Pak.j.stat.oper.res. Vol.21 No.4 2025 pp 583-599 DOI:https://doi.org/10.18187/pjsor.v21i4.4876

Pakistan Journal of Statistics and Operation Research

$L-$moment Based Regional Frequency Analysis of Annual Maximum Relative Humidity Across Pakistan

Amina Shahzadi[1*], Khizra Yousaf[1]

*Corresponding author

1. Department of Statistics, Government College University, Lahore, Pakistan, aminashahzadi@gcu.edu.pk

Abstract

Climate change has significantly influenced regional weather patterns in Pakistan, leading to an increase in the frequency and intensity of extreme weather events, such as heatwaves, droughts, and floods. This study aims to investigate the regional frequency analysis of extreme relative humidity across Pakistan, identify potential geographically related trends, and estimate possible frequencies associated with extreme relative humidity in distinct regions. Relative humidity data for 24 meteorological stations were obtained from Pakistan Meteorological Department. $L-$moment based discordance measure was used to evaluate discordant sites in the region, with Gawadar and Karachi exhibiting the highest discordance values. Three homogeneous regions were delineated based on geographical and statistical characteristics, along with heterogeneity statistics derived from $L-$moments. The $L-$moment ratio diagram and goodness–of–fit statistic were employed to identify appropriate distributions for each region. In the first region, the generalized extreme value and Pearson type III distributions were found to be most effective. The second region was best represented by the generalized extreme value and Weibull distributions, whereas the third region was most accurately characterized by the generalized log-normal and polynomial density–quantile III distributions. The findings offer valuable insights into the spatial patterns of humidity extremes, which may potentially contribute to the development for climate change adaptation strategies and flood risk management efforts.

Key Words: $L-$moment; Regional frequency analysis; Relative humidity.

1. Introduction

Climate change has had a significant impact on weather patterns in Pakistan, with the country ranking fifth among the most affected nations due to extreme weather events between 1999 and 2018 (Ajani & van der Geest, 2021). Evidence for climate change in Pakistan over the past 54 years includes an annual temperature increase of 0.09°C and an average annual precipitation increase of 478 mm (Devadoss et al., 2021). Relative humidity, a crucial component of the climate system, has also exhibited significant changes, with implications for various sectors, including agriculture, water resources, and human health (Ali, 2018). According to recent studies, Pakistan has experienced significant changes in relative humidity over the past few decades. The mean annual humidity in Pakistan experienced an increase of 2.94% from 1988 to 2014, with the rate of increase in humidity being 0.97% during this period (Ali, 2018). Long-term trend analysis revealed a significant decreasing trend in the annual maximum relative humidity, particularly in the Sindh and Punjab provinces (Ali et al., 2021).

Frequency analysis of extreme humidity events and estimation of their return periods are crucial for infrastructure

designs, disaster management planning, agricultural planning, weather forecasting, public health awareness and industrial applications . Previous studies on the statistical modeling of humidity along with other meteorological variables highlight the use of various statistical models, such as time series models, extreme value theory, regression models, and probability distributions, to analyze and forecast extreme humidity events (e.g., Adedayo, 2016; Bangash et al., 2019; Kumar et al., 2021; Solaimani et al., 2024; Yao, 1974) These studies have demonstrated the effectiveness of these approaches in quantifying the frequency and intensity of humidity conditions, which can inform adaptation strategies and decision-making.

$L-$moments based regional frequency analysis (RFA) is a statistical method used to determine the probability and intensity of extreme events such as floods, thunderstorms, extreme rainfall or extreme humidity by combining data from multiple locations with similar meteorological or hydrological conditions (Hosking & Wallis, 1997). It offers several advantages over at-site frequency analysis, particularly for variables with limited data availability. RFA offers a more comprehensive framework for understanding the spatial pattern of extreme events (Martinez-Villalobos & Neelin, 2023). There is extensive literature on the application of regional frequency analysis to model various types of meteorological and hydrological conditions such as temperature (Khan et al., 2020; Souaissi et al., 2023), extreme rainfall (Hassan et al., 2021; Mahbod et al., 2024a; Shahzadi et al., 2013), flood (Hussain & Pasha, 2009; Pilon & Adamowski, 1992; Seckin et al., 2011; Vivekanandan, 2015), stream flows (Ahmad et al., 2016; Riggs, 1973), droughts (Clausen & Pearson, 1995; Parvizi et al., 2022) and wind speed (Ahmad et al., 2024) etc.

The existing literature on the statistical modeling of extreme relative humidity lacks $L-$moments based RFA to better understand and prepare for such events. Therefore, it is necessary to extend this approach to analyze extreme relative humidity events in Pakistan, which can have significant impacts on agriculture, water resources, and human health. This study aims to fill this gap by conducting a regional frequency analysis of annual maximum relative humidity across different sites in Pakistan. This will provide a methodology for analyzing extreme humidity events both at the at-site and at the regional scale through regional frequency analysis, while providing a spatial and temporal perspective on the variability of these events. This research will contribute to a better understanding of extreme relative humidity events in Pakistan, enabling more effective preparation and response strategies across various sectors such as meteorology, agriculture planning, and weather forecasting.

The remainder of this paper is organized as follows. Section 2 describes the study region of extreme relative humidity (RH). $L-$moments are defined in Section 3, leading to the RFA methodology in Section 4. Section 5 presents data analysis and discussions. Finally, Section 6 concludes the study.

2. Study area and data description

Relative humidity (RH) data utilized in this study were obtained from the Pakistan Meteorological Department. RH data were collected from 24 meteorological sites across Pakistan, spanning the period from 1981 to 2022. Details of these 24 sites and their geographical characteristics are presented in Table 1. For each site, we used the annual maximum RH, resulting 42 extreme events for each of the 24 sites.

The study region experiences an active monsoon influence, especially in eastern and northeastern Pakistan, including Kashmir. Figure 1 shows the sites selected for the 24 study stations. The climatic conditions of the study region are tropical in summer and slightly cooler in winter; spring and autumn are transitional seasons. In the first phase of the monsoon period, the community faces unstable weather because of the rise in ground temperatures; however, the lower troposphere experiences the effects of winter. This leads to thunder, dust, and showers. On the other hand, there exists a relatively arid period following the monsoon season, occurring after the cessation of monsoons and prior to the arrival of westerly systems that bring winter precipitation. The highest humidity was recorded during the monsoon season in July, August, and September because monsoon winds cause moisture. In winter, there is a moderate occasional rise in humidity, mainly because of westerly systems that have a minor influence on regional humidity. These northern mountainous areas have much higher humidity than the areas near the southern plains.

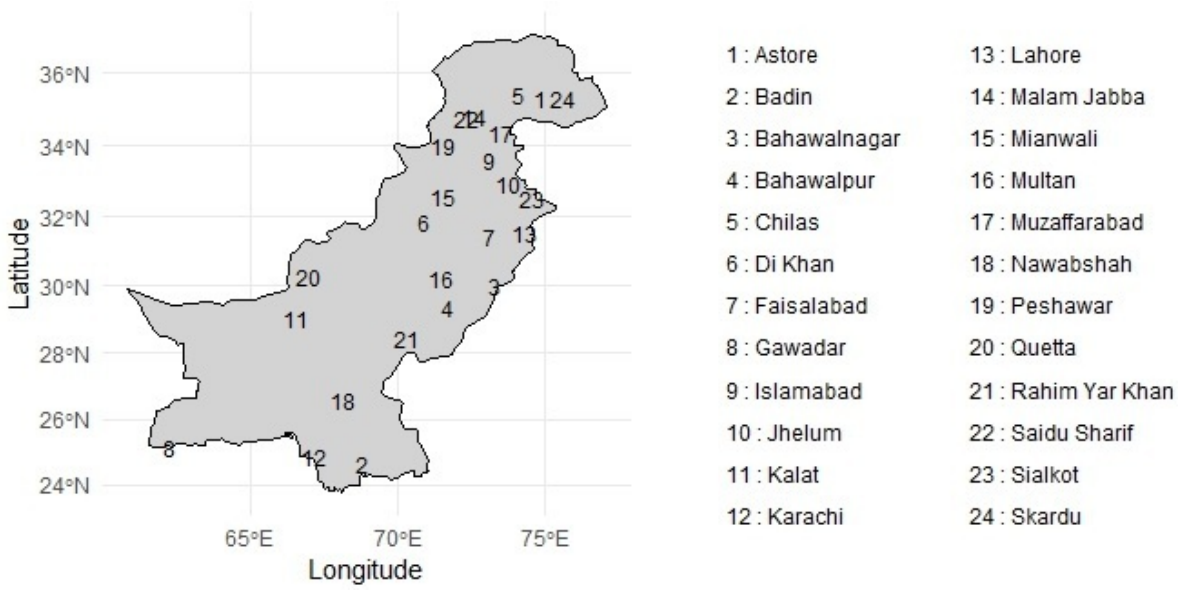

Figure 1: Humidity sites in Pakistan.

Table 1: Humidity sites and at-site characteristics

Site No.	Site name	Latitude	Longitude	Elevation	Mean annual RH
1	Astore	35.33	74.85	3674.75	55.18
2	Badin	24.65	68.83	6.15	46.43
3	Bhawalnagar	30.00	73.30	155.04	32.14
4	Bhawalpur	29.35	71.71	121.60	30.51
5	Chilas	35.41	74.10	3130.36	48.32
6	DI Khan	31.82	70.90	258.73	33.13
7	Faisalabad	31.40	73.10	182.92	35.70
8	Gawadar	25.13	62.32	13.62	70.74
9	Islamabad	33.60	73.10	543.83	45.86
10	Jhelum	32.92	73.72	384.47	43.81
11	Kalat	29.03	66.59	1765.85	33.24
12	Karachi	24.89	67.16	53.05	64.29
13	Lahore	31.50	74.30	209.49	39.70
14	Malam Jabba	34.80	72.57	2115.61	52.75
15	Mianwali	32.58	71.55	294.32	37.71
16	Multan	30.20	71.50	118.02	30.85
17	Muzaffarabad	34.36	73.47	2291.99	56.61
18	Nawabshah	26.58	68.17	43.94	33.63
19	Peshawar	33.99	71.52	986.57	41.79
20	Quettta	30.26	66.96	1813.74	34.07
21	Rahim Yar Khan	28.41	70.29	91.46	29.25
22	Saidu Sharif	34.75	72.35	2115.61	52.75
23	Sialkot	32.51	74.56	259.97	43.11
24	Skardu	35.30	75.60	3918.33	54.42

3. $L-$moments

$L-$momentsare a robust statistical method for characterizing the shape of a probability distribution. They provide a method to estimate the parameters of a distribution that is more resistant to the influence of outliers and sample size variations compared to traditional approaches like moment-based and maximum likelihood estimation (Anghel et al., 2023; Hosking, 1990; Hosking & Wallis, 1997). $L-$momentshave been widely used in RFAs for various types of extreme events, such as floods, droughts, and rainfall, as discussed in Section 1.
$L-$moments, denoted by λ_{r+1}, are defined as the linear combination of probability weighted momentss (PWMs) and are given by

$$\lambda_{r+1} = (-1)^r \sum_{k=0}^{r} p^*_{r,k}\alpha_k = \sum_{k=0}^{r} p^*_{r,k}\beta_k, \tag{1}$$

where, $p^*_{r,k} = (-1)^{r-k}\binom{r}{k}\binom{r+k}{k}$, α_k and β_k are special forms of PWMs (Greenwood et al. (1979)) defined for a random variable X with a cumulative distribution function, $F(.)$ as

$$\beta_k = E[X\{F(X)\}^k] \quad \text{and} \quad \alpha_r = E[X\{1-F(X)\}^k] \qquad k = 1, 2, 3, \ldots \tag{2}$$

$L-$moment ratios are dimensionless versions of $L-$moments and are obtained as

$$\tau_r = \lambda_r/\lambda_2, \qquad r = 1, 3, \ldots \tag{3}$$

τ_3 and τ_4 are known as $L-$skewness and $L-$kurtosis, respectively. The analogy of ordinary coefficient of variation is obtained as

$$\tau = \lambda_2/\lambda_1. \tag{4}$$

and is called $L-$CV. λ_1 and λ_2 are the $L-$location or mean and $L-$scale of the distribution, respectively. For sample $L-$moments, we use notations $l_1, l_2, t, t_r\,(r = 3,4,,\ldots)$. $L-$momentshave better discrimination ability than other standard statistical moments because of their unbiased nature and efficiency in estimating hydrological data characteristics and distribution parameters. These estimators are linear combinations of ranked observations, and are more robust than ordinary product-moment estimators against the largest observations in a sample (Hosking, 1990, 2007; Hosking & Wallis, 1997).

4. $L-$moments based regional frequency analysis

In developing countries, such as Pakistan, extrapolation is necessary because of the limited data on extreme humidity events. To overcome data limitations and model uncertainty, a regional frequency analysis for annual maximum relative humidity in Pakistan is employed, based on the index flood method of Dalrymple and Dawdy (1960) and $L-$moments developed by (Hosking, 1990). Hosking and Wallis (1997) provided a detailed explanation of RFA. This approach provides reliable and consistent design estimates for various hydrological, meteorological and agricultural tasks.

RFA assumes that data exist at N sites, site i is sampled with size n_i, and the observed values are represented by Q_{ij}, $j = 1, 2, \ldots, n_i$. Let $Q_i(F), 0 < F < 1$ be the quantile function of the frequency distribution at site i. An important assumption inherent in the index flood approach is that the stations are located in the same uniform area, implying that all N sites have identically distributed frequency distributions, except for site-specific scaling factors known as flood indices. Thus, it can be expressed as

$$Q_i(F) = \mu_i q(F), \quad i = 1, 2, \ldots, N \tag{5}$$

where μ_i is the index flood or scale factor specific to a site i. F is referred to as the nonexceedance probability that an event will not exceed a specific threshold. The return period is defined as the average time between events of a given magnitude or greater. There exists an inverse relationship between the nonexceedance probability and return period as

given by

$$\text{Return Period} = \frac{1}{1-F}. \tag{6}$$

Relationship in (6) is a fundamental concept in hydrological and environmental risk assessments. Understanding this relationship is crucial for estimating the likelihood of extreme events and managing associated risks, as demonstrated in various studies on flood events, heatwaves, and other natural phenomena (Ahmad et al., 2024; Hosking & Wallis, 1997; Shahzadi et al., 2013).

The steps followed in the regional frequency analysis are in the following subsections.

4.1. $L-$moments ratios and discordance measure

The RFA method uses data screening to identify discordant sites using the discordance measure, D_i, defined by

$$D_i = \frac{1}{3}N(\mathbf{u_i}-\bar{\mathbf{u}})^T\mathbf{A^{-1}}(\mathbf{u_i}-\bar{\mathbf{u}}) \tag{7}$$

where $\mathbf{u_i} = [t^{(i)}, t_3^{(i)}, t_4^{(i)}]^T$ is a vector of $L-$moment ratios for site i, N is the number of sites and $\bar{\mathbf{u}} = \sum_{i=1}^{N}\mathbf{u_i}/N$ and $\mathbf{A} = \sum_{i=1}^{N}(\mathbf{u_i}-\bar{\mathbf{u}})(\mathbf{u_i}-\bar{\mathbf{u}})^T$. If $D_i > 3$ for Site i, it is considered to be discordant (Hosking & Wallis, 1997).

4.2. Formation of homogeneous regions

The most crucial phase of RFA is the creation of homogeneous regions or areas. Cluster analysis based on site characteristics or geographical locations has been successfully used to create homogeneous regions in environmental applications (Burn, 1989; Hosking & Wallis, 1997; Hussain & Pasha, 2009). Subjective partitioning also helps determine homogeneous areas based on different features of stations, such as drainage area, wind speed, latitude, longitude, average annual rainfall, wastewater zone, and flood time (Hosking & Wallis, 1997; Shahzadi et al., 2013).

The geographical and climatological characteristics are crucial for the formation of homogeneous regions. For instance, Gaál et al. (2009) emphasized the importance of combining physical and geomorphological considerations with objective methods for precipitation analysis in Slovakia. Mahbod et al. (2024b) identified primary regions in Iran based on annual precipitation patterns, highlighting the significance of spatial heterogeneity in hydrological conditions. Site characteristics play a vital role in determining the homogeneous regions. Guru (2022) utilized fifteen effective variables, including nine selected through factor analysis, to divide the Mahanadi Basin into homogeneous regions. Mosaffaie (2015) identified key variables influencing flood magnitude, such as perimeter, equivalent diameter, time of concentration, length of the main waterway, and area.

$$H_j = \frac{v_j - \mu_j}{\sigma_{v_j}}, \quad j = 1, 2, 3, \tag{8}$$

where μ_j and σ_{v_j} are the mean and standard deviation of v_j. The v_j's are computed using regional averages of $L-$moments rations t, t_3, t_4, respectively such that

$$v_1 = \left(\frac{\sum_{i=1}^{N} n_i\left(t^{(i)}-t^R\right)^2}{\sum_{i=1}^{N} n_i}\right)^{1/2}, \qquad v_2 = \left(\frac{\sum_{i=1}^{N} n_i\left(t_3^{(i)}-t_3^R\right)^2}{\sum_{i=1}^{N} n_i}\right)^{1/2}, \qquad v_3 = \left(\frac{\sum_{i=1}^{N} n_i\left(t_4^{(i)}-t_4^R\right)^2}{\sum_{i=1}^{N} n_i}\right)^{1/2} \tag{9}$$

where t^R, t_3^R and t_4^R are regional average $L-$moments ratios defined by

$$t^R = \frac{\sum_{i=1}^{N} n_i t^{(i)}}{\sum_{i=1}^{N} n_i}, \qquad t_3^R = \frac{\sum_{i=1}^{N} n_i t_3^{(i)}}{\sum_{i=1}^{N} n_i}, \qquad t_4^R = \frac{\sum_{i=1}^{N} n_i t_4^{(i)}}{\sum_{i=1}^{N} n_i}. \tag{10}$$

The computations of heterogeneity measures involve a Monte Carlo simulation using a four parameter Kappa distribution, which is a generalization of many probability distributions. Kappa distribution is fitted using regional average $L-$moments ratios $1, t^R, t_3^R, t_4^R$. Subsequently, N_{sim} simulations are performed for a region comprising the same number of sites as in the original data set, with each site containing the same number of observations as in the original site data set. From the simulated values, the variable v_j, means (μ_j) and standard deviations (σ_{v_j}) are computed for heterogeneity measures defined in (8).

A region is said to be "acceptably homogeneous" for $H_j < 1$, "possibly heterogeneous" for $1 < H_j < 2$ and "definitely heterogeneous" for $H_j > 2$.

4.3. Determining an appropriate regional probability distribution

Meteorologists often focus on selecting the best probability distribution for regional frequency analysis. However, in reality, they seek a robust distribution that provides precise quantile estimates for return periods. In this study, we consider nine generalized distributions which involves 3-parameter distributions such as generalized extreme value (GEV), generalized logistic (GLO), generalized normal (GNO), Govindarajulu (GOV), generalized Pareto (GPA), Pearson Type III (PE3), polynomial denity–quantile III (PDQ3), 4–parameter asymmetric exponential power (AEP4) and a five parameter Wakeby distribution. Details of these distributions can be found in Asquith (2014), Hosking (2007), and Hosking and Wallis (1997).

$L-$moment ratio diagram and $Z^{DIST}-$statistics (Hosking & Wallis, 1997) are powerful tools for selecting the best regional probability distribution in hydrology/meteorology. $L-$moment ratio diagram plots the theoretical $L-$skewness and $L-$kurtosis values for various probability distributions, allowing a visual comparison with the sample $L-$moment ratios. This diagram is shown in Figure 4. This graphical approach helps identify the distributions that best fit the observed data.

The $Z-$statistic, often referred to as Z^{DIST}, is a goodness-of-fit measure used in conjunction with $L-$moment ratio diagram. It quantifies the difference between the sample and theoretical $L-$moment ratios for candidate distributions. For a three-parameter candidate distribution, the computation of Z^{DIST} involves fitting a four-parameter kappa distribution to the regional average $L-$moment ratios $1, t^R, t_3^R$ and t_4^R defined in (10) and simulating N_{sim} realizations of a homogeneous region with n sites from the fitted kappa distribution. Then, bias (B_4) and standard deviation (σ_4) of t_4^R are obtained, respectively, as

$$B_4 = N_{sim}^{-1} \sum_{m=1}^{N_{sim}} (t_4^{[m]} - t_4^R), \tag{11}$$

$$\sigma_4 = \left[(N_{sim} - 1)^{-1} \left\{ \sum_{m=1}^{N_{sim}} (t_4^{[m]} - t_4^R)^2 - N_{sim} B_4^2 \right\} \right]^{1/2}. \tag{12}$$

Then, for each candidate distribution, Z^{DIST} is defined as

$$Z^{DIST} = \frac{\tau_4^{DIST} - t_4^R + B_4}{\sigma_4}, \tag{13}$$

where, τ_4^{DIST} is the $L-$kurtosis of the fitted candidate distribution. A fitted distribution is declared to be appropriate if $|Z^{DIST} \leq 1.64|$ at 10% level of significance.

To test the goodness–of–fit for a $4-$parameter candidate distribution, we fit the Wakeby distribution, a more generalized distribution, to the regional average $L-$moment ratio $1, t, t_3^R, t_4^R$ and t_5^R, and simualte N_{sim} realizations of a homogeneous region from the fitted Wakeby distribution to calculate Z^{DIST} as given by

$$Z^{DIST} = \frac{\tau_5^{DIST} - t_5^R + B_5}{\sigma_5}, \tag{14}$$

where τ_5^{DIST} is 5th $L-$moment ratio of the fitted candidate distribution, and

$$B_5 = N_{sim}^{-1} \sum_{m=1}^{N_{sim}} (t_5^{[m]} - t_5^R) \tag{15}$$

and

$$\sigma_5 = \left[(N_{sim} - 1)^{-1} \left\{ \sum_{m=1}^{N_{sim}} (t_5^{[m]} - t_5^R)^2 - N_{sim} B_5^2 \right\} \right]^{1/2}. \tag{16}$$

4.4. Estimation and accuracy assessment of regional growth curve

After selecting an appropriate candidate distribution for a homogeneous, index–flood procedure (Dalrymple & Dawdy, 1960) using regional $L-$moments is a convenient and efficient method for estimating the chosen regional frequency distribution (Hosking & Wallis, 1997). Suppose there are N sites with site i having n_i sample observations Q_{ij}, $j = 1, 2, \ldots, n_i$ and $Q_i(F), 0 < F < 1$ be the quantile function of the at–site frequency distribution. The index–flood procedure states that the frequency distributions of N sites in a homogeneous region are identical, differing only by a site–specifice scaling factor, μ_i, known as the index–flood. Thus

$$Q_i(F) = \mu_i q(F) \tag{17}$$

where, $q(F)$ is the regional quantile function of the regional frequency distribution of rescaled data Q_{ij}/μ_i, referred to as the regional growth curve.

The regional and at–site quantiles at various nonexceedance probabilities (F) can be estimated using the regional $L-$moment algorithm outlined in Hosking and Wallis (1997, Table 6.1, pg. 95). This algorithm also involves the Monte Carlo simulation to assess the accuracy of the estimated quantile using relative bias, absolute relative bias and relative root mean square error (RMSE). Let $\hat{q}^{[m]}(F)$ and $\hat{Q}^{[m]}(F)$ be the estimated regional and at–site quantiles at nonexceedance probability F at the mth repetition of M simulations. Then, we define the accuracy measure in Table 2. The uniformity and consistency of quantile estimates, whether they tend to be consistently high or low, are evaluated using the measures of relative bias and relative absolute bias. The overall discrepancy between the estimated and true quantiles is measured by the relative RMSE. When the distribution of estimates is asymmetrical, additional

Table 2: Accuracy measures for at–site and regional quantile estimators

Measures	Site i quantile estimator	Regional quantile estimator
Relative bias	$B_i(F) = M^{-1} \sum_{m=1}^{M} \frac{\hat{Q}_i^{[m]}(F) - Q_i(F)}{Q_i(F)}$	$B^R(F) = N^{-1} \sum_{i=1}^{N} B_i(F)$
Relative absolute bias	$A_i(F) = M^{-1} \sum_{m=1}^{M} \frac{\lvert\hat{Q}_i^{[m]}(F) - Q_i(F)\rvert}{Q_i(F)}$	$A^R(F) = N^{-1} \sum_{i=1}^{N} A_i(F)$
Relative RMSE	$R_i(F) = \left[M^{-1} \sum_{m=1}^{M} \left\{ \frac{\lvert\hat{Q}_i^{[m]}(F) - Q_i(F)\rvert}{Q_i(F)} \right\}^2 \right]^{1/2}$	$R^R(F) = N^{-1} \sum_{i=1}^{N} R_i(F)$

valuable quantities include empirical quantiles of the distribution of estimates. These can be derived by computing the ratio of the estimated values to the true such as $\hat{q}_i(F)/q_i(F)$, and then compiling the frequency distribution of the resulting ratios across various realizations. For a specific nonexceedance probability F, it can be observed that 5% of the simulated values $\hat{q}_i(F)/q(F)$ fall below a value $L_{0.05}(F)$ whereas 5% exceeds a value $U_{0.05}(F)$. The 90% of $\hat{q}_i(F)/q(F)$ is contained within the interval

$$L_{0.05}(F) \leq \frac{\hat{q}_i(F)}{q(F)} \leq U_{0.05}(F), \tag{18}$$

or, alternatively,

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

The expression in (19) is referred to as 90% error bounds for $q(F)$.

5. Results and discussion

We have RH data on 24 sites that are listed in Table 1 along with their at-site characteristics. The `R` packages, `lmomRFA` (Hosking, 2024) and `lmomco` (Asquith, 2024) have been used for RFA of annual maximum RH. To observe the discordant sites, the $L-$moment ratios and discordance measure computed for all sites are shown in Table 3.

Table 3: $L-$moments ratios and discordance measure for all sites of study region

Site No.	Site name	l_1	t	t_3	t_4	t_5	D_i
1	Astore	77.1657	0.0535	-0.3010	0.0543	0.0311	1.3107
2	Badin	67.0279	0.0656	0.0651	0.0798	0.0424	0.7556
3	Bhawalnagar	50.1952	0.0851	0.0682	0.2201	0.0315	0.7430
4	Bhawalpur	46.6967	0.1001	0.1009	0.1800	0.0646	0.6076
5	Chilas	65.5607	0.0792	-0.0113	0.0277	-0.0239	1.5478
6	DI Khan	50.4605	0.0935	0.0749	0.1147	0.0680	0.6558
7	Faisalabad	54.1129	0.0880	0.1686	0.1905	-0.0257	0.4646
8	Gawadar	83.6712	0.0076	0.0359	0.0287	0.0314	4.1736
9	Islamabad	67.1086	0.0743	-0.1695	0.1143	-0.0060	0.6539
10	Jhelum	66.0521	0.0721	-0.1331	0.0702	0.0160	0.3656
11	Kalat	53.7205	0.0727	0.0373	0.1705	0.0652	0.1303
12	Karachi	77.8124	0.0190	0.1615	0.2090	0.1331	3.2117
13	Lahore	62.0733	0.0824	-0.0346	0.0695	-0.0007	0.5826
14	Malam Jabba	70.6443	0.0591	-0.1194	0.1388	-0.0393	0.5309
15	Mianwali	54.9002	0.0784	0.0711	0.1631	0.0918	0.0970
16	Multan	47.3871	0.1049	0.1475	0.1676	0.0673	0.9039
17	Muzaffarabad	74.5581	0.0491	-0.1433	0.1287	-0.0056	0.7488
18	Nawabshah	51.1114	0.0768	0.3499	0.2723	0.2166	1.9636
19	Peshawar	60.6045	0.0668	-0.0434	0.1782	0.0001	0.5584
20	Quettta	58.9452	0.0749	-0.0069	0.0718	0.0379	0.4977
21	Rahim Yar Khan	45.5124	0.0999	0.2299	0.2467	0.1335	1.1927
22	Saidu Sharif	70.6443	0.0591	-0.1194	0.1388	-0.0393	0.5309
23	Sialkot	67.5298	0.0716	-0.1290	0.0497	0.0073	0.5397
24	Skardu	77.1957	0.0512	-0.2922	0.0516	0.0627	1.2334

It is observed from Table 3 that Gawadar and Karachi are appeared to be discordant sites as their discordancy measures are greater than 3. The coastal cities of Karachi and Gawadar in Pakistan experience high humidity levels because of their proximity to the Arabian Sea. Karachi has different climatic patterns than other continental areas in Pakistan (Qureshi et al., 2017). Gawadar, another important port city along the Makran coast, also experiences high humidity, with an annual average relative humidity of approximately 70% (Ahmad et al., 2020). Karachi shows unexpectedly lower moisture content, whereas Gawadar likely experiences higher relative humidity, particularly in summer. These variations can be attributed to global warming, industrial activities, and regional climate patterns (Afreen et al., 2022;

Ahmad et al., 2020). Further analysis of these discordant sites, along with other sites, will be conducted to examine how the discordant sites behave in the determination of homogeneous regions.

Table 4 presents different cases for the formation of homogeneous regions. Considering the all 24 sites as a single region, we computed H_j defined in (8) with and without discordant sites in Case A and Case B, respectively, as shown in Table 4. it is clearly evident that the study area is definitely heterogeneous with and without discordant sites (Karachi and Gawadar) as $H_j > 2$ for $j = 1, 2, 3$. For formation of homogeneous regions and further analysis, we do not consider these two discordant sites. We use Wards' method of clustering (Everitt et al., 2001; Ward, 1963) to form homogeneous regions using at-site characteristics such as latitude, longitude, and elevation, as listed in Table 1 along with $L-$moments ratios λ_1, t, t_3 and t_4 in Table 3. The dendogram in Figure 2 clearly shows two main clusters that are considered as Case C with two regions named R1c and R2c. In Table 4, Case C shows that Region R1c is heterogeneous, whereas Region R2c is homogeneous. In Case D, we examine the three clusters depicted in Figure 2, designating them as Regions R1d, R2d, and R3d. As shown in Table 4, Case D identifies Region R1d as a heterogeneous region, whereas Regions R2d and R3d are characterized as homogeneous. In addition to statistical measures, Hosking and Wallis (1997) suggested for subjective partitioning to merge nearby stations into a region using geographical or at-site characteristics. For this purpose, we present a scatter plot of l_1 and t in Figure 3. In this figure, Case E is presented in Table 4 by relocating Sites 5 and 13 from Region R1d to Region R2d, subsequently designating these as Regions R1e, R2e and R3e. In Table 4, Regions R1e, R2e and R3e in Case E appear to be definitely homogeneous regions as values of H_1, H_2 and H_3 are less than 1. The regions in Case E (R1e, R2e and R3e) are the final homogeneous regions for quantile estimation.

Table 4: Formation of homogeneous regions using heterogeneity statistics (H)

	Region	Sites in a region	H_1	H_2	H_3
Case A	R1a	All 24 sites in study region with discordant sites	12.03*	8.15*	8.28*
Case B	R1b	22 sites study region without discordant sites	4.82*	7.88*	7.84*
Case C	R1c	1, 5, 9, 10, 13, 14, 17, 19, 22, 23, 24	2.00*	1.50*	1.42*
	R2c	2, 3, 4, 6, 7, 11, 15, 16, 18, 20, 21	0.63	0.74	0.29
Case D	R1d	1, 5, 9, 10, 13, 14, 17, 19, 22, 23, 24	2.00*	1.50*	1.42*
	R2d	2, 11, 20	-0.53	-0.75	-0.24
	R3d	3, 4, 6, 7, 15, 16, 18, 21	-0.64	0.22	-0.39
Case E	R1e	1, 9, 10, 14, 17, 19, 22, 23, 24	0.81	0.94	0.53
	R2e	2, 5, 11, 13, 20	-0.21	-0.66	-0.31
	R3e	3, 4, 6, 7, 15, 16, 18, 21	-0.64	0.22	-0.39

* $H_j < 1$ $(j = 1, 2, 3)$ indicates that the region is homogeneous.

To choose the best candidate probability distributions for the homogeneous regions, the regional average $L-$moment ratios of Regions R1e, R2e and R3e are drawn on the $L-$moment ratio diagram in Figure 4. We observe that $L-$moment ratio point (τ_3^R, τ_4^R) for Region R1e falls on GEV distribution curve, and is very close to PE3 and AEP4 distribution curves. For Region R2e, Figure 4 shows that the GEV and weibull (WEI) distributions are the best candidate distributions, whereas the GLO and PDQ3 distributions appear to be the best choices for Region R3e.

The goodness-of-fit Z^{DIST} statistic, defined in Subsection 4.3 is computed for each of candidate probability distributions in Table 5. The results in Table 5 demonstrate that the GEV and PE3 distributions for Region R1e, GEV and WEI distributions for Region R2e, and GLO and PDQ3 distributions for Region R3e are the most suitable distributions at a 10% level of significance, corroborating the findings presented in the $L-$moment ratio diagram in Figure 4. Table 6 presents the estimated parameters for the best candidate distributions of extreme RH in Regions R1e, R2e and R3e. A five–parameter Wakeby distribution has also been fitted using $L-$moments for all three regions in Table 6. The Wakeby distribution is employed due to its ability to imitate the forms of numerous asymmetrical distributions, including extreme value, log–normal, and Pearson type III distributions (Hosking & Wallis, 1997). This versatility

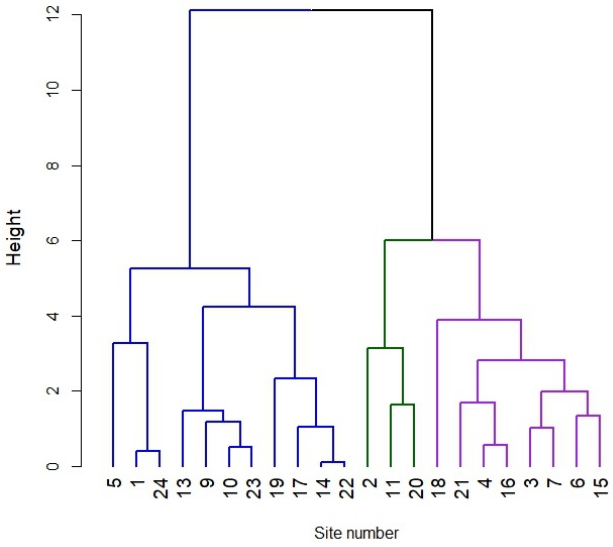

Figure 2: Dendogram of Wards's clustering method using geographical and statistical characteristics of sites serves as the primary justification for its utilization.

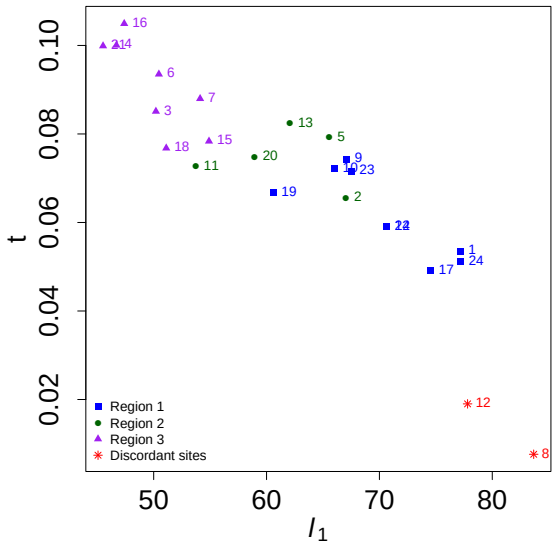

Figure 3: Scatter plot of $L-$moments l_1 and t for 24 sites of study area

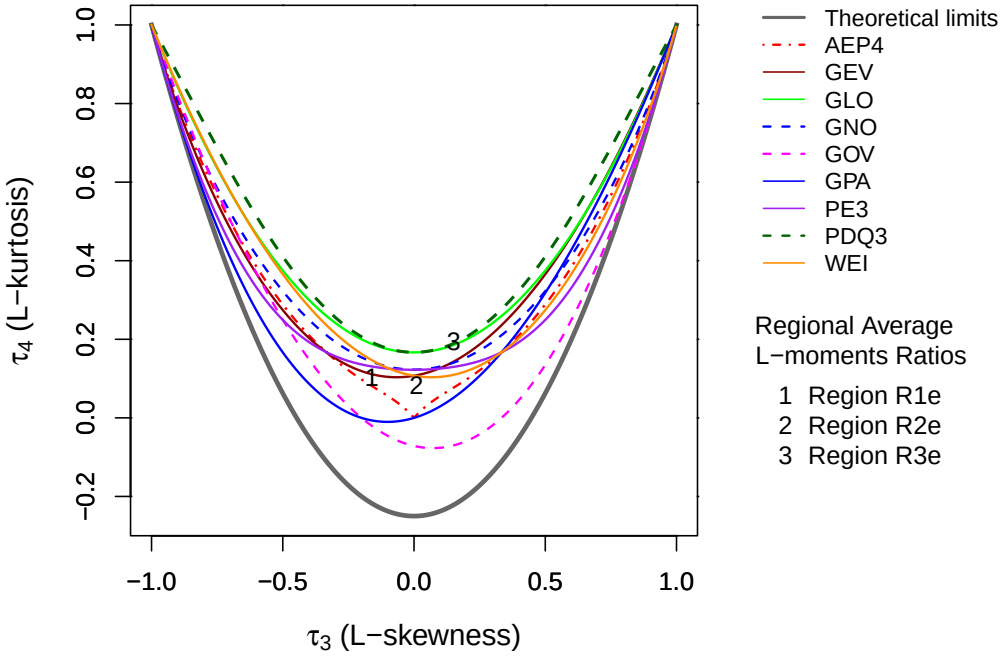

Figure 4: $L-$moments ratio diagram

The densities of the fitted regional distributions in Table 6 are plotted in Figure 5. It is clear that each region has different densities. The GEV, PE3, and Wakeby distributions are negatively skewed for Region R1e with Wakeby distribution being slightly higher in peak than others. The GEV and PE3 behave similarly for Region R1e, with very minor difference. For Region R2e, GEV and WEI are closer in shape and approximately symmetrical, but the Wakeby distribution slightly diverges in the left tail. For Region R3e, the GLO, PDQ3, and Wakeby distributions are similar in shape and show positive skewness, with very slight different in the peak and left tail for Wakeby distribution. Specifically, the distributions of the three regions exhibit distinct patterns, resulting in three clearly distinguishable regions.

Table 5: Goodness-of-fit Z^{DIST} statistic.

Region	Z^{AEP4}	Z^{GEV}	Z^{GLO}	Z^{GNO}	Z^{GOV}	Z^{GPA}	Z^{PE3}	Z^{PDQ3}	Z^{WEI}
R1e	-2.42	0.52*	4.70	2.24	-7.18	-5.86	1.58*	5.07	2.56
R2e	-1.83	1.25*	3.98	1.92	-7.43	-3.74	1.92	4.22	1.14*
R3e	-3.08	-2.49	-0.80*	-2.60	-11.08	-6.15	-3.03	-0.67*	-3.87

*Significant is set at 10% level of significance.

Table 6: Fitted regional probability distributions for three regions.

Region	Distribution	Parameters				
R1e	GEV	ξ (location)	α (scale)	κ (shape)		
		0.9782	0.1221	0.5970		
	PE3	μ (location)	σ (scale)	γ (shape)		
		1.0000	0.1130	-0.9788		
	Wakeby	ξ (location)	α (scale)	β (shape)	γ (scale)	δ (shape)
		0.7181	1.1646	4.5205	0.0996	-0.4035
R2e	GEV	ξ (location)	α (scale)	κ (shape)		
		0.9523	0.1311	0.2661		
	WEI	ξ (location)	β (scale)	δ (shape)		
		-0.6027	0.4428	3.3154		
	Wakeby	ξ (location)	α (scale)	β (shape)	γ (scale)	δ (shape)
		0.7438	0.6573	3.8237	0.1592	-0.3272
R3e	GLO	ξ (location)	α (scale)	κ (shape)		
		0.9776	0.0874	-0.1514		
	PDQ3	ξ (location)	α (scale)	κ (shape)		
		0.9754	0.1042	0.4305		
	Wakeby	ξ (location)	α (scale)	β (shape)	γ (scale)	δ (shape)
		0.7032	1.0651	5.4908	0.1227	0.0751

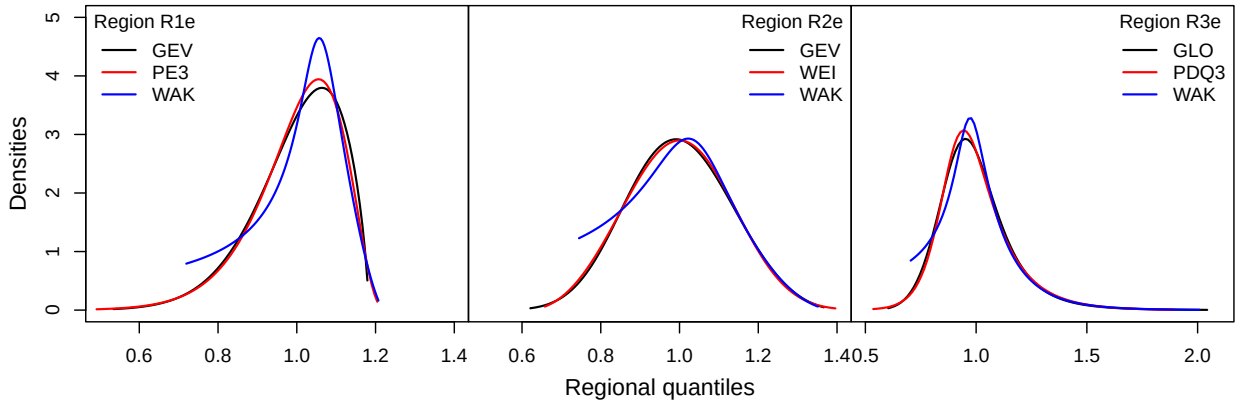

Figure 5: Density curves for regional probability distributions for Regions R1e, R2e and R3e.

The estimated regional quantiles associated with different nonexceedance probabilities (F) using best–fitted distributions, in conjunction with the Wakeby distribution, have been computed for Regions R1e, R2e and R3e in Tables 7 to 9. There is evidence of intersite dependence in all three regions. The average correlation between sites are 0.50, 0.71 and 0.41 in Regions R1e, R2e and R3e, respectively. Therefore, the simulations to compute the accuracy measures and error bounds for the regional quantiles defined in Section 4.4 have been performed using algorithm in Table 6.1 of

Hosking and Wallis (1997) with intersite dependence. The simulation results are presented in Tables 7 to 9.

Table 7: Regional growth curves and accuracy measures for Region R1e

Distribution	Statistics	Nonexceedance probability (F)									
		0.1	0.500	0.800	0.900	0.950	0.980	0.990	0.995	0.998	0.999
GEV	$\hat{q}(F)$	0.846	1.018	1.099	1.129	1.148	1.163	1.169	1.174	1.178	1.179
	$B^R(F)$	0.000	0.000	0.000	0.001	0.001	0.002	0.002	0.002	0.002	0.003
	$A^R(F)$	0.024	0.004	0.012	0.015	0.017	0.019	0.019	0.020	0.021	0.021
	$R^R(F)$	0.024	0.005	0.015	0.020	0.023	0.026	0.028	0.029	0.030	0.031
	LEB	0.808	1.010	1.074	1.097	1.110	1.119	1.124	1.126	1.127	1.128
	UEB	0.886	1.026	1.124	1.161	1.185	1.204	1.213	1.219	1.224	1.226
PE3	$\hat{q}(F)$	0.849	1.018	1.096	1.128	1.150	1.170	1.181	1.190	1.199	1.204
	$B^R(F)$	0.000	0.000	0.000	0.001	0.001	0.002	0.002	0.003	0.003	0.004
	$A^R(F)$	0.024	0.004	0.011	0.015	0.017	0.019	0.021	0.022	0.024	0.025
	$R^R(F)$	0.024	0.005	0.015	0.020	0.023	0.027	0.030	0.032	0.035	0.037
	LEB	0.811	1.009	1.071	1.095	1.111	1.125	1.131	1.136	1.140	1.142
	UEB	0.890	1.026	1.121	1.159	1.186	1.212	1.227	1.240	1.252	1.260
Wakeby	$\hat{q}(F)$	0.826	1.025	1.093	1.125	1.149	1.172	1.184	1.193	1.202	1.207
	$B^R(F)$	0.002	0.000	0.000	-0.001	-0.001	0.000	0.001	0.004	0.008	0.012
	$A^R(F)$	0.027	0.005	0.011	0.014	0.016	0.019	0.021	0.024	0.029	0.034
	$R^R(F)$	0.027	0.006	0.015	0.019	0.022	0.026	0.030	0.035	0.047	0.061
	LEB	0.783	1.015	1.069	1.096	1.115	1.130	1.134	1.132	1.117	1.101
	UEB	0.868	1.035	1.117	1.156	1.186	1.215	1.231	1.245	1.258	1.266

Table 8: Regional growth curves and accuracy measures for Region R2e

Distribution	Statistics	Nonexceedance probability (F)									
		0.100	0.500	0.800	0.900	0.950	0.980	0.990	0.995	0.998	0.999
GEV	$\hat{q}(F)$	0.830	0.998	1.114	1.174	1.221	1.271	1.300	1.325	1.351	1.367
	$B^R(F)$	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.002	0.003	0.004
	$A^R(F)$	0.025	0.005	0.012	0.018	0.023	0.029	0.033	0.038	0.043	0.047
	$R^R(F)$	0.025	0.006	0.017	0.025	0.034	0.045	0.054	0.063	0.074	0.082
	LEB	0.791	0.988	1.087	1.134	1.168	1.198	1.214	1.223	1.231	1.233
	UEB	0.872	1.008	1.142	1.216	1.277	1.345	1.389	1.427	1.469	1.495
WEI	$\hat{q}(F)$	0.827	0.999	1.114	1.172	1.219	1.271	1.305	1.335	1.371	1.396
	$B^R(F)$	0.001	0.000	0.000	0.000	0.001	0.001	0.002	0.002	0.002	0.003
	$A^R(F)$	0.025	0.005	0.012	0.018	0.023	0.028	0.032	0.035	0.039	0.041
	$R^R(F)$	0.025	0.006	0.016	0.025	0.034	0.044	0.051	0.058	0.066	0.072
	LEB	0.787	0.989	1.088	1.132	1.165	1.201	1.222	1.242	1.265	1.281
	UEB	0.868	1.010	1.141	1.214	1.274	1.342	1.388	1.429	1.478	1.513
Wakeby	$\hat{q}(F)$	0.817	1.002	1.115	1.173	1.220	1.267	1.294	1.316	1.339	1.352
	$B^R(F)$	0.002	-0.001	0.001	-0.001	-0.003	-0.002	0.001	0.006	0.017	0.029
	$A^R(F)$	0.026	0.006	0.013	0.018	0.023	0.028	0.033	0.039	0.052	0.066
	$R^R(F)$	0.026	0.008	0.018	0.026	0.033	0.043	0.052	0.065	0.093	0.128
	LEB	0.775	0.991	1.083	1.132	1.169	1.201	1.210	1.206	1.177	1.123
	UEB	0.858	1.016	1.144	1.218	1.278	1.343	1.380	1.414	1.450	1.472

To better interpret the simulation results in Tables 7 to 9, Figure 6 presents the estimated regional growth curves (RGCs), $\hat{q}(F)$, with simulation–based 90% error bounds (LEB, UEB) for the three regions. The RGCs for Region R1e are approximately equivalent for the GEV, PE3, and Wakeby distributions at all F, except for Wakeby distribution,

Table 9: Regional growth curves and accuracy measures for Region R3e

Distribution	Statistics	Nonexceedance probability (F)									
		0.100	0.500	0.800	0.900	0.950	0.980	0.990	0.995	0.998	0.999
GLO	$\hat{q}(F)$	0.814	0.978	1.113	1.206	1.302	1.441	1.558	1.687	1.879	2.043
	$B^R(F)$	-0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.002
	$A^R(F)$	0.021	0.006	0.009	0.016	0.025	0.037	0.047	0.059	0.076	0.090
	$R^R(F)$	0.021	0.008	0.012	0.024	0.040	0.067	0.093	0.126	0.182	0.236
	LEB	0.783	0.966	1.092	1.165	1.237	1.333	1.409	1.488	1.597	1.684
	UEB	0.850	0.991	1.131	1.244	1.367	1.549	1.708	1.890	2.175	2.425
PDQ3	$\hat{q}(F)$	0.819	0.975	1.113	1.212	1.313	1.447	1.550	1.653	1.789	1.893
	$B^R(F)$	-0.001	0.001	0.000	0.000	-0.001	-0.001	-0.001	-0.001	0.000	0.000
	$A^R(F)$	0.021	0.006	0.010	0.017	0.025	0.035	0.043	0.049	0.057	0.063
	$R^R(F)$	0.020	0.007	0.013	0.025	0.040	0.063	0.082	0.102	0.129	0.149
	LEB	0.787	0.963	1.092	1.173	1.249	1.347	1.419	1.492	1.588	1.661
	UEB	0.853	0.987	1.134	1.253	1.380	1.553	1.688	1.824	2.007	2.146
Wakeby	$\hat{q}(F)$	0.801	0.980	1.107	1.206	1.309	1.455	1.572	1.696	1.869	2.008
	$B^R(F)$	0.001	0.001	0.001	-0.001	-0.003	-0.005	-0.005	-0.004	0.000	0.005
	$A^R(F)$	0.023	0.007	0.010	0.017	0.026	0.039	0.052	0.067	0.093	0.115
	$R^R(F)$	0.022	0.009	0.014	0.026	0.041	0.070	0.101	0.143	0.222	0.302
	LEB	0.765	0.966	1.084	1.166	1.247	1.351	1.421	1.478	1.535	1.564
	UEB	0.837	0.994	1.129	1.250	1.383	1.580	1.751	1.945	2.240	2.494

which shows noticeable deviations at very small F. For Region R1e, the regional quantiles derived from the three distributions exhibit a gradual increase as F increases. However, these quantiles appear to stabilize. For Region R2e, the GEV, WEI and Wakeby distributions yield similar estimates of regional quantiles. These estimates show a subtle upward trend for increased F. Nevertheless, the Wakeby distribution yields slightly elevated regional quantiles for very small values of F. In Region R3e, the regional quantile estimates derived from the GLO, PDQ3, and Wakeby distributions show an upward trend for increasing F. However, the PDQ3 distribution yields marginally lower quantiles after $F = 0.980$. Moreover, for Region R3e, the error bounds for regional growth curves are wider than those for Regions R1e and R2e. The three regions exhibit distinct RGCs, each with its own unique growth rates and clear separation from the others.

To compare the selected distributions for the three regions, including the Wakeby distribution, we illustrate the bias ($B^R(F)$), absolute bias ($A^R(F)$), and RMSE ($R^R(F)$) in Figures 7 and 8, respectively. These measures are defined in Table 2 and computed in Tables 7–9. As depicted in Figure 7, the bias of the RGC derived from the examined distributions is roughly equivalent across the three regions, with the exception of the Wakeby distributions at higher F values. Notably, Wakeby distributions yield greater absolute bias in the estimated RGC when F is 0.99 or above. A similar trend is observed in the RMSE of the estimated RGC from Wakeby distributions across all three regions (Figure 8). For all three regions, the RMSE of estimated RGCs exhibits an identical pattern of decline and growth as the considered candidate distributions. In the case of Region R1e, when $F = 0.99$ or higher, the GEV and PE3 methods are more suitable for calculating regional quantiles. When estimating regional quantiles at or above F=0.99, Weibull and PDQ3 prove to be more suitable options for Regions R2e and R3e.

The index-flood method can be employed to calculate the at-site humidity quantiles for various return periods. Table 10 presents the estimated annual maximum RH for six sites, comprising two sites from each of the three regions. These estimations are made at nonexceedance probabilities of $F = 0.900, 0.980, 0.990, 0.998$ and 0.999, which correspond to return periods of 10, 50, 100, 500 and 1000 years, respectively. Analysis of annual maximum RH of sites in Region R1e (Islamabad and Peshawar) reveals an increasing trend of 0.5 for 100 and 500–year return periods. However, the annual RH for the 1000–year return period show minimal growth. This suggests that the extreme RH events in Region R1e are relatively stable for very long return periods, with only small increases expected between 100 and 1000–year events. For Region R2e, the at-site extreme RH quantiles exhibit an upward trend across various return periods, with a notably gradual increase observed between the 500–year and 1000–year events. In Region R3e, at-site quantiles of Faisalabad exhibit a sudden increase when comparing 10–year extreme RH events to 1000–year RH occurrences.

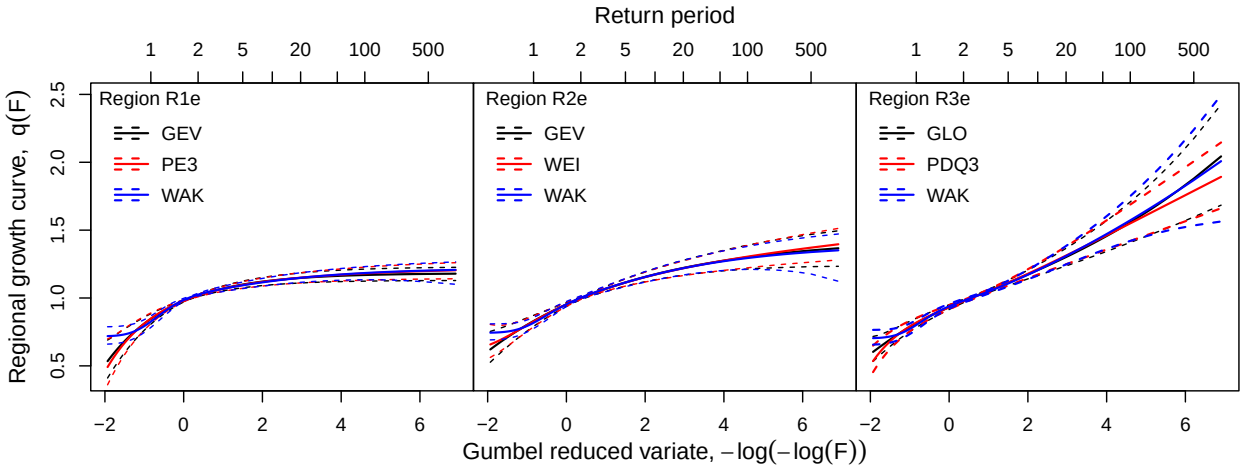

Figure 6: Estimated regional growth curves with their 90% error bounds for Regions R1e, R2e and R3e.

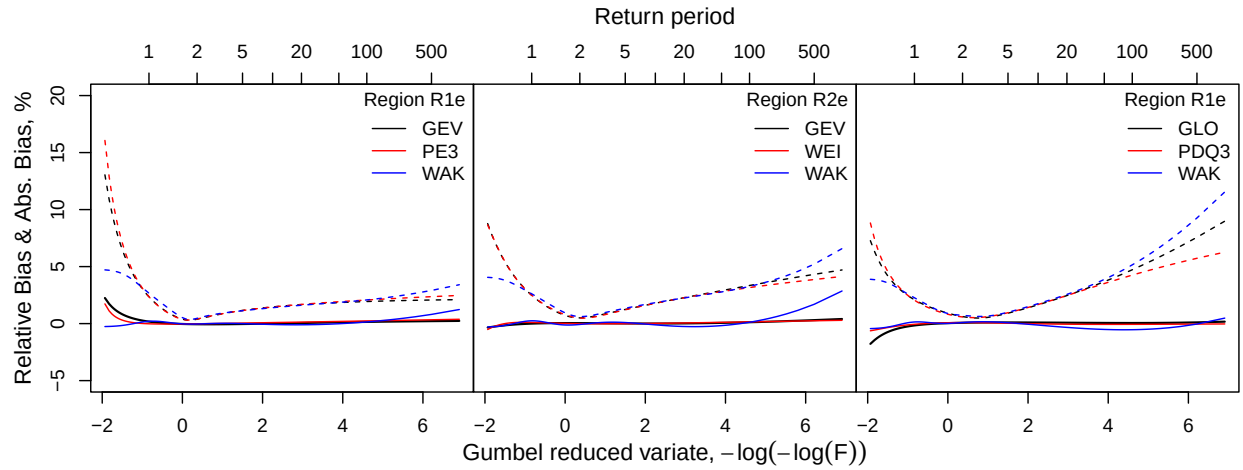

Figure 7: Relative bias (solid lines) and absolute bias (dashed lines) for regional growth curves for Regions R1e, R2e and R3e.

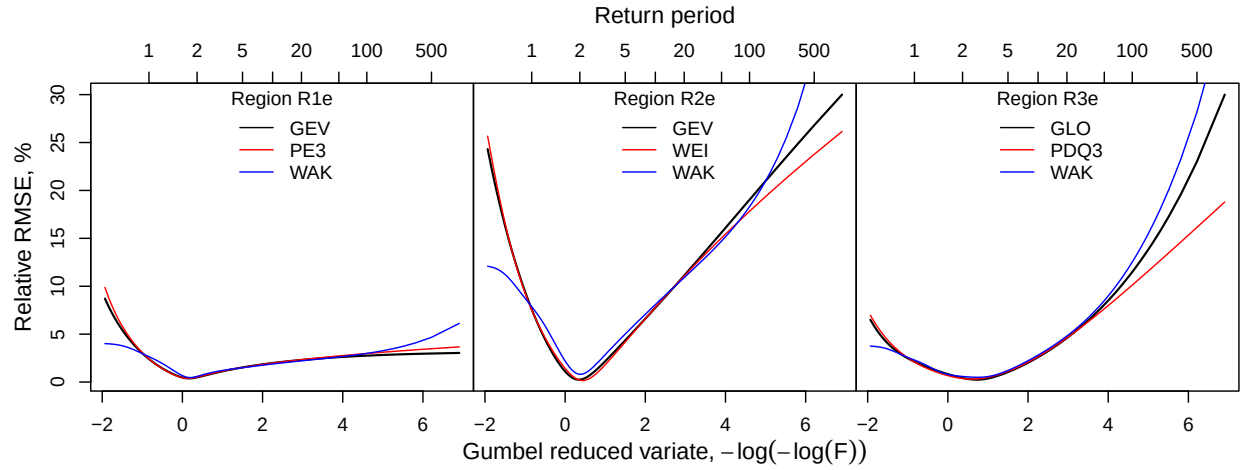

Figure 8: Relative RMSE for regional growth curves for Regions R1e, R2e and R3e.

These results offer insights into potential shifts in RH patterns, which could be indicative of broader climate trends. They also present a clearer picture of potential extreme RH events, allowing for better risk management, infrastructure and agricultural planning.

Table 10: At-site quantiles for two sites from each of three Regiona R1e, R2e and R3e.

Region	DIST	Site	Nonexceedance probability (F) & Return period (T)				
			F	0.900	0.980	0.990	0.998	0.999
			T	10	50	100	500	1000
R1e	GEV	Islamabad		75.783	78.027	78.483	79.027	79.141
		Peshawar		68.438	70.465	70.876	71.368	71.470
R2e	WEI	Lahore		72.760	78.888	80.979	85.101	86.647
		Quettta		69.093	74.912	76.898	80.813	82.280
R3e	PDQ3	Bhawalpur		56.614	67.5890	72.382	83.559	88.381
		Faisalabad		65.605	78.324	83.877	96.829	102.417

6. Conclusion

This research investigated the annual maximum humidity across 24 sites in Pakistan utilizing the $L-$moment based RFA developed by Hosking and Wallis (1997). During the data screening using a discordant measure, Gawadar and Karachi were identified as discordant sites. These sites also failed to contribute effectively to the formation of homogeneous regions and were thus excluded. The remaining 22 sites were grouped into three homogeneous regions: R1e, R2e and R3e. This grouping was achieved through Ward's clustering of site-specific characteristics and statistics, complemented by subjective partitioning based on the scatter plot of λ_1 and t. $L-$moment ratio diagram and Z^{DIST} statistic indicated that GEV and PE3 were the most suitable regional distributions for Region R1e. For Region R2e, GEV and Weibull distributions were identified as the best choices. For Region R3e, GLO and PDQ3 were determined to be the most appropriate distributions. In addition to the selected distributions for the three regions, the Wakeby distribution was also applied due to its ability to generalize various probability distributions. The RGCs calculated from three distinct distributions of associated regions showed no significant differences, except at extremely low and high nonexceedance probabilities. Nevertheless, regarding the relative bias, relative absolute bias and relative RMSE, the Wakeby distribution is not the optimal choice at F=0.99 and higher. Considering the three distinct distributions for each region, the GEV proved to be more suitable for Region R1e, while the Weibull distribution was more appropriate for Region R2e, and PDQ3 for Region R3e at $F = 0.99$ and above. The estimated regional growth curves and quantiles, researchers and policy-makers can develop more robust models to predict relative humidity extremes, which in turn is essential in diverse fields, including conservation, agriculture, building design, and materials science.

References

Adedayo, K. D. (2016). Statistical analysis of the effects of relative humidity and temperature on radio refractivity over Nigeria using satellite data. *African Journal of Environmental Science and Technology*, *13*(1), 221–229. https://doi.org/10.5897/AJEST2016.2095

Afreen, M., UCAK, İ., & Bagdatlı, M. C. (2022). The analysis of climate variability on aquaculture production in Karachi of Pakistan. *International Journal of Engineering Technologies and Management Research*, *9*(8), 16–23. https://doi.org/10.29121/ijetmr.v9.i8.2022.1210

Ahmad, I., Fawad, M., Akbar, M., Abbas, A., & Zafar, H. (2016). Regional frequency analysis of annual peak flows in Pakistan using linear combination of order statistics. *Polish Journal of Environmental Studies*, *25*, 2255–2264. https://doi.org/10.15244/pjoes/63782

Ahmad, I., Salman, M., Almanjahieand, I. M., Alshahrani, F., ul Rehman Khan, M. S., Fawad, M., & ul Haq, E. (2024). Regional frequency analysis of extreme wind in Pakistan using robust estimation methods. *Scientific Reports*, *14*, 25882. https://doi.org/10.1038/s41598-024-75248-w

Ahmad, M., Tariq, S., Alam, K., Anwar, S., & Ikram, M. (2020). Long-term variation in aerosol optical properties and their climatic implications over major cities of Pakistan. *Journal of Atmospheric and Solar-Terrestrial Physics*, *210*, 105419. https://doi.org/10.1016/j.jastp.2020.105419

Ajani, A., & van der Geest, K. (2021). Climate change in rural Pakistan: Evidence and experiences from a people-centered perspective. In *Sustainability science* (p. 1999, Vol. 16). Springer Science+Business Media. https://doi.org/10.1007/s11625-021-01036-4

Ali, G. (2018). Climate change and associated spatial heterogeneity of Pakistan: Empirical evidence using multidisciplinary approach. In *The science of the total environment* (p. 95, Vol. 634). Elsevier BV. https://doi.org/10.1016/j.scitotenv.2018.03.170

Ali, G., Sajjad, M., Kanwal, S., Xiao, T., Khalid, S., Shoaib, F., & Gul, H. N. (2021). Spatial–temporal characterization of rainfall in pakistan during the past half-century (1961–2020). *Scientific Reports*, *11*(1), 1–14. https://doi.org/10.1038/s41598-021-86412-x

Anghel, C. G., Ilinca, C., & Stanca, S. C. (2023). Two-parameter probability distributions: Methods, techniques and comparative analysis. *Water*, *15*(19), 3435. https://doi.org/10.3390/w15193435

Asquith, W. H. (2014). Parameter estimation for the 4-parameter asymmetric exponential power distribution by the method of L–moments using r. *Computational Statistics and Data Analysis*, *71*, 955–970. https://doi.org/10.1016/j.csda.2013.05.017

Asquith, W. H. (2024). *Lmomco—L–moments, censored L–moments, trimmed L–moments, L–comoments, and many distributions* [R package version 2.5.1]. https://CRAN.R-project.org/package=lmomco

Bangash, U. H., Irfan, M., & Alam, M. M. (2019). Statistical analysis of extremes of temperature, precipitation and humidity in Swat region. *Sindh University Rsearch Journal (Science Series)*, *50*, 495–506. https://doi.org/10.26692/sujo/2018.12.0080

Burn, D. H. (1989). Cluster analysis as applied to regional flood frequency. *Journal of Water Resources Planning and Management*, *115*, 567. https://doi.org/0.1061/(ASCE)0733-9496(1989)115:5(567)

Clausen, B., & Pearson, C. P. (1995). Regional frequency analysis of annual maximum streamflow drought. *Journal of Hydrology*, *173*(1–4), 111–130. https://doi.org/10.1016/0022-1694(95)02713-y

Dalrymple, T., & Dawdy, D. (1960). *Flood-frequency analyses* (Vol. 1543-A). U.S. Geological Survey.

Devadoss, P. S. M., Agamuthu, P., Bhatti, M. S., Chenayah, S., & Hamid, F. S. (2021). Strategies for reducing greenhouse gas emissions from municipal solid waste management in Pakistan. In *Waste management & research the journal for a sustainable circular economy* (p. 914, Vol. 39). SAGE Publishing. https://doi.org/10.1177/0734242x20983927

Everitt, B. S., Landau, S., & Leese, M. (2001). *Cluster analysis* (4th). Oxford University Press, Inc., New York; Arnold, London.

Gaál, L., Faško, P., Lapin, M., & Szolgay, J. (2009). Hybrid approach to delineation of homogeneous regions for regional precipitation frequency analysis. *Journal of Hydrology and Hydromechanics*, *57*(4), 226–249. https://doi.org/10.2478/v10098-009-0021-1

Guru, N. (2022). Implication of partial duration series on regional flood frequency analysis. *International Journal of River Basin Management*, *22*(2), 167–186. https://doi.org/10.1080/15715124.2022.2114486

Hassan, M., Noreen, Z., & Ahmed, R. (2021). Regional frequency analysis of annual daily rainfall maxima in skåne, sweden. *International Journal of Climatology*, *41*, 4307–4320. https://doi.org/10.1002/joc.7074

Hosking, J. R. M. (1990). L–moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society, Series B*, *52*, 105–124.

Hosking, J. R. M. (2007). Distributions with maximum entropy subject to constraints on their L–moments or expected order statistics. *Journal of Statistical Planning and Inference*, *137*(9), 2870–2891. https://doi.org/10.1016/j.jspi.2006.10.010

Hosking, J. R. M. (2024). *Regional frequency analysis using L–moments* [R package, version 3.8]. https://CRAN.R-project.org/package=lmomRFA

Hosking, J. R. M., & Wallis, J. R. (1997). *Regional frequency analysis: An approach based on L–moments*. Cambridge University Press.

Hussain, Z., & Pasha, G. R. (2009). Regional flood frequency analysis of the seven sites of Punjab, Pakistan, using l-moments. *Water Resources Management*, *23*(10), 1917–1933.

Khan, T., Yuan, N., Hilal, M. G., Ullah, R., Alam, K., Naveed, M., & Wang, Y. (2020). Frequency analysis of annual maximum temperature with linear moments for parameter estimation and determination of best fit probability distribution for regions in Pakistan. *Journal of Earth Sciences & Environmental Studies*, *5*, 1–8.

Kumar, D., Gautam, Y. P., Kumar, V., Kumar, S., Saradhi, I. V., & Kumar, A. (2021). Statistical analysis of extreme value of meteorological elements observed for the last 31 years (1989–2019) at Narora site. *Radiation Protection and Environment*, *44*(3), 123. https://doi.org/10.4103/rpe.rpe_27_21

Mahbod, M., Mahmoodi-Eshkaftaki, M., Rafiee, M. R., & Ebrahimiat, A. (2024a). An innovative regional frequency analysis approach for robust extreme precipitation assessment in data-rich and climatically diverse regions. *Research Square Platform LLC*. https://doi.org/10.21203/rs.3.rs-4356974/v1

Mahbod, M., Mahmoodi-Eshkaftaki, M., Rafiee, M. R., & Ebrahimiat, A. (2024b). An innovative regional frequency analysis approach for robust extreme precipitation assessment in data-rich and climatically diverse regions. *Research Square Platform LLC*. https://doi.org/10.21203/rs.3.rs-4356974/v1

Martinez-Villalobos, C., & Neelin, J. D. (2023). Regionally high risk increase for precipitation extreme events under global warming. *Scientific Reports*, *13*, 5597. https://doi.org/10.1038/s41598-023-32372-3

Mosaffaie, J. (2015). Comparison of two methods of regional flood frequency analysis by using L–moments. *Water Resources*, *42*(3), 313–321. https://doi.org/10.1134/s0097807815030112

Parvizi, S., Eslamian, S., Gheysari, M., et al. (2022). Regional frequency analysis of drought severity and duration in Karkheh River Basin, Iran using univariate L–moments method. *Environmental Monitoring and Assessment*, *194*, 336. https://doi.org/10.1007/s10661-022-09977-8

Pilon, P., & Adamowski, K. (1992). The value of regional information to flood frequency analysis using the method of L–moments. In *Canadian journal of civil engineering* (p. 137, Vol. 19). NRC Research Press. https://doi.org/10.1139/l92-014

Qureshi, J., Masood, A., Hussain, M., Shafi, W., & Mahmood, S. (2017). Variation in meteorological parameters over Pakistan during April 2014. *Journal of Basic & Applied Sciences*, *13*, 185–192. https://doi.org/10.6000/1927-5129.2017.13.32

Riggs, H. C. (1973). *Regional analyses of streamflow characteristics* (Report No. 04-B3). U.S. Geological Survey. https://doi.org/10.3133/twri04B3

Seckin, N., Haktanir, T., & Yurtal, R. (2011). Flood frequency analysis of Turkey using L–moments method. *Hydrological Processes*. https://doi.org/10.1002/hyp.8077

Shahzadi, A., Akhter, A. S., & betul Saf. (2013). Regional frequency analysis of annual maximum rainfall in monsoon region of Pakistan using L–moments. *Pakistan Journal of Statistics and Operation Research*, *9*(1), 111–136. https://doi.org/10.18187/pjsor.v9i1.461

Solaimani, K., Bararkhanpour Ahmadi, S., & Shokrian, F. (2024). The spatiotemporal trend changes of extreme temperature-humidity variables and their impact on climatic comfort changes. *Ecological Indicators*, *158*, 111629. https://doi.org/https://doi.org/10.1016/j.ecolind.2024.111629

Souaissi, Z., Ouarda, T. B., & St-Hilaire, A. (2023). Regional thermal index model for river temperature frequency analysis in ungauged basins. *Environmental Modelling & Software*, *164*, 105709. https://doi.org/https://doi.org/10.1016/j.envsoft.2023.105709

Vivekanandan, N. (2015). Flood frequency analysis using method of moments and L–moments of probability distributions. In *Cogent engineering* (p. 1018704, Vol. 2). Cogent OA. https://doi.org/10.1080/23311916.2015.1018704

Ward, J., J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, *58*, 236–244.

Yao, A. Y. M. (1974). A statistical model for the surface relative humidity. *Journal of Applied Meteorology and Climatology*, *13*(1), 17–21. https://doi.org/10.1175/1520-0450(1974)013⟨0017:ASMFTS⟩2.0.CO;2