Desalination and Water Treatment
www.deswater.com
doi: 10.5004/dwt.2021.26904

215 (2021) 427–451
March

Identification of the most suitable probability distribution models for monthly and annual rainfall series in Güzelyurt Region, Northern Cyprus

Youssef Kassem[a,b,*], Hüseyin Gökçekuş[b], Rifat Gökçekuş[b]

[a]*Department of Mechanical Engineering, Engineering Faculty, Near East University, 99138 Nicosia (via Mersin 10, Turkey), Cyprus, Tel. +90 (392) 2236464; emails: yousseuf.kassem@neu.edu.tr/yousef.kassem1986@hotmail.com*
[b]*Department of Civil Engineering, Civil and Environmental Engineering Faculty, Near East University, 99138 Nicosia (via Mersin 10, Turkey), Cyprus, emails: huseyin.gokcekus@neu.edu.tr (H. Gökçekuş), rifatgokcekus@hotmail.com (R. Gökçekuş)*

Received 11 May 2020; Accepted 7 December 2020

ABSTRACT

Precise information of rainfall probability distribution is truly significant for many hydrological studies such as extreme flood analysis, drought investigations, reservoir volume studies, and time-series modeling. The objective of the present study is to identify the suitable probability functions for estimating the rainfall distribution in Güzelyurt Region, Northern Cyprus. Moreover, the aim of the study is to evaluate whether proposed distribution models have more satisfying performance than the commonly used models in previous scientific studies. Based on 33-y rainfall data collected from Meteorology Department located in Lefkoşa, this study presents a statistical analysis of the average monthly, maximum monthly, and annual rainfall characteristics in the selected region with the assistance of 37 distribution models. Goodness-of-fit tests including Kolmogorov–Smirnov (K-S) test, Anderson–Darling (A-D) test, and Chi-squared (C-s) test are used to select the best-fit probability distribution model. Easyfit software was used in this study. The results indicate that Beta, Log-Pearson 3, and exponential (2P) distributions have the lowest value of K-S, A-D, and C-s, respectively, which are considered as the best distribution functions to study the average rainfall characteristics. Furthermore, among the best score results, Burr, Wakeby, and Nakagami distributions are giving the best-fit of actual data of total rainfall based on Goodness-of-fit tests. Additionally, the best fit probability distribution model for maximum monthly rainfall at the selected region was Nakagami, Wakeby, and Cauchy distributions based on Goodness-of-fit tests. The choices of the selected distribution model may be used for forecasting hydrologic events and defining the policies regarding water resource management and a source of data for flood hazard mitigation. This study provides a significant contribution to the current understanding of predicting hydrological events for various purposes.

Keywords: Average rainfall; Distribution functions; Güzelyurt; Maximum monthly rainfall; Annual rainfall

1. Introduction

The issue of water scarcity has become a global challenge due to the rapid rise in population growth around the world, and the increasing pressure on the consumption of limited water resources. Additionally, the scarcity and pollution of groundwater are still one of the most important topics nowadays. Several scientific researchers have analyzed the impact of pollution sources on the quality of groundwater [1,2]. Due to the importance of groundwater, it has become the most important source of water supply for domestic, industrial, and agricultural sectors of many places including Northern Cyprus. Rainfall is considered as a vital source of water in Northern Cyprus [3]. According

* Corresponding author.

to Song et al. [4] and Kundu et al. [5], the availability of water depends on population growth, energy demand, and climate change.

In general, high temperatures increase the rate of evaporation of the water in the atmosphere, which leads to an increase in the air's ability to carry water [6]. This causes early and short seasons to run and an increase in dry seasons [6]. The further evaporation reduces the levels of moisture in the soil, which in turn increases the frequency of droughts in the region, and increases the likelihood of desertification [7]. Besides, a decrease in the percentage of moisture in the soil as well as in infiltration rates leads to reduction in the feeding rate in the groundwater [7]. Moreover, air temperature and rainfall are considered the major factors that affect human activities such as agriculture, which has influenced the economy of the country [8,9].

Cyprus is the third largest island in the Mediterranean Sea with an area of 9,251 km^2. It has a temperate climate (Mediterranean climate) [10]. Two-thirds of the average annual rainfall of 500 mm falls from December to February. The island suffers from drought periodically [11]. The water resources on the island are very limited and the main source of water on the island is rainfall. Generally, in Cyprus, more than two-thirds of the rainfall occur between October and April.

1.1. Literature review related to water resource in Northern Cyprus mainly the Güzelyurt region

With an increasing world population, it is getting more difficult to utilize the uncontaminated surface water and groundwater resources. In parallel with the technological development, more complex chemical and biological pollutants were released into the environment from the industry without appropriate physical, chemical, and biological treatment. Therefore, these untreated water pollutants contaminated the surface water and groundwater resource to a great extent. As a result, the quantity of water suitable for drinking and irrigation purposes decreased considerably.

The climate of Cyprus is the typical Mediterranean with very hot dry summers and cooler winters [13]. In winter, the mean daytime temperatures vary from 12°C to 15°C, while the mean maximum temperature in summer reaches 40°C [14]. The wet season extends from November to March, with most (approximately 60%) of the rain falling between December and February [14]. Generally, most of the rainfall occurs from October to April in Cyprus.

According to Agboola and Egelioglu [12], and Wright [15], there are no rivers but the area is well watered by gushing perennial springs in Cyprus particularly in Northern Cyprus. There are 38 streams in the northern part of Cyprus, 10 of which originated from the southern part of the island around the Troodos Mountains and relatively rich in water as they flow; however, dam construction by the Southern Administration prevented the flow from the streams to the northern part of Cyprus [16].

The water resources in the country are classified into groundwater, surface water (dam), and Turkey-North Cyprus water pipeline project. The sources of surface water resources in Northern Cyprus are from four main river basins [17]. Surface water resources in Northern Cyprus are mainly sourced from four main river basins [17]:

- Western Mesarya Plain, including Lefke, with a catchment area of about 640 km^2.
- Central and Eastern Mesarya Plain, including Gazimağusa, with a catchment area of about 1,520 km^2.
- North shore and Girne (Beşparmak) Mountains, with a catchment area of about 460 km^2.
- Karpaz Peninsula, with a catchment area of about 680 km^2.

The streams in Northern Cyprus originate either from the Troodos Mountains in the central part of the island or from the Beşparmak Mountains in the north. According to Türker and Hansen [17], the average annual carried from the streams is estimated to be about 108 mm^3 from the Troodos Mountains and 80 mm^3 from Beşparmak Mountains. Additionally, the total average annual surface runoff is estimated at 188 mm^3 in the Northern part of Cyprus.

According to Türker and Hansen [17], 30 large earth dams are constructed in the North part of Cyprus and the total capacity storage of these dams is around 34.4 million m^3 of which only about 25.6 million m^3 is operational by now due to siltation. The estimated annual volume provided for irrigation and groundwater recharge is approximately 8.4 million m^3, which means about 33% of the operational storage capacity is used for irrigation purposes. Furthermore, based on theoretical estimation for surface water stored volume for groundwater purposes, it is found that around 14% of the operational storage capacity is used for groundwater recharge. Moreover, it is estimated that about 30%–40% of water stored is lost due to evaporation [17]. Table S1 shows some selected dams in Northern Cyprus with their capacity and year of construction.

The Güzelyurt groundwater basin is located within the western part of Northern Cyprus. Since 1957, increasing rates of pump-age have caused a progressive decline in the groundwater levels, locally reaching 45–50 m below mean sea level [18,19]. Limited natural recharge and excessive withdrawals from approximately 250 active municipal and irrigation wells have produced not only a considerable reduction in the aquifer storage but also degradation of groundwater quality due to saltwater intrusion and bedrock contamination [18]. The aquifer is the main source of potable water. The total basin area of the aquifer is around 460 km^2 of which 1/3 of this area is under the control of the Greek Cypriot Community and 2/3 is under the control of Turkish Cypriot Community Authorities. Most of the aquifers in the northern part of Cyprus are unconfined (phreatic) made up of the river or coastal alluvial deposits, mainly silts, sands, and gravels. The main aquifers in the north part of Cyprus are Girne mountain aquifer – which is located in Beşparmak Mountains close to the north coast, Güzelyurt aquifer – located in western Mesarya, and Gazimağusa aquifer – located in Southeastern Mesarya. The aquifers are mainly being recharged by rainfall and river flows (in a very limited period of the year) and are more or less all showing trends of depletion due to reduced recharging, frequent droughts, and increased abstraction mainly by farmers in their effort to increase their production level.

The aquifers are characterized into 13 groups within eight hydrological regions, which have various capacities of water storage as shown in Table S2. The Güzelyurt aquifer, which is the largest coastal aquifer in the northwestern of the island, provides water not only for irrigation requirements in the region but also for the municipal needs of Lefkoşa and Gazimağusa cities. According to Gökçekus and Doyuran [19], the capacity of the Güzelyurt aquifer is found to be 920 million cubic meters and recent studies demonstrate that the aquifer has is depleted and the average groundwater level reaches 70 m below the mean sea level in some local areas. The second important aquifer is the Mount Aquifer, which runs across the northern coast of the island with a thin strip of 1.5 km wide. The surface of this underground reservoir is about 40 km^2, with an average annual renewal of 10.5 million cubic meters. At the watershed of the Güzelyurt groundwater basin various lithological units of the Troodos massif (middle-Upper Cretaceous), Lapathos group (Oligocene-Lower Miocene), Dhali Group (Middle-Upper Miocene), and Mesaoria Group (Upper Miocene-Upper Pliocene) constitute the bedrock. The basin itself comprises flanglomerates (Plestocene) and Holocene deposits. The Pre-tertiary Troodos Massif rocks are the oldest units exposed within the watersheds of the Güzelyurt groundwater basin. In general, the Troodos Massif is a huge igneous body, which is exposed in the central part of Cyprus. It is made up of Troodos Plutonic Series, Sheeted Dyke Complex, and Pillow Lava Series.

According to Türker and Hansen [17], about 100% of water demand provided for public needs is from groundwater in the Island. However, the availability of the groundwater resources is by now rather limited due to over-abstraction of groundwater and due to limited natural recharge from rainfall, which has resulted in the depletion of available freshwater within the aquifers. Due to the over-abstraction, seawater intrusion has occurred in several coastal areas and has in some cases reached an alarming stage (in some areas the chloride concentration has in some cases reached as high as 7,000 ppm) [20–22].

1.2. Literature review related to rainfall distribution analysis

No theoretical distribution can be considered that it can characterize exclusively the annual rainfall profile [23]. Thus, the analysis of rainfall/precipitation data mainly depends on its distribution type. Many researchers have studied the precipitation (rainfall) characteristics using different distribution functions in different parts of the world [24–55] (Table S3). For instance, Yuan et al. [47] analyzed the annual maximum hourly rainfall characteristics for 15 locations in Japan using the Expanded Automated Meteorological Data Acquisition System (EA) weather data of 20 y (1981–2000). The results showed that the Log-Pearson type 3 distributions provided the best fit to the actual data for most locations in Japan. Parchure and Gedam [52] determined the best-fit probability distribution for extreme storms using two-parameter and three-parameter distribution functions. The results indicated that a three-parameter generalized extreme value was considered as the best distribution function to study the extreme storm series in the Mumbai region, India. Sharma and Singh [34] utilized four distribution functions to analyze the characteristics of rainfall in the Beirut region, Lebanon, using daily rainfall of 25 y (1991–2015). The results showed that Gumbel Maximum and Logistic distributions were able to provide the best fit to the actual data for the selected region.

Moreover, according to the authors' review, only three scientific studies [23,54,55], statistically analyzed the time series of rainfall in different locations in Cyprus. Michaelides et al. [23] studied the characteristics of the annual rainfall frequency distribution in Cyprus using a Gamma distribution function. Stamatatou et al. [54] used various distribution functions (Generalized extreme values, Gumbel, and Generalized Pareto Distribution, Gamma, Exponential, and Log-normal) to analyze the characteristics of annual maximum rainfall depth and storm duration in Limassol station in Cyprus. The results indicated that generalized extreme value distribution was selected for modeling annual maximum rainfall depth and storm duration in the selected station. Zaifoglu et al. [55] utilized conventional cluster analysis with the time series clustering approaches to perform regional frequency analysis of annual maximum daily precipitation/rainfall using L-moments in Northern Cyprus. The results found that the Pearson Type III, generalized logistic, and generalized normal distributions were chosen as the best fit for different subregions in Northern Cyprus.

Based on previous scientific studies, it can be concluded that:

- The most commonly used frequency distributions are Normal, Log-Normal, Log-Pearson Type 3, Exponential, Gumbel maximum, Gumbel minimum, Generalized extreme value, Weibull, Generalized Pareto, Generalized logistic, Gamma, Generalized Gamma, and Logistic.
- Few studies have analyzed the average, maximum, and annual hourly/daily/monthly precipitation/rainfall data in Cyprus.

1.3. Scope of the study

According to the authors' review, there is no study in Cyprus about identifying the probability distribution functions that are best suited for the estimation of average monthly, maximum monthly, and annual rainfall in the Güzelyurt region, Northern Cyprus. The existing very few ones have analyzed the rainfall characteristics in different regions using various distribution models including Gamma, Generalized extreme values, Gumbel, and Generalized Pareto, Exponential, Pearson Type III, generalized logistic, and Log-normal) [23,54,55]. To the best of our knowledge, there are no detailed studies in the selected region regarding analyzing the monthly and annual rainfall characteristics using 37 distribution functions. Therefore, the main aim of the study is to identify the suitable probability functions for estimating the rainfall distribution in the selected region. Moreover, the objective is to evaluate whether proposed distribution models (Beta, Burr, Burr (4P), Cauchy, Dagum (4P), Erlang, Erlang (3P), Exponential (2P), Gamma (3P), Generalized Gamma (4P), Generalized Logistic, Inverse Gaussian, Inverse Gaussian (3P), Log-Gamma, Log-Logistic, Log-Logistic (3P), Log-normal (3P), Nakagami, Pareto 2, Rayleigh, Rayleigh (2P), Wakeby, Weibull (3P)) have more satisfying performance than the commonly used models (Log-Pearson 3, Generalized extreme values, Generalized Pareto, Generalized Logistic, Weibull, Gumbel, Log-normal, Generalized Gamma, Exponential, Normal, Logistic and Gamma) in previous scientific studies. Furthermore, the best models for the selected region are compared with other regions in the world that have the same best distribution model to find similarities in the statistical characteristics of rainfall. Fig. 1 illustrates the analysis procedure of this study.

2. Materials and methods

2.1. Data and study area

Güzelyurt is located in the northwestern part of Cyprus (Fig. 2) at a latitude of 35° 12′ 3.528″N, the longitude of 32° 59′ 26.808″E. The amount of rainfall is ranged from 300 mm in the plains to 1,200 mm on the Troodos range located entirely in the southwest part of the island, with part of its drainage area flanking into the northern part of the island, specifically replenishing the groundwater resources of the Güzelyurt aquifer, which is one of the main water supplying aquifers. Therefore, it is important to evaluate the extreme rainfall for the specific probability of occurrence.

2.2. Probability distributions

The choice of the probability distribution models is important to select the best-fit probability distribution for a specific location. In this section, Table 1 presents the selected distribution models, which were used to analyze the characteristic of rainfall in the Güzelyurt region, Northern Cyprus. The method of maximum-likelihood was utilized to estimate the parameters of the selected models. The Easyfit software with a CPU-Intel Xeon E5-16XX, 8 core, 64GB ram, and 64-bit operating system were used to obtain the parameters of the distribution functions.

2.3. Goodness-of-fit test

To check the validity of the specified probability distribution model, goodness-of-fit test statistics are utilized.

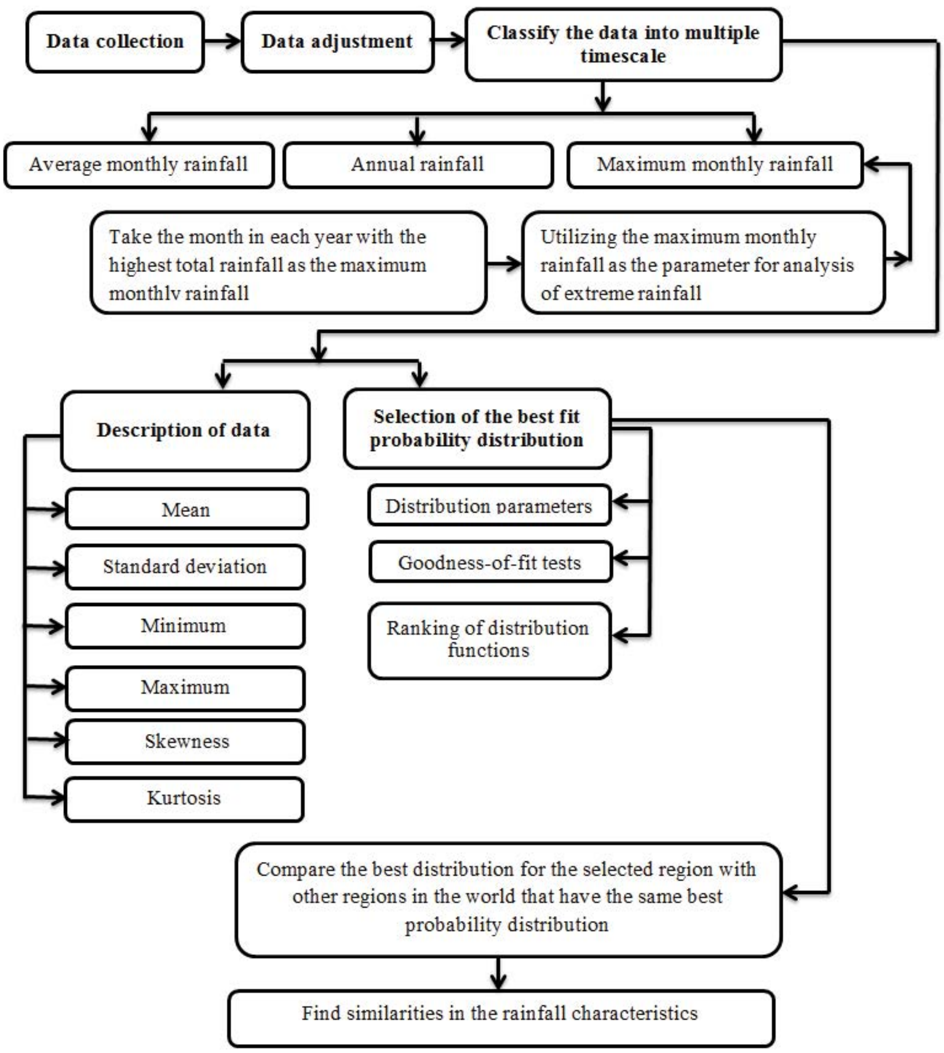

Fig. 1. Flowchart of the analysis procedure of the present study.

Table 1
Probability density and cumulative distribution of used distribution functions

Distribution function	Probability density function	Cumulative distribution function
Beta	$f(R)=\frac{1}{B(\alpha_1,\alpha_2)}\frac{(R-a)^{\alpha_1-1}(b-R)^{\alpha_2-1}}{(b-a)^{\alpha_1+\alpha_2-1}}$	$F(R)=I_z(\alpha_1,\alpha_2)$
Four-Parameter Burr	$f(R)=\frac{\alpha k\left(\frac{R-\gamma}{\beta}\right)^{\alpha-1}}{\beta\left(1+\left(\frac{R-\gamma}{\beta}\right)^{\alpha}\right)^{k+1}}$	$F(R)=1-\left(1+\left(\frac{R-\gamma}{\beta}\right)^{\alpha}\right)^{-k}$
Three-Parameter Burr	$f(R)=\frac{\alpha k\left(\frac{R}{\beta}\right)^{\alpha-1}}{\beta\left(1+\left(\frac{R}{\beta}\right)^{\alpha}\right)^{k+1}}$	$F(R)=1-\left(1+\left(\frac{R}{\beta}\right)^{\alpha}\right)^{-k}$
Cauchy	$f(R)=\left(\pi\sigma\left(1+\left(\frac{R-\mu}{\sigma}\right)^{2}\right)\right)^{-1}$	$F(R)=\frac{1}{\pi}\arctan\left(\frac{R-\mu}{\sigma}\right)+0.5$
Four-Parameter Dagum	$f(R)=\frac{\alpha k\left(\frac{R-\gamma}{\beta}\right)^{\alpha k-1}}{\beta\left(1+\left(\frac{R-\gamma}{\beta}\right)^{\alpha}\right)^{k+1}}$	$F(R)=1-\left(1+\left(\frac{R-\gamma}{\beta}\right)^{-\alpha}\right)^{-k}$
Three-Parameter Dagum	$f(R)=\frac{\alpha k\left(\frac{R}{\beta}\right)^{\alpha k-1}}{\beta\left(1+\left(\frac{R}{\beta}\right)^{\alpha}\right)^{k+1}}$	$F(R)=1-\left(1+\left(\frac{R}{\beta}\right)^{-\alpha}\right)^{-k}$
Three-Parameter Erlang	$f(R)=\frac{(R-\gamma)^{m-1}}{\beta^{m}\Gamma(m)}\exp\left(-\frac{R-\gamma}{\beta}\right)$	$F(R)=\frac{\Gamma_{(R-\gamma)/\beta}(m)}{\Gamma(m)}$
Two-Parameter Erlang	$f(R)=\frac{(R-\gamma)^{m-1}}{\beta^{m}\Gamma(m)}\exp\left(-\frac{R}{\beta}\right)$	$F(R)=\frac{\Gamma_{(R)/\beta}(m)}{\Gamma(m)}$
Two-Parameter Exponential	$f(R)=\lambda\exp(-\lambda(R-\gamma))$	$F(R)=1-\exp(-\lambda(R-\gamma))$
One-Parameter Exponential	$f(R)=\lambda\exp(-\lambda R)$	$F(R)=1-\exp(-\lambda R)$
Three-Parameter Gamma	$f(R)=\frac{(R-\gamma)^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)}\exp\left(-\left(\frac{R-\gamma}{\beta}\right)\right)$	$F(R)=\frac{\Gamma_{(R-\gamma)/\beta}(\alpha)}{\Gamma(\alpha)}$
Two-Parameter Gamma	$f(R)=\frac{R^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)}\exp\left(-\left(\frac{R}{\beta}\right)\right)$	$F(R)=\frac{\Gamma_{R/\beta}(\alpha)}{\Gamma(\alpha)}$

Table 1 Continued

Distribution function	Probability density function	Cumulative distribution function
Generalized Extreme Value	$f(R)=\begin{cases}\frac{1}{\sigma}\exp\left(-\left(1+k\frac{R-\mu}{\sigma}\right)^{-1/k}\right)\left(1+k\frac{R-\mu}{\sigma}\right)^{-1-1/k} & k\neq 0\\ \frac{1}{\sigma}\exp\left(-\frac{R-\mu}{\sigma}-\exp\left(-\frac{R-\mu}{\sigma}\right)\right) & k=0\end{cases}$	$F(R)=\begin{cases}\exp\left(-\left(1+k\frac{R-\mu}{\sigma}\right)^{-1/k}\right) & k\neq 0\\ \exp\left(-\exp\left(-\frac{R-\mu}{\sigma}\right)\right) & k=0\end{cases}$
Four-Parameter Generalized Gamma	$f(R)=\frac{k(R-\gamma)^{k\alpha-1}}{\beta^{k\alpha}\Gamma(\alpha)}\exp\left(-\left(\frac{R-\gamma}{\beta}\right)^{k}\right)$	$F(R)=\frac{\Gamma_{((R-\gamma)/\beta)^k}(\alpha)}{\Gamma(\alpha)}$
Three-Parameter Generalized Gamma	$f(R)=\frac{k(R)^{k\alpha-1}}{\beta^{k\alpha}\Gamma(\alpha)}\exp\left(-\left(\frac{R}{\beta}\right)^{k}\right)$	$F(R)=\frac{\Gamma_{((R)/\beta)^k}(\alpha)}{\Gamma(\alpha)}$
Generalized Logistic	$f(R)=\begin{cases}\frac{\left(1+k\frac{R-\mu}{\sigma}\right)^{-1-1/k}}{\sigma\left(\left(1+k\frac{R-\mu}{\sigma}\right)^{-1/k}\right)^{2}} & k\neq 0\\ \frac{\exp\left(-\frac{R-\mu}{\sigma}\right)}{\sigma\left(1+\exp\left(-\frac{R-\mu}{\sigma}\right)\right)^{2}} & k=0\end{cases}$	$F(R)=\begin{cases}\frac{1}{\left(1+k\frac{R-\mu}{\sigma}\right)^{-1/k}} & k\neq 0\\ \frac{1}{1+\exp\left(-\frac{R-\mu}{\sigma}\right)} & k=0\end{cases}$
Generalized Pareto	$f(R)=\begin{cases}\frac{1}{\sigma}\left(-\left(1+k\frac{R-\mu}{\sigma}\right)^{-1-1/k}\right) & k\neq 0\\ \frac{1}{\sigma}\exp\left(-\frac{R-\mu}{\sigma}\right) & k=0\end{cases}$	$F(R)=\begin{cases}1-\left(1+k\frac{R-\mu}{\sigma}\right)^{-1-1/k} & k\neq 0\\ 1-\exp\left(-\frac{R-\mu}{\sigma}\right) & k=0\end{cases}$
Maximum Extreme Value Type 1	$f(R)=\frac{1}{\sigma}\exp\left(-\frac{R-\mu}{\sigma}-\exp\left(-\frac{R-\mu}{\sigma}\right)\right)$	$F(R)=\exp\left(-\exp\left(-\frac{R-\mu}{\sigma}\right)\right)$
Minimum Extreme Value Type 1	$f(R)=\frac{1}{\sigma}\exp\left(\frac{R-\mu}{\sigma}-\exp\left(-\frac{R-\mu}{\sigma}\right)\right)$	$F(R)=1-\exp\left(-\exp\left(-\frac{R-\mu}{\sigma}\right)\right)$
Three-Parameter Inverse Gaussian	$f(R)=\sqrt{\frac{\lambda}{2\pi(R-\gamma)}}\exp\left(-\frac{\lambda(R-\gamma-\mu)^2}{2\mu^2(R-\gamma)}\right)$	$F(R)=\Phi\left(\sqrt{\frac{\lambda}{R-\gamma}}\left(\frac{R-\gamma}{\mu}-1\right)\right)+\Phi\left(-\sqrt{\frac{\lambda}{R-\gamma}}\left(\frac{R-\gamma}{\mu}+1\right)\right)\exp\left(\frac{2\lambda}{\mu}\right)$
Log-Gamma	$f(R)=\frac{(\ln(R))^{\alpha-1}}{R\beta^{\alpha}\Gamma(\alpha)}\exp\left(\frac{-\ln(R)}{\beta}\right)$	$F(R)=\frac{\Gamma_{(\ln(R)/\beta)^k}(\alpha)}{\Gamma(\alpha)}$
Logistic	$f(R)=\frac{\exp\left(-\frac{R-\mu}{\sigma}\right)}{\sigma\left\{1+\exp\left(-\frac{R-\mu}{\sigma}\right)\right\}^{2}}$	$F(R)=\frac{1}{1+\exp(-R)}$

Two-Parameter Inverse Gaussian	$f(R)=\sqrt{\frac{\lambda}{2\pi(R-\gamma)}}\exp\left(-\frac{\lambda(R-\mu)^2}{2\mu^2R}\right)$	$F(R)=\Phi\left(\sqrt{\frac{\lambda}{R-\gamma}}\left(\frac{R}{\mu}-1\right)\right)+\Phi\left(-\sqrt{\frac{\lambda}{R-\gamma}}\left(\frac{R}{\mu}+1\right)\right)\exp\left(\frac{2\lambda}{\mu}\right)$
Log-Logistic	$f(R)=\left(\frac{\left(\frac{\beta}{\alpha}\left(\frac{R}{\alpha}\right)^{\beta-1}\right)}{\left(1+\frac{R}{\alpha}\right)^{\beta}}\right)^2$	$F(R)=\frac{1}{\left(1+\frac{R}{\alpha}\right)^{-\beta}}$
Three-Parameter Log-normal	$f(R)=\frac{1}{(R-\gamma)\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{\ln(R-\gamma)-\mu}{\sigma}\right)^2\right]$	$F(R)=\Phi\left[\frac{\ln(R-\gamma)-\mu}{\sigma}\right]$
Two-Parameter Log-normal	$f(R)=\frac{1}{R\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{\ln(R)-\mu}{\sigma}\right)^2\right]$	$F(R)=\frac{1}{2}+\mathrm{erf}\left[\frac{\ln(R)-\mu}{\sigma\sqrt{2}}\right]$
Log-Pearson 3	$f(R)=\frac{1}{R\lvert\beta\rvert\Gamma(\alpha)}\left(\frac{\ln(R)-\gamma}{\beta}\right)^{\alpha-1}\exp\left(-\frac{\ln(R)-\gamma}{\beta}\right)$	$F(R)=\frac{\Gamma_{(\ln(R)-\gamma)/\beta}(\alpha)}{\Gamma(\alpha)}$
Nakagami	$f(R)=\frac{2m^m}{\Gamma(m)\Omega^m}R^{2m-1}e^{\left(-\frac{m}{\Omega}G^2\right)}$	$F(R)=\frac{\gamma\left(m,\frac{m}{\Omega}R^2\right)}{\Gamma(m)}$
Normal	$f(R)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{R-\mu}{2\sigma^2}\right)$	$F(R)=\frac{1}{2}\left[1+\mathrm{erf}\left(\frac{R-\mu}{\sigma\sqrt{2}}\right)\right]$
Two-Parameter Rayleigh	$f(R)=\frac{R-\gamma}{\sigma^2}\exp\left(-\frac{1}{2}\left(\frac{R-\gamma}{\sigma}\right)^2\right)$	$F(R)=1-\exp\left(-\frac{1}{2}\left(\frac{R-\gamma}{\sigma}\right)^2\right)$
One-Parameter Rayleigh	$f(R)=\frac{R}{\sigma^2}\exp\left(-\frac{1}{2}\left(\frac{R}{\sigma}\right)^2\right)$	$F(R)=1-\exp\left(-\frac{1}{2}\left(\frac{R}{\sigma}\right)^2\right)$
Wakeby	$R(F)=\xi+\frac{\alpha}{\beta}\left(1-(1-F)^{\beta}\right)-\frac{\gamma}{\delta}\left(1-(1-F)^{\delta}\right)$	
Three-Parameter Weibull	$f(R)=\left(\frac{\alpha}{\beta}\right)\left(\frac{R-\gamma}{\beta}\right)^{\alpha-1}\exp\left(-\left(\frac{R-\gamma}{\beta}\right)^{\alpha}\right)$	$F(R)=1-\exp\left(-\left(\frac{R-\gamma}{\beta}\right)^{\alpha}\right)$
Two-Parameter Weibull	$f(R)=\left(\frac{\alpha}{\beta}\right)\left(\frac{R}{\beta}\right)^{\alpha-1}\exp\left(-\left(\frac{R}{\beta}\right)^{\alpha}\right)$	$F(R)=1-\exp\left(-\left(\frac{R}{\beta}\right)^{\alpha}\right)$

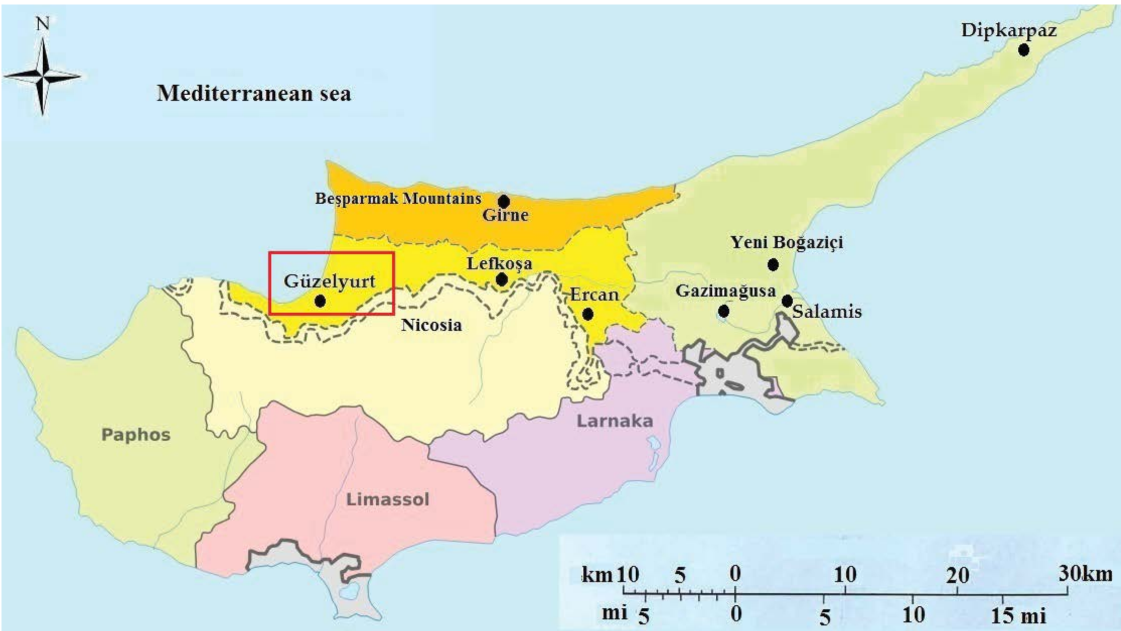

Fig. 2. Geographical location of Güzelyurt, Northern Cyprus.

Kolmogorov–Smirnov (K-S) test, the Anderson–Darling (A-D) test, and Chi-squared (C-s) test are the most well-known empirical distribution function tests [43,49]. These tests are widely used to find the best distribution [43,49,56,58].

Kolmogorov–Smirnov (K-S) test:

$$D = \max_{1 \le i \le n}\left(F(x_i) - \frac{i-1}{n}, \frac{i}{n} - F(x_i) \right) \tag{1}$$

where

$$F_n(x) = \frac{1}{n} \times (\text{Number of observation} \le x) \tag{2}$$

Anderson–Darling (A-D) test:

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1) \times \left[\ln F_X(x_i) + \ln\left(1 - F_X(x_{n-i+1})\right) \right] \tag{3}$$

where $F_X(x_i)$ is the cumulative distribution function of the proposed distribution at x_i, for $i = 1,2,\ldots,n$.

Chi-squared (C-s) test:

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \tag{4}$$

where O_i is the observed frequency for bin i, and E_i is the expected frequency for bin i calculated by:

$$E_i = F(x_2) - F(x_1) \tag{5}$$

where F is the cumulative distribution function of the probability distribution being tested, and x_1, x_2 are the limits for bin i.

3. Results and discussions

3.1. Description of rainfall data

In this section, the monthly rainfall (R) data are analyzed statistically. The statistical characteristics including arithmetic mean (Mean) standard deviation (SD), coefficient of variation in percent (CV), maximum (Max.), skewness (S), and kurtosis (K), of monthly rainfall for the selected region, are summarized in Table 2. It is found that the mean values of monthly rainfall are within the range of 0–159 mm. The maximum value of monthly rainfall occurred in February 2003 with a value of 159 mm (Fig. 3).

Besides, as shown in Fig. 3, the minimum value of 0 mm was recorded in the summer season (June, July, and August) for whole years. Moreover, Fig. 4 shows the value of annual rainfall for each year. It is found that the maximum and minimum values of the annual rainfall are recorded in 1994 and 2017 with a value of 484.8 and 128.7 mm, respectively.

3.2. Selecting the best-fit distribution model for average and annual rainfall

The distribution parameters were calculated using monthly and total rainfall with the maximum likelihood method. The best distribution among the 37 distribution function for the selected location was evaluated based on the Kolmogorov–Smirnov (K-S) test, Anderson–Darling (A-D) test, and Chi-squared (C-s) test. Generally, the distribution with the lowest value of K-S, A-D, and C-s will be selected to be the best model for the rainfall distribution in the selected region. The estimated distribution parameters for all selected models are tabulated in Table S4. Additionally, Tables S5 and S6 present the goodness-of-fit statistics for each distribution for average daily rainfall and total amount of rainfall, respectively, along with

Table 2
Statistical estimators of the mean monthly rainfall for the period 1981–2017

Year	Mean	SD	CV	Max.	S	K
1981	23.82	28.12	118.07	82.7	1.08	0.23
1982	18.34	25.06	136.61	80.2	1.58	2.35
1983	21.47	21.05	98.07	69	0.93	0.91
1984	18.44	21.1	114.41	49.51	0.5	–1.72
1985	18.95	25.23	133.14	84.2	1.77	3.41
1986	24.6	36.6	148.82	131.1	2.57	7.46
1987	26.03	32.49	124.82	89.3	0.99	–0.48
1988	38.4	45.8	119.3	116.5	0.88	–0.75
1989	11.75	15.46	131.55	50.3	1.52	2.48
1990	12.19	20.41	167.4	70.5	2.44	6.62
1991	26.5	37.5	141.54	128.3	2.07	4.83
1992	25	34.7	138.77	99.8	1.36	0.69
1993	17.23	19.8	114.89	51.8	0.78	–1.16
1994	40.4	34.8	86.25	99.3	0.25	–1.04
1995	12.33	12.21	99.04	37	0.86	0.01
1996	23.27	26.43	113.57	65.4	0.62	–1.36
1997	23.18	25.05	108.08	82.5	1.44	1.89
1998	18.32	26.11	142.54	86.4	1.84	3.58
1999	21.56	30.09	139.58	102.9	1.96	4.63
2000	29.22	33.58	114.92	99.8	1.06	0.05
2001	19.66	26.45	134.56	83.21	1.7	2.31
2002	31.68	34.6	109.23	107.11	0.99	0.28
2003	33	46.5	140.7	159	2.04	4.77
2004	27.7	40.6	146.18	135.3	1.89	4.02
2005	17.66	17.81	100.83	46.31	0.37	–1.58
2006	18.78	22.6	120.34	70.6	1.23	1.03
2007	25.25	29.84	118.21	84.1	1.02	0
2008	11.28	17.27	153.17	57.7	2.13	4.6
2009	29.94	33	110.22	90.71	0.96	–0.64
2010	29.3	50.8	173.16	154.7	1.94	2.93
2011	26.03	23	88.36	51.9	–0.08	–2.11
2012	37.7	43.7	115.86	135	1.12	0.6
2013	14.28	16.97	118.9	53.1	1.27	1.04
2014	14.7	14.51	98.73	35.2	0.39	–1.74
2015	26.32	24.92	94.69	67.5	0.65	–1.06
2016	22.25	30.19	135.67	96.6	1.62	2.31
2017	10.73	14.09	131.42	41.6	1.44	1.13
Average	23.19	21.62	93.24	54.73	0.39	–1.55

a ranking of the distribution models. The rank of the best five models based on goodness-of-fit tests (K-S, A-D, and C-s) for average monthly and annual rainfall data are summarized in Table 3. Based on the K-S tests, the Beta distribution has the lowest value, which is considered as the best distribution function to study the average monthly rainfall characteristics of Güzelyurt. Besides, Log-Pearson 3 is among the distribution giving the best fits to investigate the average rainfall distribution based on the A-D tests. Moreover, exponential (2P) is the best overall model according to the C-s test for the selected location.

Furthermore, as shown in Table S5, the inverse Gaussian distribution function cannot be used to analyze the distribution of average monthly rainfall in the selected region based on K-S and A-D tests. Additionally, Gumbel Min is less flexible and generally not suitable for analyzing the characteristics of average monthly rainfall based on the C-s test as shown in Table S5. It can be concluded that Beta, Log-Person 3 and Exponential (2P) are considerably more flexible and suitable to study the average monthly rainfall characteristics in the selected region as shown in Table 3.

Moreover, Table 3 presents the ranking of the best five distribution models for analyzing the annual rainfall during the investigation period based on the K-S, A-D, and C-s tests. It is observed that Burr is considered as the best model to study the total rainfall distribution of Güzelyurt based on the K-S test. Based on A-D and C-s tests, Wakeby and Nakagami are among the distribution giving the best fits to investigate the total rainfall distribution, respectively.

Figs. S1–S4 illustrate the frequency histograms and probability plots of average monthly and annual rainfall for the selected region.

3.3. Selecting the best-fit distribution model for extreme rainfall estimation

In general, the maximum monthly rainfall for Northern Cyprus always occurs in the winter season of December to February. Based on the monthly rainfall data, the maximum monthly rainfall occurs in January, February, November, and December in the selected region during the investigation period. The maximum monthly rainfall during the period of 1981–2017 is shown in Fig. 5. It is observed that the maximum monthly rainfall is recorded in 2003 with a value of 159 mm, while the minimum one occurred in 2014 with a value of 35.2 mm.

In this study, 37 different probability distribution functions were used to predict the probability distribution of the occurrence of maximum monthly rainfall. Table S7 lists the distribution parameters for all selected models for maximum monthly rainfall. Besides, Figs. S5 and S6 illustrate the frequency histograms and probability plots of rainfall of the selected location.

Additionally, the goodness-of-fit statistics for each distribution model along with a ranking of the distribution models are tabulated in Table S8. Based on the K-S tests, Nakagami distribution has the lowest value, which is considered as the best distribution function to study the average rainfall characteristics (Table 4). Additionally, based on the A-D tests, Wakeby is among the distribution giving the best fits to investigate the maximum daily rainfall distribution (Table 4). Moreover, Cauchy is the best overall model according to the C-s test for the selected location (Table 4). Also, it is observed that the Pareto 2 and Generalized Pareto distribution functions cannot be used to investigate the average rainfall in the studied location based on C-s and A-D tests, as shown in Table S8.

4. Discussion

The findings of this study are important for agricultural planning, management and other socio-economic

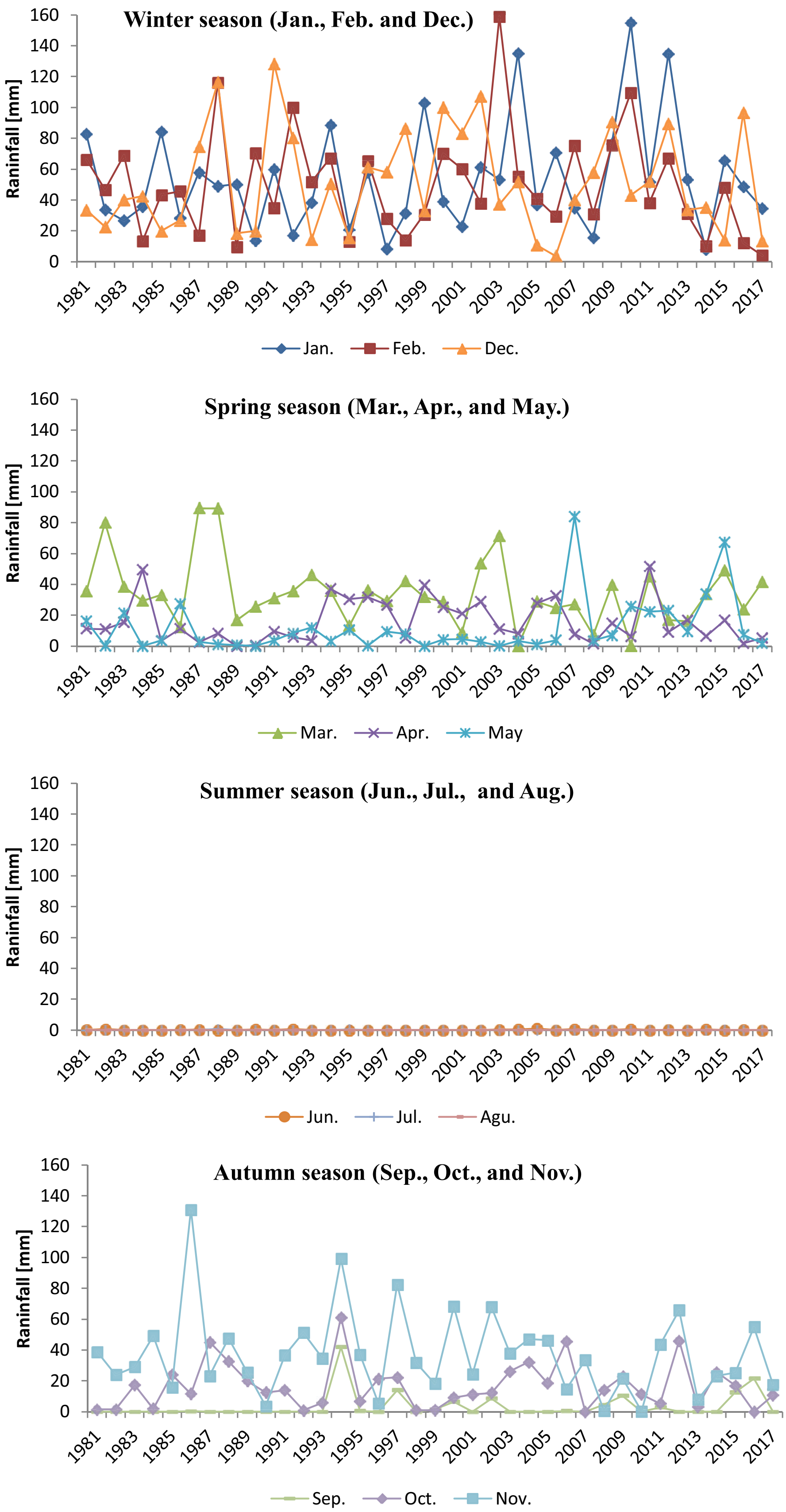

Fig. 3. Monthly rainfall during the investigation period (1981–2017).

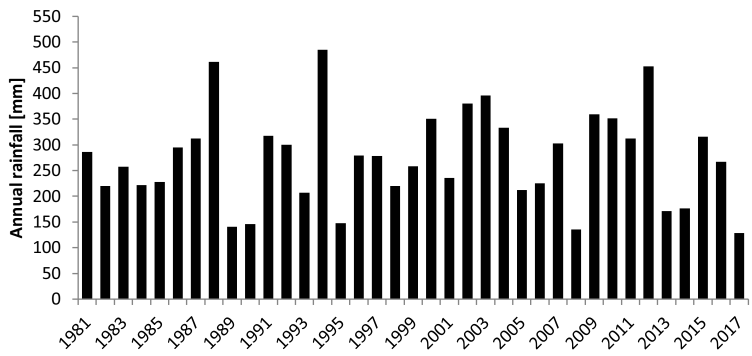

Fig. 4. Annual rainfall during the investigation period.

Table 3
Ranking of the best five distribution models based on the goodness-of-fit statistics

Rainfall data	Distribution	K-S		Distribution	A-D		Distribution	C-s	
		Statistic	Rank		Statistic	Rank		Statistic	Rank
Average monthly	Beta	0.14362	1	Log-Pearson 3	0.36036	1	Exponential (2P)	0.06054	1
	Log-Pearson 3	0.1501	2	Wakeby	0.37498	2	Erlang (3P)	0.06351	2
	Gen. Extreme Value	0.15168	3	Gen. Pareto	0.37498	3	Gamma	0.06985	3
	Wakeby	0.15512	4	Gen. Extreme Value	0.44902	4	Inv. Gaussian	0.07059	4
	Gen. Pareto	0.15512	5	Burr	0.46154	5	Burr	0.09486	5
Annual	Burr	0.07732	1	Wakeby	0.29879	1	Nakagami	0.53549	1
	Gen. Extreme Value	0.0816	2	Gen. Logistic	0.37738	2	Log-normal (3P)	0.6523	2
	Normal	0.08412	3	Burr	0.37822	3	Gen. Logistic	1.1775	3
	Gen. Logistic	0.08535	4	Normal	0.39018	4	Gamma (3P)	1.2737	4
	Log-Logistic (3P)	0.08731	5	Gen. Extreme Value	0.39429	5	Log-Pearson 3	1.3381	5

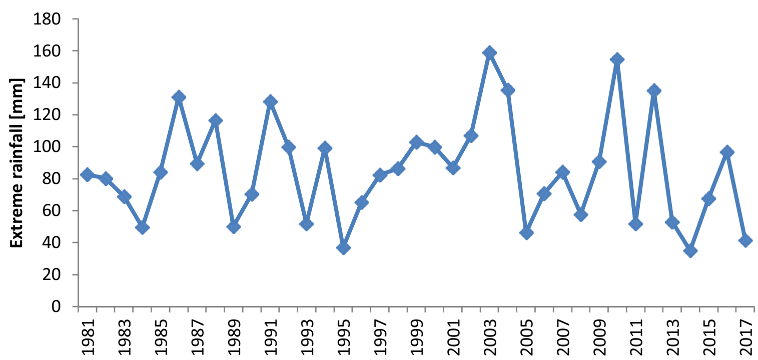

Fig. 5. Extreme rainfall for the selected region.

Table 4
Ranking of the best five distribution models based on the goodness-of-fit statistics for maximum monthly rainfall

Distribution	K-S		Distribution	A-D		Distribution	C-s	
	Statistic	Rank		Statistic	Rank		Statistic	Rank
Nakagami	0.07436	1	Wakeby	0.20183	1	Cauchy	0.09775	1
Wakeby	0.07454	2	Gen. Extreme Value	0.24554	2	Nakagami	0.64303	2
Normal	0.08257	3	Nakagami	0.26162	3	Weibull	0.64437	3
Gen. Extreme Value	0.08306	4	Gamma	0.26273	4	Burr	0.96495	4
Weibull	0.08538	5	Log-Pearson 3	0.26661	5	Logistic	1.2797	5

Table 5
Best distribution model (Rank #1) of the present study and previous scientific studies

Ref.	Region, Country	Used distribution model	Data (period)	Goodness-of-fit test	Best distribution model	Descriptive of used data in mm			
						Mean	Max.	Min.	SD
Current study	Güzelyurt, Northern Cyprus	37 distribution models	Mean monthly rainfall (1981–2017)	K-S, A-D, and C-s	Beta, Log-Pearson 3, and Exponential (2P)	23.6	54.7	0.0	21.14
			Annual rainfall (1981–2017)	K-S, A-D, and C-s	Burr, Wakeby and Nakagami	281.7	48,408	131.3	92.2
			Maximum monthly rainfall (1981–2017)	K-S, A-D, and C-s	Nakagami, Wakeby, and Cauchy	85.1	159.0	35.2	32.4
[42]	Kuantan, Malaysia	Beta, Exponential, Gamma, Generalized extreme value, Generalized Pareto, Log-Pearson 3, and Wakeby	6 h rainfall	K-S and A-D	Wakeby	16.42	–	–	39.87
[49]	Bhola, Bangladesh	Normal, Log-Normal, Pearson Type 3, Log-Pearson Type 3, Exponential, Gumbel, Generalized extreme value, Weibull and Generalized Pareto	Maximum monthly rainfall (1984–2013)	K-S, A-D, and C-s	Generalized extreme value	600.6	174.8	–	–
[49]	Bogra, Bangladesh	Normal, Log-Normal, Pearson Type 3, Log-Pearson Type 3, Exponential, Gumbel, Generalized extreme value, Weibull and Generalized Pareto	Maximum monthly rainfall (1984–2013)	K-S, A-D, and C-s	Log-Pearson Type 3	483.9	151.1	–	–

[56]	Guntur, Hyderabad, Jaipur, Kohima, Kurnool, Patna, India	Normal, Log-normal, Gamma, Inverse Gaussian, Generalized extreme value, Gumbel, Student-t, Beta and Weibull	Monthly rainfall (1901–2002)	BIC	Beta	65.6–196.3	374.5–694.2	0.0	53.9–177.6
[57]	Campinas, Brazil	Wakeby, Kappa and Generalized extreme value	Annual maximums of daily, 2-day and 3-day rainfall amount (1890–2012)	K-S	Generalized extreme value	79.0–121.5	144.7–252	50.0–66.8	21.6–36.1
[58]	Kalam, Oghi, Daggar, Puran Besham Qilla, Pakistan	Normal, Log-normal, Log-Pearson type 3, and Gumbel	24-h annual maximum rainfall (1961–2010)	K-S, A-D and C-s	Log-Pearson Type 3	62.37–89.62	114.3–196.85	19.3–45.72	17.81–31.72
[59]	Anagar Solapur, Satara and Kolhapour, India	Beta, Cauchy, Dagum, Exponential, Gamma, Generalized Pareto, Gumbel, Laplace, Log-logistic, Log-normal, Normal, Weibull and Rayleigh	Monthly rainfall dataset (1966–2000)	K-S, A-D and C-s	Beat	68.3–186.7	–	–	76.8–223.3
[59]	Sangli, India	Beta, Cauchy, Dagum, Exponential, Gamma, Generalized Pareto, Gumbel, Laplace, Log-logistic, Log-normal, Normal, Weibull and Rayleigh	Monthly rainfall dataset (1966–2000)	K-S and C-s	Dogum	75.5	–	–	71.6

activities, which are directly concerned with the rainfed agricultural system. Furthermore, it can be seen that no single distribution can accurately describe the rainfall distribution based on the summary of previous scientific studies (Section 1.2). The selection of the model depends on the available data of rainfall and the statistical tools utilized for model selection.

Since the previous studies have not considered 37 models, apples-to-apples comparison to these results is not straightforward. The dataset and duration used in previous scientific studies are also different. Nevertheless, the best models for the selected region are compared with other regions in the world that have the same best distribution model to find similarities in the statistical characteristics of rainfall.

Among the previous studies (Table 5), Alam et al. [49] found that 36% and 26% of the locations in Bangladesh were fit by Generalized extreme value and Log-Pearson type 3, respectively, for maximum monthly rainfall based on non-parametric goodness-of-fit tests (K-S, A-D, and C-s). Sukrutha et al. [56] found that the Beta function was the best-fitted distribution for analyzing the monthly precipitation data (1901–2001) for all selected locations except Gandhinagar in India based on Bayesian Information Criterion (BIC). Blain and Meschiatti [57] found that generalized extreme value was the best distribution to describe the characteristics of annual maximums of daily, 2 d, and 3 d rainfall data collected from Campinas station, Brazil. In addition, based on K-S, A-D, and C-s tests, Log-Pearson type 3 was found to be the best model for annual maximum rainfall in all the selected stations except Mardan station in Pakistan [58].

It can be concluded that none of the previous studies (section 1.2 and Table 5) have found that Nakagami, Cauchy, and Burr distributions to be an adequate fit for the rainfall data. Besides, Wakeby distribution has been considered by Hassan et al. [42] for fitting the hourly rainfall data in Malaysia. Various distribution models may be suitable for average, maximum and annual rainfall/flood frequency estimations at the same region; therefore, selecting the best model for rainfall/flood frequency analysis at the region depends on the frequency regime of the data series. Therefore, the authors hope that the current results spur future studies to consider these distributions for fitting rainfall data in Cyprus. Additionally, the recommended distributions may be used to estimate the flood frequency distribution (return periods for various flood extremes and droughts) in the region.

5. Conclusions

This study examined the selection of the best fit probability distribution for various cases of rainfall (average monthly rainfall, annual rainfall, and maximum monthly rainfall) in the Güzelyurt region, Northern Cyprus. 37 different probability distribution models and three goodness-of-fit tests were employed. It is found that Beta, Log-Pearson 3, and exponential (2P) distributions have the lowest value of K-S, A-D, and C-s, which is considered as the best distribution function to study the average rainfall characteristics of Güzelyurt. Additionally, in the case of annual rainfall, the results demonstrated that Burr, Wakeby, and Nakagami distributions are giving the best fit of actual data of total rainfall based on goodness-of-fit tests. Furthermore, the best fit probability distribution model for maximum monthly rainfall at the selected region was Nakagami, Wakeby, and Cauchy distributions.

Moreover, this study is the first one, which finds the Nakagami, Cauchy, and Burr distributions to be an adequate fit for the rainfall data. Finally, it is concluded that the results of the current study can be utilized to improve urban infrastructure planning and design such as drainage systems in the selected region. Also, it can be used to develop better models of flooding risk and damage in the future. Moreover, this study can help policymakers to plan initiatives that could result in saving lives and assets. Additionally, knowledge of the pattern of heavy rainfall will be useful in various disciplines such as environment, agriculture, construction, and building structures.

Acknowledgments

The authors would like to thank the Faculty of Civil and Environmental Engineering and Faculty of Engineering especially the Mechanical and Civil Engineering Departments at Near East University for their support and encouragement.

References

[1] M. Clemens, G. Khurelbaatar, R. Merz, C. Siebert, M.V. Afferden, T. Rödiger, Groundwater protection under water scarcity; from regional risk assessment to local wastewater treatment solutions in Jordan, Sci. Total Environ., 706 (2020) 136066.
[2] A. Rezaei, H. Hassani, S. Hassani, N. Jabbari, S.B.F. Mousavi, S. Rezaei, Evaluation of groundwater quality and heavy metal pollution indices in Bazman basin, southeastern Iran, Groundwater Sustainable Dev., 9 (2019) 100245.
[3] H. Gökçekuş, Y. Kassem, J. Aljamal, Analysis of different combinations of meteorological parameters in predicting rainfall with an ANN approach: a case study in Morphou, Northern Cyprus, Desal. Water Treat., 177 (2020) 350–362.
[4] M. Song, R. Wang, X. Zeng, Water resources utilization efficiency and influence factors under environmental restrictions, J. Cleaner Prod., 184 (2018) 611–621.
[5] S. Kundu, D. Khare, A. Mondal, Future changes in rainfall, temperature and reference evapotranspiration in the central India by least square support vector machine, Geosci. Front., 8 (2017) 583–596.
[6] C.D. Ahrens, R. Henson, Meteorology Today: an Introduction to Weather, Climate, and the Environment, Cengage Learning, Boston, MA, 2016.
[7] P. Panagos, P. Borrelli, J. Poesen, C. Ballabio, E. Lugato, K. Meusburger, C. Alewell, The new assessment of soil loss by water erosion in Europe, Environ. Sci. Policy, 54 (2015) 438–447.
[8] T. Iizumi, N. Ramankutty, How do weather and climate influence cropping area and intensity?, Global Food Secur., 4 (2015) 46–50.
[9] C.P. McMullen, J.R. Jabbour, Climate Change Science Compendium 2009, United Nations Environment Programme, Nairobi, 2009.
[10] O. Mehmet, H. Bicak, Modern and Traditional Irrigation Technologies in the Eastern Mediterranean, IDRC International Development Research Centre, 2002.
[11] A. Sofroniou, S. Bishop, Water scarcity in Cyprus: a review and call for integrated policy, Water, 6 (2014) 2898–2928.
[12] E.A. Pearce, C.G. Smith, The Times Books World Weather Guide, Times Books/Random House, New York, 1990.

[13] I. Iacovides, Water Resources in Cyprus: Endowments and Water Management Practices, Water Resour. Allocation, 1 (2011) 11–21.
[14] O.P. Agboola, F. Egelioglu, Water scarcity in North Cyprus and solar desalination research: a review, Desal. Water Treat., 43 (2012) 29–42.
[15] G.R.H. Wright, Ancient Building in Cyprus, E.J. Brill, Leiden, 1992.
[16] D. Fanta, Conceptual Model Development for Güzelyurt aquifer, North Cyprus (Master dissertation), Middle East Technical University. Retrieved February 30, 2020. Available at: http://etd.lib.metu.edu.tr/upload/12618937/index.pdf
[17] U. Türker, B.R. Hansen, River Basin Management and Characterization of Water Bodies in North Cyprus, 10th International Congress on Advances in Civil Engineering, 7–19 October 2012, Middle East Technical University, Ankara, Turkey.
[18] H. Gökçekuş, V. Nourani, Lessons from Groundwater Quantity and Quality Problems In the Güzelyurt Coastal Aquifer, North Cyprus, In Advances in Environmental Research, Vol. 64, Nova Science Publishers, Inc., USA, 2018.
[19] H. Gökçekus, V. Doyuran, Evaluation of hydrodynamic data from Güzelyurt Aquifer (Turkish Republic of Northern Cyprus) through cluster analysis, Turk. J. Earth Sci., 4 (1995) 1–10.
[20] M. Ergil, The salination problem of the Guzelyurt aquifer, Cyprus, Water Res., 34 (2000) 1201–1214.
[21] D. McKinney, P. Brown, Yeşilköy/Ayios Andronikos Aquifer Groundwater Model Technical Report, USGS/SAVE Project, Nicosia, Cyprus, 2007.
[22] H. Gökcekus, A. Iravanian, U. Türker, G. Oguz, S. Sözen, D. Orhon, Massive freshwater transport: a new dimension for integrated water-wastewater management in North Cyprus, Desal. Water Treat., 132 (2018) 215–225.
[23] S. Michaelides, F. Tymvios, T. Michaelidou, Spatial and temporal characteristics of the annual rainfall frequency distribution in Cyprus, Atmos. Res., 94 (2009) 606–615.
[24] H.N. Phien, T.J. Ajirajah, Applications of the log Pearson type-3 distribution in hydrology, J. Hydrol., 73 (1984) 359–372.
[25] R. Tortorelli, W. Asquith, R. Alan, Depth-Duration Frequency of Precipitation for Oklahoma, US Geological Survey, Water Resources Investigation Report, 1999, pp. 99–4232.
[26] Z. Şen, A.G. Eljadid, Rainfall distribution function for Libya and rainfall prediction, Hydrol. Sci. J., 44 (1999) 665–680.
[27] A. Ogunlela, Stochastic Analysis of Rainfall Events in Ilorin, Nigeria, J. Agric. Res. Dev., 1 (2001) 39–50.
[28] M. Zalina, M. Desa, V.T.A. Nguyen, A. Kassim, Selecting a probability distribution for extreme rainfall series in Malaysia, Water Sci. Technol., 45 (2002) 63–68.
[29] D.Q. Tao, V.T. Nguyen, A. Bourque, On Selection of Probability Distributions for Representing Extreme Precipitations in Southern Quebec, In: Annual Conference of the Canadian Society for Civil Engineering, 5–8th June, 2002, pp. 1–8.
[30] C. Lee, Application of rainfall frequency analysis on studying rainfall distribution characteristics of Chia-nan plain in Southern Taiwan, J. Crop Environ. Bioinf., 2 (2005) 31–38.
[31] F. Yusof, Z. Daud, V. Nguyen, S. Jamaludin, Z. Yusop, Fitting the best–fit distribution for the hourly rainfall amount in the wilayah Persekutuan, J. Teknol., 46 (2007) 49–58.
[32] X. Kwaku, O. Duke, Characterization and frequency analysis of one day annual maximum and two to five consecutive days maximum rainfall of Accra, Ghana, ARPN J. Eng. Appl. Sci., 2 (2007) 27–31.
[33] L.S. Hanson, R. Vogel, The Probability Distribution of Daily Rainfall in the United States, World Environmental and Water Resources Congress, Honolulu, Hawaii, United States, 2008.
[34] M. Sharma, J. Singh, Use of probability distribution in rainfall analysis, N. Y. Sci. J., 3 (2010) 40–49.
[35] F. Johnson, K. Haddad, A. Rahman, J. Green, Application of Bayesian GLSR to estimate sub-daily rainfall parameters for the IDF revision project, hydrology and water resources symposium, 19–22 Nov, 2012, Australia, Sydney.
[36] J. Green, K. Xuereb, F. Johnson, G. Moore, The Revised Intensity–Frequency–Duration (IFD) Design Rainfall Estimates for Australia—An Overview, Proceeding of 34th Hydrology and Water Resources Symposium 19–22 November, 2012, Sydney, Australia.
[37] M. Khudri, F. Sadia, Determination of the best fit probability distribution for annual extreme precipitation in Bangladesh, Eur. J. Sci. Res., 103 (2013) 391–404.
[38] A.A. Mamoon, N.E. Joergensen, A. Rahman, H. Qasem, Derivation of new design rainfall in Qatar using L-moment based index frequency approach, Int. J. Sustainable Built Environ., 3 (2014) 111–118.
[39] M. Montaseri, B. Amirataee, H. Rezaei, Assessment of goodness of fit methods in determining the best regional probability distribution of rainfall data, Int. J. Eng., 27 (2014) 1537–1546.
[40] S. Mandal, B.U. Choudhury, Estimation and prediction of maximum daily rainfall at Sagar Island using best fit probability models, Theor. Appl. Climatol., 121 (2014) 87–97.
[41] A.M. Subyani, N.S. Al-Amri, IDF curves and daily rainfall generation for Al-Madinah city, western Saudi Arabia, Arab. J. Geosci., 8 (2015) 11107–11119.
[42] Z. Hassan, S. Shamsudin, S. Harun, Choosing the best fit distribution for rainfall event characteristics based on 6H-Ietd within Peninsular Malaysia, J. Teknol., 75 (2015) 145–157.
[43] A.A. Mamoon, A. Rahman, Selection of the best fit probability distribution in rainfall frequency analysis for Qatar, Nat. Hazards, 86 (2016) 281–296.
[44] M.T. Amin, M. Rizwan, A.A. Alazba, A best-fit probability distribution for the estimation of rainfall in northern regions of Pakistan, Open Life Sci., 11 (2016) 432–440.
[45] T. Mohamed, A. Ibrahim, Fitting probability distributions of annual rainfall in Sudan, SUST J. Eng. Comput. Sci., 17 (2016) 35–39.
[46] A.I. Agbonaye, O.C. Izinyon, Best-fit probability distribution model for rainfall frequency analysis of Three Cities in South Eastern Nigeria, Nigerian J. Environ. Sci. Technol., 1 (2017) 34–42.
[47] J. Yuan, K. Emura, C. Farnham, M.A. Alam, Frequency analysis of annual maximum hourly precipitation and determination of best fit probability distribution for regions in Japan, Urban Climate, 24 (2018) 276–286.
[48] M. Alam, C. Farnham, K. Emura, Best-fit probability models for maximum monthly rainfall in Bangladesh using Gaussian mixture distributions, Geosciences, 8 (2018) 138.
[49] M. Alam, K. Emura, C. Farnham, J. Yuan, Best-fit probability distributions and return periods for maximum monthly rainfall in Bangladesh, Climate, 6 (2018) 9.
[50] R.K. Meena, S. Dubey, S.M. Basak, Use of different probability distribution in rainfall analysis for RRS, Pundiari, Cooch Behar, Int. J. Curr. Microbiol. Appl. Sci., 8 (2019) 1325–1337.
[51] H. Baghel, H. Mittal, P. Singh, K. Yadav, S. Jain, Frequency analysis of rainfall data using probability distribution models, Int. J. Curr. Microbiol. Appl. Sci., 8 (2019) 1390–1396.
[52] A.S. Parchure, S.K. Gedam, Probability distribution analysis of extreme rainfall events in a flood-prone region of Mumbai, India, Arab. J. Geosci., 12 (2019), https://link.springer.com/article/10.1007/s12517-019-4473-2#citeas.
[53] Y. Kassem, H. Gökçekuş, Water resources and rainfall distribution function: a case study in Lebanon, Desal. Water Treat., 177 (2020) 306–321.
[54] N. Stamatatou, L. Vasiliades, A. Loukas, The effect of sample size on bivariate rainfall frequency analysis of extreme precipitation, Proceedings, 7 (2018) 19.
[55] H. Zaifoglu, B. Akintug, A.M. Yanmaz, Regional frequency analysis of precipitation using time series clustering approaches, J. Hydrol. Eng., 23 (2018) 05018007.
[56] A. Sukrutha, S. Dyuthi, S. Desai, Probability distribution for monthly precipitation data in India, ArXiv, (2017) 1–23, https://www.researchgate.net/publication/319057222_Probability_distribution_for_monthly_precipitation_data_in_India.
[57] G.C. Blain, M.C. Meschiatti, Using multi-parameters distributions to assess the probability of occurrence of extreme rainfall data, Rev. Bras. Eng. Agríc. Ambiental, 18 (2014) 307–313.

Supplementary information

Table S1
Dams in Northern Cyprus constructed after 1987 [12]

District	Dams	Year of construction	Capacity (10^3) (m^3)	Irrigation area (ha)
Gazimağusa	Gönendere	1987	940	150
	Geçitkale	1989	1,360	240
	Mersinlik	1989	1,140	170
	Tatlisu	1989	156	50
	Ergazi	1989	400	84
Güzelyurt	Akdeniz	1988	1,470	–
	Gemikonağı	1988	4,120	–
Girne	Geçitköy	1989	1,800	161
	Zeytinlik	1989	50	–
	Karsiyaka	1989	25	–
	Arapköy 1	1990	440	40
	Arapköy 2	1990	600	65
	Beşparmak	1992	775	67
	Dağyolu	1994	392	82
Lefkoşa	Değirmenlik	1990	297	30
	Hamitköy	1992	529	95
	Serdarli	1992	391	56
	Lefkoşa	1994	517	40

Table S2
Aquifers' capacities in Northern Cyprus

Aquifers	Recharge (10^6) (m^3)	Sustainable yield (10^6) (m^3)	Withdrawals (10^6) (m^3)
Güzelyurt	37	37	57
Akdeniz	15	15	1.5
Lefke-G. Kona ği Y. dalga	15.5	6	6
Yesilirmak	7	1.5	1.5
Girne	11.5	11.5	11.5
Mountains Gazimağusa	2	2	8.5
Beyarmudu	0.5	0.5	0.5
Çayön-Güvercinlik-Türkmenköy	2	2	2
Lefkoşa-Serdarli	0.5	0.5	0.5
Yesilköy	1.6	1.6	3
Girne Coast	5	5	5
Yedikonuk-Büyükkonuk	0.3	0.3	0.3
Dipkarpaz	1.5	1.5	1.5
Korucam	1.5	1.2	1.2
Others	2	2	2
Total	89.1	74.1	103

Table S3
Summary of previous applications of probability distributions in rainfall/flood analysis

Reference	Data used	Best fit distribution	Country
[23]	Annual rainfall	Gamma	Cyprus
[24]	Annual maximum rainfall data (durations of 1 h to 31 d)	Log-Pearson type 3	Thailand, India, Laos, and the USA
[25]	15, 30, and 60 min and 1, 2, 3, 6, 12, and 24 h and 1, 3, and 7 d of rainfall	Generalized logistic	Oklahoma, USA
[26]	Monthly total rainfall data	Gamma	Libya
[27]	Daily and monthly annual maximum rainfall data	Log-Pearson type 3	Nigeria
[28]	Annual maximum rainfall data (1 h duration)	Generalized extreme value	Malaysia
[29]	Annual maximum rainfall series (5 min and 1 h duration)	Generalized extreme value	Southern Quebec, Canada
[30]	Annual maximum rainfall data (24 h duration)	Log-Pearson type 3	Taiwan
[31]	Annual maximum rainfall data (1 h duration)	Mixed-exponential	Malaysia
[32]	Annual maximum rainfall series (1-, 2-, 3-, 4- and 5 d durations)	Log-normal	Ghana
[33]	Daily rainfall	3-parameter Pearson-III distribution and 4-parameter Kappa distribution	United States
[34]	Annual maximum rainfall series (24 h duration)	Log-normal	Pantnagar, India
[35]	5 min to 72 h durations	Generalized extreme value	Australia
[36]	1–12 h, 1–7 d rainfall	Generalized extreme value	Australia
[37]	Annual maximum rainfall	Generalized extreme value and four parameters generalized gamma	Bangladesh
[38]	Annual maximum rainfall 24 h duration	Pearson type 3	Qatar
[39]	Monthly and annual rainfall data	Pearson type 3	Northwest of Iran
[40]	Annual, seasonal and monthly maximum daily rainfall	Normal for annual, post-monsoon, and summer seasons. Log-normal, Weibull, and Pearson 5 for pre-monsoon, monsoon, and winter seasons, respectively	Sagar Island
[41]	24 h annual maximum rainfall data	Extreme value type 1 and Log-Pearson type 3	Al-Madinah City, Saudi Arabia
[42]	6 h rainfall	Generalized Pareto, Wakeby and Generalized Extreme value	Peninsular Malaysia
[43]	Annual maximum rainfall series (average of 36 y)	Generalized extreme value	Qatar
[44]	24 h annual maximum rainfall	Log-Pearson type 3	Northern regions of Pakistan
[45]	Annual rainfall	Normal and Gamma distribution	Sudan
[46]	Annual maximum series of daily rainfall data	Generalized extreme value	Southeastern Nigeria
[47]	Annual maximum hourly rainfall	Log-Pearson type 3	Japan
[48]	Extreme values for precipitation	Gaussian/normal	Bangladesh
[49]	Maximum monthly rainfall	Pearson type 3 and Log-Pearson type 3	Bangladesh
[50]	Daily rainfall	Gamma	Cooch Behar
[51]	Daily maximum rainfall data	Log-normal and Gumbel	Udaipur district
[52]	Extreme rainfall events	Generalized extreme value, generalized Pareto and Frechet	Mumbai, India
[53]	Monthly rainfall	Gumbel Maximum and Logistic	Lebanon
[54]	Annual maximum rainfall	Generalized extreme value	Cyprus
[55]	Annual maximum daily rainfall	Pearson Type III, generalized logistic, and generalized normal	Northern part of Cyprus

Table S4
Distribution parameters for selected distribution functions for average rainfall for whole years (1981–2017)

Mean rainfall for whole year (1981–2017)		Total rainfall during the investigation period	
Distribution	Parameters	Distribution	Parameters
Beta	$\alpha_1 = 0.25951$ $\alpha_2 = 0.34872$ $a = 0.5$ $b = 54.7$	Beta	$\alpha_1 = 1.1179$ $\alpha_2 = 1.5445$ $a = 131.3$ $b = 484.8$
Burr	$k = 4{,}375.8$ $\alpha = 0.91159$ $\beta = 2.2469\text{E}{+}5$	Burr	$k = 5.6154$ $\alpha = 3.7541$ $\beta = 478.96$
Burr (4P)	$k = 0.55329$ $\alpha = 0.7451$ $\beta = 1.7549$ $\gamma = 0.5$	Burr (4P)	$k = 0.31323$ $\alpha = 0.64005$ $\beta = 1.7009$ $\gamma = 131.3$
Cauchy	$\sigma = 13.264$ $\mu = 14.412$	Cauchy	$\sigma = 52.104$ $\mu = 289.92$
Dagum	$k = 0.00294$ $\alpha = 180.51$ $\beta = 55.684$	Dagum	$k = 358.34$ $\alpha = 6.9414$ $\beta = 91.416$
Dagum (4P)	$k = 0.00824$ $\alpha = 86.678$ $\beta = 55.775$ $\gamma = 0.5$	Dagum (4P)	$k = 2.4519$ $\alpha = 0.38221$ $\beta = 0.64521$ $\gamma = 131.3$
Erlang	$m = 1$ $\beta = 18.914$	Erlang	$m = 9$ $\beta = 30.181$
Erlang (3P)	$m = 1$ $\beta = 23.19$ $\gamma = 0.45004$	Erlang (3P)	$m = 20$ $\beta = 20.716$ $\gamma = -123.72$
Exponential	$\lambda = 0.04233$	Exponential	$\lambda = 0.00355$
Exponential (2P)	$\lambda = 0.04324$ $\gamma = 0.5$	Exponential (2P)	$\lambda = 0.00665$ $\gamma = 131.3$
Gamma	$\alpha = 1.2491$ $\beta = 18.914$	Gamma	$\alpha = 9.3328$ $\beta = 30.181$
Gamma (3P)	$\alpha = 0.46877$ $\beta = 45.08$ $\gamma = 0.5$	Gamma (3P)	$\alpha = 19.569$ $\beta = 20.716$ $\gamma = -123.73$
Gen. Extreme Value	$k = -0.01927$ $\sigma = 18.102$ $\mu = 13.515$	Gen. Extreme Value	$k = -0.22457$ $\sigma = 89.583$ $\mu = 246.57$
Gen. Gamma	$k = 0.83978$ $\alpha = 1.0987$ $\beta = 18.914$	Gen. Gamma	$k = 0.99356$ $\alpha = 9.1934$ $\beta = 30.181$
Gen. Gamma (4P)	$k = 17.419$ $\alpha = 0.02623$ $\beta = 57.808$ $\gamma = 0.5$	Gen. Gamma (4P)	$k = 2.9896$ $\alpha = 0.30497$ $\beta = 291.9$ $\gamma = 131.3$
Gen. Logistic	$k = 0.1576$ $\sigma = 11.832$ $\mu = 20.468$	Gen. Logistic	$k = 0.03346$ $\sigma = 52.347$ $\mu = 278.78$
Gen. Pareto	$k = -0.45541$ $\sigma = 44.061$ $\mu = -6.6491$	Gen. Pareto	$k = -0.87048$ $\sigma = 281.58$ $\mu = 131.13$
Gumbel Max	$\sigma = 16.482$ $\mu = 14.112$	Gumbel Max	$\sigma = 71.888$ $\mu = 240.17$
Gumbel Min	$\sigma = 16.482$ $\mu = 33.138$	Gumbel Min	$\sigma = 71.888$ $\mu = 323.16$
Inv. Gaussian	$\lambda = 29.51$ $\mu = 23.625$	Inv. Gaussian	$\lambda = 2{,}628.7$ $\mu = 281.67$
Inv. Gaussian (3P)	$\lambda = 19.701$ $\mu = 26.855$ $\gamma = -3.2301$	Inv. Gaussian (3P)	$\lambda = 36{,}824.0$ $\mu = 674.97$ $\gamma = -393.3$
Log-Logistic	$\alpha = 0.86471$ $\beta = 9.8786$	Log-Gamma	$\alpha = 254.01$ $\beta = 0.02198$
Log-Logistic (3P)	$\alpha = 0.78596$ $\beta = 10.946$ $\gamma = 0.5$	Log-Logistic	$\alpha = 4.627$ $\beta = 261.83$
Log-Pearson 3	$\alpha = 5.1207$ $\beta = -0.69943$ $\gamma = 6.0146$	Log-Logistic (3P)	$\alpha = 18.715$ $\beta = 977.13$ $\gamma = -698.81$
Logistic	$\sigma = 11.654$ $\mu = 23.625$	Log-Pearson 3	$\alpha = 17.851$ $\beta = -0.08293$ $\gamma = 7.0647$
Log-normal	$\sigma = 1.5154$ $\mu = 2.433$	Logistic	$\sigma = 50.833$ $\mu = 281.67$
Log-normal (3P)	$\sigma = 1.4582$ $\mu = 2.4711$ $\gamma = -0.13202$	Log-normal	$\sigma = 0.34562$ $\mu = 5.5843$
Nakagami	$m = 0.65801$ $\Omega = 967.74$	Log-normal (3P)	$\sigma = 0.12864$ $\mu = 6.5525$ $\gamma = -425.17$
Normal	$\sigma = 21.138$ $\mu = 23.625$	Nakagami	$m = 2.5158$ $\Omega = 87{,}608.0$
Pareto	$\alpha = 0.31988$ $\beta = 0.5$	Normal	$\sigma = 92.2$ $\mu = 281.67$
Pareto 2	$\alpha = 100.56$ $\beta = 2{,}275.2$	Pareto	$\alpha = 1.4147$ $\beta = 131.3$
Rayleigh	$\sigma = 18.85$	Pareto 2	$\alpha = 138.81$ $\beta = 38{,}831.0$
Rayleigh (2P)	$\sigma = 28.212$ $\gamma = -10.758$	Rayleigh	$\sigma = 224.74$
Wakeby	$\alpha = 44.061$ $\beta = 0.45541$ $\gamma = 0$ $\delta = 0$ $\xi = -6.6491$	Rayleigh (2P)	$\sigma = 141.12$ $\gamma = 104.01$
Weibull	$\alpha = 0.63321$ $\beta = 21.745$	Wakeby	$\alpha = 472.13$ $\beta = 2.2562$ $\gamma = 17.636$ $\delta = 0.40015$ $\xi = 107.27$
Weibull (3P)	$\alpha = 0.72873$ $\beta = 18.376$ $\gamma = 0.5$	Weibull	$\alpha = 3.317$ $\beta = 308.21$
Log-Gamma	No fit	Weibull (3P)	$\alpha = 2.1256$ $\beta = 207.81$ $\gamma = 97.411$

Table S5
Results of goodness-of-fit and ranking of distribution functions based on goodness-of-fit for average rainfall for whole years (1981–2017)

Distribution	K-S		Distribution	A-D		Distribution	C-s	
	Statistic	Rank		Statistic	Rank		Statistic	Rank
Beta	0.14362	1	Log-Pearson 3	0.36036	1	Exponential (2P)	0.06054	1
Log-Pearson 3	0.1501	2	Wakeby	0.37498	2	Erlang (3P)	0.06351	2
Gen. Extreme Value	0.15168	3	Gen. Pareto	0.37498	3	Gamma	0.06985	3
Wakeby	0.15512	4	Gen. Extreme Value	0.44902	4	Inv. Gaussian	0.07059	4
Gen. Pareto	0.15512	5	Burr	0.46154	5	Burr	0.09486	5
Gen. Logistic	0.16436	6	Weibull	0.46715	6	Exponential	0.09536	6
Weibull	0.16639	7	Gen. Gamma	0.48709	7	Log-Logistic	0.105	7
Burr	0.16871	8	Gen. Logistic	0.53096	8	Gen. Gamma	0.11154	8
Log-normal (3P)	0.17525	9	Gumbel Max	0.53485	9	Pareto 2	0.11301	9
Gumbel Max	0.17912	10	Log-normal (3P)	0.56992	10	Wakeby	0.17753	10
Pareto 2	0.17959	11	Exponential	0.58261	11	Gen. Pareto	0.17753	11
Log-normal	0.18033	12	Inv. Gaussian (3P)	0.58463	12	Log-normal	0.18093	12
Gen. Gamma	0.18208	13	Log-normal	0.58481	13	Gen. Extreme Value	0.20294	13
Exponential	0.18476	14	Normal	0.58692	14	Log-normal (3P)	0.20373	14
Normal	0.19007	15	Pareto 2	0.6027	15	Weibull	0.21382	15
Weibull (3P)	0.19539	16	Log-Logistic	0.63974	16	Log-Pearson 3	0.21998	16
Gen. Gamma (4P)	0.19689	17	Logistic	0.72245	17	Erlang	0.2264	17
Inv. Gaussian (3P)	0.19715	18	Rayleigh (2P)	0.72988	18	Cauchy	0.24404	18
Dagum	0.19848	19	Beta	0.81948	19	Gumbel Max	0.2575	19
Erlang (3P)	0.19883	20	Erlang (3P)	0.88389	20	Beta	0.34419	20
Exponential (2P)	0.20035	21	Cauchy	0.93371	21	Log-Logistic (3P)	0.34786	21
Dagum (4P)	0.2007	22	Erlang	0.96061	22	Normal	0.35885	22
Logistic	0.20373	23	Gamma	0.9632	23	Logistic	0.46609	23
Log-Logistic	0.20862	24	Gumbel Min	1.1339	24	Gen. Logistic	0.50005	24
Rayleigh (2P)	0.20867	25	Log-Logistic (3P)	1.9675	25	Inv. Gaussian (3P)	0.51591	25
Gamma (3P)	0.21684	26	Nakagami	1.9863	26	Rayleigh (2P)	0.92738	26
Log-Logistic (3P)	0.22144	27	Pareto	3.2493	27	Rayleigh	1.8694	27
Gamma	0.22677	28	Exponential (2P)	3.3369	28	Nakagami	2.169	28
Erlang	0.24013	29	Weibull (3P)	4.1672	29	Gumbel Min	3.1508	29
Cauchy	0.24242	30	Gamma (3P)	4.3524	30	Pareto	N/A	
Gumbel Min	0.25587	31	Dagum	4.5253	31	Weibull (3P)	N/A	
Burr (4P)	0.25883	32	Dagum (4P)	4.5474	32	Gamma (3P)	N/A	
Nakagami	0.28059	33	Rayleigh	4.5483	33	Dagum	N/A	
Pareto	0.31064	34	Burr (4P)	4.5487	34	Dagum (4P)	N/A	
Rayleigh	0.31322	35	Gen. Gamma (4P)	4.5837	35	Burr (4P)	N/A	
Inv. Gaussian	0.31625	36	Inv. Gaussian	7.3891	36	Gen. Gamma (4P)	N/A	
Log-Gamma	No fit		Log-Gamma	No fit		Log-Gamma	No fit	

Table S6
Results of goodness-of-fit and ranking of distribution functions based on goodness-of-fit for annual rainfall during the investigation period (1981–2017)

Distribution	K-S		Distribution	A-D		Distribution	C-s	
	Statistic	Rank		Statistic	Rank		Statistic	Rank
Burr	0.07732	1	Wakeby	0.29879	1	Nakagami	0.53549	1
Gen. Extreme Value	0.0816	2	Gen. Logistic	0.37738	2	Log-normal (3P)	0.6523	2
Normal	0.08412	3	Burr	0.37822	3	Gen. Logistic	1.1775	3
Gen. Logistic	0.08535	4	Normal	0.39018	4	Gamma (3P)	1.2737	4
Log-Logistic (3P)	0.08731	5	Gen. Extreme Value	0.39429	5	Log-Pearson 3	1.3381	5
Inv. Gaussian (3P)	0.08744	6	Log-Logistic (3P)	0.39544	6	Log-Logistic (3P)	1.5917	6
Nakagami	0.08918	7	Logistic	0.39936	7	Beta	1.7568	7
Weibull	0.09055	8	Nakagami	0.41665	8	Gumbel Min	2.2937	8
Log-normal (3P)	0.09082	9	Log-normal (3P)	0.42189	9	Burr	2.346	9
Wakeby	0.09217	10	Weibull	0.42337	10	Gen. Extreme Value	2.3534	10
Logistic	0.09549	11	Gamma (3P)	0.43989	11	Weibull	2.3622	11
Gamma (3P)	0.09555	12	Log-Pearson 3	0.4508	12	Inv. Gaussian (3P)	2.438	12
Erlang (3P)	0.09616	13	Inv. Gaussian (3P)	0.48312	13	Erlang (3P)	2.4805	13
Log-Pearson 3	0.10024	14	Weibull (3P)	0.52243	14	Logistic	3.0587	14
Beta	0.10567	15	Gen. Gamma	0.52714	15	Weibull (3P)	3.9955	15
Weibull (3P)	0.10591	16	Gamma	0.54083	16	Normal	4.0488	16
Gamma	0.10923	17	Erlang (3P)	0.57159	17	Rayleigh (2P)	4.0579	17
Gen. Gamma	0.11076	18	Rayleigh (2P)	0.57196	18	Gamma	4.0803	18
Gen. Pareto	0.11296	19	Log-normal	0.72039	19	Gen. Gamma	4.1177	19
Rayleigh (2P)	0.11461	20	Cauchy	0.72429	20	Wakeby	4.1652	20
Gumbel Min	0.12598	21	Log-Gamma	0.81103	21	Inv. Gaussian	4.3052	21
Inv. Gaussian	0.12892	22	Log-Logistic	0.81345	22	Log-normal	4.4501	22
Cauchy	0.1302	23	Erlang	0.84891	23	Gumbel Max	4.8852	23
Log-normal	0.13199	24	Inv. Gaussian	0.88213	24	Cauchy	6.0936	24
Gumbel Max	0.13468	25	Gumbel Max	1.0251	25	Log-Gamma	6.2887	25
Log-Gamma	0.13971	26	Gumbel Min	1.4881	26	Erlang	6.5129	26
Log-Logistic	0.15114	27	Gen. Gamma (4P)	1.8066	27	Exponential (2P)	6.7738	27
Erlang	0.15358	28	Rayleigh	2.1297	28	Log-Logistic	6.9149	28
Rayleigh	0.1569	29	Beta	2.5791	29	Pareto	7.5665	29
Gen. Gamma (4P)	0.16354	30	Exponential (2P)	4.246	30	Rayleigh	7.9162	30
Exponential (2P)	0.22006	31	Pareto	6.1436	31	Gen. Gamma (4P)	10.329	31
Pareto	0.28493	32	Exponential	7.8538	32	Burr (4P)	13.98	32
Burr (4P)	0.35555	33	Pareto 2	7.91	33	Dagum	19.318	33
Exponential	0.37259	34	Burr (4P)	9.0215	34	Pareto 2	28.531	34
Pareto 2	0.37411	35	Gen. Pareto	11.505	35	Exponential	33.733	35
Dagum	0.43707	36	Dagum (4P)	18.769	36	Gen. Pareto	N/A	
Dagum (4P)	0.50285	37	Dagum	23.162	37	Dagum (4P)	N/A	

Table S7
Distribution parameters for selected distribution functions

Distribution	Parameters
Beta	α_1 = 1.0065 α_2 = 1.606 a = 34.98 b = 159.0
Burr	k = 3.5844 α = 3.3361 β = 131.27
Burr (4P)	k = 0.12384 α = 7.0890E+8 β = 3.6034E+9 γ = –3.6034E+9
Cauchy	σ = 19.408 μ = 81.938
Dagum	k = 0.62928 α = 5.3207 β = 93.317
Dagum (4P)	k = 48.987 α = 608.61 β = 15,976.0 γ = –16008.0
Erlang	m = 6 β = 12.395
Erlang (3P)	m = 3 β = 18.53 γ = 22.861
Exponential	λ = 0.01175
Exponential (2P)	λ = 0.02003 γ = 35.2
Gamma	α = 6.8684 β = 12.395
Gamma (3P)	α = 3.3607 β = 18.53 γ = 22.861
Gen. Extreme Value	k = –0.08953 σ = 28.859 μ = 70.84
Gen. Gamma	k = 1.0034 α = 6.9151 β = 12.395
Gen. Gamma (4P)	k = 2.8697 α = 0.24858 β = 106.87 γ = 35.2
Gen. Logistic	k = 0.11366 σ = 18.144 μ = 81.69
Gen. Pareto	k = –0.59176 σ = 76.468 μ = 37.094
Gumbel Max	σ = 25.328 μ = 70.514
Gumbel Min	σ = 25.328 μ = 99.753
Inv. Gaussian	λ = 584.73 μ = 85.134
Inv. Gaussian (3P)	λ = 924.59 μ = 100.65 γ = –15.516
Log-Gamma	α = 121.69 β = 0.03592
Log-Logistic	α = 4.1605 β = 77.579
Log-Logistic (3P)	α = 5.0691 β = 92.657 γ = –12.064
Log-Pearson 3	α = 72.346 β = –0.04658 γ = 7.7407
Logistic	σ = 17.91 μ = 85.134
Log-normal	σ = 0.39082 μ = 4.3707
Log-normal (3P)	σ = 0.31172 μ = 4.5945 γ = –18.655
Nakagami	m = 1.7989 Ω = 8,274.4
Normal	σ = 32.484 μ = 85.134
Pareto	α = 1.2351 β = 35.2
Pareto 2	α = 102.57 β = 8,588.3
Rayleigh	σ = 67.927
Rayleigh (2P)	σ = 47.365 γ = 26.311
Wakeby	α = 99.946 β = 3.2428 γ = 35.282 δ = –0.1621 ξ = 31.216
Weibull	α = 2.9493 β = 93.2
Weibull (3P)	α = 1.6929 β = 60.225 γ = 31.239

Table S8
Results of goodness-of-fit and ranking of distribution functions based on goodness-of-fit for maximum monthly rainfall

Distribution	K-S		Distribution	A-D		Distribution	C-s	
	Statistic	Rank		Statistic	Rank		Statistic	Rank
Nakagami	0.07436	1	Wakeby	0.20183	1	Cauchy	0.09775	1
Wakeby	0.07454	2	Gen. Extreme Value	0.24554	2	Nakagami	0.64303	2
Normal	0.08257	3	Nakagami	0.26162	3	Weibull	0.64437	3
Gen. Extreme Value	0.08306	4	Gamma	0.26273	4	Burr	0.96495	4
Weibull	0.08538	5	Log-Pearson 3	0.26661	5	Logistic	1.2797	5
Burr	0.08618	6	Gen. Gamma	0.27069	6	Gen. Logistic	1.5194	6
Gamma	0.08817	7	Burr	0.28012	7	Normal	1.5726	7
Gen. Gamma	0.0902	8	Weibull (3P)	0.28245	8	Dagum	1.7812	8
Gen. Pareto	0.09092	9	Log-normal (3P)	0.29499	9	Erlang (3P)	2.8362	9
Gen. Logistic	0.09346	10	Inv. Gaussian (3P)	0.29705	10	Gumbel Min	2.8613	10
Log-Pearson 3	0.09447	11	Rayleigh (2P)	0.29725	11	Beta	3.0976	11
Rayleigh (2P)	0.09543	12	Dagum	0.30075	12	Gen. Extreme Value	3.0986	12
Log-normal (3P)	0.09594	13	Dagum (4P)	0.30895	13	Wakeby	3.2942	13
Dagum	0.0964	14	Gamma (3P)	0.31285	14	Gamma	3.3005	14
Log-Logistic (3P)	0.09947	15	Gen. Logistic	0.31321	15	Gen. Gamma	3.3217	15
Inv. Gaussian (3P)	0.09948	16	Log-Logistic (3P)	0.33307	16	Rayleigh (2P)	3.6221	16
Logistic	0.1	17	Log-normal	0.34033	17	Erlang	3.8125	17
Weibull (3P)	0.10047	18	Weibull	0.39943	18	Burr (4P)	3.9917	18
Dagum (4P)	0.10336	19	Log-Gamma	0.40125	19	Log-Logistic	4.3953	19
Gumbel Max	0.10639	20	Gumbel Max	0.42166	20	Log-Gamma	4.8903	20
Inv. Gaussian	0.10809	21	Log-Logistic	0.43312	21	Log-Logistic (3P)	5.1431	21
Gamma (3P)	0.10935	22	Inv. Gaussian	0.44089	22	Log-Pearson 3	5.1993	22
Beta	0.10941	23	Normal	0.44407	23	Pareto	5.2072	23
Log-normal	0.11047	24	Logistic	0.51118	24	Dagum (4P)	5.2095	24
Log-Gamma	0.12171	25	Cauchy	0.66826	25	Log-normal (3P)	5.2129	25
Cauchy	0.12528	26	Burr (4P)	0.80399	26	Inv. Gaussian (3P)	5.264	26
Rayleigh	0.12629	27	Beta	1.0363	27	Gumbel Max	5.4138	27
Log-Logistic	0.13119	28	Rayleigh	1.1749	28	Log-normal	5.4162	28
Gumbel Min	0.14536	29	Erlang (3P)	1.291	29	Gamma (3P)	5.4906	29
Burr (4P)	0.17003	30	Gumbel Min	2.1003	30	Inv. Gaussian	5.5802	30
Gen. Gamma (4P)	0.18	31	Erlang	2.4318	31	Weibull (3P)	5.6437	31
Exponential (2P)	0.18851	32	Exponential (2P)	2.529	32	Rayleigh	6.2207	32
Erlang (3P)	0.19223	33	Gen. Gamma (4P)	5.4112	33	Exponential (2P)	14.795	33
Erlang	0.22123	34	Pareto	5.9708	34	Exponential	21.985	34
Pareto	0.26445	35	Exponential	6.6722	35	Pareto 2	22.175	35
Exponential	0.33865	36	Pareto 2	6.8205	36	Gen. Gamma (4P)	N/A	
Pareto 2	0.34288	37	Gen. Pareto	7.6952	37	Gen. Pareto	N/A	

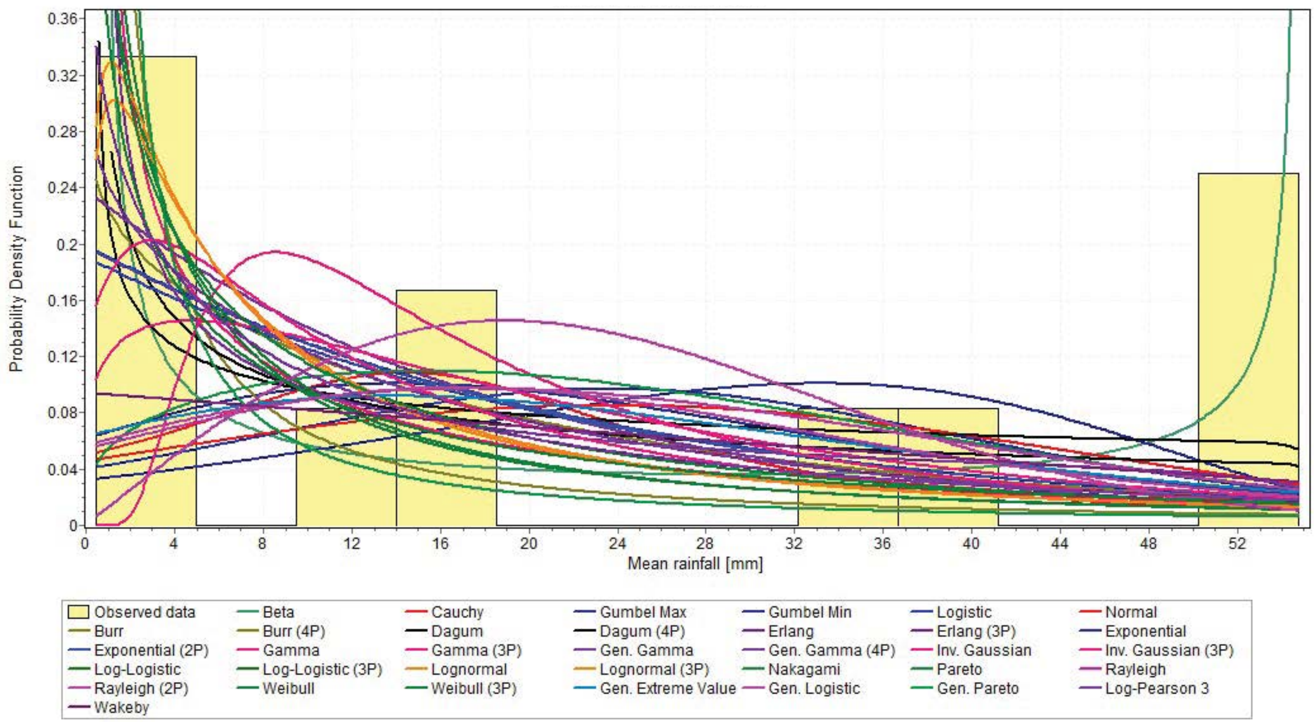

Fig. S1. Frequency histograms and probability density function plots of average rainfall (1981–2017).

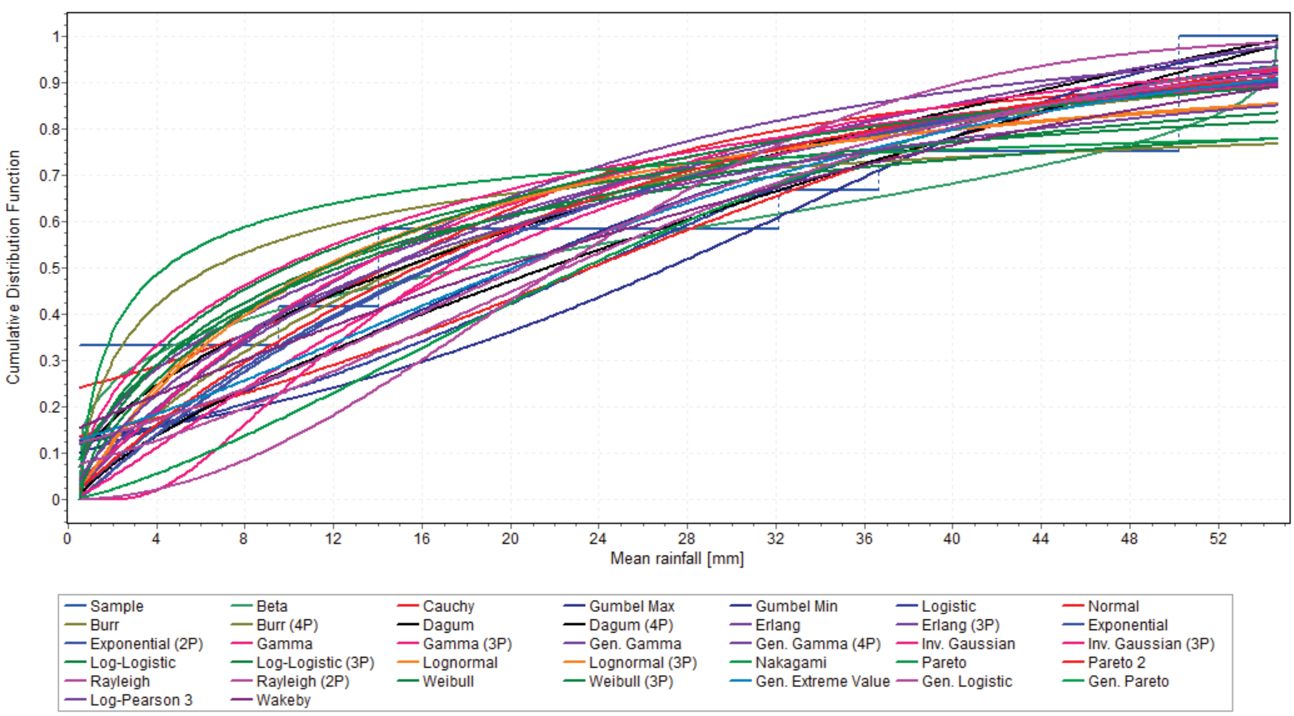

Fig. S2. Frequency histograms and cumulative distribution function plots of average rainfall (1981–2017).

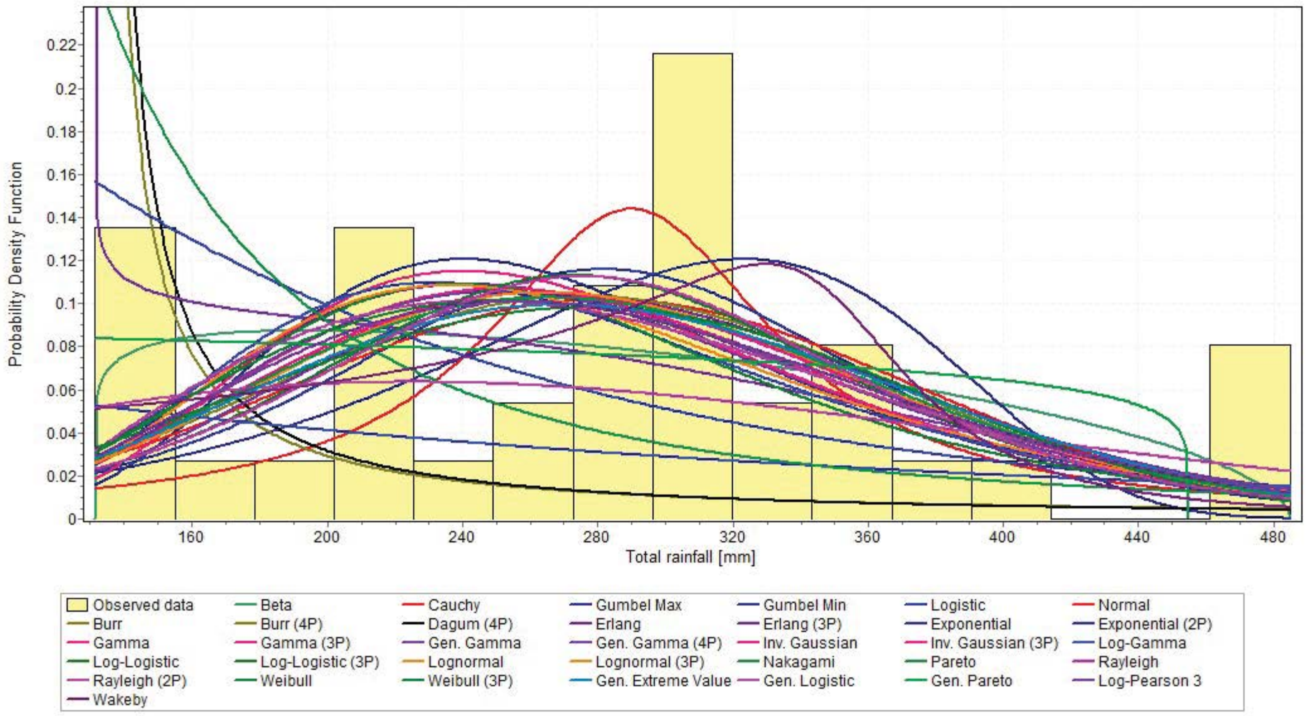

Fig. S3. Frequency histograms and probability density function plots of annual rainfall during the investigation period.

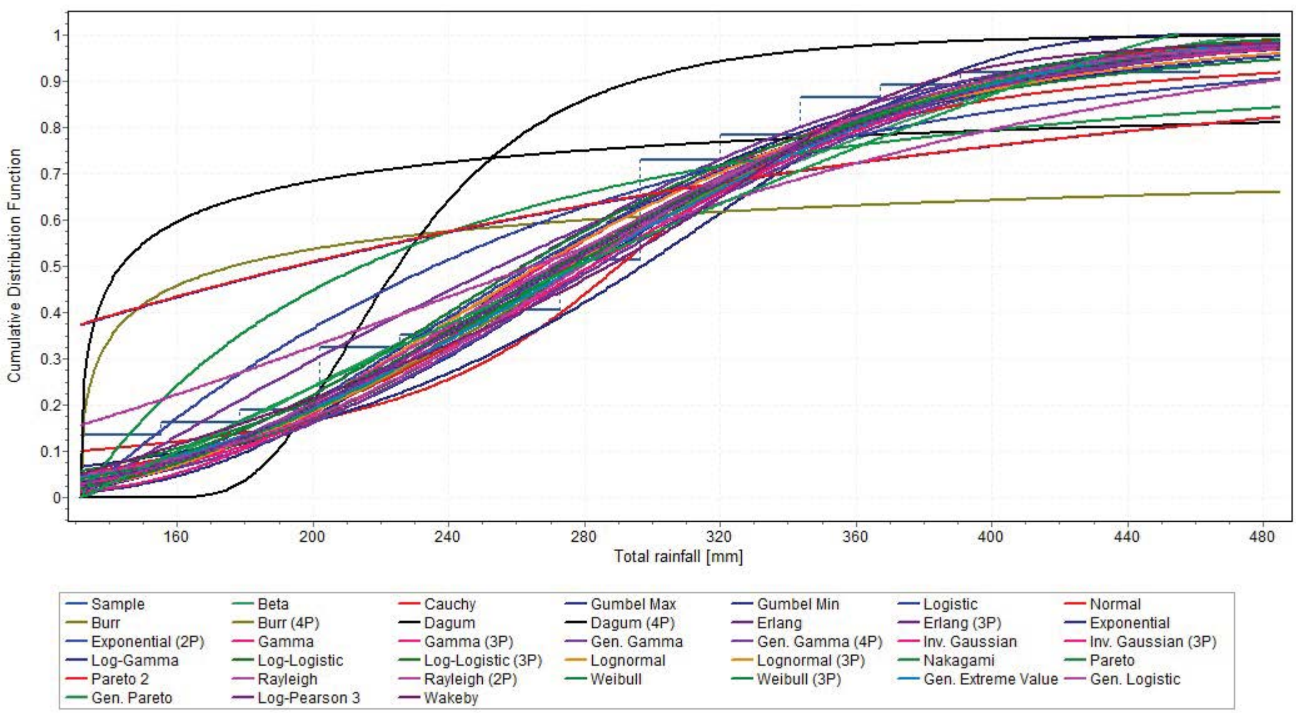

Fig. S4. Frequency histograms and cumulative distribution function plots of annual rainfall during the investigation period.

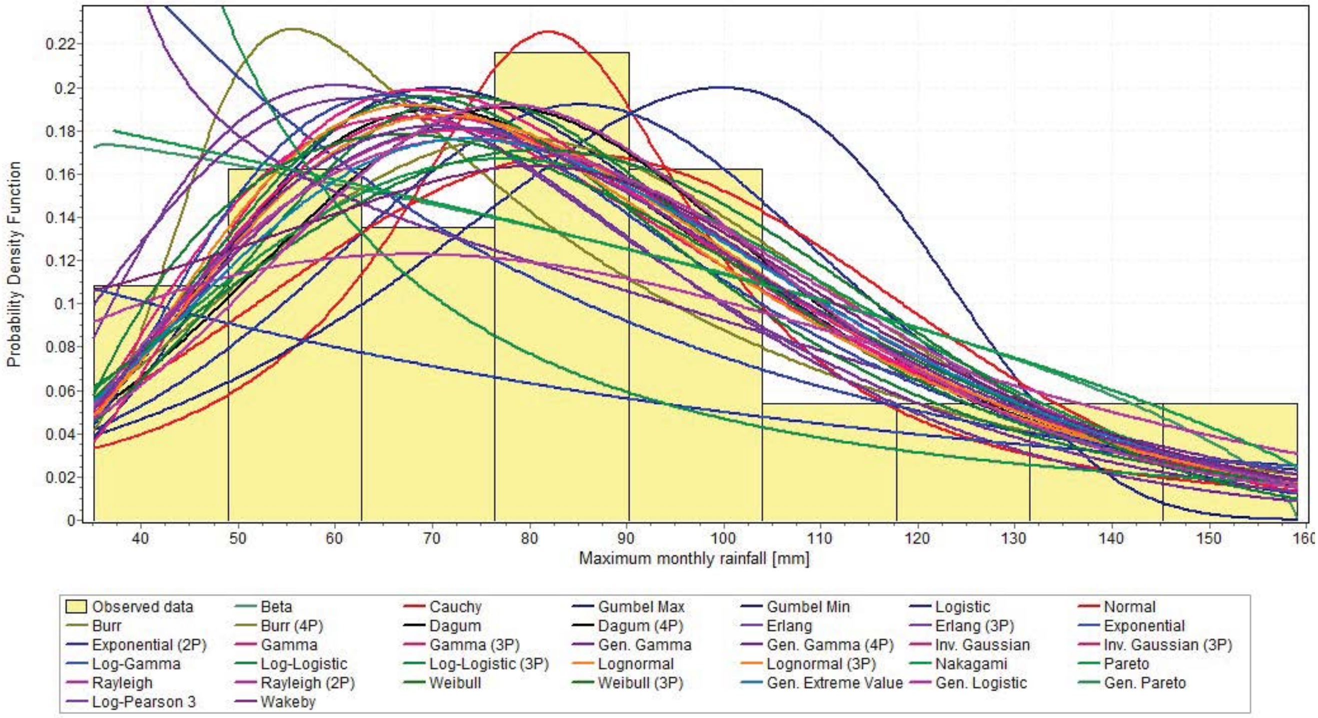

Fig. S5. Frequency histograms and probability density function plots of maximum monthly rainfall.

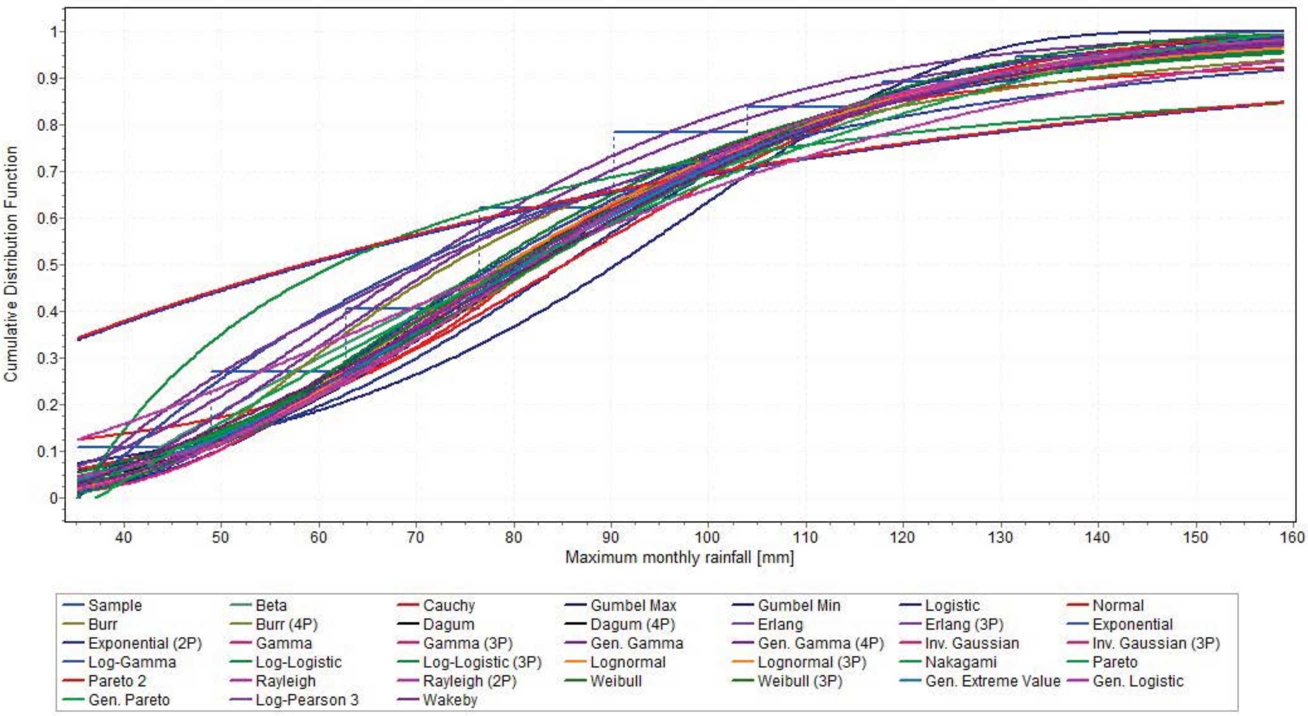

Fig. S6. Frequency histograms and cumulative distribution function plots of maximum monthly rainfall.