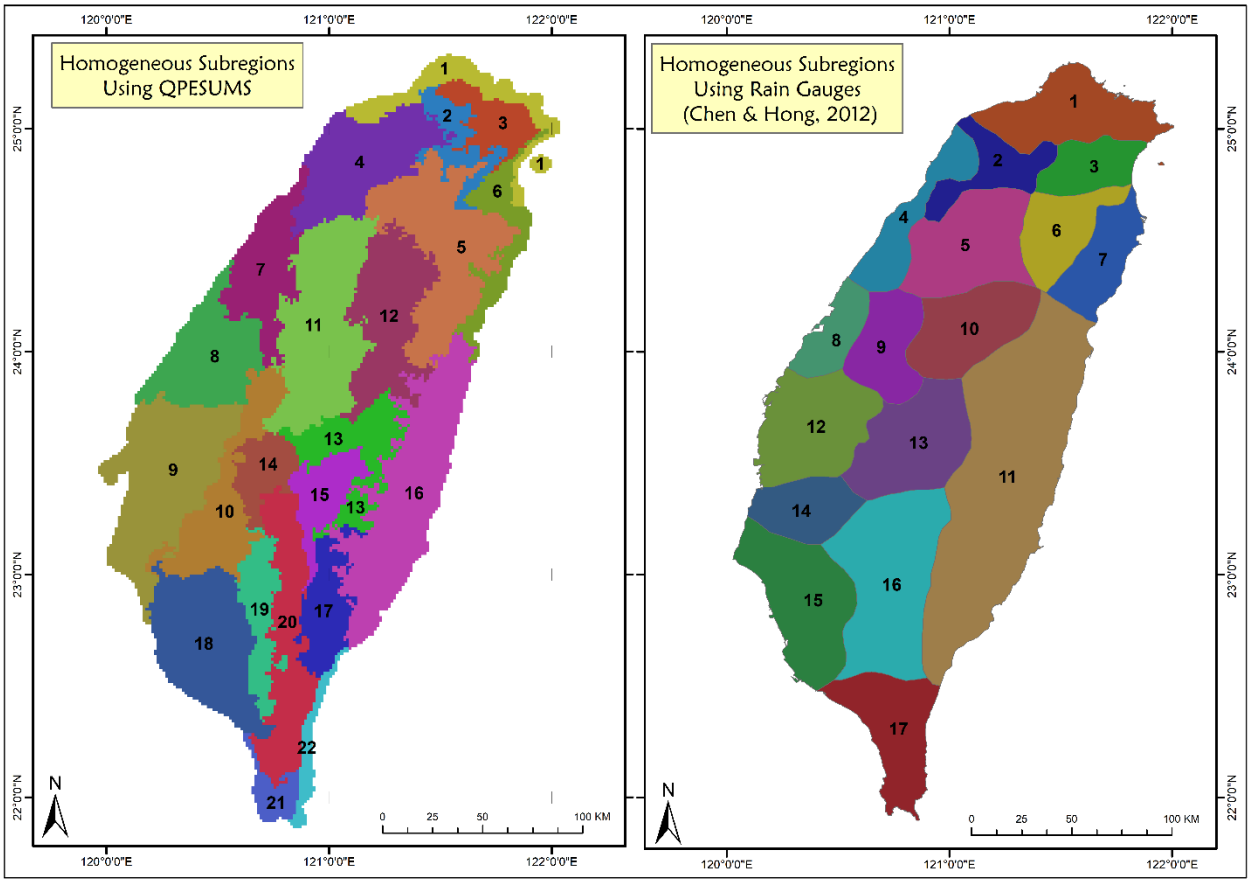


Figure 5. The division of 22 homogeneous subregions using QPESUMS in Comparison with Chen & Hong, (2012)'s homogeneous subregions using Rain Gauges in Taiwan

Geographically, the 22 sub-regions are divided into 5 regions: the northern region (sub-regions 1-6), the western region (sub-regions 7-10), the center region (sub-regions 11-15), the eastern region (sub-regions 16 & 17), and the southern region (sub-region 18-22). Table 1 shows the total number of grids in each sub-region, as well as their L-moment ratio. According to Table 1, the biggest sub-region is sub-region 16, which is located in the eastern region and has a total grid of 2231 grids, whilst the smallest sub-region is sub-region 22, which is located in the southern region and has a total grid of 248 grids. In contrast, the northern and southern areas of Taiwan have complicated distribution patterns of homogenous sub-regions. The results of heterogeneity tests for annual rainfall data are also summarized in Table 1. It can be seen that all regions are 'acceptably

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2022 doi:10.20944/preprints202207.0325.v1

homogeneous', with H < 1. These results demonstrate that the 22 regions are sufficiently homogeneous.

Table 1. L-moment ratios and heterogeneity measure of 22 subregions

Region/Sub-region	Number of Grids	L-moment ratios			*H1*
		t	t_3	t_4	
North					
1	711	0.33	0.24	0.17	-2.01
2	453	0.22	0.21	0.11	-0.62
3	509	0.19	0.18	0.15	-0.67
4	1359	0.25	0.28	0.17	-0.14
5	1806	0.28	0.15	0.10	-1.86
6	606	0.24	0.20	0.16	-1.27
West					
7	980	0.32	0.34	0.17	-5.83
8	1186	0.26	0.19	0.10	-5.33
9	1961	0.27	0.26	0.20	-7.71
10	1150	0.26	0.25	0.14	0.85
Center					
11	1938	0.32	0.29	0.08	-4.97
12	1344	0.25	0.11	0.10	-4.90
13	826	0.23	0.00	0.06	-6.01
14	474	0.22	0.07	0.05	-1.19
15	567	0.17	0.11	0.20	-6.87
East					
16	2231	0.25	0.08	0.06	-5.21
17	652	0.25	0.17	0.09	0.07
South					
18	1629	0.19	0.09	0.12	-3.63
19	609	0.16	0.10	0.12	-6.49
20	1283	0.16	0.05	0.08	-3.25
21	265	0.25	0.20	0.07	-2.55
22	248	0.35	0.30	0.14	0.92

3.3. Selection of Distribution Models

The determination of appropriate regional frequency distributions is also important in the RFA. An inappropriate selection of distributions may result in a large overestimation or underestimation of extreme rainfall. Rainfall maxima are often positively skewed, having a comparatively longer right tail. Three-parameter frequency distributions are much more flexible than two-parameter distributions, which may enhance fit in the right upper tail region. Three-parameter frequency distributions are primarily very popular, and they are frequently used in the context of a regional frequency analysis (RFA) method.

The L-moment ratio diagram is widely used as the first visual assessment tool for selecting a regional frequency distribution from sample data collected in a particular location. The sample L-moment ratios, regional average L-moment ratios, and theoretical L-moment ratio curves of candidate distributions are displayed in an L-skewness and L-kurtosis space. The theoretical plotting positions of the candidate distributions are given using polynomial approximations.

L-moment ratio diagrams have been used to show averages of L-skewness and L-kurtosis in 22 homogeneous regions. Then, theoretical curves for different distributions have been shown, as well (Figure 5). As long as the point that is the regional average is close to the curve that corresponds to a certain distribution, this distribution will be a good choice for the parent distribution in the region. As shown in the L-moment ratio diagrams

 doi:10.20944/preprints202207.0325.v1

(Figure 6), in 11 of Taiwan's 22 subregions, the regional average L-moment ratios do not approach or are exactly on the theoretical curves of five candidate distributions. As a result, [12] suggest that the Wakeby (WAK) robust distribution is the most appropriate.

More precisely, the z value indicates the degree to which the regional mean distribution fits the five alternative distributions the best. According to the z value in Table 2, the best fit parent distribution of half of 6 subregions in the northern region, specifically subregions 1, 3, and 6, is dominated by the GEV distribution. Only subregion 8 in the western region has a z value that does not fit any of the five possible distributions. while the Wakeby (WAK) distribution best fits all subregions in the east and center regions. Finally, Generalized Pareto (GPA) best fits as parent distribution in the southern region. Moreover, table two also shows the distribution parameters of the best-fit parent distribution for each subregion. A way to figure out how well a distribution fits is to look at how far samples points are from the curve for that distribution. Additionally, the optimal distribution determined at this stage will be used to calculate the regional growth curve for each homogenous region.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2022 doi:10.20944/preprints202207.0325.v1

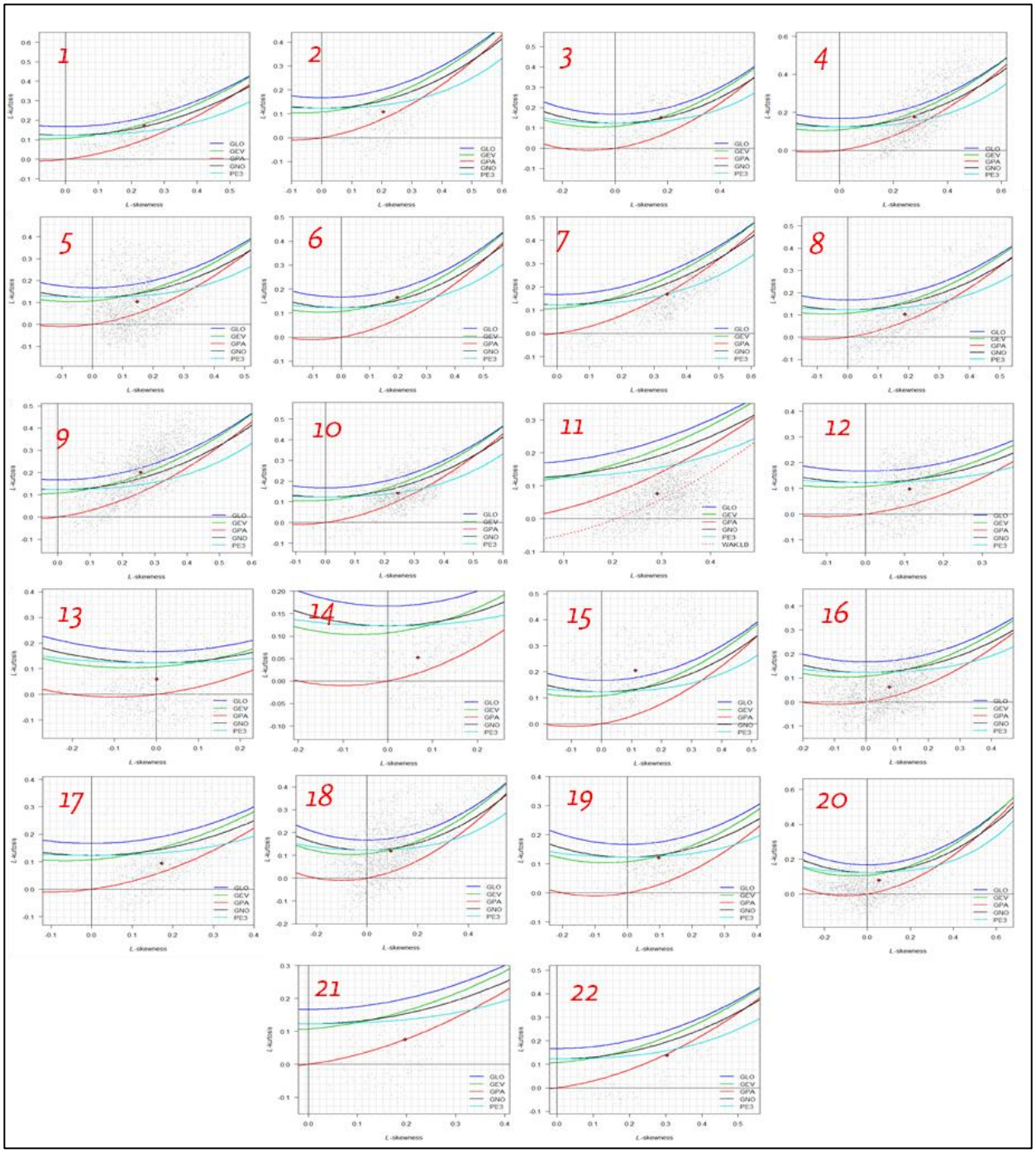


Figure 6. L-moment ratio diagram (LMRD) of 22 homogeneous regions in Taiwan (the gray dots are site sample L-moment ratios, the red points are the regional average L-moment ratios, and colored curves represent theoretical L-moment ratio curves of 5 candidate distributions)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2022 doi:10.20944/preprints202207.0325.v1

Table 2. GoF measures, Best fit Distribution, and parameter distribution of 22 homogenous Sub-Regions in Taiwan

Reg	Goodness-of-fit Measures					Best Fit Distribution	Parameter estimates for distributions accepted at the 0.90 level				
	GLO	*GEV*	*GNO*	*PE3*	*GPA*		ξ/μ	α/σ	K/β	γ	δ
North											
1	7.44	**0.61**	-2.28	-7.59	-16.25	*GEV*	0.71	0.42	-0.11		
2	18.58	11.54	9.70	5.99	-4.92	*WAK*	0.48	0.68	0.32	0.00	0.00
3	7.45	**-0.08**	-1.17	-3.90	-16.92	*GEV*	0.84	0.27	-0.01		
4	13.00	5.05	**-0.08**	-9.14	-15.97	*GNO*	0.88	0.38	-0.58		
5	34.65	17.03	16.09	11.85	-20.90	*WAK*	0.30	1.04	0.49	0.00	0.00
6	5.87	**-1.60**	-3.35	-7.00	-18.90	*GEV*	0.79	0.33	-0.05		
West											
7	19.51	14.45	8.66	-1.41	**-0.68**	*GPA*	0.37	0.62	-0.02		
8	31.21	18.77	16.39	11.10	-9.58	*WAK*	0.35	0.58	2.62	0.58	-0.17
9	**1.52**	-7.88	-13.01	-22.24	-33.28	*GLO*	0.89	0.24	-0.26		
10	19.89	11.13	7.17	**-0.05**	-10.72	*PE3*	1.00	0.50		1.48	
Center											
11	63.27	53.04	45.54	32.36	25.38	*WAK*	0.32	0.74	0.10	0.00	0.00
12	30.61	13.51	14.22	11.95	-21.70	*WAK*	0.24	2.03	7.43	0.70	-0.34
13	35.70	16.95	21.79	21.79	-16.88	*WAK*	0.18	2.61	8.57	0.92	-0.69
14	29.02	16.95	18.71	18.19	-6.53	*WAK*	0.36	0.92	1.90	0.42	-0.30
15	-11.39	-20.36	-19.99	-21.20	-38.83	*WAK*	0.40	2.43	5.54	0.20	0.11
East											
16	56.21	32.30	34.29	32.08	-15.90	*WAK*	0.26	1.17	5.37	0.83	-0.50
17	24.58	14.92	13.61	10.26	-6.62	*WAK*	0.36	0.56	2.58	0.58	-0.21
South											
18	23.65	**1.61**	5.38	3.62	-36.65	*GEV*	0.86	0.30	0.12		
19	12.97	1.83	2.73	**1.61**	-20.65	*PE3*	1.00	0.29		0.60	
20	37.51	16.77	20.29	19.72	-23.01	*WAK*	0.50	1.00	2.47	0.25	-0.16
21	20.74	14.78	13.42	10.58	**1.01**	*GPA*	0.41	0.79	0.34		
22	13.01	9.74	7.04	2.33	**0.68**	*GPA*	0.28	0.77	0.07		

3.4. Estimation of Regional Quantiles

Following successful selection of the most appropriate regional frequency distribution, the regional quantiles for various return periods are estimated using the parameter estimates for the distributions in Table 2. The index-flood method is used to estimate the quantile estimations at varied non-exceedance probabilities and recurrence intervals for 22 sub-regions in Taiwan in the current study. The index flood technique, which normalizes data at each site by dividing by the sample mean or median, is one of the most often used ways for regional frequency analysis.

Table 3 shows the desired rainfall quantile estimates with different return periods for each homogeneous region. It can be seen that the selected regional distributions for each sub-region predict quantiles that are very comparable. This suggests that each of the candidate distributions can be used in the regional frequency analysis of radar-based rainfalls in Taiwan, regardless of which distribution was chosen. Additionally, Table 3 demonstrates that the variation in the range of quantile estimations grew from 0.80-0.99 in 2-year to 1.71-3.35 in 100-year return period. This result explains that the variation in quantile estimate would increases as the return duration increases.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2022 doi:10.20944/preprints202207.0325.v1

Table 3. Quantile estimates with different return periods for each homogeneous sub-region

Region/Sub-region	Return Period					
	2	**5**	**10**	**20**	**50**	**100**
North						
1	0.86	1.39	1.78	2.18	2.75	3.21
2	0.91	1.34	1.59	1.80	2.01	2.13
3	0.94	1.25	1.46	1.65	1.91	2.11
4	0.88	1.29	1.61	1.93	2.39	2.77
5	0.91	1.46	1.74	1.94	2.12	2.21
6	0.92	1.31	1.58	1.85	2.21	2.50
West						
7	0.80	1.38	1.83	2.28	2.89	3.35
8	0.91	1.38	1.67	1.92	2.22	2.41
9	0.89	1.29	1.60	1.95	2.49	3.00
10	0.88	1.35	1.66	1.97	2.36	2.65
Center						
11	0.82	1.43	1.86	2.26	2.75	3.09
12	0.94	1.38	1.62	1.82	2.02	2.13
13	0.99	1.38	1.55	1.65	1.73	1.77
14	0.98	1.36	1.54	1.67	1.81	1.89
15	0.98	1.19	1.37	1.55	1.82	2.03
East						
16	0.96	1.40	1.62	1.78	1.91	1.98
17	0.92	1.37	1.64	1.87	2.13	2.29
South						
18	0.97	1.27	1.44	1.60	1.78	1.90
19	0.97	1.23	1.38	1.52	1.68	1.79
20	0.99	1.25	1.38	1.50	1.63	1.71
21	0.90	1.39	1.67	1.89	2.12	2.25
22	0.80	1.45	1.92	2.37	2.93	3.33

The regional growth curve, which was based on the regional distribution, was plotted for a set of return periods (Figure 7). This growth curve depicts the variation of the regional quantile (growth factor) q(F) versus the not-exceeded probability F or the return period T (T = 1 / (1 – F)). According to the variation of the growth curves for 22 homogeneous regions in Figure 6, sub-regions 1, 7, 11, and 22 have higher regional growth curves in the higher return period than other sub-regions, whereas sub-regions 13, 19, and 20 have the lowest. There are no results that show which distribution has a higher or lower growth factor. The level of growth factor, on the other hand, is influenced by site characteristics, particularly the annual maximum series and elevation. Sub-regions 13, 19, and 20 are sub-regions with high AMS and a higher average elevation.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2022 doi:10.20944/preprints202207.0325.v1

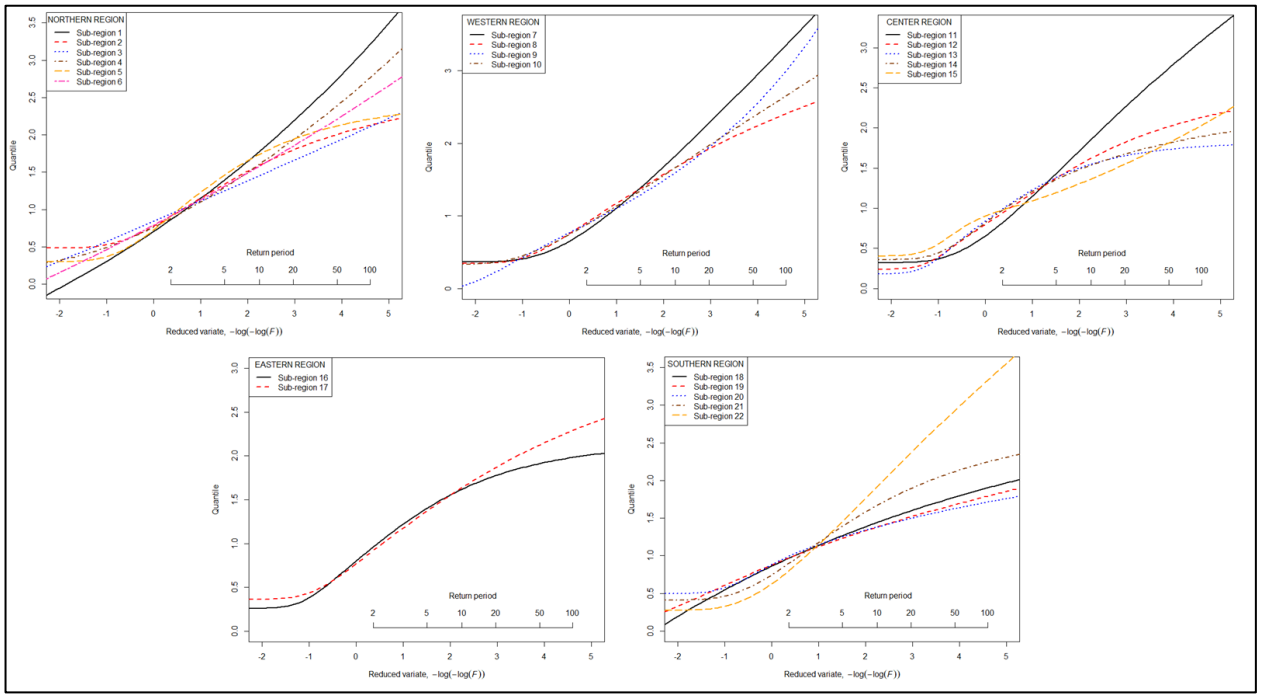


Figure 7. Regional growth curves of 22 homogeneous regions in Taiwan

3.5. Accuracy Result

Although the estimates obtained through regional frequency analysis are reliable, we still require specific measures to assess the accuracy of these estimates. Table 4 displays the relative regional RMSE values and 90 % error bounds. Table 4 shows that the RMSE and error bounds increased in line with the increase in the return period. It is clear that the uncertainty of regional quantiles grows with the return period. In the upper tail, high RMSE values and error bounds are observed, indicating the unreliability of quantities with return period T > 100 year.

Table 4 further demonstrates that the root mean square errors of the predicted regional growth curve and the quantiles for 22 homogenous regions range from 0.002 for a 2-year return period (sub-region 18) to 0.158 for a 100-year return period (sub-region 18), respectively (sub-region 9). The RMSE values indicate that the estimate of extreme rainfall is reliable up to a return period of 100 years. It will be necessary to collect more historical data for estimates of longer return periods in order to improve the reliability of the quantile estimation. Additionally, Table 4 includes the 90 percent error boundaries, demonstrating that the quantile estimations are sufficiently accurate when return periods are shorter than 100 years. This result is also consistent with [48]'s regional frequency analysis of precipitation extremes in China's Hanjiang River Basin. They discovered that when the return time exceeds 1000 years, the RMSE values are relatively large, implying that quantile estimations are sufficiently reliable when the return period is less than 100 years, and then become unreliable as the return period increases.

Table 4. RMES and 90 % error bounds of 22 homogeneous sub-regions with different return periods

Reg	2-year			5-year			10-year			20-year			50-year			100-year		
	RMSE	Error Bounds		RMSE	Error Bounds		RMSE	Error Bounds		RMSE	Error Bounds		RMSE	Error Bounds		RMSE	Error Bounds	
		Lower	Upper		Lower	Upper		Lower	Upper		Lower	Upper		Lower	Upper		Lower	Upper
North																		
1	.013	.847	.865	.011	1.377	1.393	.006	1.774	1.792	.027	2.182	2.224	.075	2.750	2.864	.126	3.214	3.408
2	.004	.899	.910	.006	1.339	1.353	.011	1.591	1.617	.014	1.789	1.830	.018	1.981	2.038	.026	2.083	2.161
3	.005	.932	.941	.004	1.243	1.250	.005	1.451	1.464	.013	1.654	1.678	.030	1.914	1.967	.047	2.110	2.187
4	.011	.865	.879	.006	1.286	1.295	.008	1.608	1.622	.028	1.934	1.975	.064	2.394	2.485	.099	2.768	2.909
5	.003	.906	.914	.006	1.461	1.473	.007	1.736	1.755	.008	1.926	1.949	.021	2.086	2.121	.036	2.157	2.213
6	.007	.904	.915	.006	1.301	1.309	.005	1.578	1.591	.017	1.851	1.882	.042	2.214	2.286	.067	2.497	2.608
West																		
7	.011	.785	.801	.009	1.369	1.383	.007	1.830	1.844	.035	2.282	2.332	.090	2.887	3.018	.146	3.351	3.566
8	.005	.901	.910	.004	1.374	1.386	.008	1.664	1.684	.011	1.916	1.944	.014	2.203	2.241	.021	2.384	2.444
9	.017	.871	.890	.012	1.275	1.289	.005	1.596	1.608	.032	1.945	1.989	.091	2.493	2.614	.158	2.995	3.207
10	.008	.871	.882	.002	1.342	1.347	.009	1.665	1.680	.020	1.971	2.002	.036	2.364	2.417	.049	2.654	2.726
Center																		
11	.009	.808	.820	.005	1.425	1.439	.017	1.858	1.883	.029	2.257	2.305	.043	2.746	2.828	.052	3.088	3.196
12	.004	.934	.943	.004	1.373	1.384	.005	1.618	1.636	.006	1.810	1.829	.013	1.993	2.025	.020	2.094	2.138
13	.004	.986	.995	.004	1.379	1.390	.005	1.543	1.559	.005	1.643	1.660	.011	1.710	1.737	.017	1.734	1.771
14	.005	.967	.982	.005	1.352	1.367	.009	1.520	1.553	.012	1.647	1.691	.014	1.783	1.823	.021	1.855	1.916
15	.006	.967	.977	.006	1.183	1.195	.005	1.357	1.373	.012	1.549	1.571	.035	1.816	1.870	.061	2.035	2.134
East																		
16	.003	.956	.965	.003	1.398	1.407	.004	1.612	1.625	.009	1.759	1.775	.018	1.885	1.913	.024	1.945	1.983
17	.006	.909	.923	.005	1.365	1.380	.010	1.631	1.660	.013	1.859	1.895	.016	2.103	2.155	.023	2.254	2.329
South																		
18	.002	.964	.967	.002	1.262	1.267	.002	1.439	1.446	.004	1.595	1.606	.009	1.777	1.795	.013	1.900	1.926
19	.003	.967	.972	.002	1.226	1.232	.003	1.378	1.389	.006	1.509	1.529	.010	1.670	1.699	.014	1.782	1.820
20	.003	.987	.995	.003	1.244	1.254	.005	1.372	1.390	.006	1.485	1.505	.008	1.611	1.639	.013	1.693	1.736
21	.004	.891	.904	.003	1.384	1.395	.006	1.662	1.683	.014	1.874	1.920	.025	2.083	2.167	.034	2.198	2.313
22	.010	.783	.801	.009	1.438	1.459	.009	1.915	1.940	.032	2.362	2.428	.076	2.919	3.068	.118	3.318	3.545

doi:10.20944/preprints202207.0325.v1

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2022 doi:10.20944/preprints202207.0325.v1

3.6. Spatial Mapping of Extreme Rainfall

Spatial rainfall mapping is a useful tool for depicting rainfall frequency for different return periods across the study area. It is a map of the region in which the colors represent the intensity of expected rainfall in various areas of the region. In this analysis, the maps were generated using GIS technique. Extreme rainfall for 24-hour duration and return period (2, 5, 10, 20, 50, and 100-year) were created by importing all the point values and convert into grid/raster values. the rainfall extreme values are obtained from the index-flood approach which multiplies the regional growth factor of different return periods by the average annual maximum series of 22787 QPESUMS grids as index-flood.

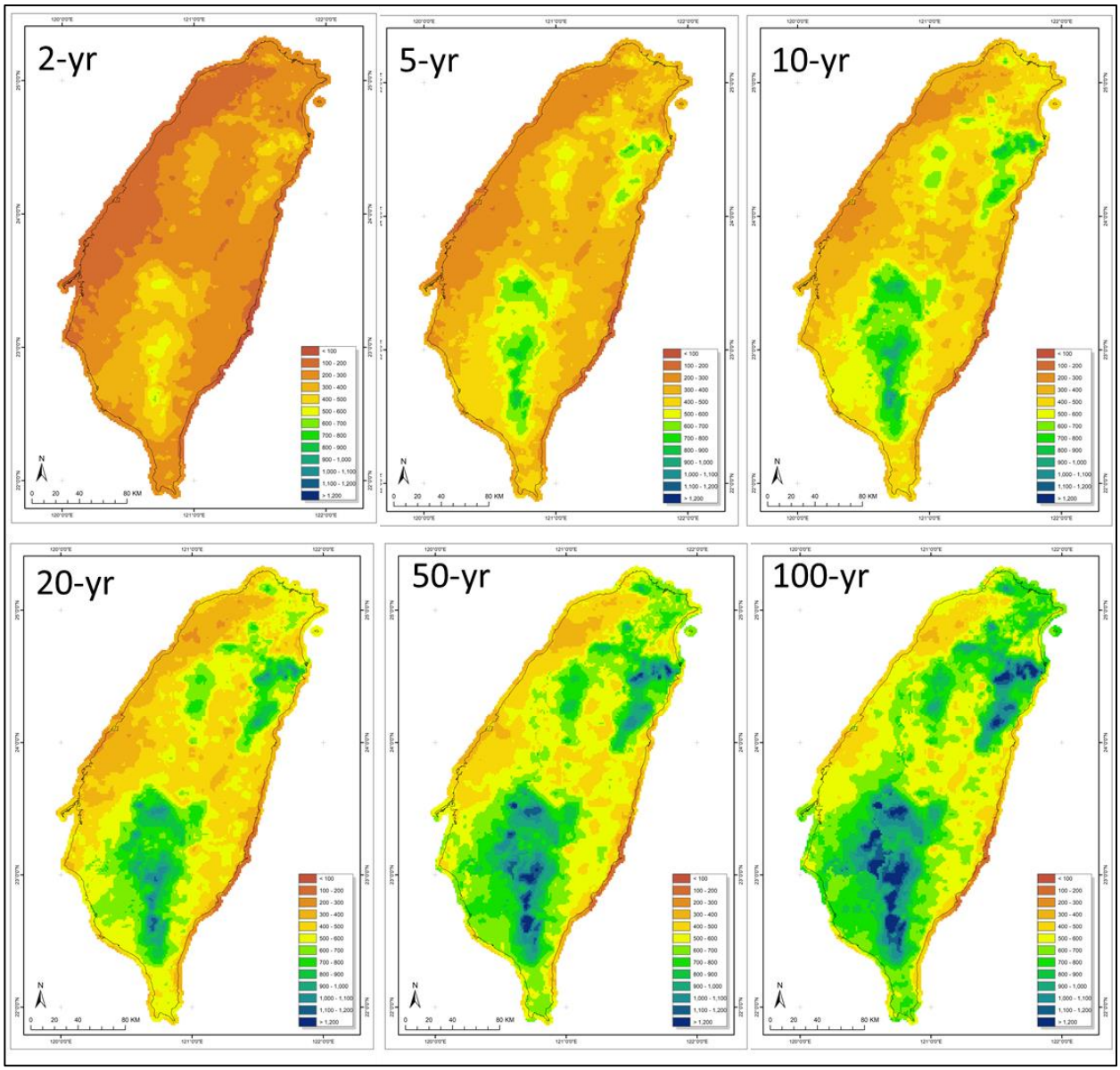


Figure 8. Spatial patterns of extreme rainfall for different return periods

Figure 8 illustrates the spatial patterns of extreme rainfall for different return periods. There are two significant regions with the highest rainfall extremes especially for 100-year return period with extreme value more than 1200 mm/day. One is the eastern part of northern region (mainly sub-region 5) which is in Yilan County, and the other is in the southern region (sub-region 10, 14, 15, 19, and 20) which is part of Tainan City, Chiayi County, Kaohsiung City, and Pingtung County. These two places are frequently struck by typhoons. The closure of the central mountain range has a profound effect on the spatial distribution, frequency, and intensity of typhoon extreme rainfall. [49] investigated,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2022 doi:10.20944/preprints202207.0325.v1

utilizing CWB hourly precipitation data from stations, and discovered that the typhoon extreme rainfall pattern is phase locked to the central mountain range (CMR), and that topographic forcing plays a critical role in typhoon extreme rainfall.

3.7. Regional and Grid-based Estimates for Watershed Scale

Extreme rainfall patterns are commonly characterized using at-site and regional frequency analysis. In previous part, it has been proven that regional frequency analysis has been able to estimate and display extreme rain for the entire Taiwan area. However, in such a case, a comparison with at-site analysis is advised, particularly in smaller scale regions. In this part, we were analyzed the difference and similarity between regional and grid-based (at-site) estimate at several return periods at Touqian Watershed (Figure 9).

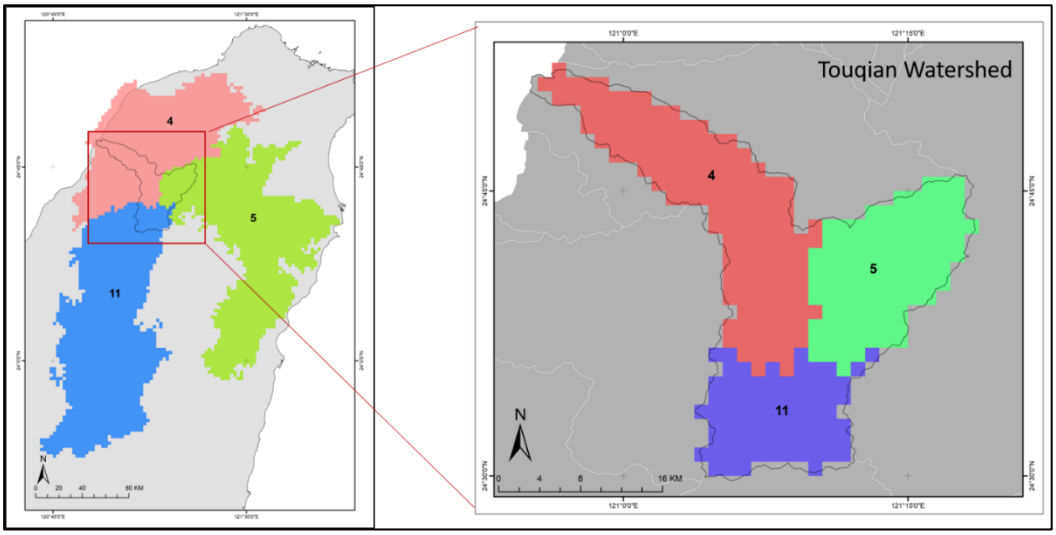


Figure 9. Touqian Watershed as region for comparing regional and grid-based (at-site) analysis

As seen in Figure 9, there are three homogeneous sub-regions within the Touqian watershed, including sub-region 4, 5, and 11, with a total of 322 grids or 566 km^2 within drainage area and a total of 5103 grids for the 3 homogeneous subregions. More specifically, of the 322 grids within the Touqian watershed, 145 grids are located in subregion 4 (10.67 % of total SR-4 grids), 101 grids are located in subregion 5 (5.59 % of total SR-5 grids), and 76 grids are located in subregion 11 (3.92 % of total of total SR-11 grids).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2022 doi:10.20944/preprints202207.0325.v1

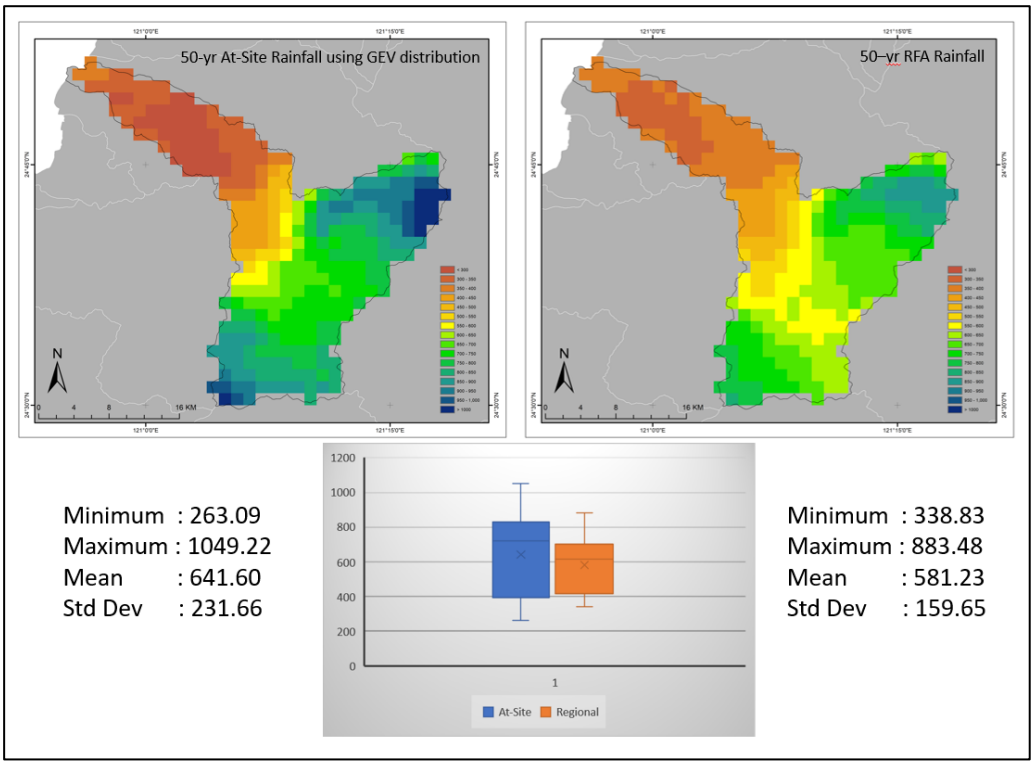
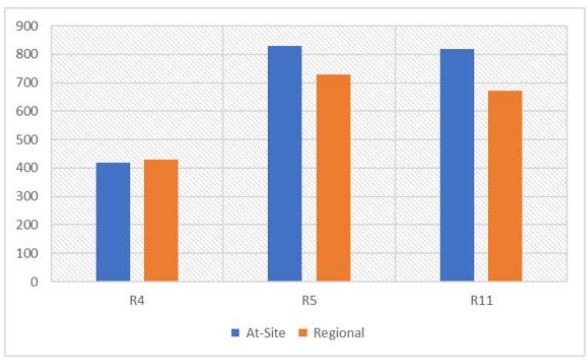


Figure 10. Extreme rainfall for at-site frequency analysis using the GEV distribution and regional frequency analysis at 50-year return period

Spatial distribution of extreme rainfall and statistics for a 50-year return period at-site frequency analysis using the GEV distribution and for a 50-year return period using regional frequency analysis is depicted in Figure 10. It's not clear which frequency distribution is best for estimating extreme hydrological events, but the GEV distribution is theoretically a good fit for the annual maximum series of extreme hydrological events. As can be observed from the at-site analysis, rainfall varies from 263.09 to 1049.22 mm/day, with an average of 641.6 mm/day. Meanwhile, regional analysis indicates a narrower rainfall range, with a minimum of 338.83 mm/day and a maximum of 883.48 mm/day, with a lower average of 581.23 mm/day. Spatially, extreme rainfall is comparatively larger when at-site analysis is used, particularly in the upper part of the Touqian watershed, which is part of subregion 5 and 11. This result indicating that regional frequency analysis may yield more precise estimates of rainfall quantiles than at-site analysis not only for larger area but also for smaller-scale area.

Figure 11 illustrates the extent to which the two analyses' estimates of extreme rainfall differ. As can be observed, the difference in the value of extreme rainfall estimates is not significant in subregion 4, however the difference is rather significant in subregion 5 and 11. The average extreme rainfall for at-site and regional analyses in subregion 5 was 829.55 mm/day and 729.84 mm/day, respectively, whereas in subregion 11, it was 817.7 mm/day and 670.99 mm/day.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2022 doi:10.20944/preprints202207.0325.v1



Region	At-Site	Regional
4	418.39	430.67
5	829.55	729.84
11	817.70	670.99

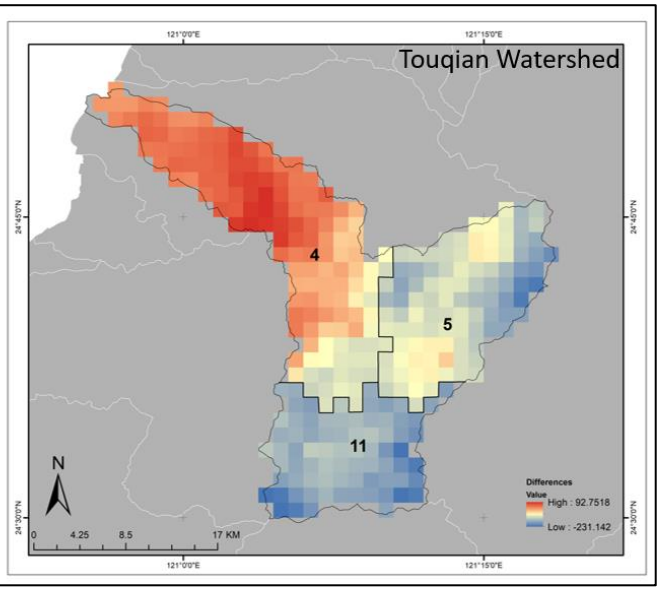


Figure 11. Comparison of the average extreme rainfall for at-site and regional frequency analysis at 50-year return period.

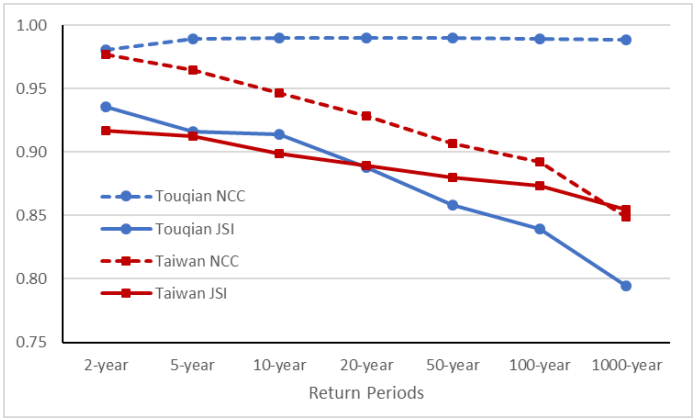


Figure 12. Similarity index for Touqian and Taiwan

Table 5. Similarity index between grid-based and regional frequency analysis for Touqian watershed and Taiwan

Return	Touqian		Taiwan	
Period	NCC	JSI	NCC	JSI
2-year	0.980	0.935	0.977	0.917
5-year	0.989	0.916	0.965	0.912
10-year	0.990	0.914	0.946	0.899
20-year	0.990	0.888	0.929	0.890
50-year	0.990	0.858	0.907	0.880
100-year	0.989	0.839	0.892	0.874
1000-year	0.988	0.795	0.849	0.855

Figure 12 and Table 5 depict similarity measures between the grid-based and regional approaches at Touqian Watershed and for the whole Taiwan as a reference for different return periods. Based on that figure and table, the NCC values for Touqian are around 0.9 which means there was a higher correlation between grid-based and regional. So was the correlation for the whole of Taiwan. The difference is that for the Touqian

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2022 doi:10.20944/preprints202207.0325.v1

watershed (small-scale), the correlation tends to be stable while for the Taiwan (big-scale), the correlation tends to decrease as the return period increases, although it was not very significant (from 0.97 at 2-year to 0.84 at 1000-year). Likewise, the JSI value which has a range from 0.79 to 0.93 which explains that there was a high similarity between grid-based and regional approaches for both small-scale and big-scale. Both scales also show a decrease in the level of similarity along with the increase in the rainfall return period. Interestingly, the similarity value for Touqian is lower than for Taiwan at a higher return period. This result is in line with the accuracy results using Monte Carlo which demonstrates increasing uncertainty in higher return periods. [50] also suggest, for higher return periods, the estimation of extreme rainfall should be treated with caution and longer rainfall data is required. By using grid-based as the basis for assessing regional frequency analysis, these results show that the regional approach in determining extreme rainfall is very suitable for large-scale applications and even better for smaller scales such as watershed areas.

3.8. Small-scale Spatial Investigation

A spatial investigation in a more specific area is carried out to determine the accuracy of regional frequency analysis. According to [51], Certain inaccuracies are to be expected when applying the approach to different geographical locations with reduced spatial variance. The spatial investigation in this study was carried out by establishing a region of interest in the northern part of Taiwan (Figure 13). The northern part of Taiwan was chosen as the region of interest because it has several small homogeneous subregions and a more dynamic rainfall pattern.

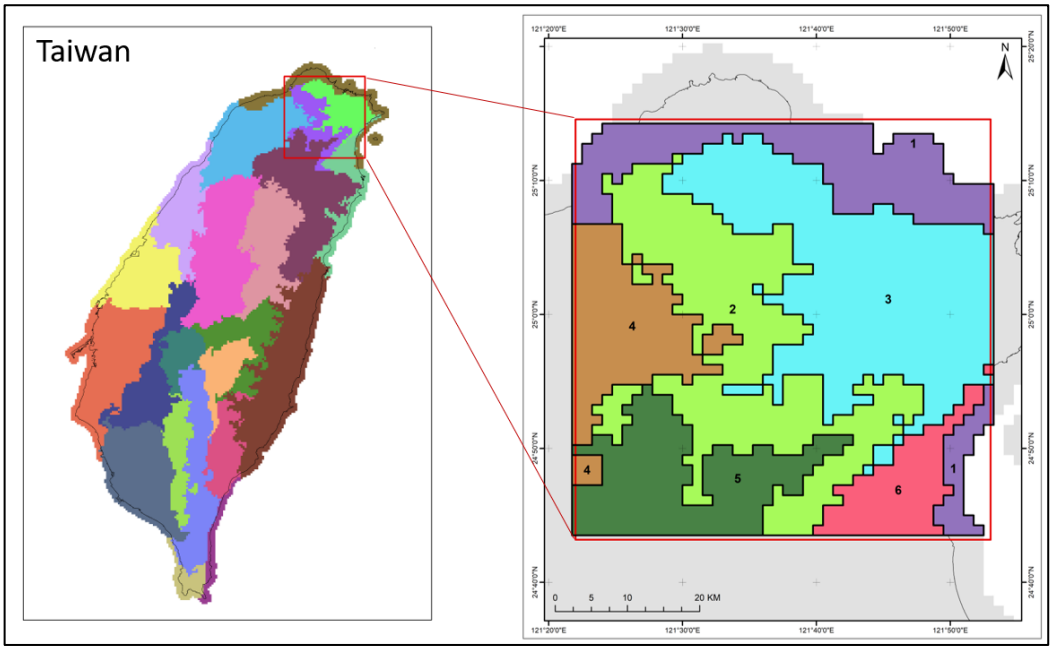


Figure 13. Region of Interest in Northern part of Taiwan used for spatial investigation

Furthermore, Figure 14 and Table 5 compare the average extreme rainfall value, standard deviation, and coefficient of variation of the RoI's six homogeneous sub-regions and the entire RoI for the 5-year return period. The coefficient of variation value for each sub-region within the RoI is markedly different, and when compared to the CV value for the entire RoI, the CV value for each sub-region is lower. This demonstrates that the regionalization was accurate and consistent.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2022 doi:10.20944/preprints202207.0325.v1

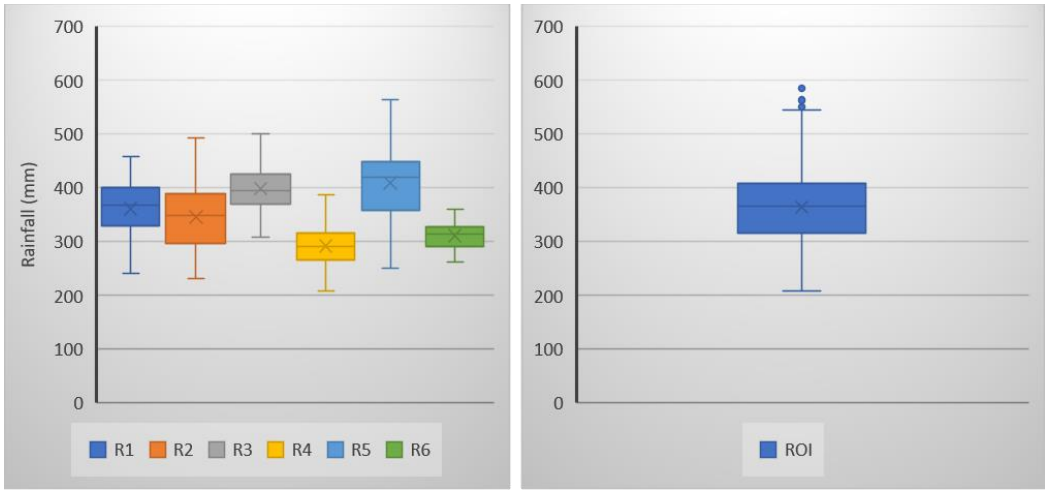


Figure 14. Extreme rainfall of 6 sub-regions within region of interest and the whole region of interest at 5-year

Table 5. Extreme rainfall statistics of 6 sub-regions within region of interest and the whole region of interest

Sub-region	Region of Interest			Whole ROI		
	Mean	Std Dev	CV	Mean	Std Dev	CV
1	361.51	49.78	13.77	363.95	67.23	18.47
2	345.39	59.00	17.08			
3	399.03	38.45	9.64			
4	291.57	35.03	12.01			
5	409.32	63.94	15.62			
6	311.17	23.14	7.44			

5. Conclusions

Regional frequency analysis of extreme rainfall using the L-moment approach was carried out in Taiwan using the QPESUMS daily rainfall dataset with high resolution from 2006 to 2020. Using K-means clustering, the entire island of Taiwan was divided into 22 homogeneous sub-regions based on extreme rainfall (series mean annual maximum, and 5-year rainfall) and location index (including latitude, longitude, and altitude). With higher resolution of radar rainfall, the clustering produces homogeneous regions that are not only similar, but also more detailed with clear boundaries between regions, especially in the center/mountain range area where the number of rain stations was still very limited. With the critical D_i increased to 6, 171 of the 22,787 grids (0.75 %) were classified as discordant grids. The grid with a discordant value greater than 6 located in Taiwan's southeastern coastal region has a significant impact on the region's homogeneity. Wakeby (WAK), Generalized Extreme Value (GEV), and Generalized Pareto (GPA) distributions of extreme rainfall are best suited for most subregions based on L-moment ratio diagrams and z-values (Goodness-of-fit measures). The Wakeby distribution is the most dominant and best-fitting distribution, particularly in the central and eastern regions. Estimation of rainfall quantiles based on parameters according to the best-fitted distribution for each sub-region resulted in an increase in the range of variation from 0.80-0.99 in 2-year to 1.71-3.35 in 100-year return period. This result explains that the variation in quantile estimate would increase as the return duration increases. According to the variation of the growth curves for 22 homogeneous regions, sub-regions 1, 7, 11, and 22 have higher regional growth curves in the higher return period than the other sub-regions, while sub-regions 13, 19, and 20 have the lowest. The relative regional RMSE values and 90 % error bounds

show that the uncertainty of regional quantiles increases with return period demonstrating that the quantile estimations are sufficiently accurate when return periods are shorter than 100 years. The index-flood method was used to create the spatial distribution of rainfall quantiles for several return periods (2 years, 5 years, 10 years, 20 years, 50 years, and 100 years), which considers the regional growth curve as a regional scale and the annual maximum rainfall as site specific. The spatial distribution of extreme rainfall reveals that two areas receive significantly more rainfall than others. The eastern part of the northern region (primarily sub-region 5) is in Yilan County, and the southern region (sub-regions 10, 14, 15, 19, and 20) is in Tainan City, Chiayi County, Kaohsiung City, and Pingtung County. A comparison of at-site (grid-based) and regional frequency analysis at the Touqian Watershed revealed that regional frequency analysis may yield more precise estimates of rainfall quantiles than at-site analysis, not only for larger areas but also for smaller-scale areas. A region of interest has been established to determine whether the regionalization is accurate in smaller and more complex sub-regions in the northern area. The coefficient of variation value for each sub-region within the RoI is significantly different, and the CV value for each sub-region is lower when compared to the CV value for the entire RoI. This demonstrates how accurate and consistent the regionalization was.

Author Contributions: Conceptualization, C.C., R.R.; methodology, C.C., R.R., S.W., C.H.; software, R.R..C.H.; validation, C.C., S.W.; formal analysis, C.C., R.R., S.W.; investigation, C.C., R.R.; resources, C.C.; data curation, R.R.; writing—original draft preparation, C.C., R.R.; writing—review and editing, C.C., R.R., S.W; visualization, R.R.; supervision, C.C., S.W., C.H. All authors have read and agreed to the published version of the manuscript

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data used in this study are duly available from the first authors on reasonable request.

Acknowledgments: This research was funded by the National Taipei University of Technology.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Overeem, A. Climatology of Extreme Rainfall from Rain Gauges and Weather Radar, Wageningen University, 2009.
2. Alias, N. E. B. Improving Extreme Precipitation Estimates Considering Regional Frequency Analysis, Kyoto University, 2014. https://doi.org/10.14989/doctor.k18562.
3. Liu, J.; Doan, C. D.; Liong, S.-Y.; Sanders, R.; Dao, A. T.; Fewtrell, T. Regional Frequency Analysis of Extreme Rainfall Events in Jakarta. *Nat. Hazards* **2015**, *75*, 1075–1104. https://doi.org/10.1007/s11069-014-1363-5.
4. Feng, Z.; Leung, L. R.; Hagos, S.; Houze, R. A.; Burleyson, C. D.; Balaguru, K. More Frequent Intense and Long-Lived Storms Dominate the Springtime Trend in Central US Rainfall. *Nat. Commun.* **2016**, *7* (May), 1–8. https://doi.org/10.1038/ncomms13429.
5. Fawad, M.; Yan, T.; Chen, L.; Huang, K.; Singh, V. P. Multiparameter Probability Distributions for At-Site Frequency Analysis of Annual Maximum Wind Speed with L-Moments for Parameter Estimation. *Energy* **2019**, *181*, 724–737. https://doi.org/10.1016/j.energy.2019.05.153.
6. Harka, A. E.; Jilo, N. B.; Behulu, F. Spatial-Temporal Rainfall Trend and Variability Assessment in the Upper Wabe Shebelle River Basin, Ethiopia: Application of Innovative Trend Analysis Method. *J. Hydrol. Reg. Stud.* **2021**, *37*, 100915. https://doi.org/10.1016/j.ejrh.2021.100915.
7. Kim, D. I. K.; Han, D.; Lee, T. Reanalysis Product-Based Nonstationary Frequency Analysis for Estimating Extreme Design Rainfall. *Atmosphere (Basel).* **2021**, *12* (2), 191. https://doi.org/10.3390/atmos12020191.
8. Tung, Y. S.; Wang, C. Y.; Weng, S. P.; Yang, C. D. Extreme Index Trends of Daily Gridded Rainfall Dataset (1960–2017) in Taiwan. *Terr. Atmos. Ocean. Sci.* **2022**, *33* (8). https://doi.org/10.1007/s44195-022-00009-z.
9. Mamoon, A. A.; Rahman, A. *Uncertainty Analysis in Design Rainfall Estimation Due to Limited Data Length: A Case Study in Qatar*; Elsevier Inc., 2019. https://doi.org/10.1016/B978-0-12-815998-9.00004-X.
10. Su, B.; Kundzewicz, Z. W.; Jiang, T. Simulation of Extreme Precipitation over the Yangtze River Basin Using Wakeby Distribution. *Theor. Appl. Climatol.* **2009**, *96*, 209–219. https://doi.org/10.1007/s00704-008-0025-5.
11. Gaál, L.; Kyselý, J.; Szolgay, J. Region-of-Influence Approach to a Frequency Analysis of Heavy Precipitation in Slovakia. *Hydrol. Earth Syst. Sci.* **2008**, *12* (3), 825–839. https://doi.org/10.5194/hess-12-825-2008.

12. Hosking, J. R. M.; Wallis, J. R. *Regional Frequency Analysis*; Cambridge University Press, 1997. https://doi.org/10.1017/cbo9780511529443.
13. Greenwood, J. A.; Landwehr, J. M.; Matalas, N. C.; Wallis, J. R. Probability Weighted Moments: Definition and Relation to Parameters of Several Distributions Expressable in Inverse Form. *Water Resour. Res.* **1979**, *15* (5), 1049–1054. https://doi.org/10.1029/WR015i005p01049.
14. Hosking, J. R. M. L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics. *J. R. Stat. Soc. Ser. B* **1990**, *52* (1), 105–124. https://doi.org/10.1111/j.2517-6161.1990.tb01775.x.
15. Eldardiry, H.; Habib, E. Examining the Robustness of a Spatial Bootstrap Regional Approach for Radar-Based Hourly Precipitation Frequency Analysis. *Remote Sens.* **2020**, *12* (22), 1–19. https://doi.org/10.3390/rs12223767.
16. Li, M.; Li, X.; Ao, T. Comparative Study of Regional Frequency Analysis and Traditional At-Site Hydrological Frequency Analysis. *Water (Switzerland)* **2019**, *11* (3). https://doi.org/10.3390/w11030486.
17. Gaál, L.; Kyselý, J. Comparison of Region-of-Influence Methods for Estimating High Quantiles of Precipitation in a Dense Dataset in the Czech Republic. *Hydrol. Earth Syst. Sci.* **2009**, *13* (11). https://doi.org/10.5194/hess-13-2203-2009.
18. Malekinezhad, H.; Nachtnebel, H. P.; Klik, A. Comparing the Index-Flood and Multiple-Regression Methods Using L-Moments. *Phys. Chem. Earth, Parts A/B/C* **2011**, *36*, 54–60. https://doi.org/https://doi.org/10.1016/j.pce.2010.07.013.
19. Gado, T. A.; Hsu, K.; Sorooshian, S. Rainfall Frequency Analysis for Ungauged Sites Using Satellite Precipitation Products. *J. Hydrol.* **2017**, *554*. https://doi.org/10.1016/j.jhydrol.2017.09.043.
20. Modarres, R.; Sarhadi, A. Statistically-Based Regionalization of Rainfall Climates of Iran. *Glob. Planet. Change* **2011**, *75*, 67–75. https://doi.org/10.1016/j.gloplacha.2010.10.009.
21. Santos, M.; Fragoso, M.; Santos, J. A. Regionalization and Susceptibility Assessment to Daily Precipitation Extremes in Mainland Portugal. *Appl. Geogr.* **2017**, *86*, 128–138. https://doi.org/10.1016/j.apgeog.2017.06.020.
22. Yin, Y.; Chen, H.; Xu, C. Y.; Xu, W.; Chen, C.; Sun, S. Spatio-Temporal Characteristics of the Extreme Precipitation by L-Moment-Based Index-Flood Method in the Yangtze River Delta Region, China. *Theor. Appl. Climatol.* **2016**, *124* (3–4), 1005–1022. https://doi.org/10.1007/s00704-015-1478-y.
23. Hsin-Fu Yeh, H. F.; Chang, C. F. Using Standardized Groundwater Index and Standardized Precipitation Index to Assess Drought Characteristics of the Kaoping River Basin, Taiwan. *Water Resour. Regime Water Bodies* **2019**, *46* (5), 670–678. https://doi.org/10.1134/S0097807819050105.
24. Wu, C. C.; Kuo, Y. H. Typhoons Affecting Taiwan: Current Understanding and Future Challenges. *Bull. Am. Meteorol. Soc.* **1999**, *80* (1), 67–80. https://doi.org/10.1175/1520-0477(1999)080<0067:TATCUA>2.0.CO;2.
25. Tung, Y.-S.; Wang, C.-Y.; Weng, S.-P.; Yang, C.-D. Extreme Index Trends of Daily Gridded Rainfall Dataset (1960–2017) in Taiwan. *Terr. Atmos. Ocean. Sci.* **2022**, *33* (1). https://doi.org/10.1007/s44195-022-00009-z.
26. Mulaomerović-Šeta, A.; Blagojević, B.; Imširović, Š.; Nedić, B. Assessment of Regional Analyses Methods for Spatial Interpolation of Flood Quantiles in the Basins of Bosnia and Herzegovina and Serbia. In *Lecture Notes in Networks and Systems, vol 316*; Springer, 2022. https://doi.org/10.1007/978-3-030-90055-7_35.
27. Overeem, A.; Buishand, T. A.; Holleman, I. Extreme Rainfall Analysis and Estimation of Depth-duration-frequency Curves Using Weather Radar. *Water Resour. Res.* **2009**, 45 (10). https://doi.org/10.1029/2009WR007869.
28. Sarmadi, F.; Shokoohi, A. Regionalizing Precipitation in Iran Using GPCC Gridded Data via Multivariate Analysis and L-Moment Methods. *Theor. Appl. Climatol.* **2015**, *122*, 121–128. https://doi.org/10.1007/s00704-014-1292-y.
29. Goudenhoofdt, E.; Delobbe, L.; Willems, P. Regional Frequency Analysis of Extreme Rainfall in Belgium Based on Radar Estimates. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 5385–5399. https://doi.org/10.5194/hess-21-5385-2017.
30. Chen, H.; Chen, C. Warning Models for Landslide and Channelized Debris Flow under Climate Change Conditions in Taiwan. *Water (Switzerland)* **2022**, *14* (5), 695. https://doi.org/10.3390/w14050695.
31. Chang, F. J.; Chiang, Y. M.; Tsai, M. J.; Shieh, M. C.; Hsu, K. L.; Sorooshian, S. Watershed Rainfall Forecasting Using Neuro-Fuzzy Networks with the Assimilation of Multi-Sensor Information. *J. Hydrol.* **2014**, *508*, 374–384. https://doi.org/10.1016/j.jhydrol.2013.11.011.
32. Chiou, P. T.; Chen, C.-R.; Chang, P.-L.; Jian, G.-J. Status and Outlook of Very Short Range Forecasting System in Central Weather Bureau, Taiwan. *Appl. with Weather Satell. II* **2005**, *5658*, 185. https://doi.org/10.1117/12.601195.
33. Chang, R. C.; Tsai, T. S.; Yao, L. Intelligent Rainfall Monitoring System for Efficient Electric Power Transmission. *Lect. Notes Electr. Eng.* **2013**, *253 LNEE*, 773–782. https://doi.org/10.1007/978-94-007-6996-0_81.
34. Neykov, N. M.; Neytchev, P. N.; VanGelder, P. H. A. J. M.; Todorov, V. K. Robust Detection of Discordant Sites in Regional Frequency Analysis. *Water Resour. Res.* **2007**, *43* (6). https://doi.org/10.1029/2006WR005322.
35. Hosking, J. R. M. Regional Frequency Analysis Using L-Moments (R Package). 2019. https://cran.r-project.org/package=lmomRFA
36. Lin, G. F.; Chen, L. H. Identification of Homogeneous Regions for Regional Frequency Analysis Using the Self-Organizing Map. *J. Hydrol.* **2006**, *324* (1–4), 1–9. https://doi.org/10.1016/j.jhydrol.2005.09.009.
37. Alem, A. M.; Tilahun, S. A.; Moges, M. A.; Melesse, A. M. *A Regional Hourly Maximum Rainfall Extraction Method for Part of Upper Blue Nile Basin, Ethiopia*; Elsevier Inc., 2019. https://doi.org/10.1016/B978-0-12-815998-9.00009-9.
38. Rao, A. R.; Srinivas, S. S. *Regionalization of Watersheds: An Approach Based on Cluster Analysis*; Springer Science & Business Media, 2008. https://doi.org/10.1007/978-1-4020-6852-2.
39. Abdi, A.; Hassanzadeh, Y.; Talatahari, S.; Fakheri-Fard, A.; Mirabbasi, R. Regional Drought Frequency Analysis Using L-Moments and Adjusted Charged System Search. *J. Hydroinformatics* **2017**, *19* (3), 426–442. https://doi.org/10.2166/hydro.2016.228.

40. Wright, M. J.; Houck, M. H.; Ferreira, C. M. Discriminatory Power of Heterogeneity Statistics with Respect to Error of Precipitation Quantile Estimation. *J. Hydrol. Eng.* **2015**, *20* (10). https://doi.org/10.1061/(ASCE)HE.1943-5584.0001172.
41. Khan, S. A.; Hussain, I.; Hussain, T.; Faisal, M.; Muhammad, Y. S.; Shoukry, A. M. Regional Frequency Analysis of Extremes Precipitation Using L-Moments and Partial L-Moments. *Adv. Meteorol.* **2017**. https://doi.org/10.1155/2017/6954902
42. Busababodhin, P.; Seo, Y. A.; Park, J. S.; Kumphon, B. LH-Moment Estimation of Wakeby Distribution with Hydrological Applications. *Stoch. Environ. Res. Risk Assess.* **2016**, *30*, 1757–1767. https://doi.org/10.1007/s00477-015-1168-4.
43. Rahman, M.; Hassan, M. R.; Buyya, R. Jaccard Index Based Availability Prediction in Enterprise Grids. *Procedia Comput. Sci.* **2010**, *1* (1), 2707–2716. https://doi.org/10.1016/j.procs.2010.04.304.
44. Chang, P. L.; Zhang, J.; Tang, Y. S.; Tang, L.; Lin, P. F.; Langston, C.; Kaney, B.; Chen, C. R.; Howard, K. An Operational Multi-Radar Multi-Sensor QPE System in Taiwan. *Bull. Am. Meteorol. Soc.* **2021**, *102* (3), 555–577. https://doi.org/10.1175/BAMS-D-20-0043.1.
45. Hu, C.; Xia, J.; She, D.; Xu, C.; Zhang, L.; Song, Z.; Zhao, L. A Modified Regional L-Moment Method for Regional Extreme Precipitation Frequency Analysis in the Songliao River Basin of China. *Atmos. Res.* **2019**, *230*, 104629. https://doi.org/10.1016/j.atmosres.2019.104629
46. Chen, L. H.; Hong, Y. T. Regional Taiwan Rainfall Frequency Analysis Using Principal Component Analysis, Self-Organizing Maps and L-Moments. *Hydrol. Res.* **2012**, *43* (3), 275–285. https://doi.org/10.2166/nh.2012.032.
47. Szolgay, J.; Parajka, J.; Kohnová, S.; Hlavčová, K. Comparison of Mapping Approaches of Design Annual Maximum Daily Precipitation. *Atmos. Res.* **2009**, *92*, 289–307. https://doi.org/10.1016/j.atmosres.2009.01.009.
48. Hao, W.; Hao, Z.; Yuan, F.; Ju, Q.; Hao, J. Regional Frequency Analysis of Precipitation Extremes and Its Spatio-Temporal Patterns in the Hanjiang River Basin, China. *Atmosphere (Basel).* **2019**, *10* (3). https://doi.org/10.3390/atmos10030130
49. Su, S. H.; Kuo, H. C.; Hsu, L. H.; Yang, Y. T. Temporal and Spatial Characteristics of Typhoon Extreme Rainfall in Taiwan. *J. Meteorol. Soc. Japan* **2012**, *90* (5), 721–736. https://doi.org/10.2151/jmsj.2012-510.
50. Malekinezhad, H.; Zare-Garizi, A. Regional Frequency Analysis of Daily Rainfall Extremes Using L-Moments Approach. *Atmosfera* **2014**, *27* (4), 411–427. https://doi.org/10.1016/S0187-6236(14)70039-6.
51. Wang, Z.; Zeng, Z.; Lai, C.; Lin, W.; Wu, X.; Chen, X. A Regional Frequency Analysis of Precipitation Extremes in Mainland China with Fuzzy C-means and L-moments Approaches. *Int. J. Cliamtology* **2017**, *37*, 429–444. https://doi.org/10.1002/joc.5013.