T.C.
IŞIK UNIVERSITY
SCHOOL OF GRADUATE STUDIES

MASTER THESIS
DEPARTMENT OF ENGINEERING AND NATURAL SCIENCES CIVIL ENGINEERING PROGRAM

Süreyya KALYONCU

FLOOD ANALYSIS IN THE ÇORUH RIVER BASIN

SUPERVISOR
Prof. Dr. Bihrat ÖNÖZ

İSTANBUL, June 2025

T.C.
IŞIK UNIVERSITY
SCHOOL OF GRADUATE STUDIES

MASTER THESIS
DEPARTMENT OF ENGINEERING AND NATURAL SCIENCES CIVIL ENGINEERING PROGRAM

Süreyya KALYONCU

(22CIVL5002)

FLOOD ANALYSIS IN THE ÇORUH RIVER BASIN

SUPERVISOR

Prof. Dr. Bihrat ÖNÖZ

İSTANBUL, June 2025

T.C.
IŞIK UNIVERSITY
SCHOOL OF GRADUATE STUDIES

MASTER'S THESIS
DEPARTMENT OF ENGINEERING AND NATURAL SCIENCES CIVIL ENGINEERING PROGRAM

Süreyya KALYONCU

(22CIVL5002)

FLOOD ANALYSIS IN THE ÇORUH RIVER BASIN

Date:

Thesis Supervisor: Prof. Dr. Bihrat ÖNÖZ / Işık University

Jury Members:

Prof. Dr. Emine Beyhan YEĞEN / MEF University

Asst. Prof. Dr. Bora AKŞAR / Işık University

İSTANBUL, June 2025

ÖZET

ÇORUH NEHRİ HAVZASINDA TAŞKIN ANALİZİ

Taşkın analizi, sürdürülebilir su kaynakları yönetimi ve afet riskinin azaltılmasında kritik bir rol oynar. Taşkınların sıklığını, büyüklüğünü ve zamanlamasını anlamak, insan hayatını korumak, ekonomik kayıpları en aza indirmek ve dayanıklı yerleşimler planlamak için esastır. Çoruh Nehri Havzası gibi karmaşık hidrolojik davranışa sahip bölgelerde, taşkın analizi taşkın kontrol yapıları tasarlamak, erken uyarı sistemleri geliştirmek ve güvenli imar uygulamaları belirlemek için bilimsel bir temel sağlar. Dahası, istatistiksel modelleme yoluyla tasarım taşkınları doğru bir şekilde tahmin etmek, özellikle hidrometeorolojik olayların değişkenliğini ve uç noktalarını yoğunlaştıran iklim değişikliği bağlamında bilinçli kararlar alınmasına yardımcı olur. Bu nedenle, kapsamlı taşkın analizi yalnızca yerel ve bölgesel su yönetimi stratejilerini desteklemekle kalmaz, aynı zamanda uzun vadeli iklim adaptasyon planlamasına da katkıda bulunur.

Çoruh Nehri'nde Taşkın Analizi adlı tez çalışmasında, Türkiye'nin kuzeydoğusunda konumlanan Çoruh Nehri üzerinde çalışılmış, istatistiksel yöntemler kullanılarak taşkın analizi gerçekleştirilmiştir.

Çalışma hazırlanırken DSİ'nin sağladığı istasyon verilerinden yararlanılmış; analizin güvenilir olabilmesi için 10 yıldan fazla gözlem süresine sahip ve inşa edilen barajlardan etkilenmemiş 23 farklı istasyon seçilmiş olup, analiz bu istasyonların maksimum akış verileri kullanılarak hazırlanmıştır.

23 istasyonun her biri için istatistiksel momentler, L momentler hesaplanmış; Gumbel, Ekstrem Değer Dağılımı, Normal Dağılım, Log-Normal Dağılım, 3 parametreli Log Normal Dağılımı ve 3 parametreli Log Pearson dağılımları ilgili verilere uyarlanmış ve dağılımların uygunlukları Kolmogorov-Smirnov ve Ki kare testi gibi farklı istatistiksel testler kullanılarak kontrol edilmiştir. Ayrıca uyarlanan her bir dağılım için 2, 5, 10, 25, 50, 100 ve 500

yıllık geri dönüş periyotları için taşkın debileri tahmin edilip, grafikler ile görselleştirilmiştir.

İstasyonlarda meydana gelen taşkınların bir trende sahip olup olmadıkları Mann Kendall testi ile araştırılmış, istasyonların çarpıklıklarına bakılarak ve Wiltshire metodu kullanılarak Çoruh Nehri homojen bölgelere ayrılmış ve istasyonlarda gerçekleşen taşkın verilerinin mevsimsellikleri de açısal mevsimsellik analizi kullanılarak tespit edilmiştir.

Anahtar Kelimeler: Taşkın Analizi, Çoruh Nehri

ABSTRACT

FLOOD ANALYSIS IN THE ÇORUH RIVER BASIN

In this thesis, the flood characteristics of the Coruh River located in the northeast of Türkiye were investigated by statistical methods. In the study, annual maximum flow data obtained from DSI were used; 23 flow observation stations that were not under the influence of dams and had at least 10 years of observation period were evaluated.

Statistical moments, L-moments and coefficients derived from these moments were calculated for each station and the structure of the flow data was analyzed in detail. Then, different probability distributions such as Gumbel, Generalized Extreme Value (GEV), Normal (N), Log-Normal (LN), Log Normal III (LN3) and Log Pearson Type III were applied and the suitability of each distribution to the data was tested with Kolmogorov-Smirnov and Chi-Square tests. Flood flow rates were estimated for 2, 5, 10, 25, 50, 100 and 500 year return periods according to the appropriate distributions. In addition, the seasonal timing of floods was assessed and the impact of climatic variability in the region was examined.

Flood analysis contributes directly to disaster risk reduction by providing an estimate of the magnitude of possible floods. The importance of these analyses becomes even more evident when the loss of life and property, damage to infrastructure systems and economic losses caused by floods are considered. Accurate flood forecasts play a critical role in the safe planning of settlements and the development of disaster management strategies. In addition, changes in flood timing provide important clues for understanding the effects of climate change at the local scale. Therefore, detailed and multidimensional analysis of flood characteristics is a fundamental approach that contributes to both understanding regional hydrological behavior and ensuring sustainable water resources management.

The research findings show that comprehensive analyses conducted with different probability distributions in the assessment of flood characteristics allow for a more accurate interpretation of regional hydrological behavior. The obtained results enable scientifically based estimates of flood hazard and provide a strategic basis for planned management of water resources.
Keywords: Flood Analysis, Çoruh River Basin

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to all the individuals and institutions who supported me in the preparation of this thesis study, which provided me with the opportunity to increase my knowledge and continue my academic development during my master's degree.

First of all, I would like to express my endless gratitude to my valuable advisor Prof. Dr. Bihrat ÖNÖZ, who guided me by sharing her knowledge and experiences with me at every stage of my thesis study. Thanks to her academic guidance, constructive criticisms and guidance since the determination of my thesis topic, she not only provided me with the opportunity to complete this study within a scientific framework, but also contributed greatly to the development of my academic perspective. I am grateful to her for patiently guiding me at every stage and increasing my motivation with her belief in me.

I would also like to express my gratitude to Asst. Prof. Ehsan ETMİNAN and other valuable instructors who trained me with their knowledge and experiences throughout my undergraduate and graduate education and contributed to the development of my equipment in the field of engineering. In particular, the course contents and academic approaches that combine theoretical knowledge with practical application have formed the basis of this study.

I would like to express my endless gratitude to my dear family, whose financial and moral support I have always felt throughout my life and who have shared my successes with as much pride and joy as I have. Thanks to their dedication, sacrifice and trust in me, I was able to find the strength to complete this process. I am grateful for always being there for me and believing in me.

In this process where academic life can sometimes be challenging and tiring, I would like to thank my life partner Gülcan İBİŞ, who has kept my morale and motivation high, with whom I can share my thoughts, who has made many sacrifices for me and supported me. The contributions of her, with whom I have exchanged ideas especially during the working processes and with whom we

have continued our academic development together, are very valuable and important to me.

Finally, I would like to thank everyone who has contributed to the creation of this thesis study, in which I aim to make a scientific contribution in the field of engineering.

Süreyya KALYONCU

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS LIST

MD: Mean Date

K-S: Kolmogorov-Smirnov Test

N: Normal Distribution

LN: Log-Normal Distribution

LN3: 3 Parameters Log-Normal Distribution

LP3: Log-Pearson Type III Distribution

GEV: Gumbel Distribution

DSİ: Devlet Su İşleri

HPP: Hydroelectric Power Plant

IPPC: International Panel on Climate Change

A-D: Anderson-Darling Test

PWM: Probability Weighted Moments

CHAPTER 1

1. INTRODUCTION

In surface water resources such as streams, the flow and water level vary both continuously and randomly over time. Floods, one of the most extreme examples of these changes, are among the extreme events in terms of hydrology and largely develop outside the control of natural processes (Bayazıt and Önöz, 2008). Floods are important not only as natural events but also as disasters that harm human life and property. Throughout history, people have built various engineering structures and developed preventive strategies in order to protect themselves from floods (Şen, 2009). Today, significant progress has been made in this area thanks to technological developments, especially computer-aided calculation techniques and statistical modeling methods.

The importance of flood analyses is not limited to predicting whether a flood will occur. At the same time, these analyses play a critical role in estimating the magnitude of the effects that will occur in the event of a flood and determining the structural or non-structural measures that can be taken against these effects. For example, correctly estimating the magnitude of a flood has a direct impact on determining the design parameters of a dam, planning the arrangements to be made in the floodplain, and determining the settlement areas at risk of flooding. Due to the random nature of hydrological events, it is necessary to use statistical methods in the analysis of these events. In this context, frequency analyses, in particular," allow the estimation of flood flow rates with a certain recurrence period based on past observations (Bayazıt, 1998).

The effects of climate change have further increased the importance of flood analyses. Changes in the precipitation regime caused by global warming have increased the frequency and severity of floods in many regions. In the Sixth Assessment Report published by the IPCC in 2021, it was emphasized that increasing temperatures and changing atmospheric conditions around the world

have led to serious increases in extreme hydrological events. Hirabayashi et al. (2013) stated that an increase in the risk of fluvial floods is expected in the coming years, especially in Asia, Africa, and South America. Studies conducted specifically for Turkey also confirm similar trends; it is reported that flood events have become more frequent and intense in recent years, especially in the Black Sea and Mediterranean Regions (Bayazıt and Önöz, 2008).

All these developments show that floods have ceased to be just a natural event and have become a complex disaster type that needs to be managed with multidisciplinary analyses. Therefore, flood analyses include a comprehensive evaluation process that requires the joint contribution of many disciplines such as hydrology, meteorology, statistics and disaster management.

1.1 THE PURPOSE OF THE THESIS

The main purpose of this thesis is to analyze flood behaviors on a station basis in the Çoruh River Basin, which is located in the eastern part of Turkey and has great hydrological importance. Floods are natural disasters that cause both loss of life and property and also threaten the ecological balance. Therefore, correctly analyzing the spatial and temporal behaviors of floods and predicting them at certain return intervals are of critical importance in terms of both engineering applications and disaster management.

In this context, data from 23 different flow observation stations in the Çoruh River Basin were analyzed using long-term daily flow data provided by the State Hydraulic Works (DSİ). For each station, various probability distributions such as Gumbel, Log-Normal, Pearson Type III, which are widely used in the literature in flood frequency analysis, were applied and maximum flow values corresponding to certain return intervals (such as 5, 10, 25, 50, 100 and 500 years) were estimated. In addition, in order to ensure the reliability of these estimates, appropriateness tests such as Kolmogorov–Smirnov (K–S) and Chi-Squared tests were performed.

In the study, not only the flood magnitudes but also the occurrence times of the floods during the year were analyzed. In this direction, the periods of occurrence of floods were determined based on four seasons and the effect of seasonality on flood formation was evaluated. In addition, the changes in flood characteristics at the stations over the years were examined with time series analyses and the relationship of these changes with possible climatic or environmental effects was interpreted.

The ultimate goal of this study is to contribute to the understanding of regional flood risks, to determine the most appropriate probability models that can be used in flood prediction and thus to provide a scientific basis for water resources management, engineering designs and flood prevention strategies. The findings obtained will be an important reference for local governments and engineering practices in future flood planning and early warning systems.

1.2 CHARACTERISTICS OF THE ÇORUH RIVER BASIN

The Çoruh River Basin is located in the northeast of Turkey, covering large parts of the provinces of Artvin, Erzurum and Bayburt, and small parts of the provinces of Kars and Erzincan. The river has an area of approximately 20,488 km², which corresponds to 2.61% of Türkiye's total surface area. The river is surrounded by the Eastern Black Sea Mountains to the north, the Giresun Mountains to the west, the Otlukbeli, Dumlu, Kargapazarı, Güllü and Allahüekber Mountains to the south, and the Yalnızçam Mountains and the Georgian border to the east.

The Çoruh River Basin is the main stream of the river, taking its source from the Mescit Mountains within the borders of Erzurum province. The river passes through the cities of Bayburt, İspir, Yusufeli and Artvin, crosses the Georgian border and flows into the Black Sea south of Batumi. The total length of the Çoruh River Basin is 431 km, 411 km of which is within the borders of Turkey. One of the fastest flowing rivers in Türkiye, the Çoruh carries approximately 5.8 million m³ of sediment per year.

The climate of the river bears the characteristics of the Black Sea and Eastern Anatolia regions. For this reason, it is a region that receives plenty of rainfall. In high altitude areas, heavy snowfall is seen in the winter months, which leads to seasonal changes in the flow of the rivers. The annual average precipitation is around 540 mm.

The Çoruh River Basin is also of great importance in terms of hydroelectric energy production. There are many hydroelectric power plants on the river, such as the Deriner, Yusufeli, Artvin, Borçka and Muratlı dams. These dams play critical roles in terms of flood control and water management, as well as energy production.

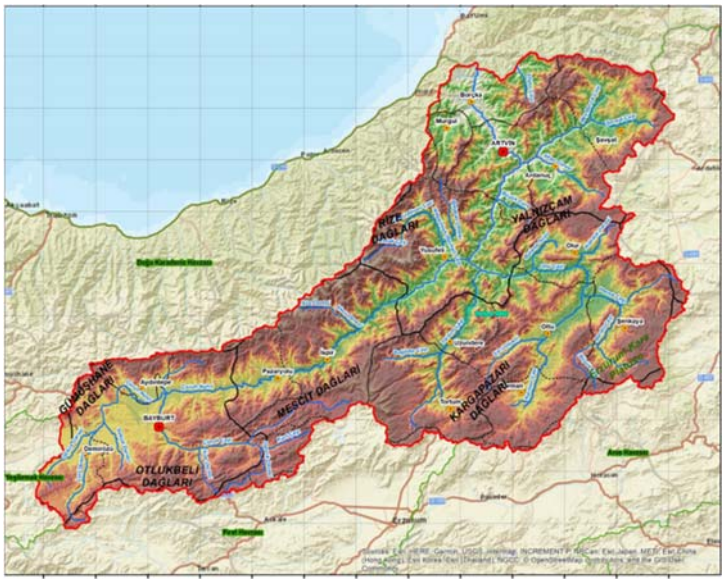


Figure 1.1 Çoruh River Basin

1.3 LITERATURE REVIEW

It is important to recognize water resources, which have a great place in our lives and are important for all living things. In order to minimize the damage caused by flood disasters that have been going on for years and to prevent loss of life and property, people have developed different techniques over the years, and in recent years, they have tried to predict future floods and foresee disasters.

1.3.1 Determining the Most Appropriate Statistical Distribution for Data

Finding the most appropriate statistical distribution for flood flows is of great importance in terms of predicting future floods. One of the oldest studies on this subject is Benson (1968). During this study, 10 different stations in the USA were examined and flood estimates were made on them. In the results obtained by testing the flood analyses, it was found that the Log-Pearson Type III (LP3) distribution best predicted flood values.

In another study conducted in Italy, Cicioni et al (1973) conducted flood analyses on 108 different stations in their country and tested these analyses with Chi-squared, Kolmogorov-Smirnov and Anderson Darling tests; as a result, they revealed that the Log Normal (LN), 3-parameter Log Normal (LN3) and General Extreme Value (GEV) distributions were the most appropriate distributions.

In another study, Beard (1974) conducted flood analyses for 300 different stations in the USA and selected the most appropriate distributions by looking at the 1000-year flood flow rates of each station. The most appropriate distributions selected in this study were the LP3 and LN distributions.

McMahon and Srikanthan (1981) conducted flood analyses on 172 stations in Australia, compared different distributions using the L moment method, and concluded that the LP3 distribution was the most appropriate distribution.

Similarly, in the study conducted by Vogel at al. (1993), 383 stations in the USA were examined, flood analyses were performed, and different distributions were compared using L moment diagrams. As a result of this study, it was concluded that the LP3, LN3, and GEV distributions were the most appropriate distributions for flood analysis.

In the study conducted by Önöz and Bayazıt (1995), 19 stations selected from different parts of the world were examined, flood analyses were performed and tested with Chi-squared, K-2, Probability Plot Correlation Coefficient (PPCC) and A-D tests. As a result of this study, it was concluded that the most suitable distribution for flood analysis was the GEV distribution.

1.3.2 Trend Analysis

One of the methods used to determine whether there is an increase or decrease in flood flow rates in water resources over the years is trend analysis. In the study conducted by Van Belle et al. (1984), they developed parametric tests to determine whether there is an increase or decrease in flood flow rates. In another study conducted by Hirsch (1984), the Mann-Kendall test was developed, and the seasonal Kendall test was introduced.

Kahya et al. (2004) examined the 31-year flood data of 26 different river basins in Turkey and applied trend analyses. After the applied trend analysis, a decreasing trend was observed in the western basins of Turkey, while no significant trend was found in the eastern basins.

In another study conducted in Turkey, Yıldız et al. (2004) analyzed the flows and trends in rivers and examined the effects of these trends on hydroelectric energy production.

Dinpashoh et al. (2011) applied Mann Kendall trend analysis on evapotranspiration data in Iran and reported that a time-dependent increase in evaporation trends was observed.

The principle of dividing stations into regions by considering their hydrological similarities is called regional flood analysis. The first study we will mention in this field is the one carried out by Lettenmaier and Potter (1985). In their study, they applied regional frequency analysis by looking at flood flow rates. They also examined the relationship between flood flow rates and the river.

1.3.3 Regional Flood Analysis

In another study carried out by Önöz (1991), regional frequency analysis was carried out on the stations in the Yeşilırmak Basin and regional frequency curves were created using various statistical distributions.

Gedikli (1994) carried out regional flood analyses on the stations in the Fırat and Tigris basins in his study and as a result, he divided the region into two sub-regions and obtained homogeneous flood zones.

Saf (1995) applied regional flood analysis to the stations located on the Western Mediterranean Basin in his study and as a result, he divided the region into three homogeneous sub-regions as Lower-Western Mediterranean, Mediterranean and Upper-Western Mediterranean.

Ayker (1995) applied regional flood analysis using stations with 40 years of data located on the Büyük Menderes Basin in his study.

In the study conducted by Bayazıt and Önöz (1994), they obtained flood envelope curves for river basins in Turkey by adding data up to 2000 to their studies conducted by DSİ based on data up to 1990. These flood envelope curves can also be used in estimating floods.

1.3.4 Seasonality Analysis

In the study conducted by Quarda et al. (2006) in Canada, three flood seasonality analyses were applied on the determined stations, and this method was compared with the traditional method.

Another study conducted in Canada was conducted by Burn (1997). In this study, seasonality analyses were applied on the stations located in Manitoba and Saskatchewan regions of Canada. As a result of the study, basins with similar hydrological characteristics were determined.

In the study conducted by McCuen and Beighley (2003), seasonality analyses were performed on the stations in Maryland, America and as a result, they observed that there was a very large difference between the flow obtained from the seasonality analysis and the annual flow; based on this, it was mentioned that using seasonality analysis to complete the deficiencies in the stations where annual measurements were missing could yield better results.

CHAPTER 2

2. METHODOLOGY

In this study, the methods suggested in the literature were used by using the annual flood data of 23 stations. All the methods and techniques used in the study are explained in the following section.

2.1 STATISTICAL ANALYSIS

The annual maximum flow rate of streams is a quantity that can vary each year depending on meteorological and hydrological factors. By examining the maximum flow values recorded in previous years, future flood flow rates can be estimated with the help of probability distribution functions suitable for these values. The parameters in the probability distribution functions reflect the basic statistical properties of the distributions. In this context, three basic properties come to the fore: the center of the distribution (mean or median), the spread around the center (standard deviation or coefficient of variation) and the skewness (asymmetry) of the distribution.

Two of the most commonly used estimation methods in flood analyses are classical statistical moments and L-moments. These methods allow for a better understanding of the structure of the data and for making decisions on the selection of the appropriate probability distribution. In addition, one of the approaches used in flood estimation analyses is parametric statistical methods. In this study, different probability distributions were evaluated to be used in flood estimation and their suitability was tested with various statistical tests and analyses were performed.

2.1.1 Statistical Moments of Random Variables

If we accept the part of a random variable between the x-axis and the probability distribution function as a mass, the moments of this mass taken by some axes are accepted as statistical moments (Bayazıt, 1996). The formula we will use to find the m^{th} order statistical moment of a random variable is given in equation (2.1) below.

$$\mu_m = E[(X-\mu)^m] = \int_{-\infty}^{\infty} (x-\mu_x)^m f(x)\ dx \tag{2.1}$$

The X value in the formula represents the random variable, the μ_x value represents the mean, and the E(…) value represents the expected value.

The parameters in the probability distribution functions are the values that reflect the statistical properties of the distributions. There are three different features that we will touch on here: the center of the distribution, the size of the spread in the center and its skewness. The two most commonly used methods in flood analysis estimation methods are statistical moments and L-moments.

2.1.1.1 Center Parameters

The first of the center parameters is the mean. The value in the middle of the cluster of random values is called the center of the distribution or mean. The mean is defined as in the equation (2.2):

$$\mu_x = \int_{-\infty}^{\infty} x\, f(x)\, dx \tag{2.2}$$

The mean of the continuous random variable X with n elements is defined as in the equation (2.3) (Bayazıt and Yeğen Oğuz, 2005).

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{2.3}$$

The second of the central parameters is the median. The median is found by sorting all the data in the random variable from smallest to largest and calculating the middle value. In a distribution with n data points, the median formula differs depending on whether n is even or odd:

If n is odd:

$$\mathrm{Med}_x = x_{\left(\frac{n+1}{2}\right)} \tag{2.4}$$

If n is even:

$$\mathrm{Med}_x = \frac{1}{2}\left(x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}\right) \tag{2.5}$$

When we compare the median and mean, we see that the median is less affected by outliers than the mean. If a distribution is symmetric, the median and mean are equal, but if the distribution is skewed to the right or left, this situation changes. In right-skewed distributions, the median is smaller than the mean, while in left-skewed distributions, the median is larger than the mean (Bayazıt and Yeğen Oğuz, 2005).

2.1.1.2 Spreading Parameters

The first parameter is variance. Parameters that indicate the size of the spread around the center point of a random distribution are called spread parameters. The most commonly used parameter when measuring the size of the spread in distributions is variance. Variance is obtained by taking the squares of the differences between the data values and the mean and adding each of them. Variance is also referred to as the second-order moment of the probability distribution and is defined as in the equation (2.6):

$$\mathrm{Var}(x) = \sigma_x^2 = E[(X - \mu_x)^2] = \int_{-\infty}^{\infty} (x - \mu_x)^2 f(x)\, dx \tag{2.6}$$

The variance of a continuous random variable with n elements can be calculated as follows:

$$\mathrm{Var}(x) = s_x^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 \tag{2.7}$$

If the number of variables n in the equation is less than 30, n-1 is used instead of n in the denominator (Bayazıt and Önöz, 2008; Bayazıt and Yeğen Oğuz, 2005).

The second parameter is standart deviation. A parameter that measures the spread is the standard deviation, which is equal to the square root of the variance. The standard deviation is calculated as in the equation (2.8).

$$S_x = \sqrt{\mathrm{Var}(x)} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{2.8}$$

The larger the standard deviation of a probability distribution, the larger the spread of that distribution. Although the standard deviation is used to determine the spread of a random variable, comparing the standard deviations is not enough to compare the spreads of two different random variables. It is more accurate to use the coefficient of variation when comparing the spreads of two different random variables. The coefficient of variation of a random variable is found by dividing its standard deviation by its mean. The coefficient of variation is defined as in the equation (2.9).

$$C_{vx} = \frac{\sigma_x}{\mu_x} \tag{2.9}$$

The coefficient of variation of a random variable with n elements is expressed as in the below equation (2.10) (Bayazıt and Önöz, 2008; Bayazıt and Yeğen Oğuz, 2005).

$$C_{vx} = \frac{S_x}{\bar{x}} \tag{2.10}$$

2.1.1.3 Skewness Parameters

Another parameter that is the third-degree statistical moment of the random variable is the skewness. The skewness helps us determine the symmetry of the random variable around the center.

We use the following equation to determine the skewness coefficient:

$$C_{sx} = \frac{\mu_x^3}{\sigma_x^3} \tag{2.11}$$

When determining the skewness coefficient of a continuous random variable with a value of n, we use the formula below formula (2.12).

$$C_{sx} = \frac{n}{(n-1)(n-2)} \cdot \frac{\sum_{i=1}^{n}(x_i - \bar{x})^3}{S_x^3} \tag{2.12}$$

The skewness coefficient, which helps us find symmetry, is a dimensionless parameter. If the skewness coefficient of a distribution is 0, we can say that the distribution is symmetrically distributed around the center. If the skewness coefficient is greater than zero, we can say that the distribution is skewed to the right; if it is less than zero, we can say that the distribution is skewed to the left (Bayazıt and Önöz, 2008; Bayazıt and Yeğen Oğuz, 2005).

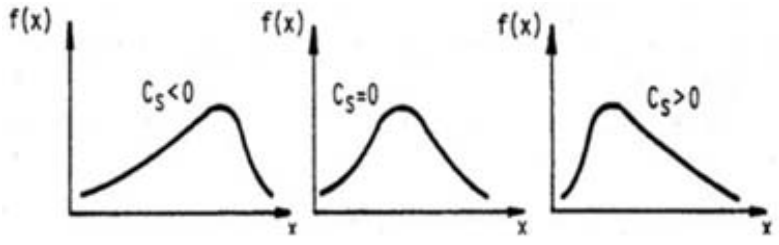


Figure 2.1 Representation of negative, symmetric and positive skewness (Bayazıt and Yeğen Oğuz,2005).

2.1.1.4 L-Moments

L moments, which are linear functions of Probability Weighted Moments (PWM), are used to express the main properties of distributions (Vogel, McMohon and Chiew, 1993). We obtain PWM with the following formula (2.13):

$$\beta_r = E\{x[F_x(x)]^r\} \tag{2.13}$$

Here, F(x) is the additional distribution function of X, while r is the order of PWM. When the r value is 0, β_0 is equal to the mean of the distribution. Another formula used for calculating β_r is given below in (2.14).

$$\beta_r = \frac{1}{N}\sum_{j=1}^{N-r}\frac{\binom{N-j}{r}x_j}{\binom{N-1}{r}} \tag{2.14}$$

By looking at this equation, the first 4 order formulas are found as follows:

$$\beta_0 = \frac{1}{N}\sum_{j=1}^{N} x_j = \bar{x} \tag{2.15}$$

$$\beta_1 = \sum_{j=1}^{N-1}\frac{(N-j)x_j}{N(N-1)} \tag{2.16}$$

$$\beta_2 = \sum_{j=1}^{N-1}\frac{(N-j)(N-j-1)x_j}{N(N-1)(N-2)} \tag{2.17}$$

$$\beta_3 = \sum_{j=1}^{N-1}\frac{(N-j)(N-j-1)(N-j-2)x_j}{N(N-1)(N-2)(N-3)} \tag{2.18}$$

The x_j expression in the equations (2.15), (2.16), (2.17) and (2.18) is the j value obtained when the data in the random variable are sorted from smallest to largest.

Based on this explanation, the L moment values are calculated as follows:

$$\lambda_1 = \beta_0 \tag{2.19}$$

$$\lambda_2 = 2\beta_1 - \beta_0 \tag{2.20}$$

$$\lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0 \tag{2.21}$$

$$\lambda_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0 \tag{2.22}$$

The general equation for any value of r is as follows:

$$\lambda_{r+1} = \sum_{k=0}^{r} \beta_k (-1)^{r-k} \binom{r}{k} \binom{r+k}{k} \tag{2.23}$$

Using L moments, we can calculate the coefficient of variation C_{vx}, the coefficient of skewness C_{sx} and the kurtosis coefficient K_s that we just mentioned. We make these calculations with the following formulas:

$$\tau_2 = \frac{\lambda_2}{\lambda_1}, \quad \text{L-change parameter;} \tag{2.24}$$

$$\tau_3 = \frac{\lambda_3}{\lambda_2}, \quad \text{L-skewness parameter;} \tag{2.25}$$

$$\tau_4 = \frac{\lambda_4}{\lambda_2}, \quad \text{L-Kurtosis parameter;} \tag{2.26}$$

2.2 PROBABILITY DISTRIBUTIONS FREQUENTLY USED IN FLOOD ANALYSIS

In order to access meaningful information about the frequency distributions obtained using the observation results, we need to adapt these data to probability distribution functions. Probability distribution functions have a certain number of parameters and these parameters can be calculated using the available data. The probability distribution functions used in this study are

functions that are stated to be suitable for flood analysis (Bayazıt and Önöz, 2008).

2.2.1 Normal Distribution Family

The first probability distribution functions we will talk about are the most commonly used distribution family in statistics, called Gaussian or Normal. We will touch on 3 different distributions from this family. These are the Normal distribution, Lognormal Distribution and 3-parameter Lognormal distribution, respectively.

2.2.1.1 Normal Distribution

The most commonly used normal distribution in statistics is also one of the distributions that are most compatible with the data in hydrology. According to the Central Limit Theorem in statistics, if the random variable X consists of the sum of n different data, the probability distribution function of X approaches the normal distribution as the number n increases.

The normal probability distribution function of the random variable X is found with the following formula (2.27).

$$f(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2\right] \tag{2.27}$$

The μ_x in the expression is the distribution mean and σ_x is the standard deviation. The parameters of the normal probability distribution are its mean and standard deviation. Since normal distributions are symmetric, they cannot be skewed to the right or left and their skewness coefficients are 0. The L-moments of the normal distribution are given below in the equations (2.28), (2.29), (2.30) and (2.31).

$$\lambda_1 = \mu \tag{2.28}$$

$$\lambda_2 = 0.546, \quad \sigma = \pi^{-1/2} \tag{2.29}$$

$$\tau_3 = 0 \tag{2.30}$$

$$\tau_4 = 0.1226 \tag{2.31}$$

The distribution parameters are:

$$\mu = \lambda_1 \quad \text{and} \quad \sigma = \pi^{1/2}\lambda_2 \tag{2.32}$$

Since the additional distribution function of normal distributions cannot be obtained analytically, it has been tabulated. In order to make the table uniform, it has been obtained by standardizing the random variable (z):

$$z = \frac{x - \mu_x}{\sigma_x} \tag{2.33}$$

The result of the z value in the additional distribution function is calculated with the following formula (2.34).

$$\Phi(z) = 1 - 0.5 \exp\left(-\frac{(83z + 351)z + 562}{703/z + 165}\right) \tag{2.34}$$

The inverse of the additional distribution function is:

$$\Phi^{-1}(z) = z_p = \frac{p^{0.135} - (1 - p)^{0.135}}{0.1975} \tag{2.35}$$

When calculating the formula, the inverse function for more precise values is calculated with the following equation (2.36).

$$\Phi^{-1}(z) = z_p = -\sqrt{\frac{y^2[(4y + 100)y + 205]}{[(2y + 56)y + 192]y + 131}} \tag{2.36}$$

Using the normal distribution, the quantile corresponding to the p probability value is obtained with the following equation (2.37) (Bayazıt and Önöz, 2008; Bayazıt and Yeğen Oğuz, 2005).

$$x_p = \mu_x + z_p\sigma_x \tag{2.37}$$

2.2.1.2 Lognormal Distribution

Let's call the new random variable Y obtained by taking the logarithm of the random variable X. If the variable Y exhibits a normal distribution, this distribution is called the Lognormal Distribution. Since the lognormal distribution calculated by taking the logarithm can be calculated for X random variables above 0, the lognormal distribution gives good results for positive skewness in hydrology.

$$Y = \ln(x) \tag{2.38}$$

Since Y is calculated as in the equation (2.38), the value of X can be calculated as in the following equation (2.39).

$$X = \exp(Y) \tag{2.39}$$

We can reach the additive distribution function of the lognormal distribution X with the following equation (2.40).

$$F(x) = P(X \leq x) = P[Y \leq \ln(x)]$$

$$= P\left[\frac{Y - \mu_Y}{\sigma_Y} \leq \frac{\ln(x) - \mu_Y}{\sigma_Y}\right] \tag{2.40}$$

$$= \Phi\left[\frac{\ln(x) - \mu_Y}{\sigma_Y}\right]$$

In lognormal distributions, there is a relationship between skewness and coefficient of variation, and this relationship is expressed by the following equation (2.41).

$$C_{sx} = 3C_{vx} + C_{vx}^3 \tag{2.41}$$

In lognormal distributions, as the skewness and coefficients of variation approach zero, the distribution approaches the normal distribution.

The two parameters of the lognormal distribution of the random variable X are also the first two moments of the distribution and are calculated as in the following eqautions (2.42) and (2.43).

$$\mu_x = \exp\left(\mu_y + \frac{\sigma_y^2}{2}\right) \tag{2.42}$$

$$\sigma_x^2 = \mu_x^2\left[\exp\left(\sigma_y^2\right) - 1\right] \tag{2.43}$$

Since the variable Y is the logarithmic form of the variable X, the following relationship exists between the moments of the distributions:

$$\sigma_Y = \left[\ln\left(1 + \frac{\sigma_x^2}{\mu_x^2}\right)\right]^{1/2} \tag{2.44}$$

$$\mu_Y = \ln(\mu_x) - \frac{1}{2}\sigma_Y^2 \tag{2.45}$$

The second L moment of the Lognormal distribution is calculated with the formula below and this formula is also valid for the 3-parameter Lognormal distribution.

$$\lambda_2 = \exp\left(\mu_Y + \frac{\sigma_Y^2}{2}\right)\operatorname{erf}\left(\frac{\sigma_Y}{2}\right) = 2\exp\left(\mu_Y + \frac{\sigma_Y^2}{2}\right)\left[\Phi\left(\frac{\sigma_Y}{\sqrt{2}} - \frac{1}{2}\right)\right] \tag{2.46}$$

The following formula (2.47) is used to find the X_p quantile in the lognormal distribution (Bayazıt and Önöz, 2008; Bayazıt and Yeğen Oğuz, 2005).

$$X_p = \exp\left(\mu_Y + z_p\sigma_Y\right) \tag{2.47}$$

2.2.1.3 The 3-Parameter Lognormal Distribution

As can be seen from different studies, the distribution obtained after taking the logarithm of the variable generally does not comply with the normal distribution. For this, a lower limit such as X_0 is subtracted from the variable and

its logarithm is taken, and in this case the newly obtained distribution is more compatible with the normal distribution.

$$Y = \ln(-X_0) \tag{2.48}$$

To obtain X in terms of X_0 and Y, the following equation (2.49) is used:

$$X = X_0 + \exp(Y) \tag{2.49}$$

The equations (2.50) and (2.51) showing the first two moments of the variable X via the expression with X_0 and Y are as follows:

$$\mu_x = X_0 + \exp\left(\mu_Y + \frac{1}{2}\sigma_Y^2\right) \tag{2.50}$$

$$\sigma_x^2 = [\exp(2\mu_Y + \sigma_Y^2)][\exp(\sigma_Y^2) - 1] \tag{2.51}$$

Similarly, if we calculate the skewness coefficient based on the X_0 and Y variables, the equation will be as follows:

$$c_{sx} = 3\beta + \beta^3 \tag{2.52}$$

Where $\beta = [\exp(\sigma_Y^2) - 1]^{0.5}$.

Using the moment technique to estimate the parameters of the 3-parameter log-normal distribution is not very effective, instead one should proceed using the quantile-lower bound technique used to find X_0.

$$X_0 = \frac{X_{(1)}X_{(N)} - X_{med}^2}{X_{(1)} + X_{(N)} - 2X_{med}} \tag{2.52}$$

Here in the equation (2.53), X_N represents the smallest value of the distribution, X_1 represents the largest value, and X_{med} represents the median value. The equation that needs to be looked at to determine whether X_0 is the lower or upper limit is as follows: If $X_1 + X_N - 2X_{med} < 0$, then X_0 is the upper limit, if it is greater than 0, then X_0 is a lower limit.

The other two parameters are calculated based on the calculated X_0. While the first parameter is found by calculating the first moment of the distribution, the second moment calculation is used for the second parameter. $\lambda_1 = \mu, \lambda_2$ is found with the equation (2.46) (Bayazıt and Önöz, 2008; Bayazıt and Yeğen Oğuz, 2005).

2.2.2 GEV Family

When we examine hydrological events, we generally see that data such as precipitation data, maximum flow rate, or minimum flow rate are similar to each other. Gumbel suggested in 1958 that these variables can adapt to extreme value distributions. When we look at the extreme value theory in statistics, it is accepted that as the number of independent variables in a distribution becomes infinite, the distribution of the maximum values examined converges to one of the extreme value distributions.

In this section, the Gumbel distribution and the extreme value distribution will be discussed.

2.2.2.1 Gumbel Distribution

Let's say $M_i, i = 1,2, \dots 365$ for the annual flow data of a basin and determine the X variable to take their maximum. If the number N can go to infinity and the Mi values are independent of each other and the distributions they fit are the same, the distribution that the X variable will converge to for large values will be the Gumbel distribution.

The Gumbel distribution is a 2-parameter distribution and the formula that estimates its parameters is as follows like in the equation (2.54) and (2.55).

$$\alpha = \frac{s_x\sqrt{6}}{\pi} = 0.7796 s_x \tag{2.54}$$

$$\xi = \bar{x} - 0.5772 \tag{2.55}$$

The α in the equation (2.54) is calculated with the following equation:

$$\alpha = \frac{\lambda_2}{\ln(2)} = 1.443\lambda_2 \tag{2.56}$$

The λ_2 value here in the (2.56) is the moment of the distribution, and we find the moments of the distribution with the following formulas:

$$\lambda_1 = \xi + 0.5772 \tag{2.57}$$

$$\lambda_2 = \alpha \ln(2) \tag{2.58}$$

In addition, the mean, variance and skewness parameter formulas for the distribution are as follows:

$$\mu_x = \xi + 0.5772 \tag{2.59}$$

$$\sigma_x^2 = \frac{\pi^2\alpha^2}{6} \approx 1.645\alpha^2 \tag{2.60}$$

$$c_{sx} = 1.1396 \approx 1.14 \tag{2.61}$$

The distribution's probability density function and additional distribution functions are found with the equations (2.62) and (2.63).

$$f(x) = \frac{1}{\alpha}\exp\left[-\frac{x-\xi}{\alpha} - \exp\left(-\frac{x-\xi}{\alpha}\right)\right] \quad -\infty < x < \tag{2.62}$$

$$F(x) = \exp\left[-\exp\left(-\frac{x-\xi}{\alpha}\right)\right] \tag{2.63}$$

The first of the two methods used in parameter estimation in the Gumbel distribution is the moment technique. The moment technique gives better results when the values really fit the Gumbel distribution. The maximum likelihood method is the second technique and is more suitable if the data has a limited approximation to the Gumbel distribution.

There is a high similarity between the probability density functions of the Gumbel distribution and the distribution of the lognormal distribution with a skewness number of 1.13. Since the skewness coefficient of the Gumbel

distribution is constant, the shape of the probability density function does not change.

The quantile corresponding to the p-value can be found by inverting the additional distribution function of the Gumbel distribution (Bayazıt and Önöz, 2008; Bayazıt and Yeğen Oğuz, 2005).

$$X_p = \xi - \alpha \ln[-\ln(F)] \tag{2.64}$$

2.2.2.2 Extreme Value Distribution

The additional distribution function of the GEV distribution, referred to as General Extreme Value in the literature, is as follows:

$$F(x) = \exp\left\{-\left[1 - \frac{k(x-\xi)}{\alpha}\right]^{\frac{1}{k}}\right\} \tag{2.65}$$

The distribution has 3 parameters and its parameters are: α is the scale parameter, k is the shape parameter and ξ is the location parameter. The values where k is 0 are the Gumbel distribution.

When the parameter k is greater than zero, the distribution is bounded from above and its upper limit is the $\xi + \frac{\alpha}{k}$. When k is less than 0, the distribution is bounded from below and its lower limit is the $\xi + \frac{\alpha}{k}$.

The formulas used to calculate the moments of the distribution are as follows in the (2.66) and (2.67).

$$\mu_x = \xi + \left(\frac{\alpha}{k}\right)[1 - \Gamma(1+k)] \tag{2.66}$$

$$\sigma_x^2 = \left(\frac{\alpha}{k}\right)^2 \{\Gamma(1+2k) - [\Gamma(1+k)]^2\} \tag{2.67}$$

The places indicated by Γ(.) are the moments of the distribution. The variance can only be calculated when k is greater than -0.5. The equation used to calculate the skewness coefficient is as follows:

$$C_{sx} = \frac{-\Gamma(1+3k) + 3\Gamma(1+k)\Gamma(1+2k) - 2\Gamma^3(1+k)}{[\Gamma(1+2k) - \Gamma^2(1+k)]^{3/2}} \tag{2.68}$$

While tables can be used when calculating the Γ(.) function, it can also be obtained from the following equations (2.69) and (2.70).

$$\Gamma(1+\delta) = 1 + \sum_{i=1}^{5} a_i \delta^i + \xi \tag{2.69}$$

$$a_1 = -0.5748646$$

$$a_2 = 0.9512363$$

$$a_3 = -0.6998588$$

$$a_4 = 0.4245549$$

$$a_5 = -0.1010678$$

$$|\varepsilon| \leq 5 \times 10^{-5} \tag{2.70}$$

The simpler version of this is below.

$$\Gamma(1+w) = w\Gamma(w) \quad , \quad 0 < w < 1 \tag{2.71}$$

For integer w's:

$$\Gamma(1+w) = w! \tag{2.72}$$

The L moments of the distribution and the ratios of the moments are calculated with the following equations (2.73), (2.74), (2.75) and (2.76):

$$\lambda_1 = \xi + \frac{\alpha}{k}\{1 - \Gamma(1+k)\} \tag{2.73}$$

$$\lambda_2 = \frac{\alpha}{k}(1 - 2^{-k})\Gamma(1+k) \tag{2.74}$$

Moment Ratios:

$$\tau_3 = \left\{\frac{2(1-3^{-k})}{(1-2^{-k})} - 3\right\} \tag{2.75}$$

$$\tau_4 = \frac{1-5(4^{-k})+10(3^{-k})-6(2^{-k})}{1-2^{-k}} \tag{2.76}$$

The L-moment method is used to estimate the parameters in the distribution. The estimation equations are given below in the equation (2.77).

$$k = 7.8590c + 2.9554c^2 \tag{2.77}$$

Estimation of k is zero and variance is $\mathrm{var}(k) = \frac{0.5633}{N}$.

$$\alpha = \frac{k\lambda_2}{\Gamma(1+k)(1-2^{-k})} \tag{2.78}$$

$$\xi = \lambda_1 + \frac{\alpha}{k}[\Gamma(1+k) - 1] \tag{2.79}$$

$$c = \frac{2\lambda_2}{\lambda_3 + 3\lambda_2} - \frac{\ln(2)}{\ln(3)} = \frac{2\beta_1 - \beta_0}{3\beta_2 - \beta_0} - \frac{\ln(2)}{\ln(3)} \tag{2.80}$$

The quantile formula for a random probability p is given in the equation (2.81) (Bayazıt and Önöz, 2008; Bayazıt and Yeğen Oğuz, 2005).

$$X_p = \xi + \frac{\alpha}{k}\{1 - [-\ln(F)]^k\} \tag{2.81}$$

2.2.3 Pearson Type III Family

Although there are different distributions in the Pearson Type III family, this study will focus on the Log-Pearson type III distribution, which is frequently used in hydrology.

2.2.3.1 Log Pearson Type III Distribution

The Log-Pearson III distribution is the form of the Pearson III distribution obtained by taking the logarithm of the variable. If Y is equal to the logarithm of X, we express X in terms of Y with the formula below and we call the distribution of X the Log-Pearson III distribution.

$$X = \exp(Y) \tag{2.82}$$

The probability density function of the distribution is expressed as follows in the (2.83).

$$f(x) = |\beta|[\beta(x-\xi)]^{\alpha-1}\frac{\exp\{-\beta[\ln(x)-\xi]\}}{\alpha\Gamma(\alpha)} \tag{2.83}$$

The parameters here are α shape, β scale and ξ location parameters.

In cases where β is less than zero, the distribution is bounded from above and the ξ is the upper limit of the distribution; in cases where β is greater than zero, the distribution is bounded from below and the ξ is the lower limit of the distribution.

The mean, variance and skewness parameters of the distribution are calculated as follows for values where β is less than 0 or greater than r.

For $\beta > r \; or \; \beta < 0$:

$$\mu_x = e^{\xi}\left(\frac{\beta}{\beta-r}\right)^{\alpha} \tag{2.84}$$

$$\sigma_x^2 = e^{r\xi}\left[\left(\frac{\beta}{\beta-r}\right)^{\alpha} - \left(\frac{\beta}{\beta-r}\right)^{2\alpha}\right] \tag{2.85}$$

Skewness Parameter:

$$C_{sx} = \frac{E[x^3] - 3\mu_x E[x^2] + 2\mu_x^3}{\sigma_x^3} \tag{2.86}$$

In order to find the quantile for a random p probability value, the Pearson Type III distribution is applied to the logarithm of the data and the main variable is passed with the following transformation (Bayazıt and Önöz, 2008).

$$Y = \mu_Y + \sigma_Y K(C_s) \tag{2.87}$$

$$X_p = \exp(Y_p) \tag{2.88}$$

2.3 STATISTICAL TESTS

Some tests are performed to find the probability function that best fits the distribution of station data. In this study, the methods used to test the suitability of the data to the probability distribution are: Chi-Squared test and Kolmogorov-Smirnov test.

2.3.1 Chi-Squared Test

The Chi-Squared Goodness of Fit Test is a statistical test used to evaluate the extent to which the observation values obtained from a sample fit a certain theoretical distribution. The test is used especially in categorical data and is calculated based on the sum of the ratios of the squares of the differences between the observed frequencies and the theoretically expected frequencies to the expected frequencies.

The test is conducted under the following two hypotheses:

- **Null hypothesis (H_0):** There is no significant difference between the observed distribution and the expected theoretical distribution.
- **Alternative hypothesis (H_1):** There is a significant difference between the observed distribution and the expected theoretical distribution.

The Chi-Squared test statistic is calculated with the following formula:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

Where:

O_i: Observed frequency in the i th category,

E_i: Expected frequency in the i th category,

χ^2: Chi-Squared test statistic.

In the evaluation of the test, the calculated χ^2 value is compared with the table value corresponding to the degrees of freedom ($df = k - 1$) and the selected significance level (α). If the calculated χ^2 value is greater than the table value, H_0 is rejected and it is concluded that the data is not compatible with the theoretical distribution.

2.3.2 K-S Test

The Kolmogorov-Smirnov test is used to evaluate whether an observed additive frequency distribution conforms to a theoretical distribution. In this test, the test statistic is defined as in the equations (2.89).

$$D = \max|F(x_i) - F^*(x_i)| \tag{2.89}$$

In the equation (2.89), $F(x_i)$ represents the ordinate of the selected theoretical additional distribution function corresponding to the value x_i.

$F^*(x_i)$ is calculated from the observed ordered sample data with the following formula:

$$F^*(x_i) = \frac{i}{N} \tag{2.90}$$

The D statistic represents the largest difference between the observed and theoretical distributions, and this value depends on the number of N elements in the sample.

The additional distribution function, $F(x_i)$, is calculated using equations (2.63) and (2.65) for the GEV and Gumbel distributions, respectively. Since the additional distribution function for normal, log-normal and three-parameter log-normal distributions cannot be obtained analytically, it is created in a table form using the numerical integration method.

In order to create a single table for all distributions, the random variables are converted to the standard variable (z) form. For each x_i value, first the z value is calculated and then the $F(x_i)$ value is read using additional tables (A.3).

If the parameters of the selected distribution are determined independently of the sample data, the critical values of the Δ test statistic (Δ_α), taking into account the sample size (n) and the probability of exceedance (α), can be taken from Appendices Table (A.1) prepared by Chowdhury et al. (1991). However, if the parameters of the distribution are estimated from the sample data, as is often the case in practice, the critical Δ_α values will be slightly lower. In this case, the Δ_α values will vary depending on the selected distribution. For the normal distribution and the Gumbel distribution, Appendices Table (A.2) prepared by Crutcher (1975) can be used (Bayazıt, 1996).

As a result, if the Δ statistic value obtained in the Kolmogorov-Smirnov test is smaller than the critical value corresponding to the 5% significance level, it is accepted that the selected distribution fits the observation results. Otherwise, it is concluded that the distribution is incompatible.

2.4 CORRELATION COEFFICIENT

The correlation coefficient, which takes values between -1 and 1, looks at the similarities of the changes in two variables and is expressed with r. When r is 0, it is said that there is no relationship between the changes in two variables, when it is greater than 0, it can be said that there is a positive relationship; when it is less than zero, it can be said that there is a negative relationship.

Although there are different methods used to calculate the correlation coefficient, one of them is given below.

2.4.1 Hypothesis Tests Related to Correlation Coefficient

Although the similarities of changes are measured with the correlation coefficient, the results of these measurements may not always be significant. In order to test whether the result found by the correlation coefficient is significant,

different tests have been put forward over time. In this study, the t-test was applied while measuring the significance of the correlation coefficient. In the method we call t-test, two different hypotheses are considered and the accuracy of the zero hypothesis is measured to decide which of the hypotheses is valid. While the two hypotheses are expressed as follows, the t value testing the zero hypothesis is calculated with the formula (2.92).

$$t = \frac{r \times \sqrt{n-2}}{\sqrt{1-r^2}} \tag{2.92}$$

H_0: $\rho = 0$ No relation between x and y

H_1: $\rho \neq 0$ There is a relation between a and y

2.5 TREND ANALYSIS

If the elements of a set of values show an increase or decrease depending on time, this is called a trend. There are different methods, parametric and nonparametric, to examine the distribution trends of data. The Mann-Kendall test used in this study is a nonparametric test. In cases where the observations are short-term, skewed and irregular, nonparametric tests are preferred to ignore the negative effects of these features. The most commonly used method in trend analysis is the Mann-Kendall method (Önöz, 2010).

2.5.1 Mann Kendall Trend Analysis

Nonparametric Mann Kendall analysis is independent of the distribution of the random variable. Since it is based on the calculation of the correlation coefficient τ, it is also called the Kendall T statistic. What is important in Mann Kendall is the magnitude orders of the values. It tests whether there is a trend in a time series with the null hypothesis (Bayazıt, 1996).

Kendall's Tau, τ, test is a nonparametric correlation measure that analyzes the relationships between ordered data pairs. The main purpose of this test is to determine whether there is a nonlinear but directional relationship between two

variables. This method, which is used especially in time series, examines the increasing or decreasing trends between observations and statistically reveals monotonic trends (one-way increase or decrease) in the data. In cases where the data does not comply with the normal distribution assumption and in environments where extreme values may distort the analysis results, Kendall's Tau test stands out with its reliability.

For this reason, it is frequently preferred in environmental, hydrological and meteorological studies; It allows strong and meaningful inferences to be made especially in long-term trend analyses.

The first step in applying the test is to create the x_i, x_j pairs for the values x_1, x_2 ,...x_n. The number of cases where i is smaller than j and at the same time x_i, is smaller than x_j is called P, the number of cases where xi is larger than x_j is called M, and the test statistic is calculated as $S = P - M$. Based on this, the τ coefficient we mentioned above is calculated as follows like in the equation (2.93):

$$\tau = \frac{S}{n(n-1)/2} \tag{2.93}$$

As we mentioned in the correlation coefficient, the Kendall coefficient also takes values between -1 and 1. A value of zero indicates that there is no trend among the data, while -1 or 1 indicates the existence of a trend. When the S value is positive, this trend is increasing, while when it is negative, it is decreasing.

In cases where the number of data n is less than ten, the x(-x) values, which are the probability p for S to remain small, are given in the table, while in cases where n is greater than ten, the mean and standard deviation values are mentioned below in (2.94).

$$\mu_s = 0 \quad \sigma_s = \sqrt{\frac{n(n-1)(2n+5)}{18}} \tag{2.94}$$

The Z value corresponding to these values, whose distribution is standard normal distribution, is calculated with the formula below.

$$Z = \begin{cases} \frac{S-1}{\sigma_s} & S > 0 \\ 0 & S = 0 \\ \frac{S+1}{\sigma_s} & S < 0 \end{cases} \qquad (2.95)$$

In the above cases of the equation (2.95), we talked about the conditions where there are no equivalent data in the sample. If there are equivalent values between the data, the standard deviation is calculated with the formula below.

$$\sigma_s = \sqrt{\left[n(n-1)(2n+5) - \sum_i t_i(t_i-1)(2t_i+5)\right]/18} \qquad (2.96)$$

The t_i value in the formula shows the numbers of data that are equal to each other. For example, if 3 values are the same in the sample, $t_i = 3$, if there are two different pairs that are equal to each other, $t_i = 2$ and $t_i + 1 = 2$ will be considered. When testing the hypothesis, the absolute value of the Z value found above is taken and compared with the $Z_{\alpha/2}$ value corresponding to the selected alpha significance level. If the value we find is smaller than the value corresponding to $Z_{\alpha/2}$, the zero hypothesis is accepted, and it is concluded that there is no trend in the distribution of the data. By interpreting the curves according to the results, it is possible to find out whether the data has changed over the years and whether there is a trend.

Since the Mann Kendall test is based on two hypotheses and works on a principle of accepting or rejecting the zero hypothesis, it is important to correctly determine the significance level used in the studies. The significance value of 0.05 was selected in the Mann Kendall test conducted in this study. In cases where it is said that there is no trend, the probability of a trend is 5% (Önöz, 2010). You can see the regions where the hypothesis will and will not be accepted in the Figure 3 below.

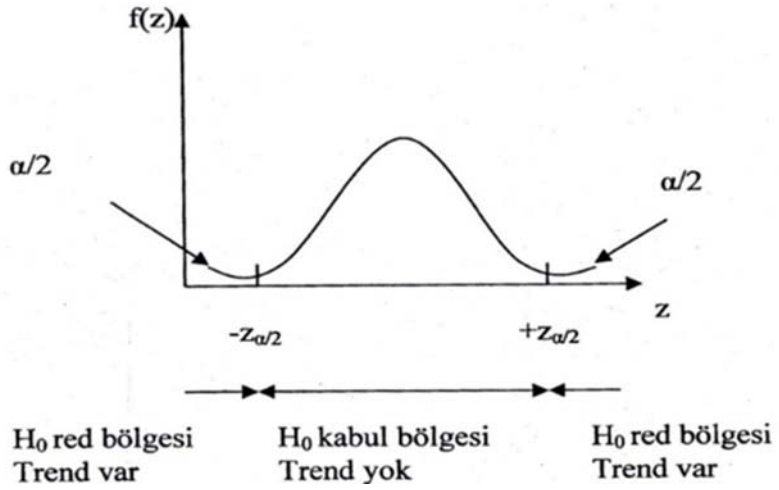


Figure 2.2 Mann-Kendall hypotheses (Cebe, 2007).

2.6 REGIONAL FLOOD ANALYSIS

When examining data such as flood flow and precipitation in hydrology, working on a single station is not sufficient because the data diversity will be low. In order to increase the data diversity, the data of different stations located in the same homogeneous region should be evaluated together. After finding a statistically homogeneous region, it is assumed that the dimensionless parameters in this region do not change. Examining all the station values in a homogeneous region together is called regional analysis.

Although there are many different methods for regional analysis, the priority here is to create a homogeneous region. When creating a homogeneous region, the situation that is looked at is whether the differences between the data at one station and the data at the other station are statistically significant.

After determining the homogeneous region, the stations in the region are analyzed and the statistical characteristics of the region are decided.

2.6.1 Determining Homogeneous Area

The most important step in regional analysis is determining homogeneous regions. Although it is thought that homogeneous regions should be

geographically close to each other, this is not always true. Even stations located next to each other may have very different characteristics. When determining a homogeneous region, instead of looking at statistical characteristics, it is necessary to look at physical characteristics and then compare statistical characteristics of stations falling within these regions to see if they belong to a homogeneous region (Bayazıt and Önöz, 2008).

The method proposed by Wiltshire in 1986 is given below.

In the first stage, it is assumed that the region where the stations are located is divided into M homogeneous regions by looking at the physical properties of the region.

In the second stage, the dimensionless coefficient of variation is calculated for each station.

$$C_{vx} = \frac{s_x}{\overline{x}} \tag{2.97}$$

The sample distributions of the calculated coefficients of variation are calculated and the significance level of the null hypothesis for this distribution is selected as 0.05.

$$\overline{C}_v^2 \left(1 + 2\overline{C}_v^2\right)/2n \tag{2.98}$$

$$\overline{C}_v \pm 1.96 \left[\frac{\overline{C}_v^2 \left(1 + 2\overline{C}_v^2\right)}{2n}\right]^{0.5} \tag{2.99}$$

All C_v values in a homogeneous region are compared. The C_v value of each station must be between the calculated maximum C_v and the minimum C_v.

In the case that the populations of the C_v distributions of two different regions are the same, the F statistic is calculated as follows:

$$F = \frac{\overline{C}_1^2}{\overline{C}_2^2} \tag{2.100}$$

n_1-1 and n_2-1 are the degrees of freedom of the F statistic and the significance of the selected grouping method is investigated using the F table. In order for the grouping to be significant, the value corresponding to the calculated F value must be greater than the value corresponding to the $F_{0.05}$ point.

In the last stage of the analysis, after determining the correct grouping method, it is investigated whether the stations in the group are homogeneous or not. When the coefficient of variation of the i_{th} station in the group C_{vi} is expressed as the average coefficient of variation of the stations in the group C_{vi} is above var, and the sampling variance of the \$ith\$ station is expressed as var(C_{vi}), the S statistic is calculated with the formula below and this value is expected to be less than the $\chi^2_{0.05}$ value in the χ^2 distribution table. Until this rule is met, some stations in the group are removed and retested.

$$S = \sum_j \frac{\left(c_{vj} - \overline{c}_{vj}\right)^2}{\mathrm{var}\left(c_{vj}\right)} \tag{2.101}$$

When all the above steps are completed, it is assumed that all groups are homogeneous.

There are different methods used to determine homogeneous regions. These include clustering analysis, principal component analysis and factor analysis (Hosking and Wallis, 1997; Bayazıt, 2006).

After determining homogeneous regions, various tests are applied to accept the regions as homogeneous and the S statistic test mentioned above was used in this research (Bayazıt and Önöz, 2008).

When determining homogeneous regions, if the regions do not comply with the tests performed, some stations are removed, new regions are opened and statistical tests are tried again and these processes are repeated until the statistical tests give the correct answer.

2.7 SEASONALITY ANALYSIS

It is necessary to know the seasonality of floods in infrastructure studies of water resources, flood analyses and many other applications (Cundelik et al., 2004). Flood seasonality, which is an information that should be used in every project using water resources, allows for a better understanding of the movements of the water resource and to manage the process accordingly.

2.7.1 Angular Seasonality Analysis

In regionalization studies where we divide basins into homogeneous regions, hydrological similarities are generally taken into account and when determining the homogeneity of the river, the characteristic features and occurrence times of floods experienced in the river are examined (Burn and Goel, 1997).

The first step in performing angular seasonality analysis is to express the annual flood observation date for each station in terms of radials. The method used to express days in terms of radials is as follows: January 1 is accepted as the 1st day of the year, December 31 is accepted as the 365_{th} day of the year and the radial equivalent of the days is found using the formula below.

θ in the equation (2.102) represents the radial angle of the flood day that occurred on the i_{th} day. When working on a station with N different annual data, the radial angles of the flood days are shown on the unit circle in order to visualize the flood dates like in the Figure 79.

$$\theta_i = (\text{Calendar Day})_i \left(\frac{2\pi}{365}\right)$$

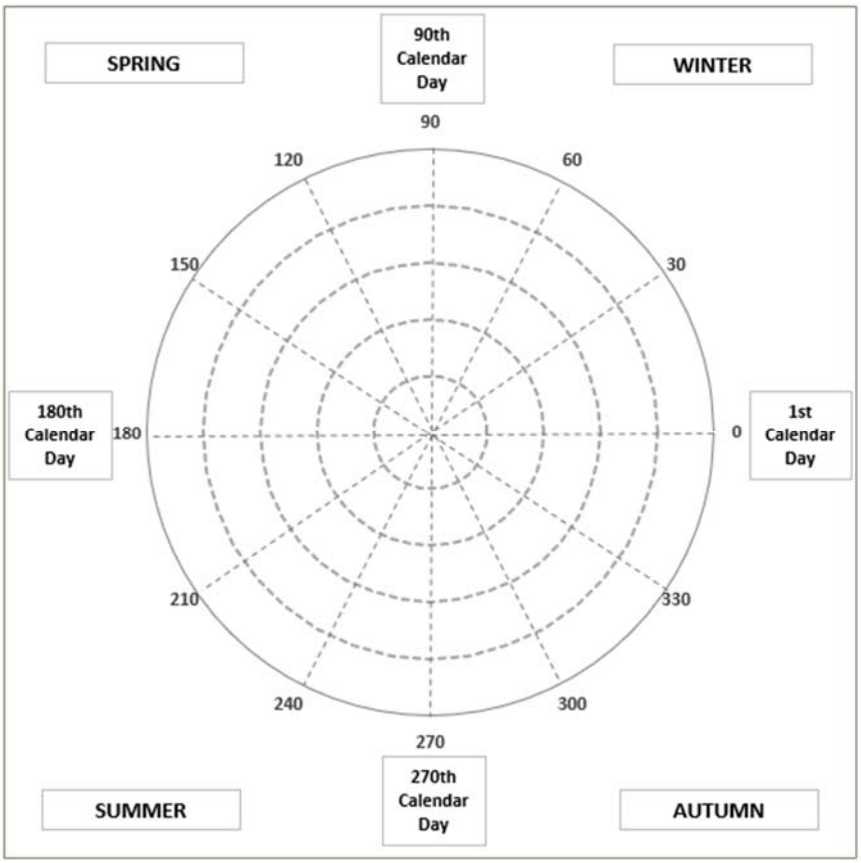


Figure 2.3 Unit Circle

We can find the x and y coordinates of the values of the radial angle of the flood days with n number of data with the following formulas.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} \cos(\theta_i) \tag{2.103}$$

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} \sin(\theta_i) \tag{2.104}$$

The $\bar{x}$ and $\bar{y}$ expressions in the equations represent the flood day (MD) of the station. When calculating the angular value of the flood time, it is calculated with two different methods by looking at the positivity/negativity of the $\frac{\bar{y}}{\bar{x}}$ expression. If this expression is positive, the equation (2.105) is used, if it is

negative, the equation (2.106) is used (Burn and Goel, 1997). The relevant formulas are as follows.

$$\bar{\theta} = \tan^{-1}\left(\frac{\bar{y}}{\bar{x}}\right) \tag{2.105}$$

$$\bar{\theta} = \tan^{-1}\left(\frac{\bar{y}}{\bar{x}}\right) + \pi \tag{2.106}$$

We can obtain the actual calendar day with the following equation:

$$MD = \bar{\theta}\left(\frac{365}{2\pi}\right) \tag{2.107}$$

The calculated MD values are the average day of the time when the flood values at the station occur. It is expected that the stations where the MD values are similar to each other will also be similar to each other in terms of hydrology.

In basins where the cause of the flood is snowmelt, the MD values are generally close to each other.

There is a parameter that shows the diversity of the floods that occur at a station and this parameter is calculated with the formula below in (2.108) (Burn and Goel, 1997).

$$r = \sqrt{(\bar{x})^2 + (\bar{y})^2} \tag{2.108}$$

The r parameter here is a dimensionless parameter and represents the distribution of flood values. The r value takes a value between 0 and 1, and as r approaches 1, it shows that the proximity of the flood dates at the station increases; and as it approaches 0, it shows that the flood dates differ from each other. From here, we can say that the increase in the r value indicates the regularity of the flood dates at the station like in the Figure.

When examining the similarities between stations, the x and y coordinates mentioned above are used. When the distances between the x and y coordinates of the average flood value of each station are taken using the Euclidean function, the result can be interpreted as the similarity parameter.

The Euclidean function we mentioned is calculated as follows:

$$d_s^{ij} = \left[\left(\overline{x_\iota} - \overline{x_J}\right)^2 + \left(\overline{y_\iota} - \overline{y_J}\right)^2\right]^{0.5} \quad (2.109)$$

The expression d_s^{ij} above is considered as the similarity parameter and the decrease of this parameter indicates that the hydrological homogeneity and similarity between the two stations increases (Burn and Goel, 1997).

CHAPTER 3

3. FLOOD ANALYSIS IN THE ÇORUH RIVER BASIN

In this section, flood analyses were carried out for selected stations on the Çoruh River Basin using the methods explained in Section 2 and the results obtained with each method were interpreted in detail.

3.1 APPLICATION AREA

The Çoruh River Basin is an important hydrological basin located in the northeast of Türkiye and also encompassing the borders of Georgia. With a surface area of approximately 19,654 km², 411 km of this basin is in Turkey and 20 km in Georgia, hosting a river system of 431 km in total (Ministry of Agriculture and Forestry, 2020). The source of the river rises from the Mescit Mountains in Erzurum, passes through Artvin to Georgia and reaches the Black Sea. The Çoruh River Basin has one of the steepest valleys in Turkey and has a high hydroelectric energy potential due to this feature. In fact, in line with this potential, many dam and hydroelectric power plant (HPP) projects have been constructed or planned by the State Hydraulic Works (DSİ) in the basin. However, since these structures change the natural flows of some stations, only the stations that are not affected by the dams and have continuous flow data for at least 10 years were included in the analysis in our study. In this context, flood analyses were carried out at 23 selected stations using DSİ data. All data used in the analyses were evaluated in terms of accuracy and timeliness, and only data sets updated until 2015 were taken as basis.

Çoruh River Basin is of great importance not only in terms of energy production, but also in terms of drinking water supply, agriculture, flood control and environmental sustainability. Since the basin is located in the transition zone

between the Black Sea climate and the Eastern Anatolian climate, it exhibits quite complex hydrological characteristics. The annual average precipitation is approximately 540 mm, and the precipitation is higher, especially in the northern parts. The high altitude regions of the basin receive heavy snowfall in the winter months; in the spring, snowmelt makes a significant contribution to the flow rates of the streams. Temperature values change with altitude, and while temperate climate effects are seen in the north, in the areas close to the Black Sea, the continental climate effect is dominant in the south. Evaporation amounts can reach up to 1,000 mm (Ministry of Environment, Urbanization and Climate Change, 2019).

Agricultural and livestock activities are also important economic inputs of the basin. Agricultural production is concentrated in river valleys and plain areas, and the main products include wheat, barley, corn, potatoes and fruit and vegetable varieties. In addition, small and large cattle breeding is common in high altitude areas. Since agriculture and livestock have a large place in the basin economy, flood control is of vital importance not only in terms of preventing physical damage but also socioeconomic losses.

Historically, the Çoruh River Basin has been the scene of many major floods. Serious flood events have been recorded in Artvin province, especially in districts such as Yusufeli, Murgul and Arhavi. These floods have sometimes caused loss of life and property in residential areas and major damage to infrastructure systems. In this context, both structural measures (such as dams, dams, flood walls) and non-structural measures (such as flood early warning systems, risk maps, public information) have been developed for flood control in the basin. In the studies carried out within the scope of Flood Management Plans, variables such as slope, soil type, land use and geological structure were evaluated throughout the basin and flood risk maps were created (Ministry of Agriculture and Forestry, 2019). In these analyses, especially the northwestern parts of Artvin and the sub-basin areas where the river passes into Georgia were identified as regions with high flood risk.

As a result, the flood analyses conducted for the 23 stations selected in this study provide important information both in terms of understanding the hydrological character of the region and in terms of developing sustainable management strategies to reduce flood risks. Considering the topographic, climatic and socioeconomic characteristics of the basin, flood analyses are of great importance not only in terms of engineering but also in terms of environmental and economic sustainability.

Considering the accuracy and currency of the data used in the analyses, official data sets published by the State Hydraulic Works (DSİ) and updated until 2015 were taken as basis. Accordingly, all hydrometric data used in the current study cover the period until 2015. In order to increase the reliability of flood analyses, stations with a data period of less than ten years were excluded from the evaluation. In addition, in order to minimize the effect of dam structures, the locations of the relevant stations were compared with the basins affected by the dams. If the data of a station belong to a date after the operation of a dam located on it and the flow regime of this station is affected by the dam, then the station was not included in the analysis. However, if it was determined that the flow regime of the station was not directly affected despite the dam structure, the station was included in the study.

As a result of the evaluations made in this context, 23 stations reflecting natural flow conditions and having long-term observation data were selected for analysis. The geographical locations of these selected stations are presented in Figure 5. The relevant visual shows the flow observation stations in the entire region, and the stations included in the study are highlighted in blue. In addition, the locations of the stations under the influence of dams are given in Figure 6, and the basic characteristics of the stations included in the study are given in detail in Table 1.

Figure 3.1 Location of the Stations in the Çoruh River Basin

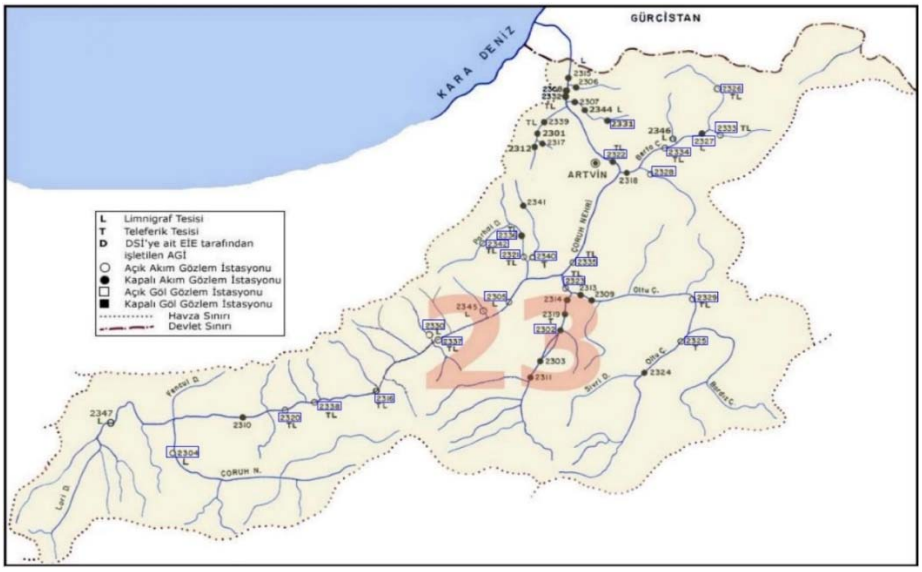


Figure 3.2 Location of the Dams in Çoruh River Basin

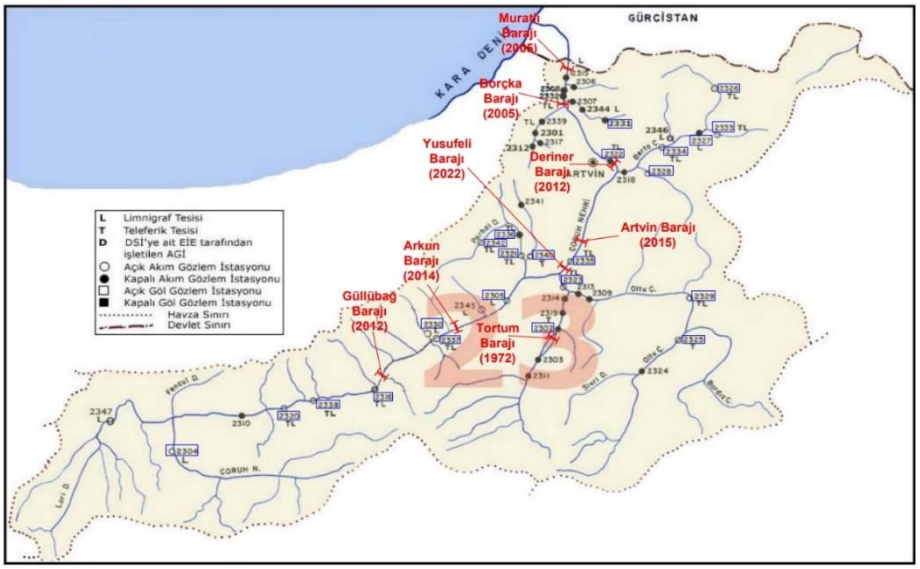


Table 3.1 Characteristics of the Stations

Name of the Station	Station Number	Observation Range	Number of Years	Rainfall Area (km²)	Elevation (m)
TORTUM Ç. TEV KÖP.	2302	1940-1968	21	1745	960
ÇORUH NEHRİ BAYBURT	2304	1941-2011	61	1734	1545
ÇORUH NEHRİ PETEREK	2305	1941-2011	48	7272	654
ÇORUH NEHRİ İSPİR KÖP.	2316	1964-2011	46	5505	1170
ÇORUH NEHRİ LALELİ	2320	1970-2011	38	4759	1365
PARHAL DERESİ DUTDERE	2321	1971-2011	39	586	705
ÇORUH NEHRİ ALTINSU	2322	1971-2000	28	18326	201
OLTU SUYU İŞHAN KÖPRÜSÜ	2323	1962-2011	46	6854	585
OLTU SUYU AŞAĞIKUMLU	2325	1973-2011	36	1762	1150
MEYDANCIK DERESİ DUTLU	2326	1981-2011	23	248	875
BERTASUYU ÇİFTEHANLAR	2327	1981-1998	17	1229	570
ARDANUÇ DERESİ FERHATLI	2328	1981-2010	29	547	365
OLTU SUYU ÇOŞKUNLAR	2329	1981-2011	30	3519	1004
ÇAMLIKAYADERESİ ÇAMLIKAYA	2330	1981-2011	30	114	995
DEVİSKEL DERESİ GÜNDOĞDU	2331	1987-2000	13	100	560
MANSURAT DERESİ ŞAVŞAT	2333	1989-2011	20	329	865
BERTASUYU BAĞLIK	2334	1989-2010	21	1473	366
ÇORUH NEHRİ İNANLI	2335	1989-2011	19	15687	435
PARHAL SUYU İKİZKAVAK	2336	1989-2007	15	541	805
ÇORUH NEHRİ ÇAMLIKAYA	2337	1988-2011	22	6634	892
ÇORUH NEHRİ PAZARYOLU	2338	1988-2011	20	5168	1265
ÖĞDEM DERESİ UYSALLAR	2340	1990-2011	20	202	682
PARHAL SUYU ALTIPARMAK	2342	1992-2011	19	318	1130

3.2 STATISTICAL VALUES OF THE STATIONS

In this section, the calculated statistical moments, L-moments and the calculations of the ratios related to these moments for 23 selected stations will be given.

3.2.1 Calculation of Statistical Moments

Based on the highest flow rates recorded at each station, statistical parameters were calculated using (2.3), (2.4), (2.5), (2.7), (2.8), (2.10), (2.12) and the results obtained are presented in the Table 2.

When the statistical parameters of the 23 stations in the Table 2 are examined, the coefficient of skewness (C_s) values show a significant difference between the stations. In general, positive skewness was observed at all stations except station 2342. This situation shows that the flow data are mostly concentrated below the average value and that high flow values occur, albeit rarely. The negative skewness value of station 2342 (-0.3) indicates that low values are more frequent and high values are rarer at this station and that the

distribution is skewed to the left. The skewness coefficient of 2.6, especially observed at station 2333, reveals that the distribution is quite skewed to the right.

The coefficient of variation (C_v) values remained below 1 at all stations; this shows that the standard deviation values are low compared to the average, and therefore the data exhibit a relatively homogeneous distribution. The highest (C_v) value was observed as 0.6 at station 2325, which reveals that the relative variability of this station is higher compared to the others. On the other hand, the (C_v) values at stations 2322 and 2321 are quite low at 0.2, respectively, and it is understood that the flow data at these stations are more consistently distributed around the mean.

Table 3.2 Statistical Characteristics of the Stations

Station Number	Maximum Flow (m^3/s)	Minimum Flow (m^3/s)	$\bar{x}$ (m^3/s)	Median	Variance	σ	C_v	C_s
2302	158,0	20,7	70,0	66,5	861,9	29,4	0,42	1,15
2304	296,0	21,0	94,3	78,0	2.337,4	48,3	0,51	1,73
2305	684,0	201,0	388,2	367,0	14.564,2	120,7	0,31	0,60
2316	558,0	121,0	257,0	234,5	9.110,0	95,4	0,37	0,98
2320	328,0	22,3	173,0	161,0	4.769,9	69,1	0,40	0,37
2321	122,0	44,4	76,3	76,3	322,6	18,0	0,24	0,55
2322	1.352,0	581,0	918,4	850,0	47.426,4	217,8	0,24	0,38
2323	504,0	97,8	235,4	225,5	8.206,7	90,6	0,38	0,99
2325	196,0	20,1	82,8	65,1	2.224,9	47,2	0,57	0,60
2326	123,0	27,6	70,9	67,2	832,7	28,9	0,41	0,23
2327	289,0	69,5	148,4	139,0	2.641,6	51,4	0,35	1,10
2328	88,9	21,5	46,3	44,5	314,8	17,7	0,38	0,54
2329	380,0	70,6	174,6	150,5	6.530,6	80,8	0,46	0,77
2330	54,4	7,0	18,9	17,0	100,9	10,0	0,53	1,69
2331	80,5	33,7	52,9	59,3	236,6	15,4	0,29	0,12
2333	89,0	16,8	33,1	30,0	249,6	15,8	0,48	2,60
2334	291,0	88,9	169,9	166,0	2.807,2	53,0	0,31	0,43
2335	1.355,0	361,0	721,6	622,0	82.283,0	286,9	0,40	0,94
2336	107,0	44,5	71,8	64,0	425,3	20,6	0,29	0,64
2337	548,0	172,0	344,9	339,0	12.607,4	112,3	0,33	0,29
2338	375,0	83,9	212,1	193,0	5.110,7	71,5	0,34	0,52
2340	32,3	9,5	19,6	17,2	34,0	5,8	0,30	0,82
2342	65,4	24,6	46,3	50,8	136,3	11,7	0,25	-0,27

As a result, important information about the flood characteristics of the stations has been obtained with the help of these parameters. In particular, stations with high skewness coefficients should be considered as critical regions where extreme flood events are observed; stations with low coefficients of variation should be addressed with different hydrological approaches due to their more stable flow characteristics.

3.2.2 Calculation of L Moments and Their Ratios

Using the annual maximum flow data of each station, statistical quantities commonly used in flood analysis were calculated with the L-moment approach. First, the sample mean (β) value represents the arithmetic mean of the maximum flow rates at each station, and this value was obtained using Equation (2.15). The second, third and fourth order moments $\beta_1, \beta_2, \beta_3$ were calculated using Equation (2.16), Equation (2.17) and Equation (2.18), respectively, and provided information about the distribution properties of the data such as variance, skewness and kurtosis. From these moments, the first (λ_1), second (λ_2), third (λ_3) and fourth (λ_4) L-moments are derived by Equations (2.19)–(2.22), respectively. These L-moments are related to the location, spread, skewness and kurtosis, respectively. In particular, λ_2 is the L-moment equivalent of the variation and provides information about the degree of spread of the sample data. L-moment ratios were calculated to better analyze the shape properties of the distribution. L-variation coefficient ($\tau_2 = \lambda_2/\lambda_1$), L-skewness coefficient ($\tau_3 = \lambda_3/\lambda_2$) and L-kurtosis coefficient ($\tau_4 = \lambda_4/\lambda_2$) are found by Equation (2.24)–(2.26). These ratios explain the level of variation, deviation of the distribution from symmetry and the effect of extreme values, respectively. All the results obtained are summarized in Table 3 and the differences between the stations are evaluated in the light of these parameters.

Table 3.3 L Moments Values

Station Number	β_0	β_1	β_2	β_3	λ_1	λ_2	λ_3	λ_4	LCv,τ_2	L-Skewness,τ_3	L-Skewness,τ_4
2302	70,02	43,20	31,85	25,55	70,02	96,85	1,95	3,78	1,38	0,02	0,04
2304	94,27	59,70	45,21	36,96	94,27	128,84	7,31	5,07	1,37	0,06	0,04
2305	388,17	228,78	165,55	130,80	388,17	547,55	8,80	6,61	1,41	0,02	0,01
2316	257,00	155,34	114,19	91,23	257,00	358,66	10,12	6,04	1,40	0,03	0,02
2320	173,03	106,40	78,17	62,34	173,03	239,65	3,59	5,65	1,39	0,01	0,02
2321	76,35	43,35	30,81	24,08	76,35	109,34	1,13	1,12	1,43	0,01	0,01
2322	918,43	523,55	372,78	291,00	918,43	1.313,31	13,83	0,74	1,43	0,01	0,00
2323	235,45	142,99	105,30	84,32	235,45	327,90	9,29	7,86	1,39	0,03	0,02
2325	82,83	55,02	41,97	34,09	82,83	110,63	4,50	0,31	1,34	0,04	0,00
2326	70,87	44,04	32,45	25,76	70,87	97,69	1,32	-0,65	1,38	0,01	-0,01
2327	148,41	88,80	64,84	51,78	148,41	208,03	4,69	7,48	1,40	0,02	0,04
2328	46,25	28,32	20,82	16,58	46,25	64,18	1,26	0,45	1,39	0,02	0,01
2329	174,61	110,66	82,96	66,93	174,61	238,56	8,42	3,06	1,37	0,04	0,01
2330	18,93	12,17	9,25	7,57	18,93	25,70	1,42	0,94	1,36	0,06	0,04
2331	52,92	31,09	22,32	17,42	52,92	74,74	0,30	-0,98	1,41	0,00	-0,01
2333	33,10	20,41	15,43	12,69	33,10	45,78	3,22	2,69	1,38	0,07	0,06
2334	169,89	100,68	72,82	57,44	169,89	239,10	2,71	2,58	1,41	0,01	0,01
2335	721,58	443,96	329,81	265,28	721,58	999,20	36,66	17,30	1,38	0,04	0,02
2336	71,77	42,02	30,44	24,03	71,77	101,52	2,30	-0,22	1,41	0,02	-0,00
2337	344,86	206,03	149,35	117,81	344,86	483,70	4,81	3,08	1,40	0,01	0,01
2338	212,10	127,03	92,48	73,44	212,10	297,16	4,78	6,71	1,40	0,02	0,02
2340	19,57	11,45	8,31	6,60	19,57	27,70	0,76	0,50	1,42	0,03	0,02
2342	46,27	26,60	18,80	14,57	46,27	65,95	-0,52	0,26	1,43	-0,01	0,00

The first L-moment λ_1 values calculated from the annual maximum flow data used in the study represent the average flood flows of each station. Stations 2322 and 2335 with high λ_1 values show values such as 918.43 m^3/s and 721.58 m^3/s, respectively, indicating that these stations are located in larger drainage areas or in regions with heavy rainfall. On the other hand, λ_1 values at stations such as 2330 and 2340 are below 20 m^3/s, reflecting rivers with smaller flow regimes. The second L-moment λ_2 reveals the spread of the distribution, i.e. the annual variability in the flow data. Examples such as station 2322, where the λ_2 value is 1,313.31 m^3/s, indicate that flood values vary greatly from year to year, thus indicating areas with high flood risk. On the other hand, the fact that λ_2 is below 30 at stations such as 2330 and 2342 reveals that the flows are more constant and regular.

The L-variation coefficient, τ_2 (LC_v), is the ratio of λ_2 to λ_1 and indicates the relative variability of the flows. The τ_2 values for all stations are between 1.34 and 1.47, indicating that the floods in the Çoruh River Basin, located in the northeast of Türkiye, are quite variable. In this context, the fact that τ_2 is close to and above 1 confirms that the annual maximum flow rates at the stations differ significantly.

The third L-moment λ_3 and its corresponding τ_3 (L-skewness) value were used to evaluate whether the distribution was symmetrical. Stations with positive τ_3 values (e.g. 2304, 2333) show a right-skewed distribution, indicating that flood flows are more likely to occur above average and extreme floods are frequently observed. Stations with negative or near-zero skewness values (e.g. 2326, 2331) have a relatively more balanced or left-skewed distribution structure, representing areas where high flows are less likely to be observed. Finally, the fourth L-moment λ_4 and τ_4 (L-kurtosis) values reveal the extreme sensitivity of the distribution. Similarly, stations with negative τ_4 values (e.g. 2331 and 2336) have a flatter distribution structure and it can be said that extreme flood events rarely occur in these regions.

3.3 GOODNESS OF FIT TESTS FOR DISTRIBUTIONS

Various statistical fit tests need to be applied to determine the level of fit of the probability distributions of the flood data obtained in the previous section to each station. Such analyses evaluate the fit of the distribution functions to the observational data in line with objective criteria and ensure that the most appropriate distribution is selected.

In this study, widely accepted statistical methods such as the Chi-Squared test and the Kolmogorov-Smirnov (K-S) test will be used in order to comparatively evaluate the distribution fits. By means of these tests, the extent to which each distribution function fits the observed data will be determined and the most appropriate distribution type to be used in regional frequency analyses will be determined.

3.3.1 Chi-Squared Test

In this study, Chi-Square (χ^2) goodness of fit test was applied to evaluate the suitability of various probability distributions for each station. The analyses were performed via EasyFit software; both the calculated test statistics and the critical values at the 5% significance level, which vary according to the distributions and sample sizes, were automatically calculated by EasyFit 3.0.

The test statistic calculated for each distribution type is compared with the critical value corresponding to that distribution. If the calculated statistic is smaller than the critical value, the null hypothesis (the assumption that the data is compatible with the distribution) is not rejected; that is, it is accepted that there is no statistically significant difference between the distribution and the data.

According to the Chi-Square test results:

Normal distribution (N) was found to be suitable at 21 stations. This distribution did not provide significant fit only at stations 2304 and 2333.

Lognormal (LN), Lognormal-3 (LN3) and Generalized Extreme Value (GEV) distributions was found to be suitable at all stations.

Gumbel distribution was found to be suitable at 20 stations; only at stations 2304, 2326 and 2342 was the calculated χ^2 value of this distribution found to be higher than the critical value. As a result, the Gumbel distribution was found to be consistent for all stations except four.

Log Pearson 3 (LP3) distribution was found to be suitable at 22 stations. It was found to be suitable at all except station 2320.

Table 3.4 Results of Chi-Squared Tests and Critical Values

Station Number		N	LN	LN3	GEV	GUMBEL	LP3
2302	Calculated Statistics	0,0005	0,3346	0,0540	0,0666	0,3980	0,0629
	Critical Statistics	3,8415	3,8415	3,8415	3,8415	3,8415	3,8415
2304	Calculated Statistics	23,6930	7,4421	8,8850	6,4871	14,6580	8,8850
	Critical Statistics	11,0700	11,0700	11,0700	11,0700	11,0700	11,0700
2305	Calculated Statistics	1,8486	0,7446	1,4458	0,4818	1,5971	0,6891
	Critical Statistics	9,4877	9,4877	9,4877	9,4877	9,4877	9,4877
2316	Calculated Statistics	1,1673	1,3785	1,0464	1,2024	0,2194	1,2263
	Critical Statistics	11,0700	11,0700	11,0700	11,0700	11,0700	11,0700
2320	Calculated Statistics	1,0403	4,0980	2,1634	1,7744	0,9130	#N/A
	Critical Statistics	11,0700	11,0700	11,0700	11,0700	11,0700	
2321	Calculated Statistics	0,9647	2,3937	2,3939	2,4659	3,4157	2,4209
	Critical Statistics	9,4877	9,4877	9,4877	9,4877	9,4877	9,4877
2322	Calculated Statistics	1,8126	1,8529	3,8007	1,6187	2,1465	1,8175
	Critical Statistics	7,8147	7,8147	7,8147	7,8147	7,8147	7,8147
2323	Calculated Statistics	2,6607	1,2184	1,2184	1,2430	2,3174	1,2257
	Critical Statistics	11,0700	11,0700	11,0700	11,0700	11,0700	11,0700
2325	Calculated Statistics	6,4437	4,7449	4,6492	5,3399	5,6893	3,9133
	Critical Statistics	7,8147	7,8147	7,8147	7,8147	7,8147	7,8147
2326	Calculated Statistics	3,1031	2,4918	2,3462	2,3514	7,8268	2,2772
	Critical Statistics	5,9915	5,9915	5,9915	5,9915	5,9915	5,9915
2327	Calculated Statistics	1,5306	1,1421	1,1492	1,1800	1,1485	1,1781
	Critical Statistics	5,9915	5,9915	5,9915	5,9915	5,9915	5,9915
2328	Calculated Statistics	3,9508	3,5651	0,7445	5,4344	5,5672	3,4859
	Critical Statistics	7,8147	7,8147	7,8147	7,8147	7,8147	7,8147
2329	Calculated Statistics	0,7185	1,4664	2,0145	1,3336	1,8273	1,4085
	Critical Statistics	7,8147	7,8147	7,8147	7,8147	7,8147	7,8147
2330	Calculated Statistics	4,2404	2,0712	0,2104	1,7780	1,0303	1,0673
	Critical Statistics	7,8147	7,8147	7,8147	7,8147	7,8147	7,8147
2331	Calculated Statistics	1,0929	0,4342	0,1458	1,0853	0,5676	0,3567
	Critical Statistics	3,8415	3,8415	3,8415	3,8415	3,8415	3,8415
2333	Calculated Statistics	4,0646	0,0225	0,1828	0,2217	0,0483	0,2336
	Critical Statistics	3,8415	3,8415	3,8415	3,8415	3,8415	3,8415
2334	Calculated Statistics	0,9184	0,3626	0,9654	1,6550	0,5337	0,9110
	Critical Statistics	5,9915	5,9915	5,9915	5,9915	5,9915	5,9915
2335	Calculated Statistics	1,4100	1,0672	0,0374	0,8830	0,8862	1,0808
	Critical Statistics	5,9915	5,9915	5,9915	5,9915	5,9915	5,9915
2336	Calculated Statistics	0,5254	1,0793	0,2430	0,1921	1,1678	1,2007
	Critical Statistics	3,8415	3,8415	3,8415	3,8415	3,8415	3,8415
2337	Calculated Statistics	0,1297	0,1869	0,2025	0,3154	0,2484	0,1740
	Critical Statistics	5,9915	5,9915	5,9915	5,9915	5,9915	5,9915
2338	Calculated Statistics	0,5407	3,6638	0,1685	0,5038	1,6298	0,1586
	Critical Statistics	3,8415	3,8415	3,8415	3,8415	3,8415	3,8415
2340	Calculated Statistics	1,3699	1,5696	1,5718	1,1314	1,5831	1,5417
	Critical Statistics	5,9915	5,9915	5,9915	5,9915	5,9915	5,9915
2342	Calculated Statistics	2,5547	0,0048	2,8328	1,9983	5,4970	2,8471
	Critical Statistics	3,8415	3,8415	3,8415	3,8415	3,8415	3,8415

Table 3.5 Results of Chi-Squared Tests

Station Number	N	LN	LN3	GEV	GUMBEL	LP3
2302	+	+	+	+	+	+
2304		+	+	+		+
2305	+	+	+	+	+	+
2316	+	+	+	+	+	+
2320	+	+	+	+	+	
2321	+	+	+	+	+	+
2322	+	+	+	+	+	+
2323	+	+	+	+	+	+
2325	+	+	+	+	+	+
2326	+	+	+	+		+
2327	+	+	+	+	+	+
2328	+	+	+	+	+	+
2329	+	+	+	+	+	+
2330	+	+	+	+	+	+
2331	+	+	+	+	+	+
2333		+	+	+	+	+
2334	+	+	+	+	+	+
2335	+	+	+	+	+	+
2336	+	+	+	+	+	+
2337	+	+	+	+	+	+
2338	+	+	+	+	+	+
2340	+	+	+	+	+	+
2342	+	+	+	+		+

3.3.2 Kolmogorov-Smirnov Test

In this part, the Kolmogorov-Smirnov (K-S) test was applied to determine the most appropriate probability distribution of flood data for each station. In the applied test; the cumulative distribution functions of the Normal (N), Log-Normal (LN), Three Parameter Log-Normal (LN3), Gumbel and Generalized Extreme Value (GEV) distributions were compared with the observational distribution. The maximum difference Δ values calculated for each distribution and the corresponding critical $\Delta_{critical}$ values are presented in Table 6.

As a result of the comparisons made, it was decided whether a distribution was suitable for the station under the condition $\Delta < \Delta_{critical}$. Distributions that met this condition were accepted as "suitable", and those that did not were evaluated as "not suitable". The distribution that fits the data for each station is summarized in Table 7.

According to the K-S test results, the majority of the tested distributions provided statistically significant fit to all 23 stations. As a result of comparing the calculated Δ values with the relevant $\Delta_{critical}$ thresholds, it was determined that the Normal distribution exceeded the threshold value only at station number 2333 and therefore was not suitable. For all other distribution and station combinations, the $\Delta < \Delta_{critical}$ condition was met. This situation reveals that especially the LN3 distribution shows a high level of general fit and indicates that this distribution should be evaluated as a priority in modeling studies.

Table 3.6 Results of K-S Tests and Critical Values

Station Number		N	LN	LN3	GEV	GUMBEL
2302	Calculated Δ	0,1486	0,1312	0,1167	0,1226	0,1269
	Critical Δ	0,2872	0,2872	0,2872	0,1938	0,2968
2304	Calculated Δ	0,1575	0,1057	0,1024	0,0958	0,1219
	Critical Δ	0,1709	0,1709	0,1709	0,1137	0,1741
2305	Calculated Δ	0,0892	0,0660	0,0678	0,0570	0,0764
	Critical Δ	0,1922	0,1922	0,1922	0,1282	0,1963
2316	Calculated Δ	0,1119	0,0643	0,0701	0,0549	0,0607
	Critical Δ	0,1962	0,1962	0,1962	0,1309	0,2005
2320	Calculated Δ	0,1040	0,1193	0,0799	0,0664	0,0652
	Critical Δ	0,2154	0,2154	0,2154	0,1441	0,2206
2321	Calculated Δ	0,0979	0,0753	0,0740	0,0688	0,0918
	Critical Δ	0,2127	0,2127	0,2127	0,1422	0,2178
2322	Calculated Δ	0,1483	0,1176	0,1257	0,1068	0,1333
	Critical Δ	0,2499	0,2499	0,2499	0,1678	0,2570
2323	Calculated Δ	0,1209	0,0547	0,0547	0,0548	0,0578
	Critical Δ	0,1962	0,1962	0,1962	0,1309	0,2005
2325	Calculated Δ	0,1850	0,1191	0,1193	0,1213	0,1255
	Critical Δ	0,2212	0,2212	0,2212	0,1480	0,2267
2326	Calculated Δ	0,1620	0,1593	0,1563	0,1410	0,1677
	Critical Δ	0,2749	0,2749	0,2749	0,1852	0,2836
2327	Calculated Δ	0,1513	0,1177	0,1102	0,1209	0,1151
	Critical Δ	0,3179	0,3179	0,3179	0,2154	0,3298
2328	Calculated Δ	0,1250	0,1192	0,1282	0,1064	0,1245
	Critical Δ	0,2457	0,2457	0,2457	0,1649	0,2525
2329	Calculated Δ	0,1417	0,0985	0,0958	0,0914	0,0984
	Critical Δ	0,2417	0,2417	0,2417	0,1621	0,2483
2330	Calculated Δ	0,1775	0,1174	0,1024	0,1064	0,1099
	Critical Δ	0,2417	0,2417	0,2417	0,1621	0,2483
2331	Calculated Δ	0,1934	0,2388	0,2814	0,1977	0,2526
	Critical Δ	0,3614	0,3614	0,3614	0,2463	0,3772
2333	Calculated Δ	0,2974	0,2113	0,1580	0,1225	0,2270
	Critical Δ	0,2941	0,2941	0,2941	0,1986	0,3041
2334	Calculated Δ	0,0983	0,0983	0,0985	0,0836	0,1098
	Critical Δ	0,2872	0,2872	0,2872	0,1938	0,2968
2335	Calculated Δ	0,1630	0,1221	0,1416	0,1277	0,1147
	Critical Δ	0,3014	0,3014	0,3014	0,2037	0,3120
2336	Calculated Δ	0,2332	0,1943	0,1409	0,1652	0,1806
	Critical Δ	0,3376	0,3376	0,3376	0,2293	0,3512
2337	Calculated Δ	0,0880	0,1021	0,0987	0,0821	0,1189
	Critical Δ	0,2809	0,2809	0,2809	0,1893	0,2900
2338	Calculated Δ	0,1287	0,1405	0,1108	0,1251	0,1302
	Critical Δ	0,2941	0,2941	0,2941	0,1986	0,3041
2340	Calculated Δ	0,1963	0,1687	0,1678	0,1658	0,1622
	Critical Δ	0,2941	0,2941	0,2941	0,1986	0,3041
2342	Calculated Δ	0,1850	0,2254	0,1899	0,1667	0,2509
	Critical Δ	0,3014	0,3014	0,3014	0,2037	0,3120

Table 3.7 Results of K-S Tests

Station Number	N	LN	LN3	GEV	GUMBEL
2302	+	+	+	+	+
2304	+	+	+	+	+
2305	+	+	+	+	+
2316	+	+	+	+	+
2320	+	+	+	+	+
2321	+	+	+	+	+
2322	+	+	+	+	+
2323	+	+	+	+	+
2325	+	+	+	+	+
2326	+	+	+	+	+
2327	+	+	+	+	+
2328	+	+	+	+	+
2329	+	+	+	+	+
2330	+	+	+	+	+
2331	+	+	+	+	+
2333		+	+	+	+
2334	+	+	+	+	+
2335	+	+	+	+	+
2336	+	+	+	+	+
2337	+	+	+	+	+
2338	+	+	+	+	+
2340	+	+	+	+	+
2342	+	+	+	+	+

3.4 PARAMETERS OF PROBABILITY DISTRIBUTIONS FOR STATION DATA

In Table 8, the parameter values obtained for each station for two-parameter probability distributions are presented. Similarly, the station-based parameter values for three-parameter distributions are shown in Table 9.

Table 3.8 Estimation of parameters of two-parameter distributions

Station Number	N		LN		GUMBEL	
	$\bar{x}$	S_x	$\bar{y}$	S_y	μ	α
2302	70,02	30,03	4,16	0,44	56,49	23,46
2304	94,27	48,75	4,43	0,47	72,33	38,01
2305	388,17	121,96	5,91	0,31	333,28	95,09
2316	257,00	96,50	5,48	0,36	213,57	75,24
2320	173,03	69,99	5,05	0,49	141,53	54,57
2321	76,35	18,20	4,31	0,23	68,16	14,15
2322	918,43	221,77	6,79	0,24	818,62	172,92
2323	235,45	91,59	5,39	0,38	194,23	71,41
2325	82,82	47,84	4,24	0,62	61,29	37,30
2326	70,86	29,50	4,17	0,44	57,59	23,01
2327	148,41	52,98	4,94	0,35	124,57	41,31
2328	46,25	18,06	3,76	0,39	38,12	14,08
2329	174,61	82,19	5,05	0,47	137,62	64,09
2330	18,93	10,21	2,82	0,49	14,34	7,96
2331	52,91	16,01	3,92	0,30	45,71	12,48
2333	33,09	16,21	3,42	0,38	25,80	12,64
2334	169,89	54,29	5,08	0,32	145,46	42,33
2335	721,58	294,71	6,51	0,38	588,94	229,78
2336	71,77	21,35	4,23	0,28	62,16	16,64
2337	344,86	114,92	5,79	0,34	293,14	89,61
2338	212,09	73,35	5,30	0,36	179,09	57,19
2340	19,57	5,98	2,93	0,29	16,88	4,66
2342	46,27	11,99	3,80	0,27	40,88	9,35

Table 3.9 Estimation of parameters of three-parameter distributions

Station Number	LN3			GEV			LP3		
	γ	$\bar{y}$	S_y	k	μ	α	α	β	γ
2302	4,67	-39,83	0,26	-0,08	57,29	25,33	10,34	-0,14	5,62
2304	-5,88	4,51	0,44	0,18	70,68	29,79	1312,10	-0,01	21,90
2305	33,35	5,81	0,34	-0,07	333,61	106,24	4152,10	-0,01	26,19
2316	66,93	5,12	795,57	0,03	211,31	75,36	151,89	0,03	1,01
2320	-365,46	6,28	0,13	-0,13	143,47	63,76	1,55	-0,40	5,68
2321	2,67	4,27	0,24	-0,10	68,42	16,23	10543,00	0,00	28,69
2322	332,39	6,30	0,38	-0,10	820,14	201,91	979,60	0,01	-0,75
2323	-0,12	5,39	0,38	0,02	192,65	71,43	2725,90	-0,01	25,38
2325	0,67	4,37	4,16	-0,01	60,29	39,53	122,22	-0,06	11,16
2326	-15,21	4,40	0,35	-0,15	58,36	28,05	55,07	-0,06	7,50
2327	-32,45	5,16	0,28	-0,01	124,39	42,65	81,87	-0,04	8,17
2328	5,79	3,60	0,46	-0,08	38,15	16,01	490,71	-0,02	12,63
2329	31,03	4,80	0,60	0,02	135,22	66,38	18070,00	0,00	69,11
2330	4,19	2,47	0,68	0,14	13,98	6,73	79,17	0,06	-1,61
2331	33,22	2,30	1,48	-0,22	46,73	15,82	236,47	-0,02	8,74
2333	13,61	2,73	0,69	0,35	25,27	7,08	3,10	0,22	2,73
2334	-57,30	5,40	0,23	-0,13	146,67	50,68	71,94	-0,04	7,88
2335	224,09	6,04	0,60	0,08	575,00	222,34	86,38	0,04	2,83
2336	36,09	3,39	0,62	0,03	61,34	17026,00	44,26	0,04	2,30
2337	-151,43	6,18	0,23	-0,16	296,50	110,09	59,51	-0,04	8,48
2338	-142,91	5,85	0,20	-0,09	179,72	65,33	18,73	-0,08	6,87
2340	0,31	2,91	0,30	0,09	16,62	4,39	10920,00	0,00	-28,49
2342	-246,23	5368,00	0,04	-0,42	13,02	42,78	10,79	-0,08	4,72

3.5 ESTIMATION OF FLOOD FLOW RATES AT VARIOUS RETURN INTERVALS

In this section, the distributed distribution models in the 2.2 section were applied to each station separately; then, flood flows corresponding to 2, 5, 10, 25, 50, 100 and 500-year return periods were estimated.

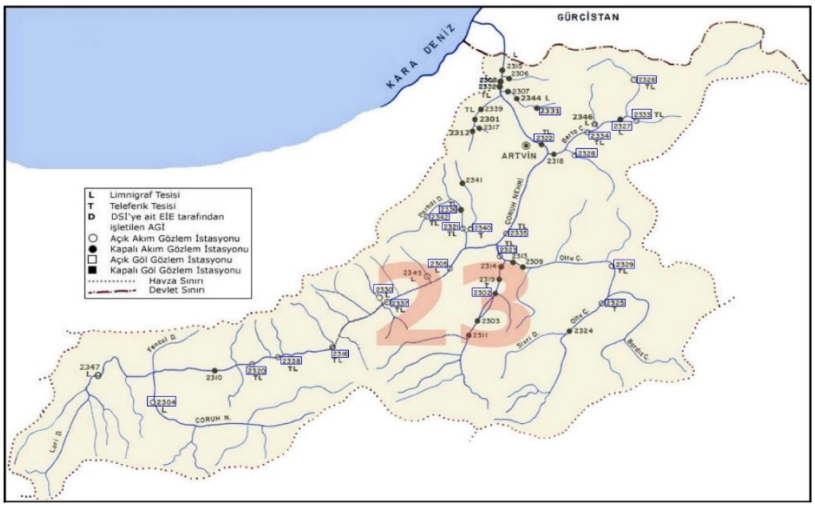


Figure 3.3 Stations Used in Flood Analysis in Çoruh River Basin

The locations of the river regions of the 23 stations used in the analyses are presented in Figure 7. In this way, the spatial data at the points where flood estimates are formed can be monitored and evaluations can be made in terms of regional analyses.

The distributions preferred in the study were selected both because they are widely used in the literature within the scope of flood analyses and because of the successful fit they provide in practice. In addition, EasyFit v3.0 software, which evaluates the statistical fit of observational data with different distributions, was also used in the distribution selection process. EasyFit is a professional statistical analysis program that can make comparative statistical fit assessments by fitting various distributions to the data with its user-friendly interface and using tests such as Kolmogorov-Smirnov and Chi-Squared. In the

preliminary analyses conducted with this software, it was observed that the six distributions selected (Normal, LN, LN3, Gumbel, GEV etc.) were among the top 10 distributions that provided the best fit at the vast majority of stations. Therefore, these distributions were selected to be used in the study in line with both theoretical justifications and software-supported empirical analysis results.

The visual outputs of the analysis results made with the relevant software are presented in the figures between Figure 8 and Figure 10. In these figures, the "Goodness of Fit" rankings obtained for each station can be seen in detail.

(a) Station 2302 **(b)** Station 2304 **(c)** Station 2305

(d) Station 2316 **(e)** Station 2320 **(f)** Station 2321

(g) Station 2322 **(h)** Station 2323 **(i)** Station 2325

Figure 3.4 Goodness-of-fit evaluations for selected stations (1/3)

Goodness of Fit - Summary

#	Distribution	Kolmogorov Smirnov		Chi-Squared	
		Statistic	Rank	Statistic	Rank
16	Power Function	0,11022	1	0,23727	2
20	Uniform	0,12182	2	N/A	
6	Kumaraswamy	0,12598	3	N/A	
3	Gen. Extreme Value	0,14096	4	2,3514	10
9	Log-Pearson 3	0,14791	5	2,2772	8
12	Lognormal (3P)	0,15633	6	2,3462	9
11	Lognormal	0,15934	7	2,4916	12
13	Normal	0,16198	8	3,1031	14
19	Triangular	0,16483	9	3,0011	5
4	Gumbel Max	0,16774	10	7,8266	16
17	Reciprocal	0,17951	11	1,7626	6
2	Exponential (2P)	0,18108	12	1,1919	4
10	Logistic	0,18468	13	3,3025	15
5	Gumbel Min	0,1962	14	2,7069	13
14	Pareto	0,20641	15	1,0115	3
1	Exponential	0,33313	16	2,0661	7
8	Levy (2P)	0,35625	17	1,8027E-4	1
15	Pareto 2	0,37915	18	2,4526	11
7	Levy	0,51022	19	8,3713	17
18	Student's t	0,99934	20	49956,0	18

(a) Station 2326

Goodness of Fit - Summary

#	Distribution	Kolmogorov Smirnov		Chi-Squared	
		Statistic	Rank	Statistic	Rank
12	Lognormal (3P)	0,11017	1	1,1492	4
4	Gumbel Max	0,11511	2	1,1483	3
11	Lognormal	0,1177	3	1,1421	2
9	Log-Pearson 3	0,11837	4	1,1781	5
3	Gen. Extreme Value	0,1206	5	1,18	6
13	Normal	0,13131	6	1,9306	7
10	Logistic	0,13974	7	1,4352	8
20	Uniform	0,17734	8	N/A	
19	Triangular	0,19804	9	2,8516	11
17	Reciprocal	0,19814	10	4,7647	14
5	Gumbel Min	0,22077	11	3,9137	12
2	Exponential (2P)	0,27616	12	3,9427	13
6	Kumaraswamy	0,2857	13	2,5098	10
8	Levy (2P)	0,33427	14	11,767	15
16	Power Function	0,34383	15	2,121	9
14	Pareto	0,35661	16	0,18888	1
15	Pareto 2	0,39484	17	13,832	17
1	Exponential	0,40258	18	13,798	16
7	Levy	0,50827	19	23,219	18
18	Student's t	0,9999	20	1,1309E+5	19

(b) Station 2327

Goodness of Fit - Summary

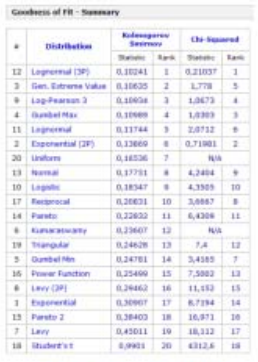
#	Distribution	Kolmogorov Smirnov		Chi-Squared	
		Statistic	Rank	Statistic	Rank
19	Triangular	0,08184	1	0,23757	1
6	Kumaraswamy	0,0864	2	N/A	
20	Uniform	0,10431	3	N/A	
3	Gen. Extreme Value	0,10637	4	3,4344	12
17	Reciprocal	0,10804	5	0,48276	2
9	Log-Pearson 3	0,11193	6	3,4859	7
11	Lognormal	0,11922	7	3,5651	8
4	Gumbel Max	0,12443	8	5,5672	13
13	Normal	0,12498	9	3,9558	10
12	Lognormal (3P)	0,12823	10	0,74454	4
10	Logistic	0,14723	11	3,7109	9
2	Exponential (2P)	0,15619	12	0,51105	3
5	Gumbel Min	0,17191	13	2,9039	6
14	Pareto	0,19724	14	4,0213	11
16	Power Function	0,21637	15	6,0046	14
8	Levy (2P)	0,30792	16	1,9896	5
1	Exponential	0,37177	17	10,119	15
15	Pareto 2	0,4218	18	16,149	16
7	Levy	0,49969	19	21,186	17
18	Student's t	0,99892	20	30749,0	18

(c) Station 2328

Goodness of Fit - Summary

#	Distribution	Kolmogorov Smirnov		Chi-Squared	
		Statistic	Rank	Statistic	Rank
19	Triangular	0,08574	1	0,13236	1
3	Gen. Extreme Value	0,09142	2	1,3336	5
9	Log-Pearson 3	0,09429	3	1,4085	6
12	Lognormal (3P)	0,09577	4	0,9145	9
4	Gumbel Max	0,09833	5	1,8273	8
11	Lognormal	0,09843	6	1,4664	7
17	Reciprocal	0,10367	7	0,91667	4
2	Exponential (2P)	0,11329	8	0,61796	2
20	Uniform	0,1347	9	N/A	
13	Normal	0,14171	10	0,71652	3
10	Logistic	0,15554	11	3,536	12
14	Pareto	0,20582	12	3,0197	11
5	Gumbel Min	0,20732	13	5,1138	13
6	Kumaraswamy	0,2266	14	4,3333	13
16	Power Function	0,24658	15	4,9428	14
8	Levy (2P)	0,28233	16	2,0932	10
1	Exponential	0,33258	17	7,9859	17
15	Pareto 2	0,33903	18	7,8211	16
7	Levy	0,44081	19	20,233	18
18	Student's t	0,9999	20	4,3577E+5	19

(d) Station 2329



Goodness of Fit - Summary

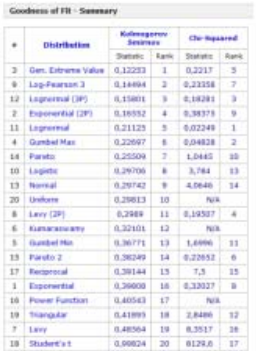
#	Distribution	Kolmogorov Smirnov		Chi-Squared	
		Statistic	Rank	Statistic	Rank
12	Lognormal (3P)	0,10241	1	0,21037	1
3	Gen. Extreme Value	0,10635	2	1,778	5
9	Log-Pearson 3	0,10656	3	1,0673	4
4	Gumbel Max	0,10989	4	1,0303	3
11	Lognormal	0,11744	5	2,0712	6
2	Exponential (2P)	0,13869	6	0,71901	2
20	Uniform	0,16536	7	N/A	
13	Normal	0,17751	8	4,2404	9
10	Logistic	0,18547	9	4,3505	10
17	Reciprocal	0,20831	10	3,6867	8
14	Pareto	0,22632	11	6,4306	11
6	Kumaraswamy	0,23607	12	N/A	
19	Triangular	0,24628	13	7,4	12
5	Gumbel Min	0,24781	14	5,4185	7
16	Power Function	0,25469	15	7,5802	13
8	Levy (2P)	0,26462	16	11,152	15
1	Exponential	0,30967	17	8,7194	14
15	Pareto 2	0,38403	18	16,971	16
7	Levy	0,45011	19	18,112	17
18	Student's t	0,9901	20	4312,6	18

(e) Station 2330

Goodness of Fit - Summary

#	Distribution	Kolmogorov Smirnov		Chi-Squared	
		Statistic	Rank	Statistic	Rank
20	Uniform	0,16912	1	0,52656	7
13	Normal	0,19344	2	1,0929	13
3	Gen. Extreme Value	0,19774	3	1,0853	12
17	Reciprocal	0,2005	4	1,2564	14
10	Logistic	0,2118	5	1,2767	15
5	Gumbel Min	0,2212	6	0,95701	11
9	Log-Pearson 3	0,22393	7	0,35671	3
11	Lognormal	0,23081	8	0,43418	6
8	Levy (2P)	0,23933	9	0,71089	10
19	Triangular	0,24651	10	1,4023	17
4	Gumbel Max	0,25261	11	0,56758	8
2	Exponential (2P)	0,27458	12	0,00148	1
6	Kumaraswamy	0,27957	13	0,05444	2
12	Lognormal (3P)	0,28143	14	0,14581	3
14	Pareto	0,2891	15	0,16868	4
16	Power Function	0,34726	16	N/A	
15	Pareto 2	0,46098	17	1,3694	16
1	Exponential	0,47105	18	1,4131	18
7	Levy	0,5618	19	0,66604	9
18	Student's t	0,99958	20	19040,0	19

(f) Station 2331



Goodness of Fit - Summary

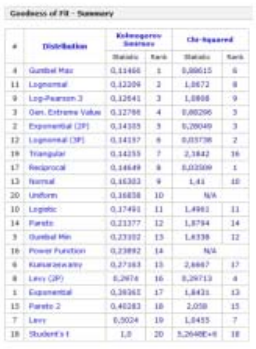
#	Distribution	Kolmogorov Smirnov		Chi-Squared	
		Statistic	Rank	Statistic	Rank
3	Gen. Extreme Value	0,12233	1	0,2217	5
9	Log-Pearson 3	0,14494	2	0,23358	7
12	Lognormal (3P)	0,15801	3	0,18291	3
2	Exponential (2P)	0,16552	4	0,38379	9
11	Lognormal	0,21125	5	0,02249	1
4	Gumbel Max	0,22697	6	0,04828	2
14	Pareto	0,25509	7	1,0445	10
10	Logistic	0,29706	8	3,784	13
13	Normal	0,29742	9	4,0646	14
20	Uniform	0,29813	10	N/A	
8	Levy (2P)	0,2989	11	0,19307	4
6	Kumaraswamy	0,32101	12	N/A	
5	Gumbel Min	0,36771	13	1,6896	11
15	Pareto 2	0,38249	14	0,22652	6
17	Reciprocal	0,38144	15	7,5	15
1	Exponential	0,39808	16	0,32027	8
16	Power Function	0,40543	17	N/A	
19	Triangular	0,41893	18	2,8486	12
7	Levy	0,48564	19	8,3517	16
18	Student's t	0,99824	20	8129,0	17

(g) Station 2333

Goodness of Fit - Summary

#	Distribution	Kolmogorov Smirnov		Chi-Squared	
		Statistic	Rank	Statistic	Rank
3	Gen. Extreme Value	0,08338	1	1,693	12
9	Log-Pearson 3	0,09025	2	0,91095	6
13	Normal	0,09833	3	0,91843	7
11	Lognormal	0,09834	4	0,9426	8
12	Lognormal (3P)	0,09849	5	0,96562	9
20	Uniform	0,1076	6	N/A	
4	Gumbel Max	0,10984	7	0,53372	5
10	Logistic	0,11393	8	0,28624	1
19	Triangular	0,12278	9	1,1158	10
17	Reciprocal	0,13	10	1,0238	9
6	Kumaraswamy	0,1486	11	0,30794	4
5	Gumbel Min	0,16229	12	1,556	11
2	Exponential (2P)	0,18751	13	0,46887	3
14	Pareto	0,28823	14	3,2791	13
8	Levy (2P)	0,32916	15	7,3787	14
16	Power Function	0,34019	16	10,96	17
1	Exponential	0,40743	17	7,5422	15
15	Pareto 2	0,41908	18	7,6839	16
7	Levy	0,51199	19	15,218	18
18	Student's t	0,99994	20	4,2542E+5	19

(h) Station 2334

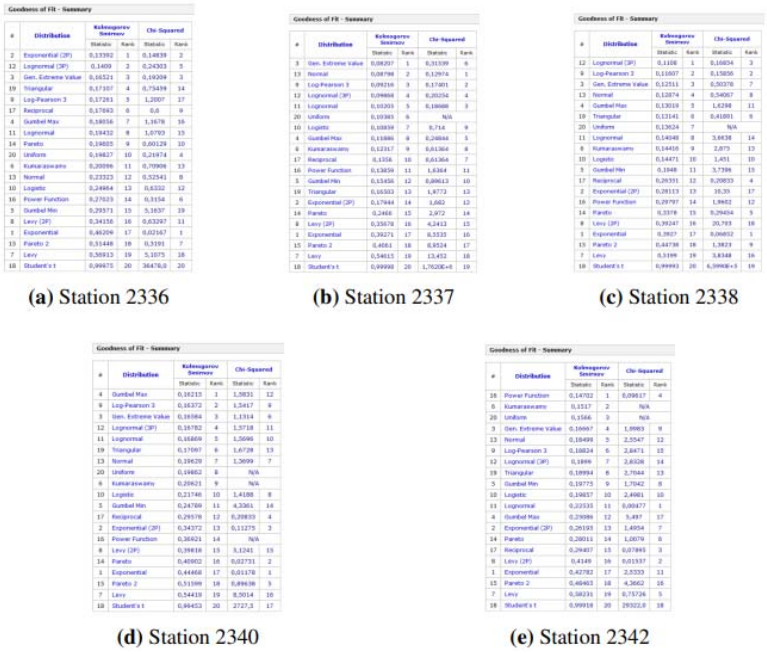


Goodness of Fit - Summary

#	Distribution	Kolmogorov Smirnov		Chi-Squared	
		Statistic	Rank	Statistic	Rank
4	Gumbel Max	0,11466	1	0,88615	6
11	Lognormal	0,12209	2	1,0672	8
9	Log-Pearson 3	0,12641	3	1,0808	9
3	Gen. Extreme Value	0,12786	4	0,88296	5
2	Exponential (2P)	0,14105	5	0,28049	3
12	Lognormal (3P)	0,14197	6	0,03738	2
19	Triangular	0,14253	7	2,1842	16
17	Reciprocal	0,14649	8	0,02509	1
13	Normal	0,16303	9	1,41	10
20	Uniform	0,16838	10	N/A	
10	Logistic	0,17491	11	1,4961	11
14	Pareto	0,21377	12	1,8794	14
5	Gumbel Min	0,23102	13	1,8338	12
16	Power Function	0,23892	14	N/A	
6	Kumaraswamy	0,27163	15	2,0667	17
8	Levy (2P)	0,2674	16	0,29713	4
1	Exponential	0,38365	17	1,8431	13
15	Pareto 2	0,40283	18	2,058	15
7	Levy	0,5024	19	1,8455	7
18	Student's t	1,0	20	5,2648E+8	18

(i) Station 2335

Figure 3.5 Goodness-of-fit evaluations for selected stations (2/3)



(a) Station 2336 (b) Station 2337 (c) Station 2338

(d) Station 2340 (e) Station 2342

Figure 3.6 Goodness-of-fit evaluations for selected stations (3/3)

Table 3.10 Estimations for various Return Period (T) of station 2302

	X(0.5)	X(0.8)	X(0.9)	X(0.96)	X(0.98)	X(0.99)	X(0.998)
Lognormal (3P)	66,36	92,33	108,34	127,56	141,28	154,59	184,58
Gen. Extreme Value	66,44	93,07	109,4	128,65	142,01	154,53	180,9
Gumbel Max	65,08	91,67	109,27	131,51	148,01	164,39	202,23
Lognormal	63,94	92,84	112,83	138,9	158,87	179,27	228,94
Normal	70,02	95,34	108,58	122,69	131,81	140,01	156,61
Log Pearson 3	66,99	94,37	110,09	127,49	138,86	149,07	169,29

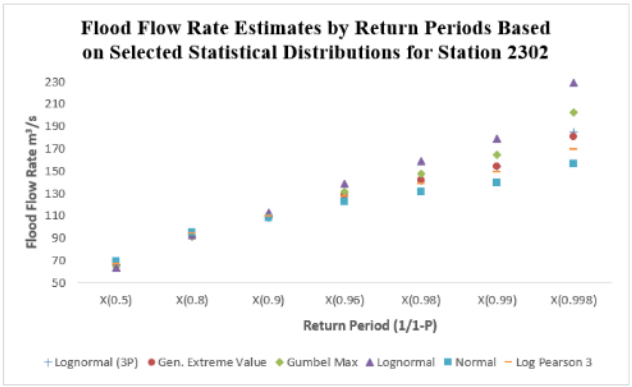


Figure 3.7 Discharge Estimates by Return Periods Based on Selected Statistical Distributions for Station 2302

When the flood flow rates estimated according to the return probabilities of station number 2302 are evaluated, the highest estimated flow rate at low return intervals (e.g. 50% and 80%) is given by the Normal distribution, while the lowest estimates are made by the Lognormal and Gumbel Max distributions, respectively. Especially at 50% probability, the Normal distribution gives the highest value with 70.02 m^3/s, while the Lognormal distribution predicts the lowest value with 63.94 m^3/s. This situation shows that the Normal distribution offers relatively less conservative estimates at low return intervals. On the other hand, at high return intervals (99% and 99.8%), the highest flood flow rate estimates are given by the Gumbel Max (202.23 m^3/s) and Lognormal (228.94 m^3/s) distributions, respectively. This shows that the Gumbel Max and Lognormal distributions tend to predict extreme values more severely. The lowest estimates are given by the Normal distribution at these high return intervals. These results reflect the sensitivity of each distribution to extreme events and point out the importance of which distribution to choose in engineering applications, especially in infrastructure designs.

Table 3.11 Estimations for various Return Period (T) of station 2304

	X(0.5)	X(0.8)	X(0.9)	X(0.96)	X(0.98)	X(0.99)	X(0.998)
Lognormal (3P)	84,73	125,69	154	190,94	219,23	248,13	318,5
Gen. Extreme Value	81,97	121,99	153,35	199,54	239,28	284,03	411,73
Gumbel Max	86,26	129,34	157,87	193,91	220,64	247,18	308,5
Lognormal	84,07	125,73	155,18	194,22	224,52	255,79	333,05
Normal	94,27	135,3	156,75	179,61	194,39	207,68	234,58
Log Pearson 3	84,44	126,31	155,52	193,78	223,13	253,15	326,17

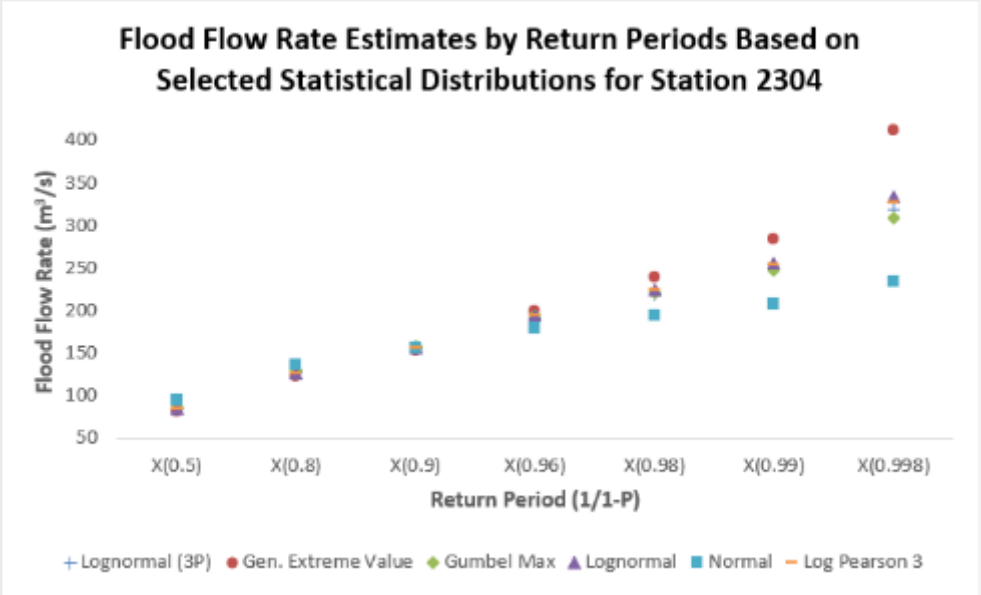


Figure 3.8 Discharge Estimates by Return Periods Based on Selected Statistical Distributions for Station 2304

When the flood flows estimated according to the return intervals of station 2304 are examined, in the low return intervals (X(0.5) and X(0.8)), the highest value is given by the Normal distribution (94.27 m^3/s and 135.3 m^3/s), while the lowest value is predicted by the Generalized Extreme Value (GEV) distribution (81.97 m^3/s and 121.99 m^3/s). This situation shows that the Normal distribution offers a cautious approach by predicting higher flood flows in short-term predictions.

In the medium and high return intervals (X(0.9)–X(0.998)), the highest prediction values are mostly given by the GEV distribution; especially at 99.8% probability, the GEV distribution predicted the highest flow with 411.73 m^3/s, which shows that it exhibits a structure that is quite sensitive to extreme events. The lowest estimates in the same ranges are generally provided by the Normal distribution (e.g. 234.58 m^3/s for X(0.998)), which reveals that it tends to underestimate extreme events.

As a result, the GEV distribution stands out especially in high return ranges and can be considered as an alternative to be considered in terms of safety in

critical engineering applications. The Normal Distribution, on the other hand, is notable for its structure that emphasizes upper limit values in short-term estimates and is not conservative.

Table 3.12 Estimations for various Return Period (T) of station 2305

	X(0.5)	X(0.8)	X(0.9)	X(0.96)	X(0.98)	X(0.99)	X(0.998)
Lognormal (3P)	368,12	480,53	553,59	644,7	711,88	778,58	934,36
Gen. Extreme Value	372,06	485,06	555,2	638,84	697,47	725,95	871,46
Gumbel Max	368,13	475,91	547,27	637,43	704,32	770,71	924,14
Lognormal	369,96	480,81	551,4	638,14	701,3	763,43	906,56
Normal	388,17	490,81	544,46	601,68	638,64	671,89	739,19
Log Pearson 3	370,56	482,37	553,15	639,67	702,36	763,79	904,41

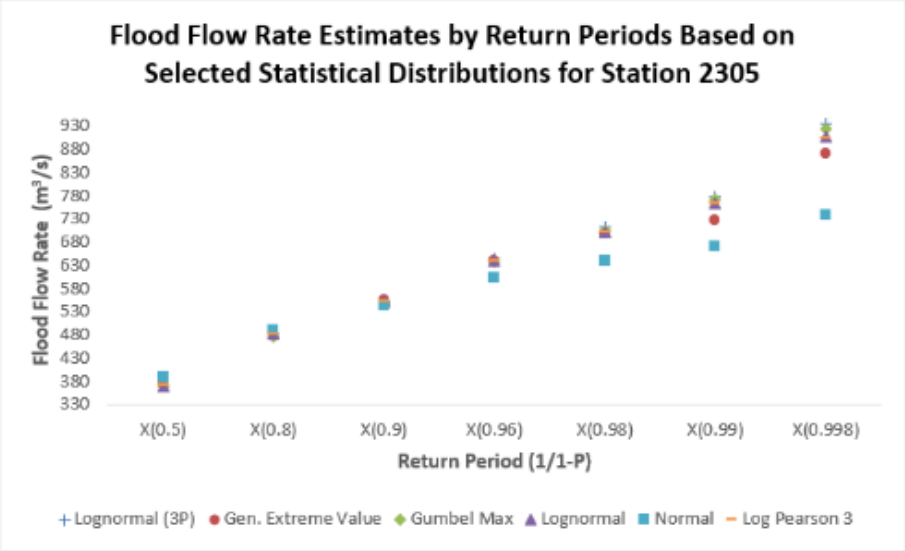


Figure 3.9 Discharge Estimates by Return Periods Based on Selected Statistical Distributions for Station 2305

When the flood flow rates calculated for the return intervals of station number 2305 are examined, the highest estimates in the low return intervals (X(0.5), X(0.8)) were made by the Normal distribution (388.17 m^3/s and 490.81 m^3/s). This situation shows that the Normal distribution offers a more cautious approach in short-term estimates and predicts the potential flood risk higher in low return intervals compared to other distributions.

Although the estimate values in the medium return intervals (X(0.9), X(0.96)) are quite close to each other, the Generalized Extreme Value (GEV) and Lognormal (3P) distributions stand out. While the Lognormal (3P) made the highest estimate with 644.7 m^3/s in the X(0.96) return interval, the GEV distribution came in second with 638.84 m^3/s. These similarities indicate that these distributions show close performances in medium-term extreme estimates.

At high return intervals (X(0.98), X(0.99), X(0.998)), the highest flow rate estimates were again made by the Lognormal (3P) distribution, and especially at the X(0.998) return interval, it provided the highest prediction with a value of 934.36 m^3/s. This distribution was followed by the Gumbel Maximum with 924.14 m^3/s and the classical Lognormal distribution with 906.56 m^3/s. These results show that the Lognormal (3P) distribution provides the most sensitive estimate against extreme events.

In summary, while the Normal distribution stands out in short-term estimates at station number 2305, the Lognormal (3P) distribution draws attention by estimating the highest flood flow rates in long-term estimates. This situation emphasizes the importance of choosing the distribution depending on the return interval when making risk assessments in engineering projects.

Table 3.13 Estimations for various Return Period (T) of station 2316

	X(0.5)	X(0.8)	X(0.9)	X(0.96)	X(0.98)	X(0.99)	X(0.998)
Lognormal (3P)	234,52	324,49	389,36	476,64	545,21	616,65	795,57
Gen. Extreme Value	239,07	326,81	386,48	463,73	522,41	581,84	723,81
Gumbel Max	241,15	326,43	382,89	454,23	507,16	556,69	681,09
Lognormal	240,82	325,67	381,34	451,22	503,03	554,7	676,1
Normal	257	338,22	380,67	425,94	455,19	481,49	534,75
Log Pearson 3	238,47	325,71	385,53	463,44	523,19	584,43	734,64

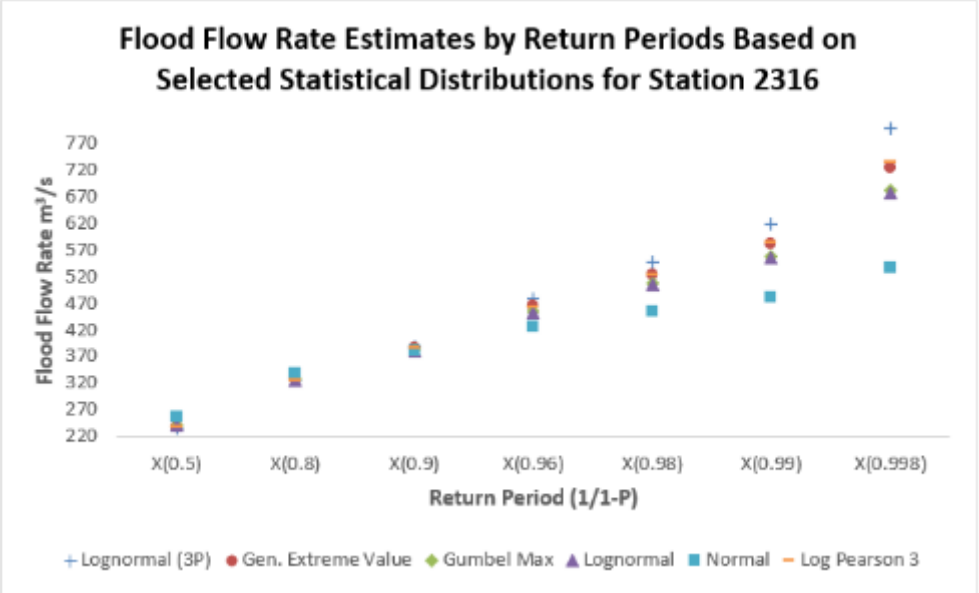


Figure 3.10 Discharge Estimates by Return Periods Based on Selected Statistical Distributions for Station 2316

According to the flood analysis results made for station number 2316 at various return intervals, the highest flood discharge estimates were systematically produced by the Lognormal (3P) distribution. Especially at 100 and 500 year return intervals, this distribution presents values such as 616.65 m^3/s and 795.57 m^3/s, thus distinguishing itself from other distributions. On the other hand, the lowest estimates were mostly given by the Normal distribution, for example, the value of 534.75 m^3/s at a probability of 0.998 draws attention. This situation shows that the Lognormal (3P) distribution reflects extreme flood events with higher values due to its tail behavior, while the Normal distribution produces relatively conservative estimates. In general, in non-symmetric and right-skewed flow data, the preference of Lognormal and its derivatives at high return intervals can yield more meaningful results.

Table 3.14 Estimations for various Return Period (T) of station 2320

	X(0.5)	X(0.8)	X(0.9)	X(0.96)	X(0.98)	X(0.99)	X(0.998)
Lognormal (3P)	168,65	229,34	263,76	302,66	329,06	353,7	406,28
Gen. Extreme Value	166,3	230,48	268,13	310,81	339,29	365,13	416,7
Gumbel Max	161,53	223,38	264,34	316,08	354,47	392,57	480,62
Lognormal	156,65	237,47	295,15	372,18	432,33	494,69	649,82
Normal	173,03	231,94	262,73	295,56	316,77	335,85	374,48
Log Pearson 3	178,03	235,81	257,5	273,46	280,31	284,61	289,44

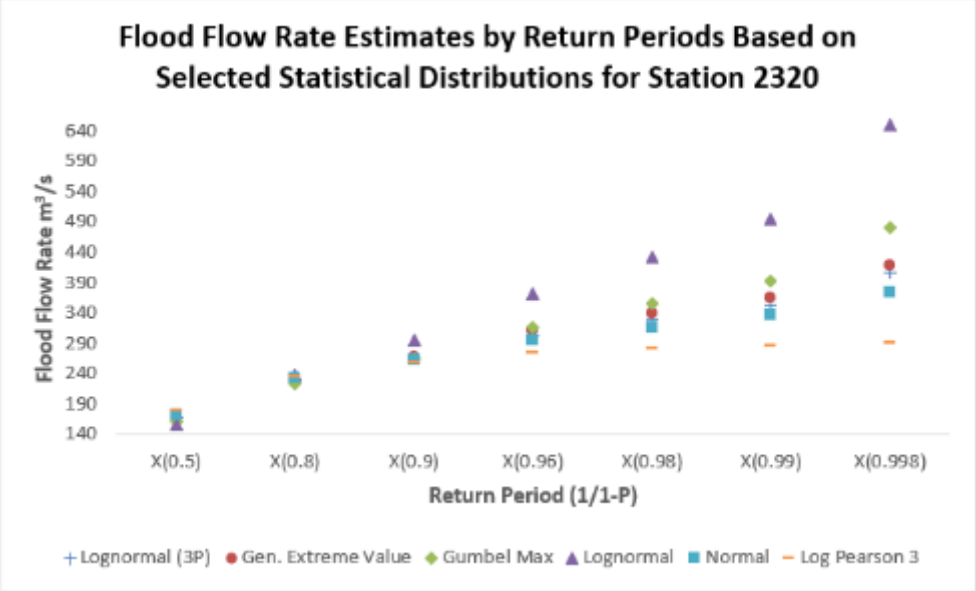


Figure 3.11 Discharge Estimates by Return Periods Based on Selected Statistical Distributions for Station 2320

In the flood Estimations made for station number 2320, the highest flow estimates were made by the Lognormal distribution at high return intervals (X(0.99) and X(0.998)). Especially at the probability of 0.998, it provided a significantly higher estimate compared to the other distributions with a value of 649.82 m^3/s. On the other hand, the lowest estimates belong to the Log Pearson 3 distribution at almost all return intervals; for example, the lowest value was recorded with an estimate of 289.44 m^3/s at the return interval of 0.998. This situation shows that the Lognormal distribution tends to overestimate extreme

values, while Log Pearson 3 provides more conservative results. The Gumbel Max distribution also produced high values, especially at medium return intervals (X(0.9)–X(0.98)), and attracted attention with its ability to represent extreme floods. These findings reveal the differences in the approach of different distributions to extreme events and show that the choice of distribution is critical in flood risk analyses.

Table 3.15 Estimations for various Return Period (T) of station 2321

	X(0.5)	X(0.8)	X(0.9)	X(0.96)	X(0.98)	X(0.99)	X(0.998)
Lognormal (3P)	74,21	90,47	100,4	112,22	120,61	128,7	146,82
Gen. Extreme Value	74,26	91,06	101,2	112,99	121,06	128,54	144
Gumbel Max	73,36	89,44	100,08	113,54	123,51	133,42	156,31
Lognormal	74,28	90,49	100,32	111,98	120,23	128,16	145,86
Normal	76,35	91,66	99,66	108,2	113,72	118,68	128,72
Log Pearson 3	74,34	90,74	100,66	112,4	120,69	128,64	146,33

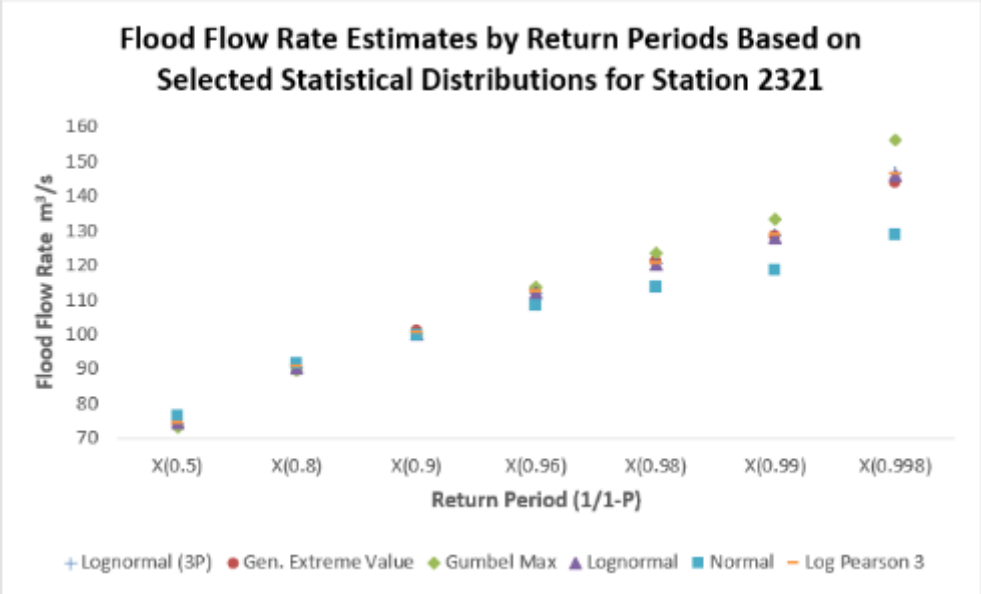


Figure 3.12 Discharge Estimates by Return Periods Based on Selected Statistical Distributions for Station 2321

In the flood analysis of station number 2321, it is seen that the estimation values in all return intervals are quite close to each other. This situation shows that the data observed at the station can be modeled similarly with various distributions. The highest estimation value was calculated as 156.31 m^3/s by the Gumbel Max distribution in the 0.998 return interval. This value is followed by the Lognormal (3P) distribution with 146.82 m^3/s and the Log Pearson 3 distribution with 146.33 m^3/s. The lowest value is generally produced by the Normal distribution, and especially in the 0.99 return interval, it lags behind the other distributions with the value of 118.68 m^3/s. This table shows that the Gumbel Max distribution predicts extreme events more, while the Normal distribution offers more moderate predictions. However, the fact that the difference between the distributions is generally low provides flexibility in model selection and suggests that the model sensitivity is relatively low at this station.

Table 3.16 Estimations for various Return Period (T) of station 2322

	X(0.5)	X(0.8)	X(0.9)	X(0.96)	X(0.98)	X(0.99)	X(0.998)
Lognormal (3P)	877,68	1086,6	1226	1403,2	1535,8	1669,2	1986
Gen. Extreme Value	892,8	1101,4	1227,2	1373,2	1472,9	1565,3	1755,6
Gumbel Max	881,99	1078	1207,7	1371,7	1493,3	1614,1	1893
Lognormal	893,07	1090	1209,6	1351,7	1452,3	1549,1	1765,3
Normal	918,43	1105,1	1202,6	1306,7	1373,9	1434,3	1556,7
Log Pearson 3	890,78	1093,1	1218,3	1369,2	1477,4	1582,6	1821,3

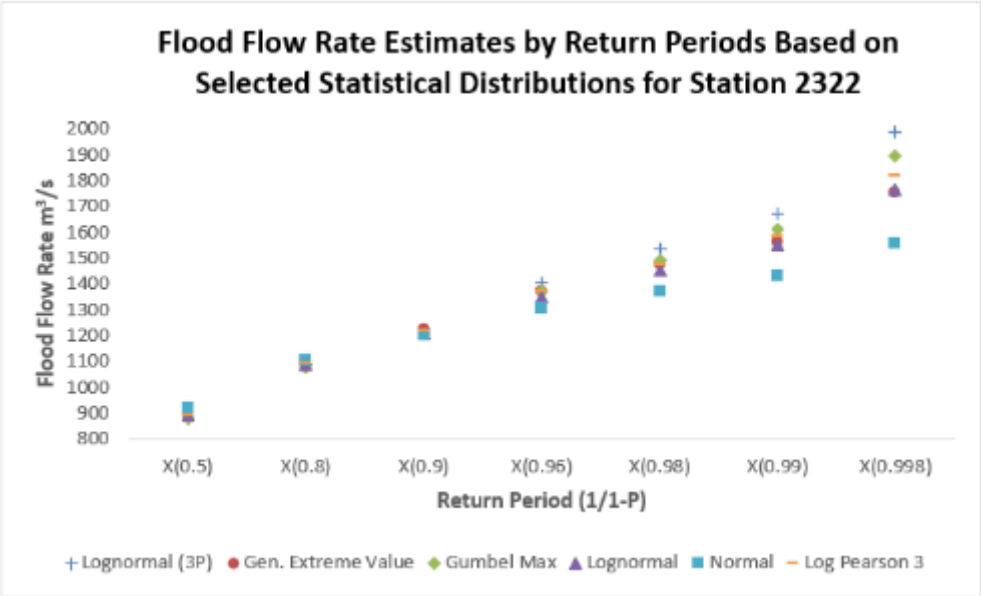


Figure 3.13 Discharge Estimates by Return Periods Based on Selected Statistical Distributions for Station 2322

In the flood analysis conducted at station 2322, it is observed that the differences in estimates between the distributions become more apparent as the return intervals increase. Especially at the return probability of 0.998, the Lognormal (3P) distribution provided the highest flow rate estimate with 1986 m^3/s and became the most sensitive distribution to extreme events. This is followed by the Log Pearson 3 with 1821.3 m^3/s and the Gumbel Max distribution with 1893 m^3/s. On the other hand, the Normal distribution generally provided the lowest estimates at this station, producing the lowest values with 1556.7 m^3/s at the return interval of 0.998. This situation shows that the Normal distribution is inadequate in representing extreme values. The wide variation and high estimate intervals reveal that the flood risk is high at this station and the choice of distribution has significant effects on the estimates. Therefore, it can be recommended to consider distributions that provide more extreme estimates, especially Lognormal (3P) and Gumbel Max, in decision-making processes.

Table 3.17 Estimations for various Return Period (T) of station 2323

	X(0.5)	X(0.8)	X(0.9)	X(0.96)	X(0.98)	X(0.99)	X(0.998)
Lognormal (3P)	219,29	301,58	356,23	425,46	477,17	529,04	651,92
Gen. Extreme Value	218,94	301,56	357,38	429,24	483,5	538,19	667,78
Gumbel Max	220,4	301,34	354,93	422,65	472,88	522,74	637,96
Lognormal	219,29	301,58	356,25	425,49	477,23	529,12	652,07
Normal	235,45	312,53	352,83	395,8	423,55	448,52	499,06
Log Pearson 3	219,82	302,86	357,59	426,45	477,57	528,56	648,34

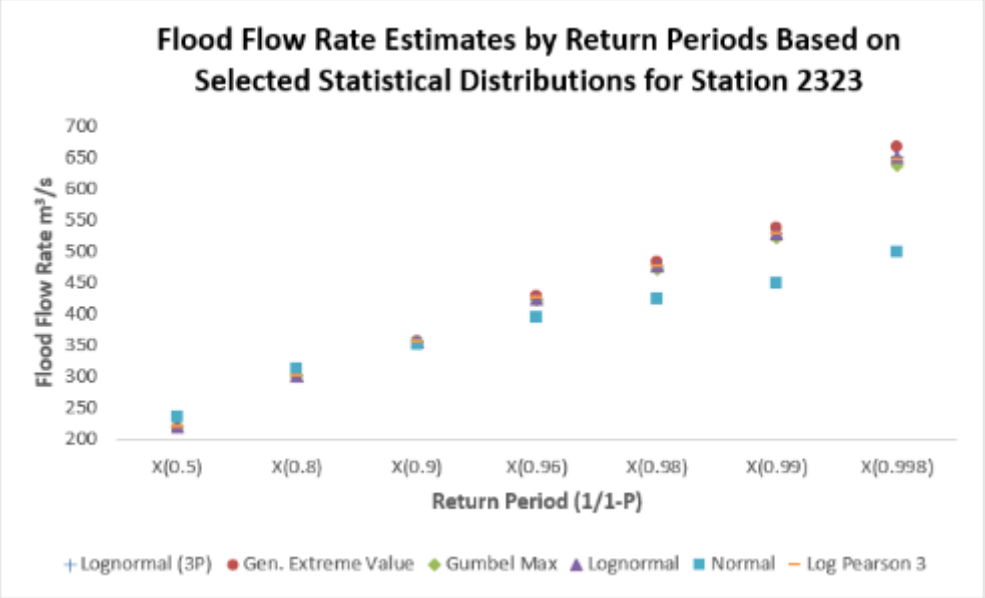


Figure 3.14 Discharge Estimates by Return Periods Based on Selected Statistical Distributions for Station 2323

It is observed that flood estimations at station number 2323 yielded very similar results between the distributions. In the 500-year return interval (X(0.998)), the highest estimate was made by the Generalized Extreme Value (GEV) distribution with 667.78 m^3/s. This value is followed by the Lognormal (3P), Lognormal and Log Pearson 3 distributions, respectively. On the other hand, the Normal distribution gave the lowest estimate (499.06 m^3/s) in this return interval. Especially in the X(0.98) and X(0.99) return probabilities, no significant difference is observed between the distributions, which shows that

the flood structure of the station can be modeled with a more homogeneous distribution. The obtained results show that the GEV distribution can be more conservative about extreme values and therefore can be a safer choice in high-risk scenarios. Comparing the results of all distributions for such stations makes the model selection more rational.

Table 3.18 Estimations for various Return Period (T) of station 2325

	X(0.5)	X(0.8)	X(0.9)	X(0.96)	X(0.98)	X(0.99)	X(0.998)
Lognormal (3P)	68,31	116,62	155,01	210,52	256,83	307,31	442,51
Gen. Extreme Value	74,76	119,26	148,52	185,26	212,36	239,11	300,43
Gumbel Max	74,97	117,24	145,23	180,6	206,83	232,88	293,06
Lognormal	69,32	116,54	152,89	204,24	246,25	291,38	409,62
Normal	82,82	123,09	144,13	166,57	181,07	194,11	220,51
Log Pearson 3	70,64	117,96	152,61	199,28	235,77	273,49	366,36

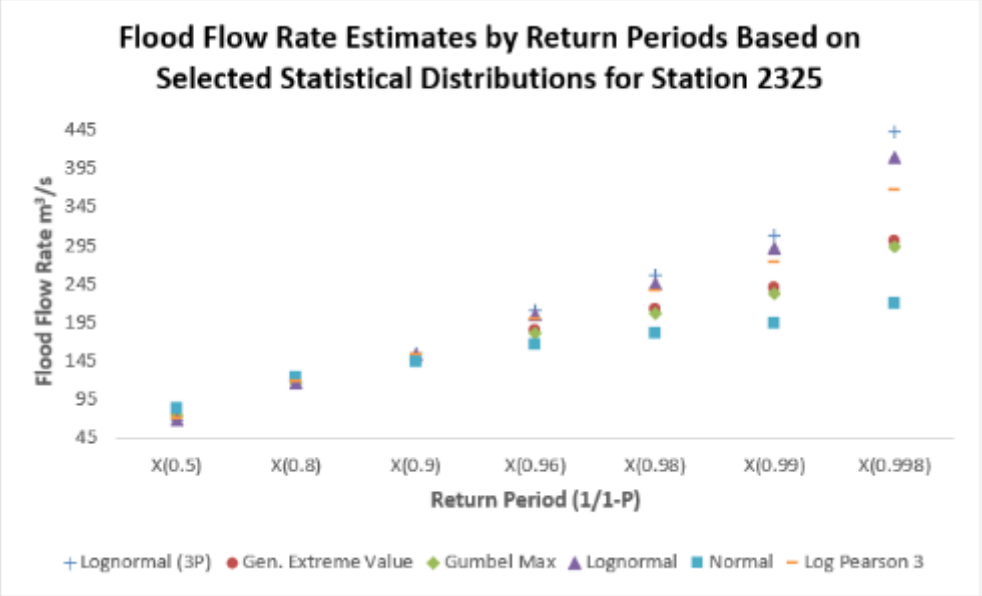


Figure 3.15 Discharge Estimates by Return Periods Based on Selected Statistical Distributions for Station 2325

In flood Estimations for station number 2325, Lognormal (3P) distribution was the distribution that predicted the highest flow rate in the 500-year return

interval with 442.51 m^3/s. This distribution was followed by Lognormal (409.62 m^3/s) and Log Pearson 3 (366.36 m^3/s), respectively. In contrast, Normal distribution generally provided the lowest Estimations at this station and predicted only 220.51 m^3/s for a return probability of 0.998. This situation shows that Normal distribution may be inadequate in tail behaviors. It can be said that Lognormal (3P) and its derivatives provide more reliable results, especially in extreme value Estimations. In addition, the estimation differences between different distributions became more pronounced as the return interval increased. In this context, it becomes clear that decision makers should pay special attention to the choice of distribution when creating flood management strategies.

Table 3.19 Estimations for various Return Period (T) of station 2326

	X(0.5)	X(0.8)	X(0.9)	X(0.96)	X(0.98)	X(0.99)	X(0.998)
Lognormal (3P)	65,92	93,63	111,7	134,29	150,97	167,56	206,39
Gen. Extreme Value	68,36	96,02	111,89	129,53	141,08	151,4	171,46
Gumbel Max	66,02	92,09	109,36	131,17	147,35	163,41	200,53
Lognormal	64,7	93,58	113,48	139,4	159,21	179,43	228,54
Normal	70,86	95,7	108,68	122,52	131,46	139,5	155,79
Log Pearson 3	66,01	94,8	113,27	135,85	152,09	167,87	203,28

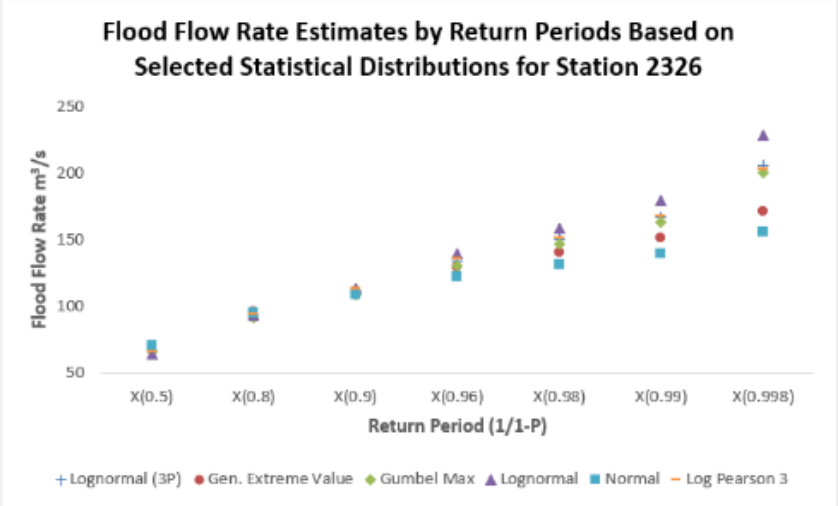


Figure 3.16 Discharge Estimates by Return Periods Based on Selected Statistical Distributions for Station 2326

In the analysis conducted for station number 2326, it is observed that all distributions have very close estimates at low return intervals (X(0.5)–X(0.9)). Especially in the X(0.5) return interval, the highest estimate is made by the Normal distribution with 70.86 m^3/s, while the lowest estimate is obtained from the Lognormal distribution with 64.7 m^3/s. However, as the return interval increases, the difference between the distributions becomes more pronounced. In the X(0.998) return interval, the highest estimate is 228.54 m^3/s, while the lowest value is again from the Normal distribution with 155.79 m^3/s. This situation shows that the Lognormal distribution estimates extreme events more, whereas the Normal distribution predicts lower risks. Since this difference can directly affect the safety coefficient used in flood management decisions, it should be taken into account especially in the planning of critical infrastructures.

Table 3.20 Estimations for various Return Period (T) of station 2327

	X(0.5)	X(0.8)	X(0.9)	X(0.96)	X(0.98)	X(0.99)	X(0.998)
Lognormal (3P)	141,56	187,38	215,94	250,51	275,35	299,55	354,54
Gen. Extreme Value	139,98	187,68	218,84	257,73	286,23	314,24	377,91
Gumbel Max	139,71	186,53	217,52	256,69	285,75	314,59	381,23
Lognormal	139,95	187,33	218,17	256,67	285,09	313,33	379,34
Normal	148,41	193	216,31	241,16	257,22	271,66	300,89
Log Pearson 3	141,8	189,62	219,13	254,3	279,14	302,97	355,63

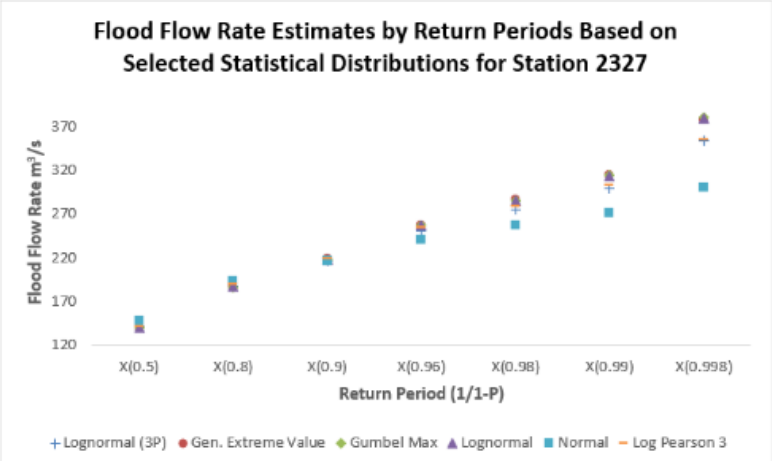


Figure 3.17 Discharge Estimates by Return Periods Based on Selected Statistical Distributions for Station 2327

In the evaluation of station number 2327, it is seen that the estimates are quite close to each other among the distributions in all return intervals. Especially in the X(0.5) return interval, the Normal distribution provides the highest estimate with 148.41 m^3/s, while the Gumbel Max distribution gives the lowest value with 139.71 m^3/s. In high return intervals (e.g. X(0.998)), the Gumbel Max distribution estimates the highest value with 381.23 m^3/s, while the Lognormal (3P) distribution gives a more moderate estimate with 354.54 m^3/s. In general, the Gumbel Max and Gen. Extreme Value distributions provide higher flood estimates in high return intervals, while the Normal distribution remains relatively more conservative. This observation shows that the choice of distribution used in high return intervals is critical in terms of flood safety and that it may be useful to prefer models that predict higher risks when planning safety measures.

Table 3.21 Estimations for various Return Period (T) of station 2328

	X(0.5)	X(0.8)	X(0.9)	X(0.96)	X(0.98)	X(0.99)	X(0.998)
Lognormal (3P)	42,34	59,67	71,79	87,74	100,03	112,66	143,64
Gen. Extreme Value	43,93	60,83	71,24	83,55	92,13	100,2	117,26
Gumbel Max	43,28	59,24	69,81	83,16	93,06	102,89	125,61
Lognormal	42,89	59,72	71,01	85,4	96,22	107,11	133,08
Normal	46,25	61,45	69,39	77,86	83,34	88,26	98,22
Log Pearson 3	43,15	60,17	71,36	85,37	95,72	106	129,96

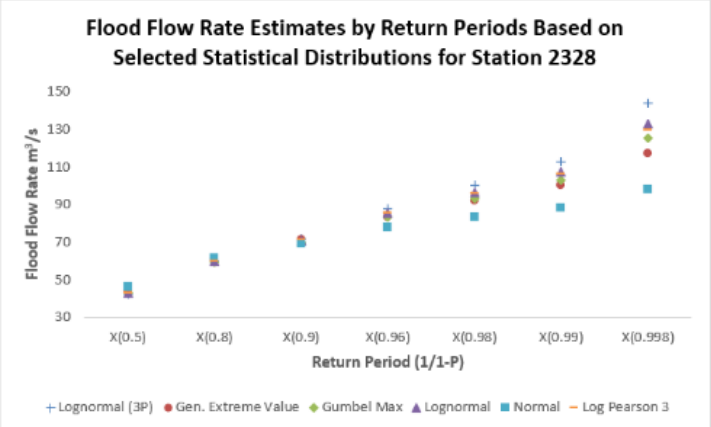


Figure 3.18 Discharge Estimates by Return Periods Based on Selected Statistical Distributions for Station 2328

When the flood Estimations obtained for station number 2328 are evaluated, it is observed that the estimated flow rates are quite close to each other in all return intervals. While the Normal distribution provided the highest estimate with 46.25 m^3/s in the X(0.5) return interval, the Lognormal (3P) distribution provided the lowest value with 42.34 m^3/s. A similar situation is observed in the high return intervals; while the Lognormal (3P) distribution provided the highest estimate with 143.64 m^3/s in the X(0.998) return interval, the Normal distribution provided a lower estimate with only 98.22 m^3/s. The general tendency at this station is that the Normal distribution provides lower estimates compared to the other models. This situation shows that in analyses requiring a more conservative approach in the assessment of flood risk, distributions that provide higher values such as Lognormal (3P) or Gumbel Max should be taken into consideration.

Table 3.22 Estimations for various Return Period (T) of station 2329

	X(0.5)	X(0.8)	X(0.9)	X(0.96)	X(0.98)	X(0.99)	X(0.998)
Lognormal (3P)	152,13	232,31	293,54	379,48	449,44	524,29	719,3
Gen. Extreme Value	159,62	236	287,35	353,12	405,56	452,19	569,05
Gumbel Max	161,11	233,74	281,84	342,6	387,68	432,42	535,83
Lognormal	156,8	232,59	285,83	356,1	410,42	466,33	603,91
Normal	174,61	243,79	279,95	318,51	343,42	365,82	411,18
Log Pearson 3	156,99	234,25	288,56	360,24	415,64	472,63	612,7

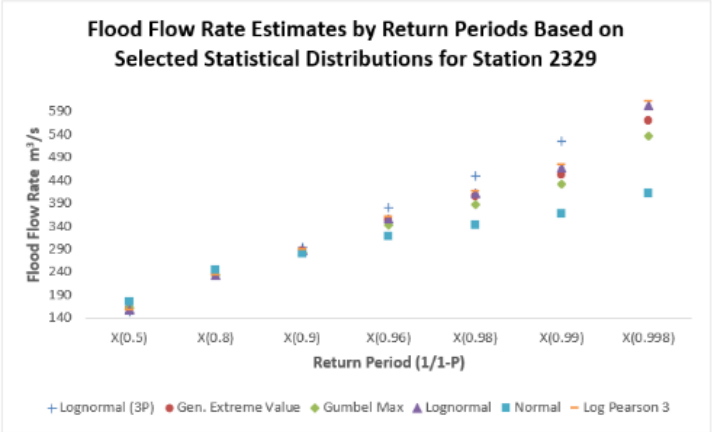


Figure 3.19 Discharge Estimates by Return Periods Based on Selected Statistical Distributions for Station 2329

When the flood flow estimates of station number 2329 are examined, it is seen that the Lognormal (3P) distribution makes significantly higher estimates than the other distributions, especially at high return intervals (X(0.98), X(0.99), X(0.998)). While the Lognormal (3P) gives the highest value with 719.3 m^3/s at the X(0.998) return interval, the Normal distribution only estimated 411.18 m^3/s. At low return intervals (X(0.5), X(0.8)), the Normal distribution again reaches the highest values and exhibits a more conservative estimate. This shows that the Lognormal (3P) distribution stands out in extreme value estimation, while the Normal distribution gives higher estimates at low return intervals. For this station in particular, if sensitivity to extreme situations is required in flood planning, the Lognormal (3P) distribution can be considered a safer option.

Table 3.23 Estimations for various Return Period (T) of station 2330

	X(0.5)	X(0.8)	X(0.9)	X(0.96)	X(0.98)	X(0.99)	X(0.998)
Lognormal (3P)	16,01	25,19	32,55	43,27	52,25	62,09	88,59
Gen. Extreme Value	16,51	25,21	31,77	41,11	48,88	57,38	80,5
Gumbel Max	17,25	26,28	32,26	39,81	45,41	50,97	63,82
Lognormal	16,75	28,28	31,35	39,44	45,75	52,27	68,45
Normal	18,93	27,53	32,02	36,82	39,91	42,7	48,3
Log Pearson 3	16,44	25,3	32,03	41,54	49,36	57,83	80,38

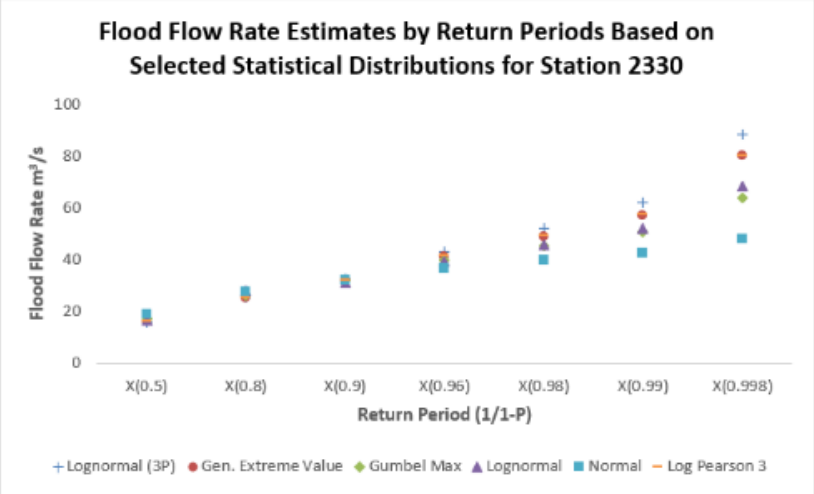


Figure 3.20 Discharge Estimates by Return Periods Based on Selected Statistical Distributions for Station 2330

In the flood Estimations made for station number 2330, Lognormal (3P), Gen. Extreme Value, Gumbel and Log Pearson 3 distributions produced values close to each other in all return intervals (X(0.5)–X(0.998)). However, the highest value in the X(0.5) return interval was given by the Normal distribution with 18.93 m^3/s. On the other hand, in extreme return intervals such as X(0.998), the highest estimate was given by the Lognormal (3P) distribution (88.59 m^3/s), while the lowest estimate belonged to the Normal distribution (48.3 m^3/s). This situation shows that the Normal distribution predicts lower values in flood analyses, especially in high return intervals, and is less sensitive to extreme events in this context. Considering that other distributions show an increase with wider variances, it can be stated that distributions such as Lognormal (3P) and Log Pearson 3 provide more cautious results in terms of risk assessment at this station.

Table 3.24 Estimations for various Return Period (T) of station 2331

	X(0.5)	X(0.8)	X(0.9)	X(0.96)	X(0.98)	X(0.99)	X(0.998)
Lognormal (3P)	43,16	67,71	99,3	165,42	240,13	342,8	733,08
Gen. Extreme Value	52,29	66,86	74,65	82,79	87,79	92,03	99,62
Gumbel Max	50,28	64,43	73,8	85,64	94,42	103,13	123,27
Lognormal	50,62	65,22	74,46	85,76	93,96	102	120,43
Normal	52,91	66,39	73,43	80,94	85,8	90,16	98,995
Log Pearson 3	50,97	66,02	75,3	86,39	94,26	101,84	118,77

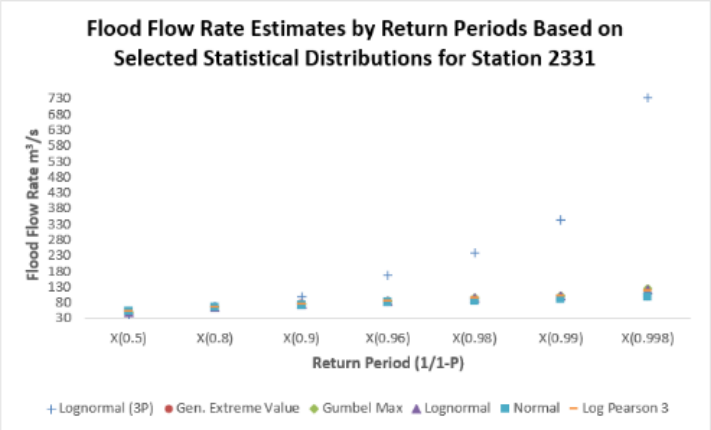


Figure 3.21 Discharge Estimates by Return Periods Based on Selected Statistical Distributions for Station 2331

In the flood analysis performed for station number 2331, while the Normal, Gen. Extreme Value and Gumbel distributions gave similar and higher estimates at low return intervals (X(0.5)–X(0.9)), the Lognormal (3P) distribution produced the lowest flow rates in these intervals. Especially, the value of 733.08 m^3/s predicted by the Lognormal (3P) distribution at the 0.998 probability level is quite high compared to all other distributions. This situation reveals that this distribution is more sensitive to extreme flood risks and makes more cautious estimates. The lowest estimate was given by the Normal distribution, which shows that the Normal distribution may be limited in representing extreme values at high return intervals. Therefore, if a sensitive approach is desired to be adopted against flood hazard for this station, the Lognormal (3P) distribution comes to the fore.

Table 3.25 Estimations for various Return Period (T) of station 2333

	X(0.5)	X(0.8)	X(0.9)	X(0.96)	X(0.98)	X(0.99)	X(0.998)
Lognormal (3P)	28,83	40,91	50,66	64,92	76,93	90,12	125,83
Gen. Extreme Value	28,05	39,27	49,59	67,19	84,62	106,75	184,41
Gumbel Max	30,43	44,76	54,24	66,22	75,11	83,94	104,33
Lognormal	30,5	41,92	49,49	59,09	66,26	73,44	90,46
Normal	33,09	46,74	53,87	61,47	66,38	70,8	79,75
Log Pearson 3	28,39	40,65	51,29	68,13	83,55	101,78	158,25

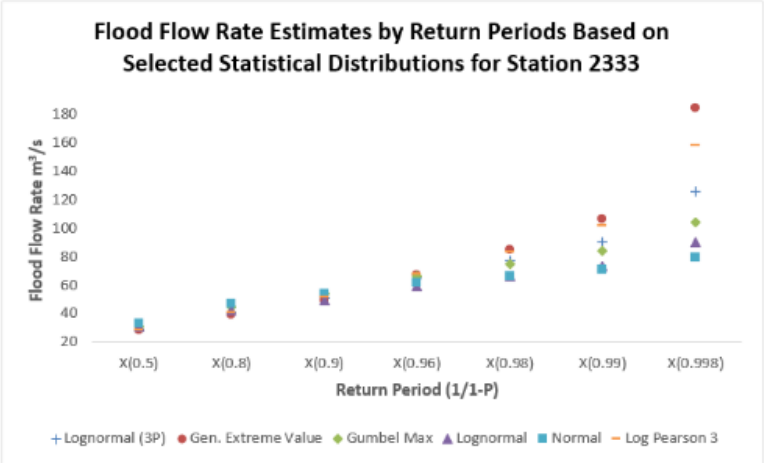


Figure 3.22 Discharge Estimates by Return Periods Based on Selected Statistical Distributions for Station 2333

The flood estimation results obtained for station number 2333 show that the highest flow rate prediction is made by the Normal distribution at low return intervals (0.5–0.9), while the lowest values are predicted by the Gen. Extreme Value and Lognormal (3P) distributions. However, it should be noted that as the return interval increases (0.98–0.998), the values predicteby the Gen. Extreme Value distribution increase rapidly and the highest flood flow rates are given by this distribution. This distribution produced the highest value at the X(0.998) level with 184.41 m^3/s. This situation reveals that the Gen. Extreme Value distribution tends to reflect extreme values more strongly. The lowest prediction is made by the Normal distribution, which shows that the Normal distribution may not adequately reflect extreme flood risks. In this context, the Gen. Extreme Value distribution may be a more appropriate choice for measures to be taken at high return intervals.

Table 3.26 Estimations for various Return Period (T) of station 2334

	X(0.5)	X(0.8)	X(0.9)	X(0.96)	X(0.98)	X(0.99)	X(0.998)
Lognormal (3P)	163,77	212,04	241,33	276,08	300,66	324,3	377,05
Gen. Extreme Value	164,79	215,49	245,04	278,35	300,45	320,4	359,93
Gumbel Max	160,97	208,95	240,72	280,85	310,63	340,18	408,48
Lognormal	161,55	211,84	244,08	283,88	312,98	341,7	408,14
Normal	169,89	215,58	239,47	264,94	281,39	296,19	326,15
Log Pearson 3	163,66	213,93	244,32	280,02	304,93	328,59	380,2

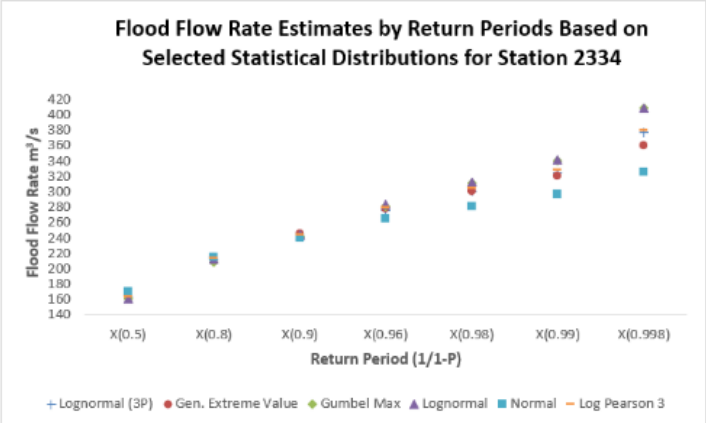


Figure 3.23 Discharge Estimates by Return Periods Based on Selected Statistical Distributions for Station 2334

When the flood flow rates calculated according to the return interval of station number 2334 are examined, no significant differences are observed between the estimates in low probability events (X(0.5)–X(0.9)). All distributions produced similar results in this range. However, as the return interval increases (X(0.98) and above), the Gumbel Max and Lognormal distributions give the highest flood values. Especially in the X(0.998) return interval, the Gumbel Max distribution makes the highest estimate with 408.48 m^3/s, while the Normal distribution makes the lowest estimate with 326.15 m^3/s. This situation shows that the Gumbel Max distribution shows a more aggressive tendency in extreme value estimates and offers a more cautious approach against higher flood risks. The analysis points out the necessity of taking into consideration such values estimated in high return intervals, especially in critical infrastructure planning.

Table 3.27 Estimations for various Return Period (T) of station 2335

	X(0.5)	X(0.8)	X(0.9)	X(0.96)	X(0.98)	X(0.99)	X(0.998)
Lognormal (3P)	642,65	920,27	1132,4	1430,2	1672,7	1932,3	2608,5
Gen. Extreme Value	657,66	928,53	1121,3	1381,4	1587	1802,4	2346,7
Gumbel Max	673,16	933,61	1106	1323,9	1485,5	1646	2016,7
Lognormal	669,47	925,34	1095,9	1312,6	1474,9	1637,9	2025,1
Normal	721,58	969,61	1099,3	1237,5	1326,8	1407,2	569,8
Log Pearson 3	660,05	929,11	1120	1375,5	1576,3	1785,9	2315,3

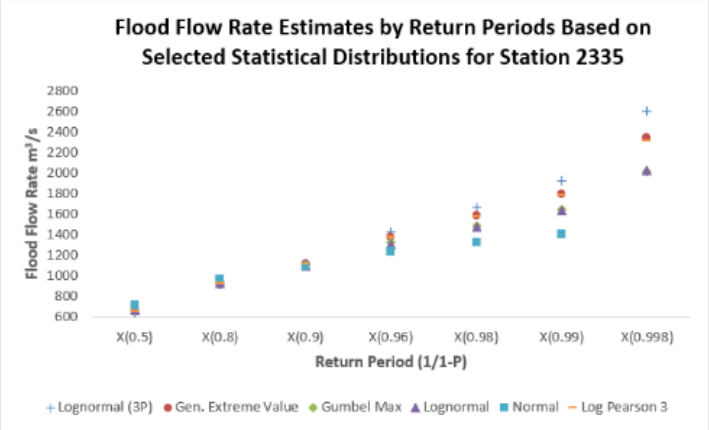


Figure 3.24 Discharge Estimates by Return Periods Based on Selected Statistical Distributions for Station 2335

According to the flood flow rates calculated for station 2335, although the distributions produced close estimates at low return intervals (X(0.5)–X(0.9)), serious deviations are observed especially at the X(0.998) return interval. While the Lognormal (3P) distribution produced the highest estimate with 2608.5 m^3/s, the Normal distribution produced the lowest estimate with 569.8 m^3/s. When this gap is examined, it is seen that the Normal distribution makes risky low estimates at high return intervals, thus under-reflecting the flood risk. On the other hand, distributions such as Lognormal (3P), Gumbel and Log Pearson 3 evaluate the flood magnitudes higher and provide a safer approach in terms of critical engineering structures and flood management. As a result, it is recommended that more cautious distributions be preferred in terms of safety in flood estimation for station 2335.

Table 3.28 Estimations for various Return Period (T) of station 2336

	X(0.5)	X(0.8)	X(0.9)	X(0.96)	X(0.98)	X(0.99)	X(0.998)
Lognormal (3P)	65,81	86,39	102,31	124,88	143,4	163,33	215,75
Gen. Extreme Value	67,69	87,75	101,37	118,97	132,32	145,82	178
Gumbel Max	68,26	87,12	99,61	115,39	127,1	138,72	165,57
Lognormal	68,96	87,3	98,75	112,62	122,6	132,33	154,46
Normal	71,77	89,73	99,12	109,14	115,61	121,42	133,2
Log Pearson 3	67,97	87,57	100,81	117,91	130,94	144,24	176,7

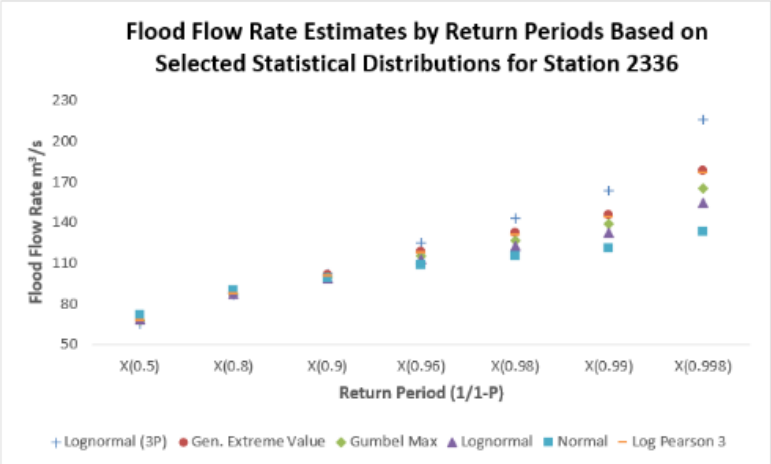


Figure 3.25 Discharge Estimates by Return Periods Based on Selected Statistical Distributions for Station 2336

When the flood flow estimates for station number 2336 are examined, it is observed that all distributions produce similar values at low return intervals (e.g. X(0.5)–X(0.9)). However, as the return interval increases, the differences between the estimated flow values also increase. Especially in the X(0.998) return interval, the Lognormal (3P) distribution produces the highest estimate with 215.75 m^3/s, while the Normal distribution gives the lowest value with 133.2 m^3/s. This situation shows that the Normal distribution tends to underestimate the flood risk at high return intervals and may be insufficient in terms of security. Log Pearson 3 and Gen. Extreme Value distributions also produced remarkable estimates at high return intervals. In this context, preferring Lognormal (3P) or similar distributions, which provide more cautious estimates in flood risk analyses for station number 2336, may produce safer results in flood management.

Table 3.29 Estimations for various Return Period (T) of station 2337

	X(0.5)	X(0.8)	X(0.9)	X(0.96)	X(0.98)	X(0.99)	X(0.998)
Lognormal (3P)	332,17	434,88	496,98	570,47	622,31	672,12	782,97
Gen. Extreme Value	335,7	443,39	504,71	572,41	616,42	655,5	730,77
Gumbel Max	325,98	427,55	494,79	579,75	642,78	705,35	849,92
Lognormal	326,11	434,34	504,53	591,91	656,26	720,09	868,94
Normal	344,86	441,59	492,15	546,06	580,89	612,22	675,64
Log Pearson 3	331,06	438,82	504,26	581,26	635,04	686,14	797,51

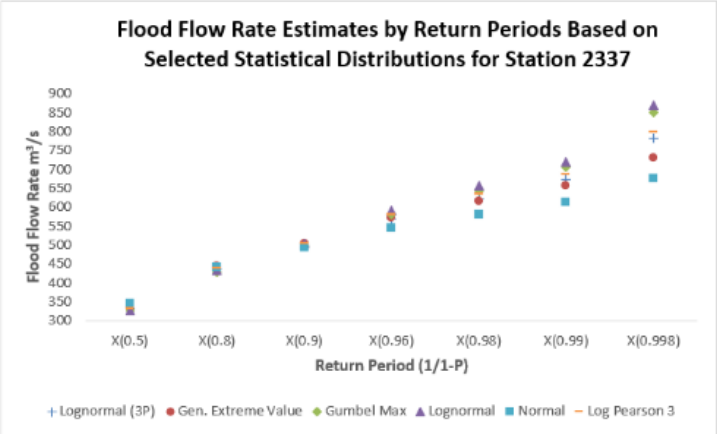


Figure 3.26 Discharge Estimates by Return Periods Based on Selected Statistical Distributions for Station 2337

When the estimated flood flow rates at station 2337 are examined, it is seen that especially at high return intervals (X(0.99) and X(0.998)), the Gumbel Max distribution gives the highest values (705.35 and 849.92 m^3/s, respectively). This shows that Gumbel Max tends to highlight extreme values. In contrast, the Normal distribution produced the lowest value (675.64 m^3/s) at the X(0.998) return interval, which reveals that the Normal distribution can be conservative in flood analyses. In the middle return intervals, there are no major differences between the distributions, and the values are close to each other. In general, while Gumbel Max has the potential to exaggerate flood risks for this station, the possibility that the Normal distribution may underestimate the risks should be taken into account. Distributions such as Lognormal and Log Pearson 3 can be preferred in decision support processes by providing balanced results.

Table 3.30 Estimations for various Return Period (T) of station 2338

	X(0.5)	X(0.8)	X(0.9)	X(0.96)	X(0.98)	X(0.99)	X(0.998)
Lognormal (3P)	205,03	269,1	307,16	351,62	382,65	412,23	477,28
Gen. Extreme Value	203,28	271,44	312,93	361,52	395,01	426,24	491,37
Gumbel Max	200,05	264,86	307,78	362	402,23	442,16	534,43
Lognormal	199,77	269,32	314,84	371,88	414,11	456,18	554,88
Normal	212,09	273,82	306,09	340,5	362,73	382,72	423,2
Log Pearson 3	199,77	269,32	314,84	371,88	414,11	456,18	554,88

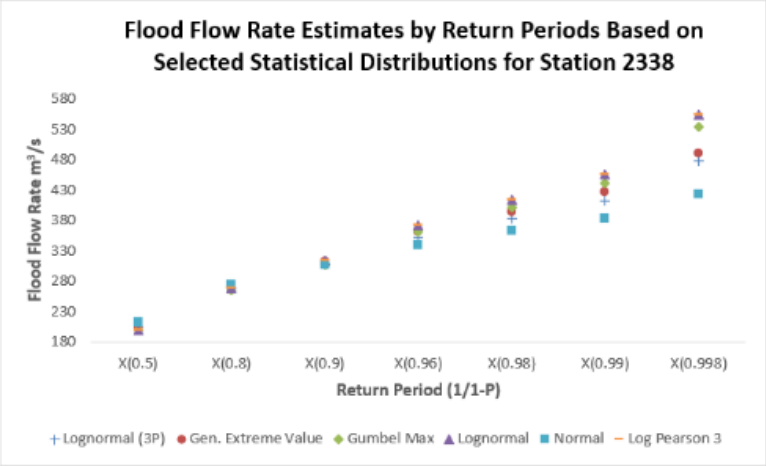


Figure 3.27 Discharge Estimates by Return Periods Based on Selected Statistical Distributions for Station 2338

When the flood estimation results of station number 2338 are examined, it is seen that very close estimates are made between the distributions at low return intervals (e.g. X(0.5)). However, as the return interval increases, the differences become more apparent. Especially in the X(0.998) return interval, the highest value was given by Log Pearson 3 and Lognormal distribution (554.88 m^3/s), while the lowest estimate was made by the Normal distribution (423.20 m^3/s). This situation shows that the Normal distribution evaluates extreme flood values more conservatively, while the Log Pearson 3 and Lognormal distributions provide more risk-oriented estimates. For this station in particular, although there are no major deviations between the distributions, it can be said that the Log Pearson 3 and Lognormal (2-parameter) distributions predict more extreme values at high return intervals. In this context, it would be appropriate to take these two distributions into account in decision-making processes in terms of risk management.

Table 3.31 Estimations for various Return Period (T) of station 2340

	X(0.5)	X(0.8)	X(0.9)	X(0.96)	X(0.98)	X(0.99)	X(0.998)
Lognormal (3P)	18,74	23,99	27,31	31,37	34,31	37,19	43,79
Gen. Extreme Value	18,26	23,67	27,56	32,85	37,07	41,53	52,94
Gumbel Max	18,59	23,88	27,38	31,8	35,08	38,34	45,87
Lognormal	18,75	24	27,3	31,32	34,23	37,08	43,59
Normal	19,57	24,61	27,24	30,05	31,86	33,49	36,79
Log Pearson 3	18,73	24,14	27,58	31,81	34,88	37,9	44,87

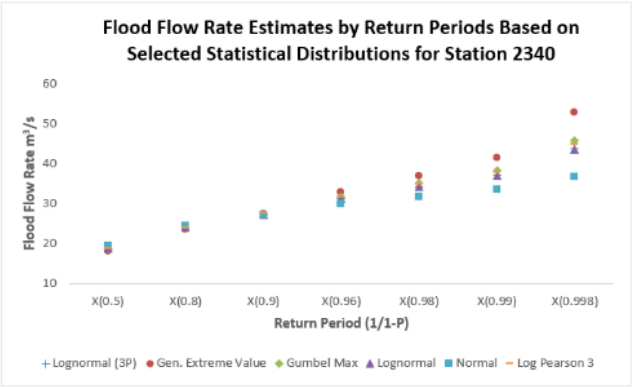


Figure 3.28 Discharge Estimates by Return Periods Based on Selected Statistical Distributions for Station 2340

In the flood Estimations performed at station number 2340, different probability distributions gave similar results. In the low return intervals (e.g. X(0.5)), the estimated values were stuck between 18–20 m^3/s. The highest estimates occurred in the X(0.998) return interval, at which point the highest flow estimate was made by the Generalized Extreme Value distribution (52.94 m^3/s), while the lowest estimate was provided by the Normal distribution (36.79 m^3/s). It is seen that the GEV and Gumbel Max distributions are sensitive to more extreme values, whereas the Normal distribution provides a more conservative estimate. When a more reliable approach is sought for extreme floods at this station, it may be more appropriate to consider the GEV and Log Pearson 3 distributions. However, it is also observed that there are no major differences between the distributions in terms of general estimation consistency.

Table 3.32 Estimations for various Return Period (T) of station 2342

	X(0.5)	X(0.8)	X(0.9)	X(0.96)	X(0.98)	X(0.99)	X(0.998)
Lognormal (3P)	46,17	56,19	61,56	67,39	71,21	74,69	81,85
Gen. Extreme Value	47,197	57,23	61,66	65,58	67,62	69,13	71,3
Gumbel Max	44,3	54,9	61,92	70,79	77,36	83,89	98,98
Lognormal	44,67	56,23	63,42	72,1	78,33	84,39	98,14
Normal	46,27	56,37	61,64	67,27	70,9	74,17	80,79
Log Pearson 3	45,95	56,83	62,56	68,56	72,33	75,62	81,93

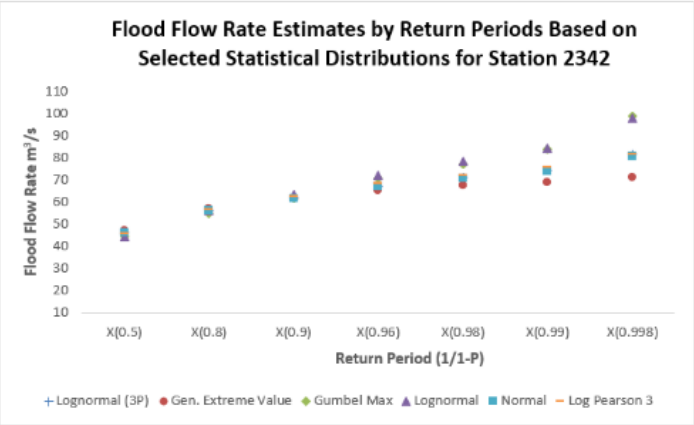


Figure 3.29 Discharge Estimates by Return Periods Based on Selected Statistical Distributions for Station 2342

The flood estimations for Station 2342 reveal that the different statistical distributions show significant differences, especially at high return periods (X(0.99) and X(0.998)). For example, while the Gumbel Max distribution for X(0.998) gives a value of 98.98 m³/s, the GEV distribution only predicts 71.3 m^3/s. This clearly demonstrates the impact of model selection on flood risk analysis. The graphical comparison shows that the distributions diverge increasingly at extreme values, supporting the need for model reliability and standardization.

3.5.1 Result

The analysis results revealed significant differences between the flow rates estimated by the selected distributions, especially at high return intervals (e.g. 500 years, X(0.998)), in the flood predictions applied to 23 different hydrological stations in the Coruh River Basin. In the study, it was observed that the Lognormal (3P) and Log Pearson III distributions gave the highest flood flow rates at many stations, while the Normal and Generalized Extreme Value (GEV) distributions gave lower predictions at some stations. These findings show that the distribution choices in the prediction of extreme flood values have a direct and significant effect on the prediction results.

As a result, it should be emphasized that in infrastructure projects, flood control and risk management applications where extreme values play a critical role, not only statistical suitability but also the behavior of the distributions in the engineering context should be taken into consideration. However, although some distributions (e.g. Normal or Lognormal) exhibit poor performance at certain stations, these distributions are widely used in the literature and are among the generally accepted assumptions, so they are included in the comparative analyses at all stations. In particular, the Normal distribution, due to its symmetrical structure, does not show suitability in flow regimes with skewness such as the Çoruh River Basin; however, it was evaluated in order to make comparisons with other distributions. Thus, it was aimed to make a more balanced and engineering-meaningful assessment by considering hydrological

and physical suitability as well as statistical compatibility in the distribution preference.

3.6 TREND ANALYSIS

In this section, annual maximum flood flow rates of 23 different stations in the Çoruh River Basin were examined and it was analyzed whether there was a significant trend in these time series. In performing the trend analyses, the Mann-Kendall rank test, which is explained in detail in section 2.5, was used. The Table 37 includes the findings of the trend analysis applied to each station, and the evaluations regarding the trend direction and statistical significance levels are presented in detail.

The given Mann-Kendall trend analysis Table 33 was made to determine the temporal trends on the annual maximum flow data observed at 23 stations belonging to the Çoruh River Basin. In the analysis, the statistical significance of the trends at the stations was evaluated using the calculated S, τ, standard deviation (σ_S), Z value and the corresponding $Z_{\alpha/2}$ value at the 5% significance level. Stations with an absolute Z value greater than $Z_{\alpha/2} = 1.96$ show a significant trend.

Table 3.33 Results of Mann Kendall Trend Analysis

Station Number	N	τ	S	σ_s	z	$Z_{\alpha/2}$	Trend	No Trend
2302	21	0,07	15,00	33,12	0,42	1,96		-
2304	61	0,15	277,00	160,70	1,72	1,96		-
2305	48	0,07	78,00	112,51	0,68	1,96		-
2316	46	0,12	128,00	105,62	1,20	1,96		-
2320	38	0,11	80,00	79,54	0,99	1,96		-
2321	39	0,02	18,00	82,67	0,21	1,96		-
2322	28	0,21	80,00	50,62	1,56	1,96		-
2323	46	-0,04	-41,00	105,62	-0,38	1,96		-
2325	36	-0,18	-116,00	73,42	-1,57	1,96		-
2326	23	-0,23	-58,00	37,86	-1,51	1,96		-
2327	17	0,13	17,00	24,28	0,66	1,96		-
2328	29	-0,36	-146,00	53,31	-2,72	1,96	+	
2329	30	0,09	38,00	56,05	0,66	1,96		-
2330	30	-0,30	-130,00	56,05	-2,30	1,96	+	
2331	13	0,21	16,00	16,39	0,92	1,96		-
2333	20	0,03	5,00	30,82	0,13	1,96		-
2334	21	-0,09	-18,00	33,12	-0,51	1,96		-
2335	19	0,17	29,00	28,58	0,98	1,96		-
2336	15	0,15	16,00	20,21	0,74	1,96		-
2337	22	-0,01	-2,00	35,46	-0,03	1,96		-
2338	20	-0,02	-4,00	30,82	-0,10	1,96		-
2340	20	-0,36	-69,00	30,82	-2,21	1,96	+	
2342	19	-0,40	-68,00	28,58	-2,34	1,96	+	

According to the table, only four stations (2328, 2330, 2340 and 2342) have a z value exceeding the threshold of 1.96, indicating the existence of a statistically significant trend. However, in order to evaluate how strongly the trends at these four stations are explained by linear relationships, dimensionless flow ratio ($Q/\bar{Q}$) time series were created for each station and linear regression curves were drawn. The coefficient of determination (R^2) values were calculated on these curves and the explanatory power of the graphical trends was tested.

The obtained results show that although there are statistically significant trends at some stations, the R^2 values are quite low, indicating that these trends are insufficient to explain a large portion of the temporal variability. For example, although a negative trend is found significant at station number 2342 with a value of $z = -2.34$, $R^2 = 0.2152$, which explains only 21% of the observed variability. This situation shows that even if it supports the existence of a trend, it is difficult to talk about a strong trend in practice. Therefore, statistical significance, graphical explanatory power and hydraulic context should be evaluated together.

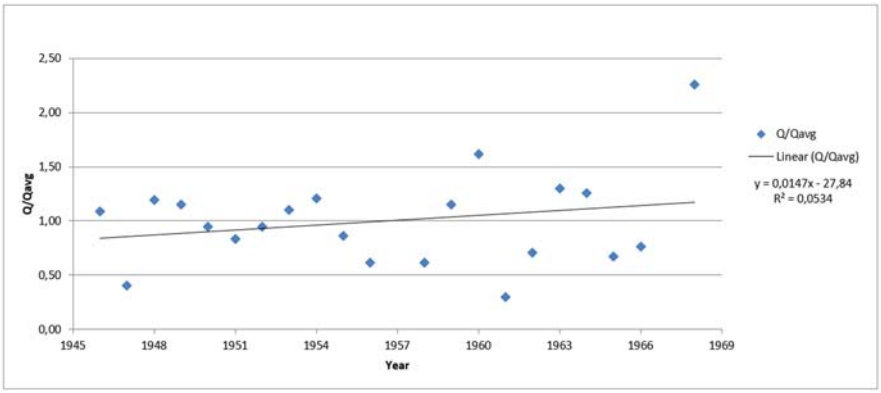


Figure 3.30 Trend Analysis for Station 2302

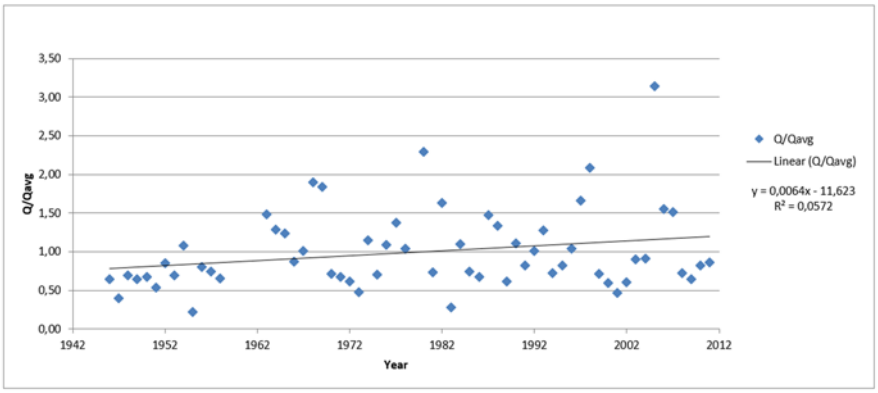


Figure 3.31 Trend Analysis for Station 2304

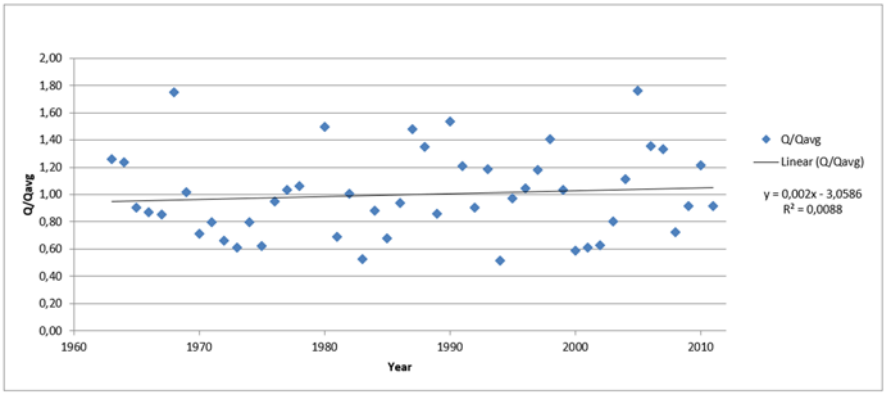


Figure 3.32 Trend Analysis for Station 2305

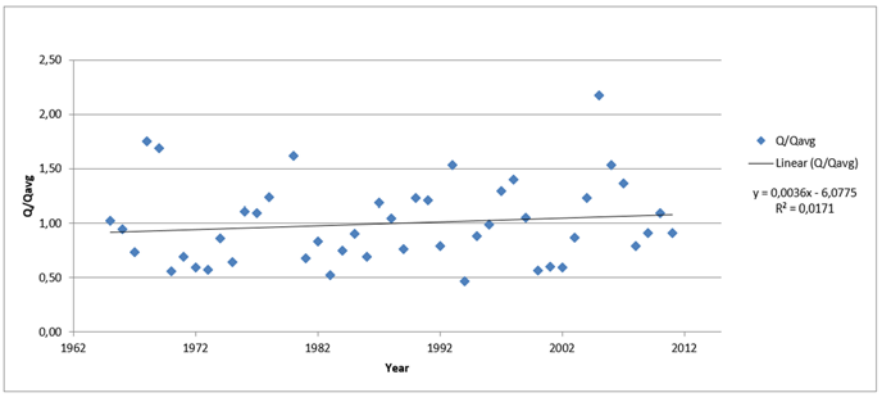


Figure 3.33 Trend Analysis for Station 2316

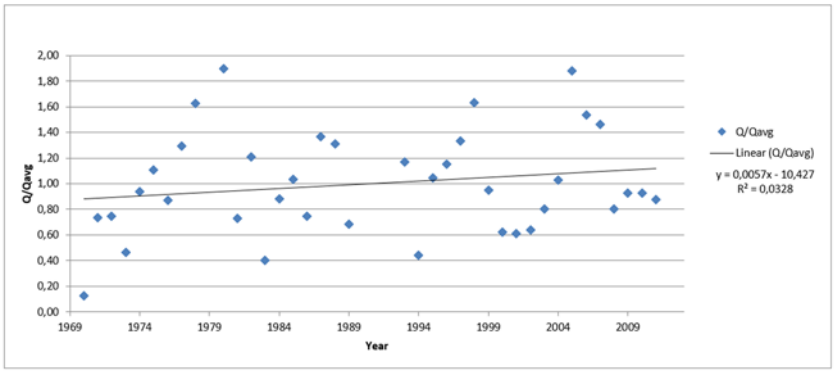


Figure 3.34 Trend Analysis for Station 2320

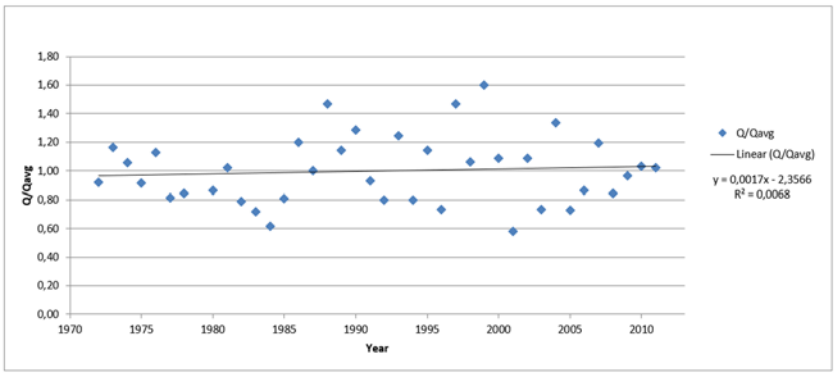


Figure 3.35 Trend Analysis for Station 2321

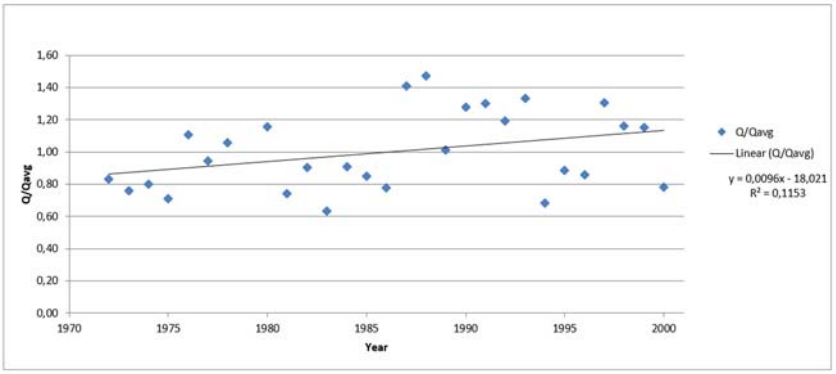


Figure 3.36 Trend Analysis for Station 2322

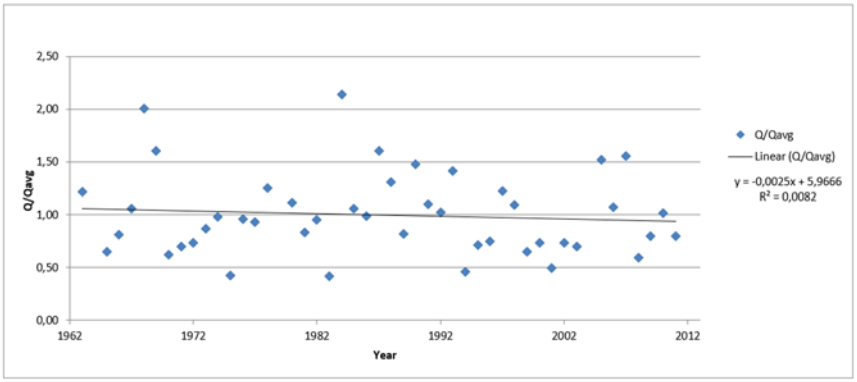


Figure 3.37 Trend Analysis for Station 2323

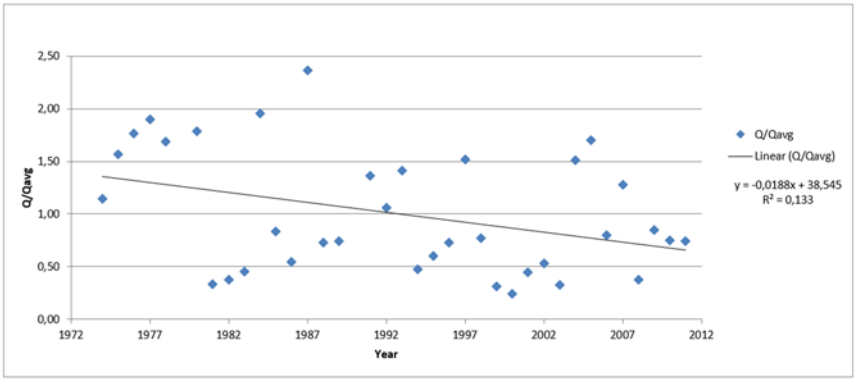


Figure 3.38 Trend Analysis for Station 2325

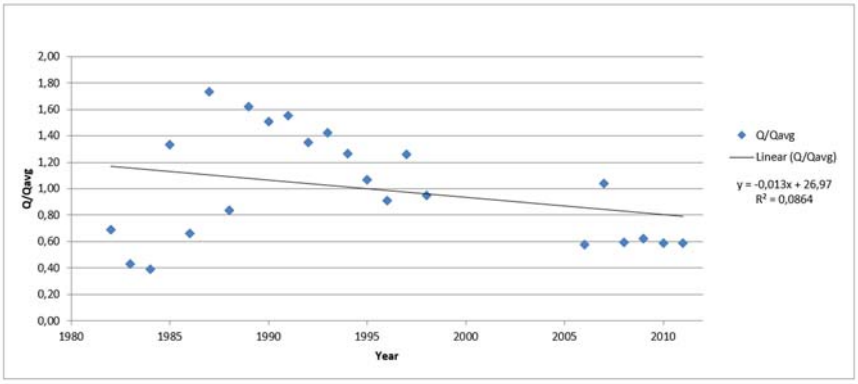


Figure 3.39 Trend Analysis for Station 2326

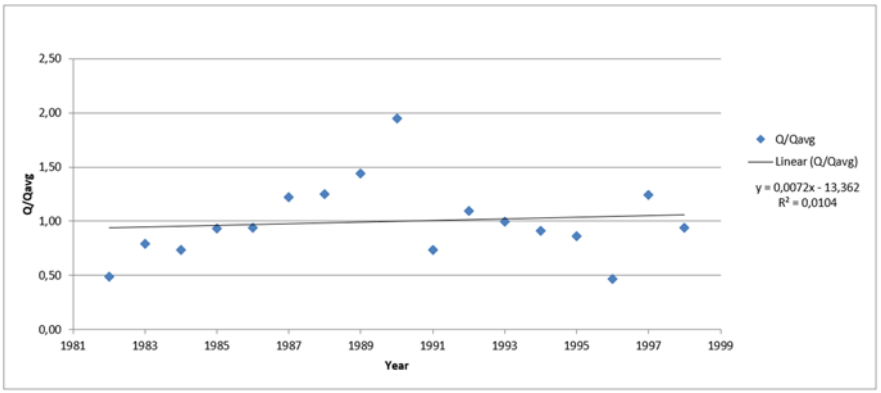


Figure 3.40 Trend Analysis for Station 2327

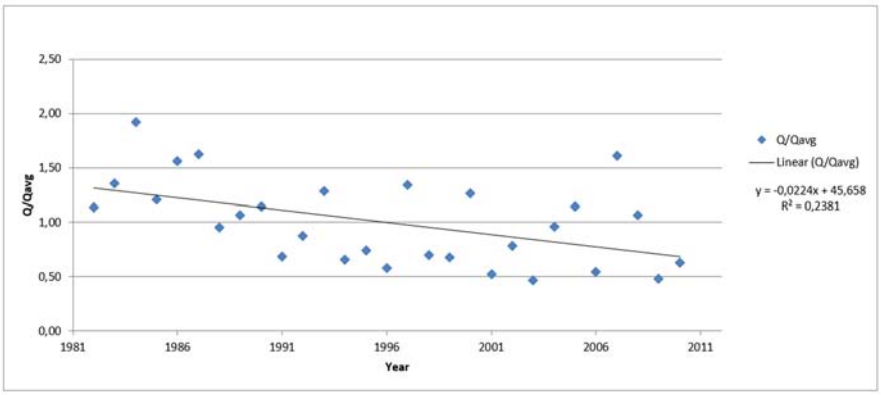


Figure 3.41 Trend Analysis for Station 2328

At station 2328, $Q/\bar{Q}$ values show a decreasing trend over the years and the slope of the regression line is negative. However, the R^2 value is 0.2381 and has a medium level of explanatory power. This value shows that the model explains only about 24% of the variability in the data. Although the Mann-Kendall test gives a positive sign in the direction of the trend, the R^2 value shows that this trend is not a strong and reliable indicator.

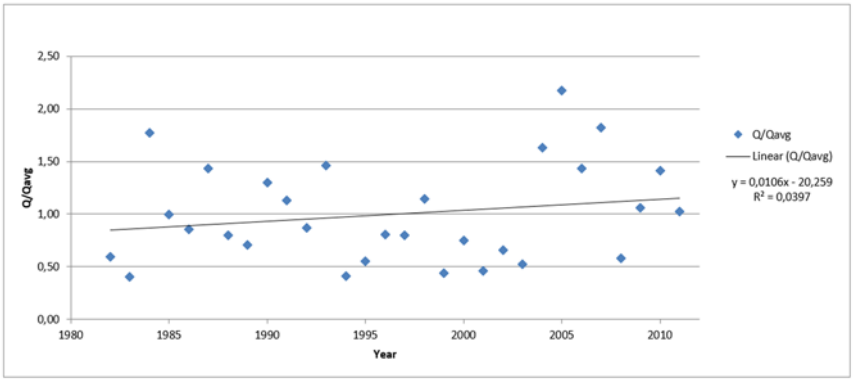


Figure 3.42 Trend Analysis for Station 2329

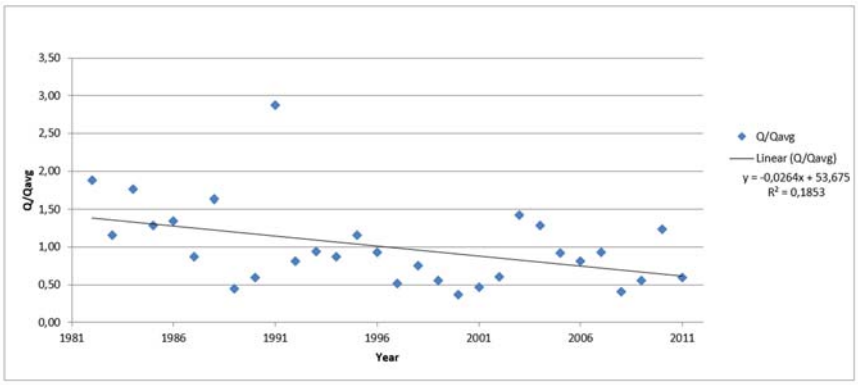


Figure 3.43 Trend Analysis for Station 2330

The slope of the regression line drawn for station 2330 is significantly negative. The R^2 value is 0.1853, which is relatively higher than some stations, but is insufficient to conclude that there is a significant linear relationship between the data. This value indicates that the model only explains about 19% of the total variability. Therefore, the trend cannot be said to be statistically significant for this station either.

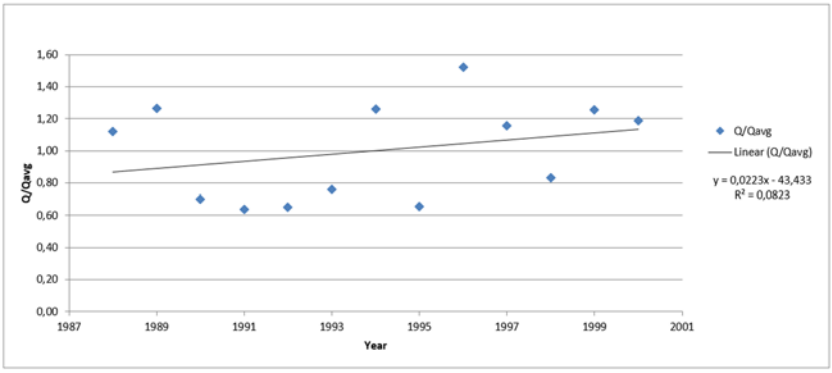


Figure 3.44 Trend Analysis for Station 2331

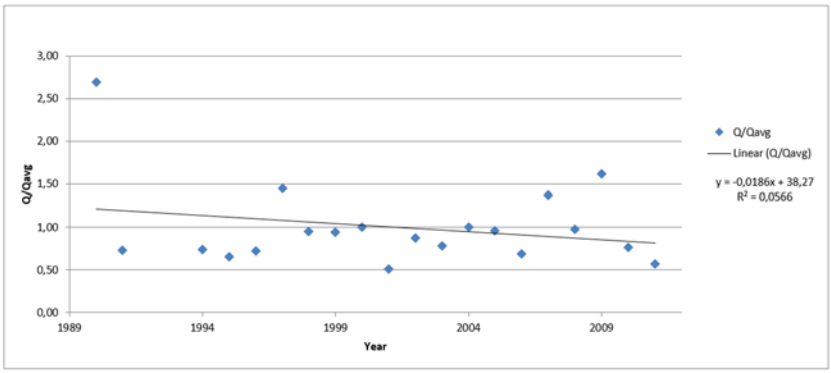


Figure 3.45 Trend Analysis for Station 2333

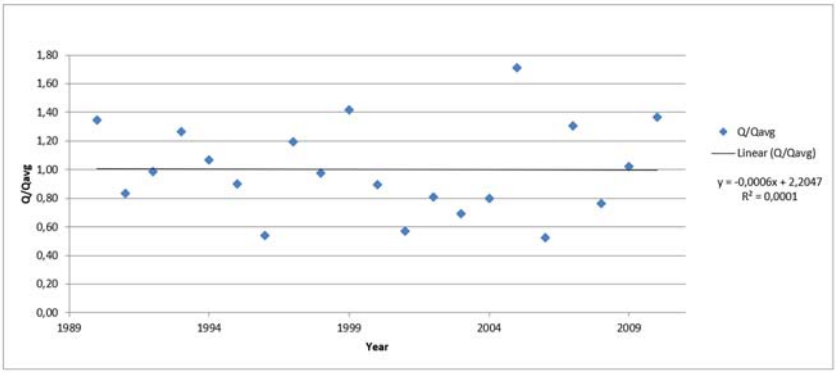


Figure 3.46 Trend Analysis for Station 2334

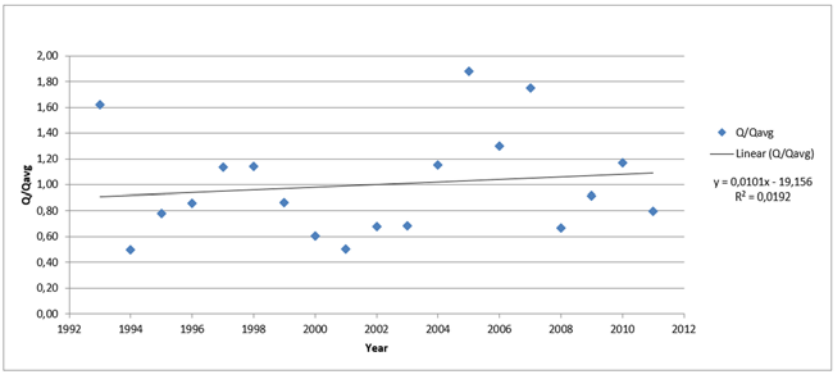


Figure 3.47 Trend Analysis for Station 2335

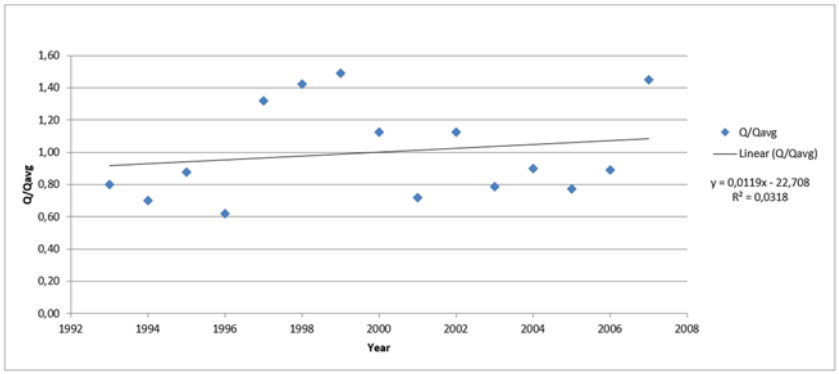


Figure 3.48 Trend Analysis for Station 2336

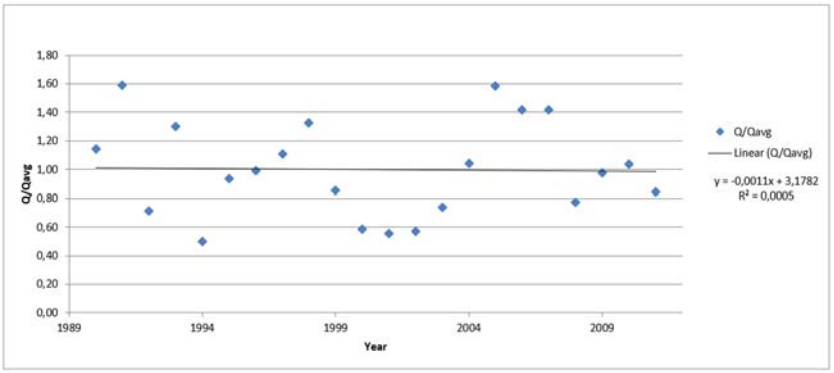


Figure 3.49 Trend Analysis for Station 2337

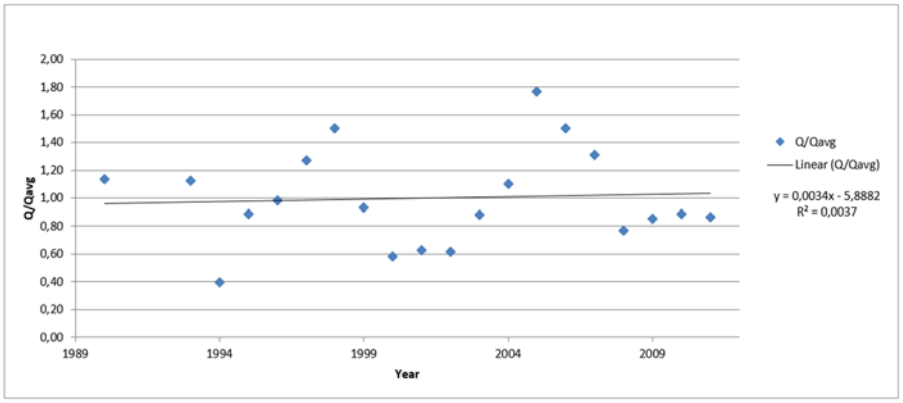


Figure 3.50 Trend Analysis for Station 2338

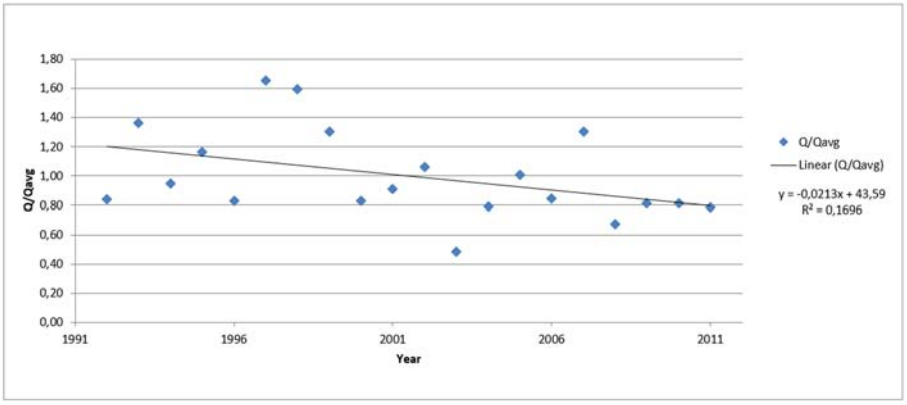


Figure 3.51 Trend Analysis for Station 2340

The slope of the regression line drawn for station number 2340 is negative, indicating that the flow rates have decreased over the years. However, the R^2 value is only 0.1696. This shows that the explanatory power of the model is quite weak and that a large part of the observed change may be random. Therefore, the trend detected here is not statistically significant.

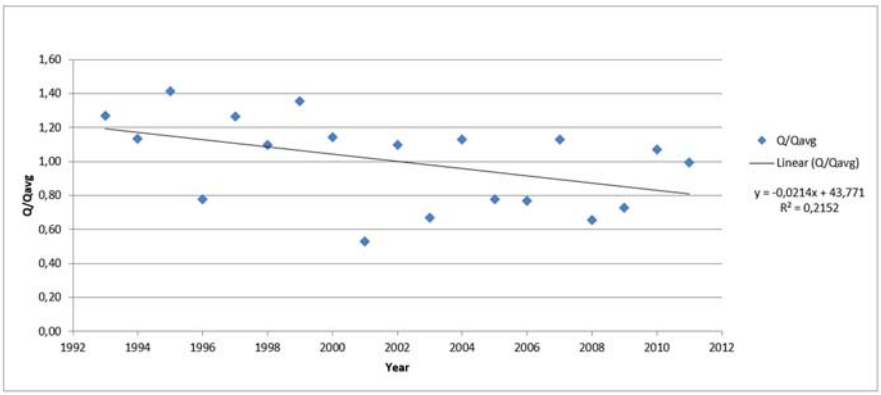


Figure 3.52 Trend Analysis for Station 2342

At station 2342, there is also a slightly decreasing trend in $Q/\bar{Q}$ values and the slope of the regression line is negative. The R^2 value is 0.2152. This low value indicates that the effect of the regression line on the variability is weak. Therefore, even if there is a visual trend at this station, this trend cannot be confirmed reliably.

3.6.1 Result

In this part of the study, trends in annual maximum flood discharges of 23 different hydrological stations were examined. First, the Mann-Kendall trend test was applied to the annual data of the stations, and test statistics such as (τ), (S), (σ_s), (Z) and $(Z_{\alpha/2})$ were calculated. Statistically significant trends were detected in four stations (2328, 2330, 2340, 2342) that met the condition $(|Z| > 1{,}96)$. However, in order to see the significance levels and trend directions of these trends more clearly, dimensionless flood discharge $((Q/\bar{Q}))$ graphs were drawn for each station and linear regression curves were created to calculate R^2 values.

Station 2328: The regression line with a negative trend yields the value of $R^2 = 0.2381$. This value shows that the model explains approximately 24% of the variation in flood discharges. The regression and Mann-Kendall test results are consistent.

Station 2330: In the regression analysis, the slope is negative and the value of $R^2 = 0.1853$ provides a moderate level of explanatory power. The trend was also found to be statistically significant with the Mann-Kendall test.

Station 2340: The trend line is negative and the value of $R^2 = 0.1696$ indicates that the model explains less than 17% of the variation in flood data. Both analyses indicate a decreasing trend at the station.

Station 2342: The obtained value of $R^2 = 0.2152$ revealed that the negative trend observed at this station has a moderate linear explanatory power. The Mann-Kendall test also confirms this trend.

As a result, the Mann-Kendall test showed a significant trend only in four stations. However, as a result of graphical evaluations of these stations, R^2 values remain at a moderate level, and it is observed that a strong linear relationship is mostly absent. These findings reveal that trends should be evaluated in longer time periods or together with regional climatic effects.

3.7 REGIONAL FLOOD ANALYSIS

One of the most fundamental problems encountered in hydrological studies is the inadequacy of long-term and reliable observation data. This situation causes significant uncertainties, especially in the application of statistical methods such as flood frequency analysis. One of the most common approaches used to overcome this problem is the use of regional analysis methods. Regional analysis enables the evaluation of stations with homogeneous characteristics by bringing them together and thus making parametric estimates on a wider data base. In recent years, this method has been widely used in many hydrological applications, especially flood estimating and flood frequency analysis. As emphasized in the literature, regional analyses produce more stable results in statistical terms and enable estimations to be made with lower error rates compared to individual station analyses. In this respect, regional analyses both reduce the effect of data inadequacy and increase the reliability of hydrological estimates.

In this part of the study, the Wiltshire method was used to create homogeneous regions based on the flood data of the stations. Homogeneity assessment, which is one of the basic stages in regional frequency analyses, was carried out on the statistical properties of the stations. In the Table 34, the number of observations for each station n, mean ($\bar{x}$), standard deviation (σ) and coefficient of variation (C_{vx}) values are given. These coefficients represent the relative variability in flood magnitude and are used as the basic parameter in homogeneity assessments.

Homogeneity analyses were carried out within the framework of the Wiltshire method with the help of the coefficient of variation, which is a comparable measure of variation between stations; in line with these analyses, it was questioned whether the stations formed a single homogeneous region. According to the results obtained, homogeneity could not be achieved in the scenario where all stations were evaluated as a single region in the first stage, therefore, the analyses were detailed by dividing them into regions.

Table 3.34 All Stations Statistics Features

Station Number	n	$\bar{x}$	σ	C_{vx}
2302	21	70,0	29,4	0,42
2304	61	94,3	48,3	0,51
2305	48	388,2	120,7	0,31
2316	46	257,0	95,4	0,37
2320	38	173,0	69,1	0,40
2321	39	76,3	18,0	0,24
2322	28	918,4	217,8	0,24
2323	46	235,4	90,6	0,38
2325	36	82,8	47,2	0,57
2326	23	70,9	28,9	0,41
2327	17	148,4	51,4	0,35
2328	29	46,3	17,7	0,38
2329	30	174,6	80,8	0,46
2330	30	18,9	10,0	0,53
2331	13	52,9	15,4	0,29
2333	20	33,1	15,8	0,48
2334	21	169,9	53,0	0,31
2335	19	721,6	286,9	0,40
2336	15	71,8	20,6	0,29
2337	22	344,9	112,3	0,33
2338	20	212,1	71,5	0,34
2340	20	19,6	5,8	0,30
2342	19	46,3	11,7	0,25

In the first stage of the regional analysis process, it was assumed that all stations in the Çoruh River Basin formed a single homogeneous region and in this direction, the homogeneity test was applied using the Wiltshire method. However, as a result of the evaluation, it was determined that there were significant differences between the stations in terms of the coefficient of variation (C_{vx}), and this situation revealed that all stations did not form a homogeneous structure under a single region.

In the final stage of the analysis, it was determined that the stations were not homogeneous even though they were initially divided into two regions. Therefore, in order to eliminate this inconsistency, the exclusionary stations in the two regions were re-evaluated. In particular, the stations with low coefficients of variation (C_{vx}) in the second region were grouped into a third group with the stations in the first region that had similar characteristics. As a result of this arrangement, the stations in the Çoruh River Basin were divided into three homogeneous regions and the homogeneity tests were re-applied for each region with the Wiltshire method.

Table 3.35 Statistics Features of Final Regions

	Station Number	n	x	σ	C_{vx}
Region 1	2305	48	388	121	0,311
	2316	46	257	95	0,371
	2320	38	173	69	0,399
	2323	46	235	91	0,385
	2326	23	71	29	0,407
	2327	17	148	51	0,346
	2328	29	46	18	0,384
	2331	13	53	15	0,291
	2334	21	170	53	0,312
	2335	19	722	287	0,398
	2337	22	345	112	0,326
	2338	20	212	71	0,337
Region 2	2302	21	70	29	0,419
	2329	30	175	81	0,463
	2333	20	33	16	0,477
	2304	61	94	48	0,513
	2325	36	83	47	0,569
	2330	30	19	10	0,530
Region 3	2321	39	76	18	0,235
	2342	19	46	12	0,252
	2336	15	72	21	0,287
	2340	20	20	6	0,298
	2322	28	918	218	0,237

Table 3.36 Comparison with sampling values of 3 regions.

Region Number		Present C_{vx}	Sampling Value
1	Maximum C_{vx}	0,41	0,45
	Minimum C_{vx}	0,29	0,24
2	Maximum C_{vx}	0,57	0,64
	Minimum C_{vx}	0,42	0,35
3	Maximum C_{vx}	0,30	0,34
	Minimum C_{vx}	0,24	0,18

The results showed that all three regions met the homogeneity criteria. Thus, the iterative classification process, which took into account the statistical characteristics of the stations based on flood data, both increased the reliability of regional estimates and allowed the emergence of a meaningful sub-region structure in terms of hydrological homogeneity. This situation is considered as an important methodological gain that strengthens the accuracy and application capability of regional frequency analyses.

3.7.1 Regional Mann-Kendall Trend Assessment

In this section, after dividing the 23 selected stations into regions, the Mann-Kendall trend test results calculated at the stations belonging to each region were analyzed. For each region, trending and untrending stations were evaluated separately and regional trend behaviors were revealed.

Region 1: The stations in this region are: 2305, 2316, 2320, 2323, 2326, 2327, 2328, 2331, 2334, 2335, 2337 and 2338. According to the results of the Mann-Kendall test, only station 2328 met the condition $Z = -2.72 < -1.96$ and a statistically significant decreasing trend was observed at this station. Since all other stations have $|Z| < 1.96$, there is no significant trend. Therefore, Region 1 generally shows a stationary structure and only one station has a weak trend.

Region 2: The stations in this region are: 2302, 2329, 2333, 2304, 2325 and 2330. As a result of the analysis, a statistically significant decreasing trend was detected at station 2330 with $Z = -2.30 < -1.96$. In addition, a trend

close to the significance limit was observed at station 2325 with $Z = -1.57$. There is no significant trend at other stations. According to these results, Region 2, although it includes decreasing trends at certain stations, generally exhibits a limited trend behavior.

Region 3: The stations in this region are: 2321, 2342, 2336, 2340 and 2322. According to the Mann-Kendall test results, statistically significant decreasing trends were determined with the values of $Z = -2.21 < -1.96$ at station 2340 and $Z = -2.34 < -1.96$ at station 2342. No significant trend was observed at the other three stations. Therefore, Region 3 is the region with the most statistically significant trend.

General Assessment: In light of the above analyses, only one of the three regions (Region 1) showed a largely stable structure, while trends were detected at a limited number of stations in Region 2. The most striking finding is the existence of statistically significant trends at two stations in Region 3. This suggests that Region 3 has shown greater change compared to other regions in terms of climatic or geographical effects.

3.8 SEASONALITY ANALYSIS

In this part of the study, angular seasonality analysis was applied to the flow data of the stations and the regularity and seasonal concentration trends regarding the timing of flood events within the year were evaluated.

3.8.1 Angular Seasonality Analysis

In this study, in order to perform seasonality analysis, the dates of the flood events recorded at each station were calculated as the corresponding day of the year (calendar day). The obtained calendar day information was converted to radians using Eq. 2.102. Here $(\text{Calendar Day})_i$ represents the order of the relevant flood event within the year, and θ represents the corresponding radian angle.

Using radian angle values, the x_i and y_i coordinates on the Cartesian plane were calculated for each flood observation (Eq. 2.103 and 2.104). For each station, the averages of these x_i and y_i values were taken to determine $\bar{x}$ and $\bar{y}$, respectively, and using these average coordinates, the direction of flood events, i.e. the time interval in which they are concentrated during the year, was estimated.

Finally, the average timing of annual floods for each station was determined with the average angle $\bar{\theta}$ (Eq. 2.105 or Eq. 2.106) corresponding to the averages of $\bar{x}$ and $\bar{y}$ and the flood day (Eq. 2.107) obtained from this angle.

In this study, the temporal distribution of floods within the year was examined for each station and seasonality analysis was performed by calculating the angular parameters of this distribution. With the method used in the study, the day of the year (MD) for each flood event, the radian value $\bar{\theta}$ that corresponds to these days, and the vector averages representing the orientation with the $\bar{x}$ and $\bar{y}$ coordinates were determined using these values. In addition, the r value, which expresses the homogeneity of the distribution of floods within the year, was also calculated.

According to the results obtained, it was determined that the majority of floods occurred between the end of April and mid-May. In particular, the 120–130th calendar day (approximately between April 30 and May 10) stands out as the flood day in most of the stations. This situation corresponds to the period when snowmelt and spring precipitation are effective in the region and is an important indicator of hydrological seasonality.

Table 3.37 Angular Parameters of the Stations

Station Number	n	$\bar{x}$	$\bar{y}$	Radian	r	MD	Flood Calendar Day
2302	21	-0,69	0,67	2,37	0,97	138	18.May
2304	61	-0,47	0,82	2,09	0,95	121	1.May
2305	48	-0,53	0,78	2,17	0,95	126	6.May
2316	46	-0,48	0,83	2,09	0,96	121	1.May
2320	38	-0,48	0,80	2,11	0,93	122	2.May
2321	39	-0,79	0,45	2,62	0,91	152	1 Jun
2322	28	-0,55	0,79	2,18	0,97	127	7.May
2323	46	-0,50	0,82	2,12	0,96	123	3.May
2325	36	-0,43	0,79	2,07	0,90	120	30 Apr
2326	23	-0,53	0,64	2,26	0,83	131	11.May
2327	17	-0,55	0,78	2,19	0,96	127	7.May
2328	29	-0,38	0,71	2,06	0,80	120	30 Apr
2329	30	-0,43	0,85	2,04	0,95	118	28 Apr
2330	30	-0,76	0,52	2,54	0,92	148	28.May
2331	13	-0,16	0,51	1,88	0,53	109	19 Apr
2333	20	-0,55	0,55	2,36	0,78	137	17.May
2334	21	-0,33	0,79	1,97	0,86	115	25 Apr
2335	19	-0,44	0,82	2,07	0,93	120	30 Apr
2336	15	-0,69	0,47	2,54	0,83	148	28.May
2337	22	-0,47	0,81	2,10	0,94	122	2.May
2338	20	-0,37	0,87	1,98	0,95	115	25 Apr
2340	20	-0,33	0,66	2,04	0,73	118	28 Apr
2342	19	-0,81	0,53	2,56	0,97	149	29.May

When the r values of the stations are examined, a high level of directionality is generally observed in terms of the distribution of flood dates within the year. The r value varies between 0 and 1, and as it approaches 1, it is understood that the flood days are concentrated around a certain date. According to the table given, the average r value is above 0.90, and this situation shows that floods are repeated in a certain period at most of the stations.

For example, r = 0.97 was calculated for station No. 2302, and the floods are concentrated on the 138th day (May 18). Similarly, a very significant directionality was observed at station No. 2342 with the r = 0.97 value, and the floods coincide with the 149th calendar day (May 29). These high r values reveal that flood events occur on a fairly consistent date over the years at these stations.

However, the lowest r value was calculated at station 2331 (r = 0.53). This situation shows that the flood dates at this station are more scattered throughout the year. Such stations may be exposed to different hydrological regimes or may have been affected by different meteorological conditions during the data period.

As a result, with this analysis, it can be said that the majority of the stations in the Çoruh River Basin are exposed to floods in the spring months and especially in early May, flood dates are mostly repeated at certain time intervals and this situation is also supported by the r coefficient. These findings are important in terms of making seasonal flood predictions and shaping flood prevention policies.

3.8.2 Unit Circles

In this part of the study, the days on which floods occurred for each of the 23 selected flow observation stations were calculated by converting them to radians and the obtained angular values were shown on the unit circle. Diagrams visualizing the locations of the flood days of each station on the unit circle are presented below.

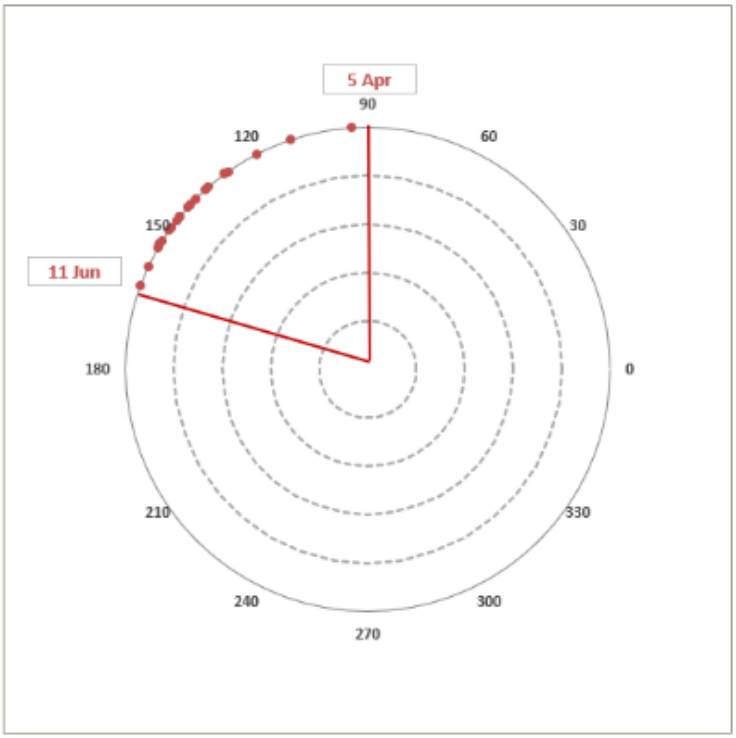


Figure 3.53 Unit Circle of the Station 2302

When the angular seasonality analysis of the flow observation station number 2302 is examined, it is seen that the floods generally occur towards the end of spring during the year. The mean flood day (Mean Date - MD) is the 138th day and this date corresponds to May 18. This historical intensity indicates the increased surface flow with the snowmelt in the region. The earliest flood observed at the station was recorded before March 21, and the latest flood was recorded near the beginning of June. The circular distribution of the points in the

graph is quite clustered, which shows that the floods are concentrated in a certain period during the year. In parallel with this pattern, the homogeneity coefficient (r) of the station was found to be quite high as 0.97, indicating a strong directionality and low variation between the dates when the floods were observed. In this context, it can be said that the floods for station number 2302 developed in the spring months largely due to snowmelt and were distributed homogeneously in time.

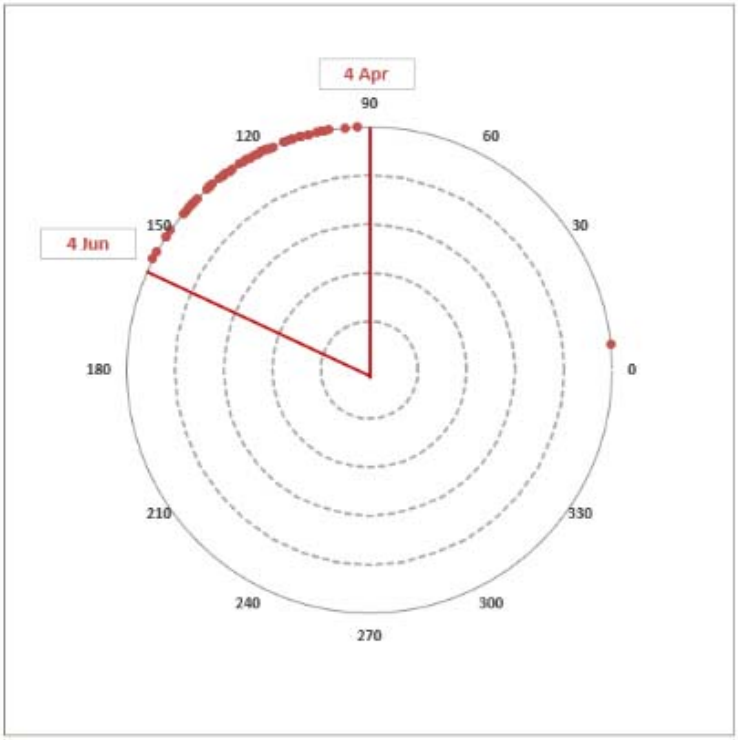


Figure 3.54 Unit Circle of the Station 2304

When the unit circle graph of station 2304 is examined, it is observed that flood events are concentrated in the spring months of the year. The average flood day (MD) is the 121st day and this date corresponds to May 1. This result indicates the periods when snowmelt and seasonal precipitation combine, and shows that the hydrological regime in the region is activated in the spring months. The clustered structure of the points on the graph shows that floods are repeated in a certain period during the year.

The homogeneity coefficient (r) was calculated as 0.95, and this value indicates a high consistency in the timing of floods. In this context, it can be said that the floods observed at station 2304 throughout the years generally occur on similar dates, and therefore temporal homogeneity is strong. This obtained pattern shows that this station provides a reliable sample in terms of seasonality analysis and that flood estimates can be interpreted with high accuracy in terms of timing.

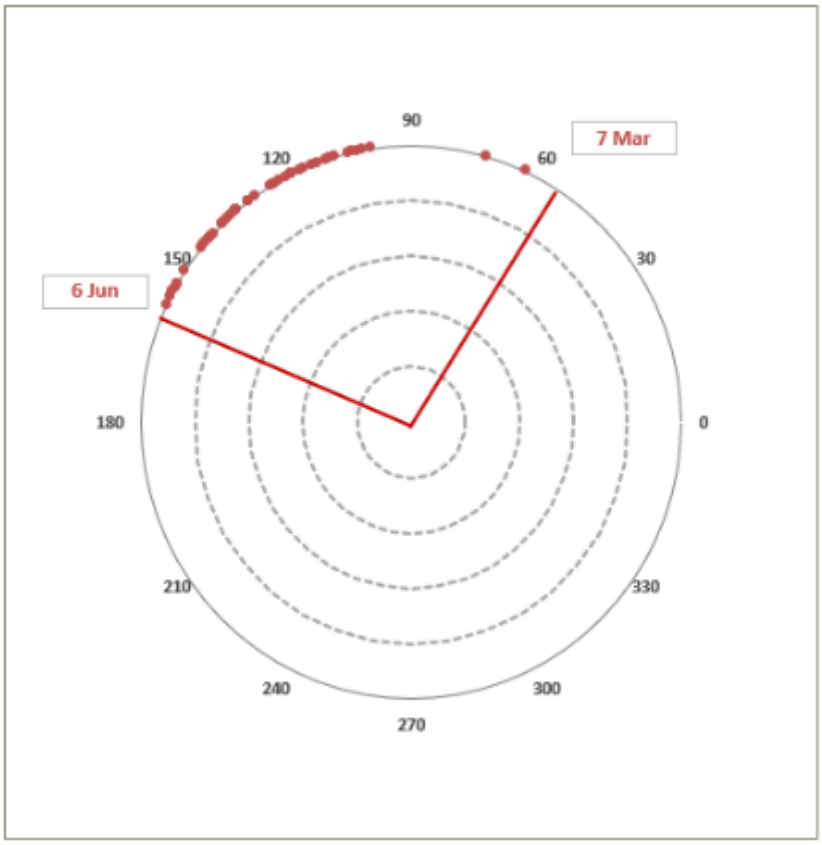


Figure 3.55 Unit Circle of the Station 2305

The unit circle graph for station number 2305 reveals that flood events mostly occur in the spring season. The average flood day (MD) was calculated as 126, which corresponds to May 6. The intensification of floods during this period shows that the flow regime in the region has been revived, especially due to the effect of snowmelt and spring precipitation.

When the locations of flood events on the circle are examined, it is seen that the observations are clustered in a certain period during the year. This situation indicates that the floods at the station show a regular seasonal behavior. The homogeneity coefficient (r) value is 0.95, which shows a very high temporal consistency. This shows that the floods that occurred throughout the years mostly occur in the same periods and the station is a strong example in terms of seasonal analysis.

It can be said that the floods at this station develop based on spring precipitation and snowmelt in terms of both timing and periodic intensity. As a result, the angular analysis performed at station 2305 reveals that seasonality is evident and provides reliable data in terms of estimating and planning.

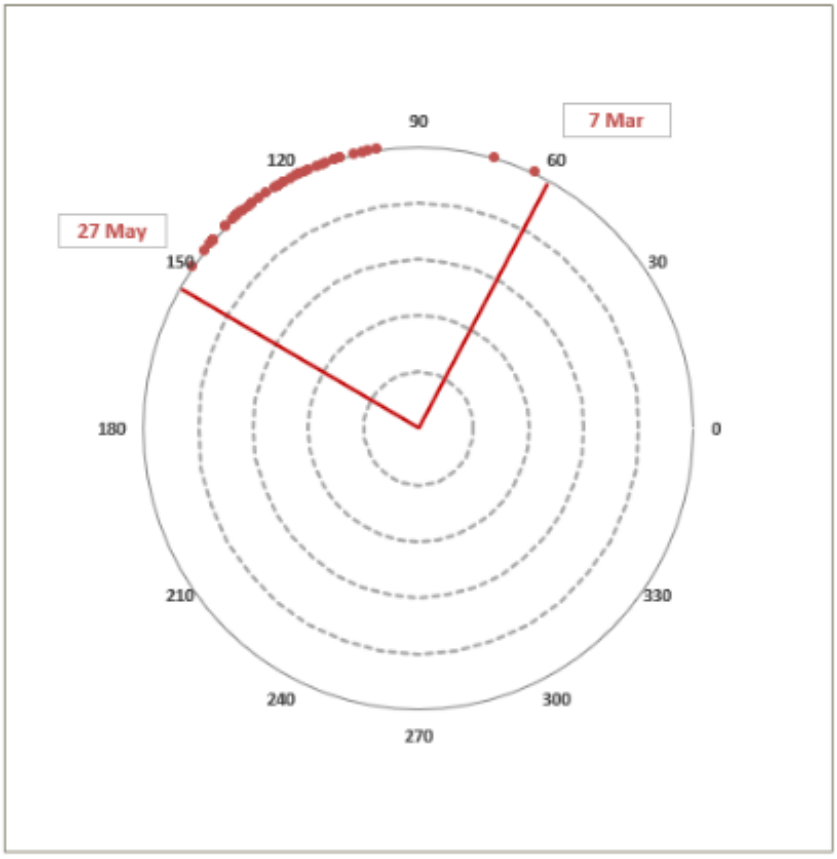


Figure 3.56 Unit Circle of the Station 2316

When the unit circle graph of the flood observations of station no. 2316 is examined, it is seen that the events are concentrated around May 1, the 121st day of the year. The average flood day (MD) was calculated accordingly and seasonaly indicates the middle of the spring season. This period corresponds to the period when snowmelt and spring precipitation are effective.

The homogeneity coefficient (r) of the station is a very high value of 0.96. This shows that the floods are consistent and accumulated in a certain period over time. The fact that the distribution on the circle is tight and dense supports that the floods occur in the same period of the year and have a distinct seasonality.

In the light of these data, it can be said that the flood events of station no. 2316 are highly dependent on the spring period, occur with the effect of increasing precipitation and snowmelt, and show a fairly homogeneous distribution in time. This station has an exemplary modeling potential in angular seasonality analysis.

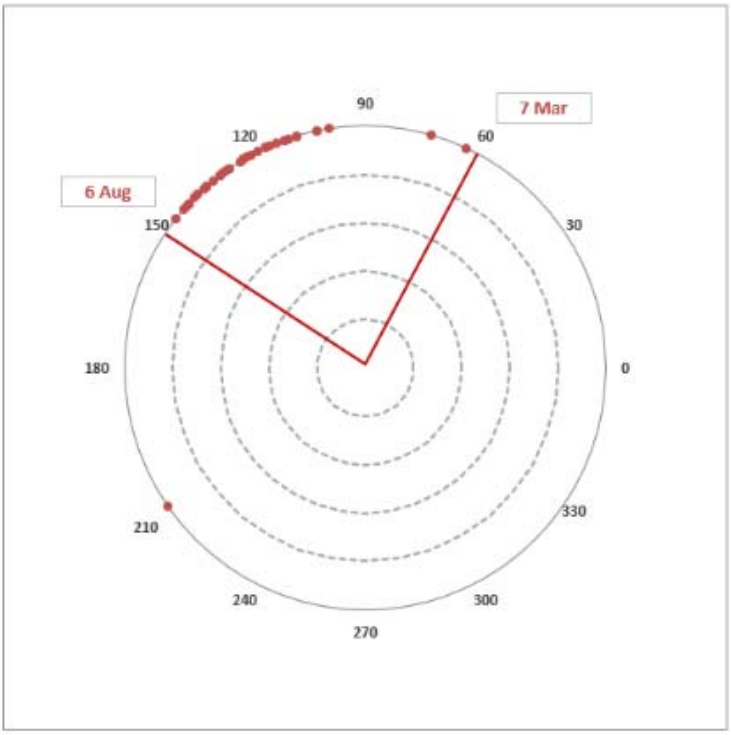


Figure 3.57 Unit Circle of the Station 2320

According to the seasonality analysis carried out at station 2320, the average occurrence day (MD) of flood events is day 122, which corresponds to May 2. At this station, where floods intensify towards the end of spring, the r value was calculated as 0.93. This high r value shows that the distribution of floods throughout the year is quite distinct and regularly concentrated in a certain period.

When the angular distribution on the unit circle is examined, it is seen that floods are mostly observed between calendar days 110 and 150, that is, they mostly coincide with the end of April and mid-May. However, in a few extreme cases, floods have also occurred in different periods such as the end of March (day 60) and the end of July (day 210). This situation shows that exceptional meteorological conditions (sudden snow melt, summer showers, etc.) may be effective.

In general, the clustering of floods in a certain period of the year reveals that snowmelt and spring precipitation play a decisive role in the formation of floods at this station. In this period when agricultural activities accelerate, flood risk poses a significant threat. Therefore, it is recommended that early warning systems be activated and agricultural planning be made according to the flood calendar during this period.

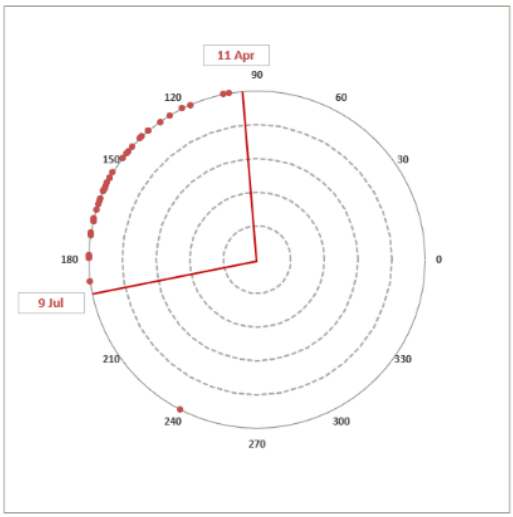


Figure 3.58 Unit Circle of the Station 2321

When the distribution of floods during the year at the flow observation station numbered 2321 was examined, it was determined that the majority of the floods occurred in the late May and early June period. The calculated mean flood day (Mean Date - MD) is the 152nd day, which corresponds to June 1. The floods that occurred during this period are most likely the result of sudden flow increases caused by the melting of snow cover accumulated in high mountain regions during the winter season in the spring.

The value of r, which is the homogeneity indicator calculated for the station, was determined as 0.91. This high value reveals that flood events are clustered in a certain time interval of the year and show seasonal regularity. In the angular distribution graph made on the unit circle, it is clearly seen that the observation values are concentrated between 120 and 180 calendar days. However, floods that occurred on later dates (approximately 240th day - early September) were also observed, although rarely. This situation shows that the flood regime is predominantly concentrated in the spring-summer transition period, but exceptional floods can last until the end of summer due to regional hydro-meteorological variability.

The analysis conducted at this station emphasizes that hydrological planning should focus especially on the end of spring and the beginning of summer and that it is critical to activate flood risk management during this time period. At the same time, since this period is also the time period when agricultural activities are intensified, early warning systems and flood control strategies should be considered in order to minimize agricultural damages from possible floods.

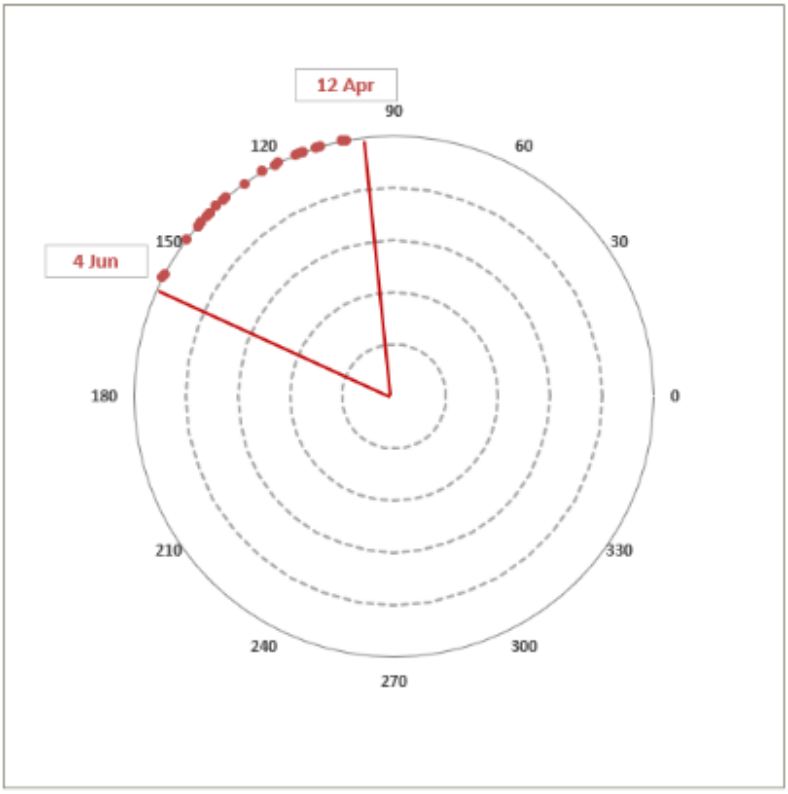


Figure 3.59 Unit Circle of the Station 2322

According to the seasonality analysis results of station 2322, the average day of occurrence of flood events is the 127th day, which corresponds to May 7. This date corresponds to the end of spring and indicates a period when the risk of flooding increases due to the effect of increasing precipitation with snow melt. The r value was calculated at a very high level of 0.97, which shows that flood events are concentrated in a very specific time interval during the year.

In the analysis conducted on the unit circle, it is observed that the majority of the data is concentrated between days 110 and 150. This situation reveals that floods occur due to meteorological and hydrological climatic conditions specific to the spring season (for example: snow melting and seasonal precipitation due to temperature increase). At the same time, this pattern shows that the hydrological regime in the region has a significant seasonality.

Expressing a high level of homogeneity, r = 0.97 indicates that there is a very low dispersion in the timing of the floods observed at station 2322, that is,

the floods occur almost at the same time every year. This feature shows that flood predictions can be made more reliably at the station in question and the effectiveness of early warning systems can be increased.

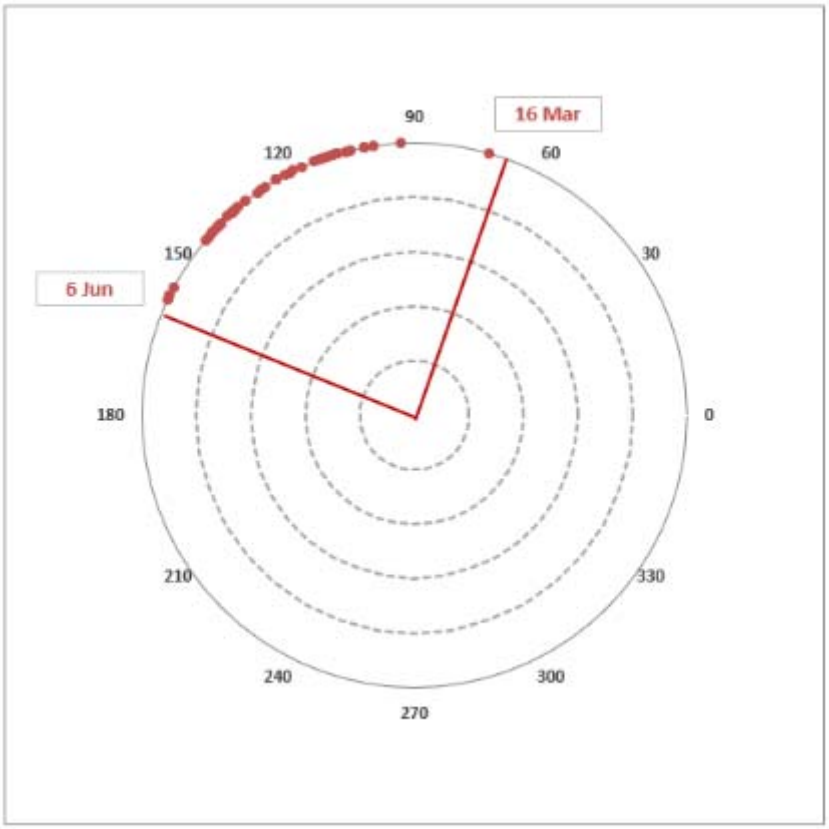


Figure 3.60 Unit Circle of the Station 2323

When the flood observations of station 2323 are examined, the average day when the floods occur is the 123rd day, which corresponds to May 3. This historical location coincides with the middle of the spring months in the region and indicates the period when seasonal precipitation increases with the effect of snow melt.

In the angular analysis results of the station, the r value was calculated as 0.96. This high homogeneity coefficient shows that flood events are concentrated in a certain period during the year and are repeated in a consistent manner in time. In other words, there is a high probability that floods will occur

on similar dates every year at station 2323. This situation reveals that flood timings have a significant seasonality.

When the unit circle graph is examined, it is seen that floods are mostly concentrated between the 110th and 150th calendar days. This distribution shows that flood events are more frequent in the spring season, especially in late April and early May. Although there are a few years in which floods occur early among the observations, the general distribution is concentrated in the spring months.

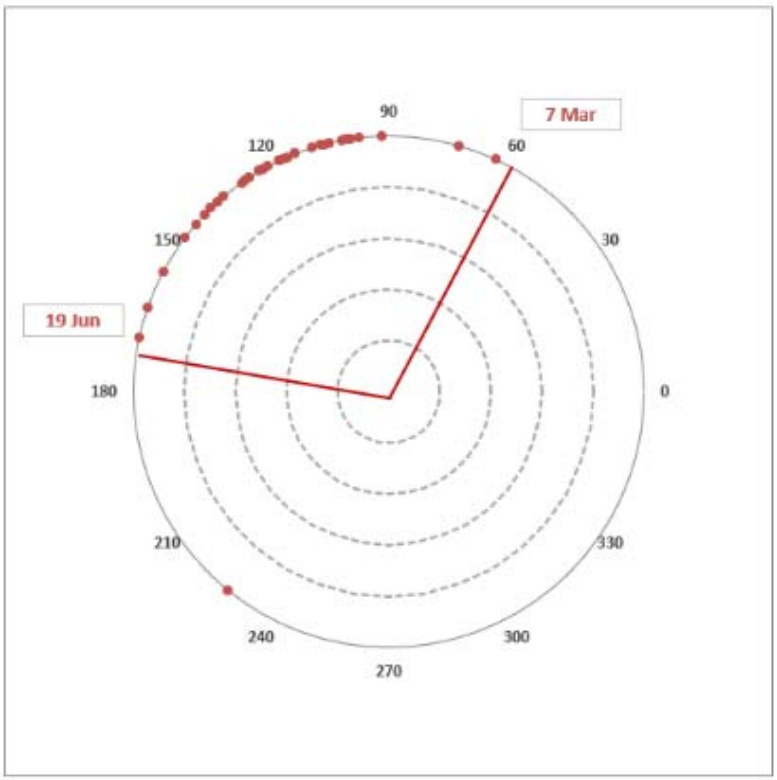


Figure 3.61 Unit Circle of the Station 2325

According to the seasonality analysis conducted for station number 2325, the average day when floods occur is the 120th day, which corresponds to April 30th in calendar terms. This time period coincides with the period when spring precipitation in the region intensifies and indicates a critical time interval when floods are triggered by the effect of snowmelt.

The r value of the station was calculated as 0.90. This high homogeneity value shows that flood events occur quite consistently and at similar times throughout the year. This reveals that flood occurrences have a clear seasonal pattern and are highly predictable.

In the unit circle graph, it is observed that the majority of the data is concentrated between calendar days 110–150. In addition, floods were recorded at earlier or later times in a few years. Especially floods occurring on dates close to the beginning of summer, such as days 210 and 240, can be associated with extreme events or unusual rainfall conditions.

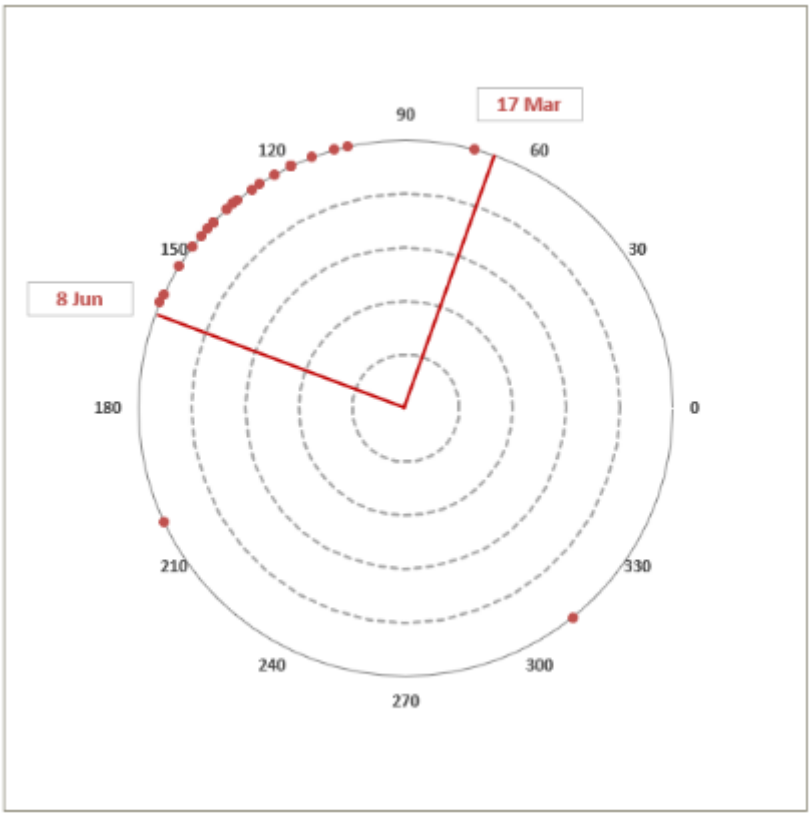


Figure 3.62 Unit Circle of the Station 2326

According to the seasonality analysis of station 2326, the average day of flood occurrence is the 131st day, which corresponds to May 11. This period coincides with the middle of the spring season and indicates a time period when

flood risk becomes evident with both snow melt and increasing precipitation regime.

The r value of the station was calculated as 0.83. This value indicates that flood dates show a largely homogeneous distribution throughout the year, that is, floods are generally concentrated in a certain period. However, it is observed in the unit circle graph that floods occur on unusual dates for a few years, especially around the 210th and 300th days. These deviations may be a result of extreme meteorological events or regional hydrological variability.

However, the majority of the floods are concentrated between calendar days 120–150, indicating regular flood occurrence in spring months during the analyzed period. This feature shows that the station has a strong seasonal flood characteristic and provides a reliable temporal pattern that can be used in early warning systems.

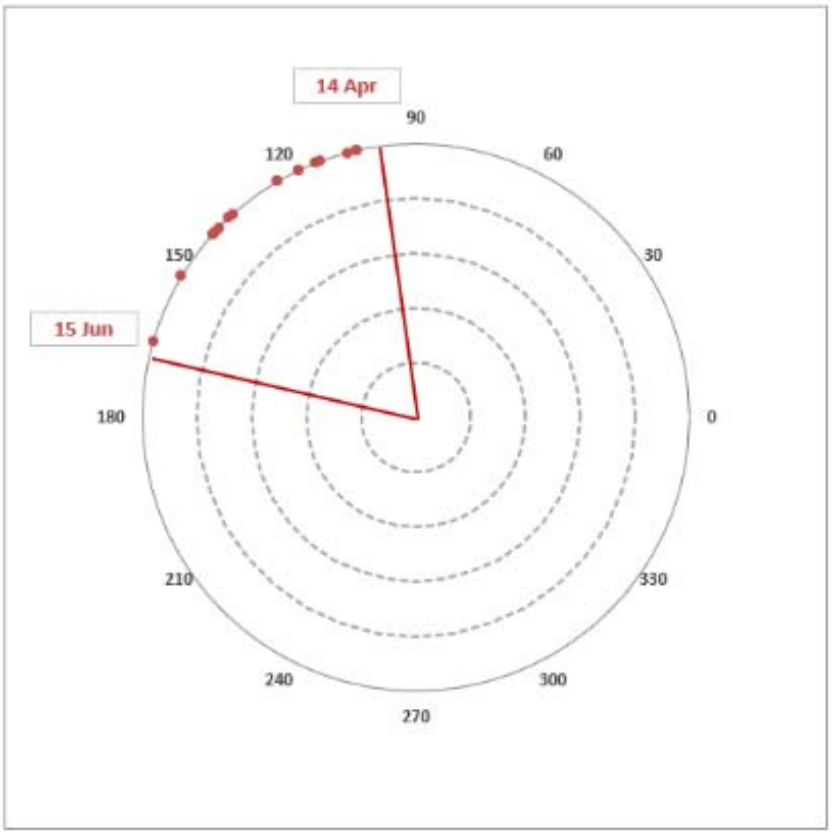


Figure 3.63 Unit Circle of the Station 2327

According to the seasonality analysis results of station 2327, the average day of occurrence of floods is the 127th day, which corresponds to approximately May 7. This date corresponds to the middle of the spring season and shows that spring precipitation, which increases with snowmelt, plays a decisive role in the formation of floods.

The seasonal homogeneity coefficient of the station, r, is quite high at 0.96. This value indicates that floods occur regularly and intensively in a certain time interval on an annual basis. It is also seen in the unit circle graph that flood days are largely concentrated between calendar days 120–140, with only a few data falling outside this range. This situation proves that there is a high level of homogeneity and seasonality in terms of the timing of floods.

The fact that the flood times of the station show such a distinct pattern reveals that hydrological processes are especially active in the spring period in this location and that this period is critical in terms of floods. Therefore, station 2327 can be considered as an important station that provides high predictability in terms of the development of flood early warning systems and regional flood management.

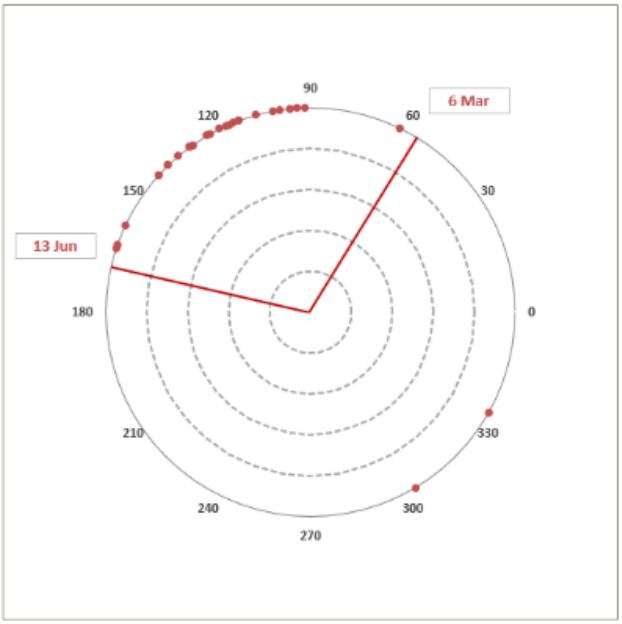


Figure 3.64 Unit Circle of the Station 2328

According to the seasonality analysis conducted at station number 2328, the average day of occurrence of floods is the 120th day, which corresponds to approximately April 30. This date indicates a period approaching the end of the spring season, suggesting that snowmelt and precipitation increases experienced during seasonal transitions are effective in flood formation.

The r value expressing the seasonal homogeneity level of the station is 0.80, which indicates a medium-high level of regularity in the distribution of floods throughout the year. In the representation made on the unit circle, the majority of floods are clustered between calendar days 90 and 160. However, it is also observed that a few observations are shifted to the autumn period of the year (for example, after day 300). This situation reveals that different meteorological conditions (for example, late autumn precipitation) may cause floods in some years at the station.

Although the timing of floods at this station is generally concentrated in the spring season, there are also occasional flood observations in other months of the year. This diversity suggests that the hydrological system in the region where the station is located may be affected by different climatic effects periodically. Therefore, station 2328 can be considered as a location that should be carefully monitored in terms of both spring and, rarely, autumn floods.

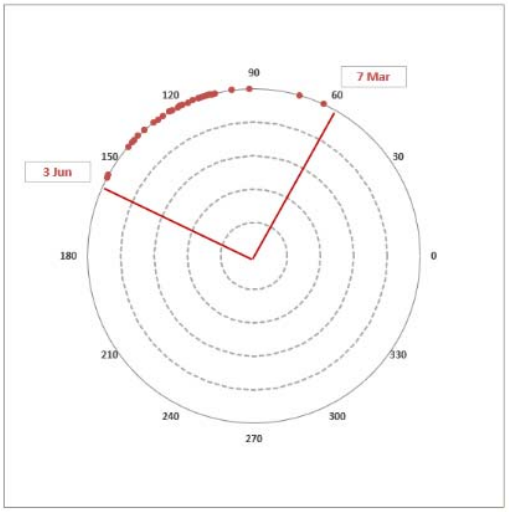


Figure 3.65 Unit Circle of the Station 2329

According to the angular seasonality analysis carried out at station number 2329, the average day of occurrence of floods is the 118th day, which corresponds to approximately April 28. This date corresponds to the middle of the spring season and represents the period when floods are triggered, especially by the effect of snowmelt and spring precipitation.

The r value indicating the homogeneity level of the station is 0.95, which is quite high. This situation shows that floods are concentrated in a certain time interval during the year and exhibit a consistent distribution in time. The distribution on the unit circle also supports this result, and flood observations are largely concentrated between days 90 and 150.

The fact that flood events occur in such a narrow time interval during the year shows that the hydrological regime in the region has high seasonal dependence. In this region where climatic conditions are distinct, especially in the spring period, increasing temperatures, snow melting and precipitation trigger flood formation. Therefore, station 2329 can be considered as a critical station with high reliability in terms of homogeneity, where spring floods are frequently observed.

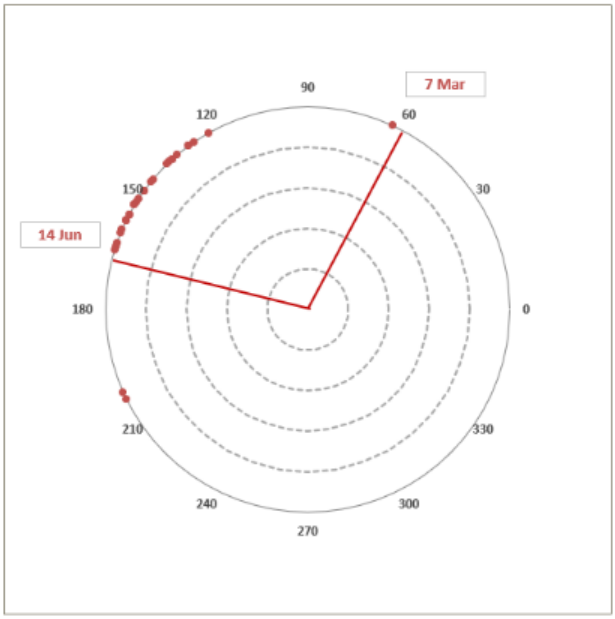


Figure 3.66 Unit Circle of the Station 2330

According to the angular seasonality analysis conducted for station 2330, the average occurrence day of flood events is day 148, which corresponds to May 28. This shows that floods are generally concentrated in a period close to the end of spring and the beginning of summer.

The r value, which is the homogeneity indicator of the station, was calculated as 0.92. This is a very high value and reveals that floods occur repeatedly in a certain period of time during the year. When the distribution on the unit circle is examined, it is seen that floods are mostly concentrated between calendar days 120 and 150. However, the fact that a few data points are distributed around day 60 and day 210 indicates that exceptional flood events were also recorded.

In the light of these data, it can be said that the majority of the floods observed at station 2330 occurred in the last part of the spring season and the beginning of summer, which corresponds to the period when snow melting is completed in the region and seasonal precipitation is effective. The high r value of the station shows that the seasonal flood dynamics are distinct and consistent, indicating that this station constitutes a reliable example in seasonality analyses.

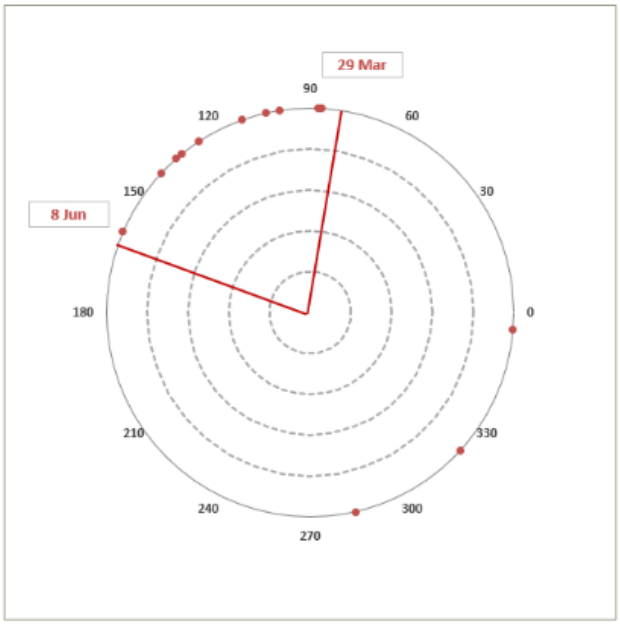


Figure 3.67 Unit Circle of the Station 2331

In the seasonality analysis conducted for station 2331, the average occurrence day of flood events was determined as the 109th day, namely April 19. This date coincides with the middle of spring and corresponds to the period when precipitation begins to increase.

However, the r value of this station was calculated as 0.53, which is quite low compared to other stations. The low r value indicates that floods are distributed throughout the year and are not concentrated in a certain season. Indeed, when the distribution on the unit circle is examined, it is understood that floods are observed not only in spring but also in summer, autumn and winter seasons. In particular, the data points located around calendar days 0, 270 and 330 reveal that floods occur in a wide time interval throughout the year.

In this context, it can be concluded that floods do not show a seasonal intensity at station 2331, i.e., seasonality is weak. This situation may be due to irregularities in local precipitation conditions or different hydrometeorological factors. Therefore, this station shows that the reliability of seasonal Estimations in predicting floods may be low.

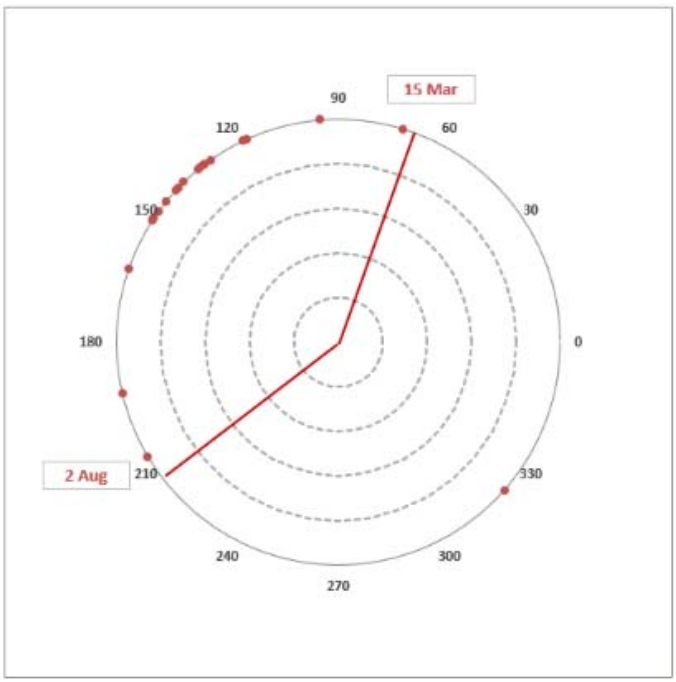


Figure 3.68 Unit Circle of the Station 2333

The average time of occurrence of flood events for station number 2333 was calculated as the 137th calendar day, i.e. May 17. This date coincides with the end of spring and indicates the period when seasonal precipitation intensifies together with snowmelt. This situation suggests that the floods in the region are largely triggered by the melting of snow accumulated during the spring-summer transition, especially in high mountain areas.

The r value indicating the degree of homogeneity of the station is 0.78, which reveals a moderate seasonal concentration in the timing of floods. When the distribution on the unit circle is examined, it is seen that the majority of the floods occur between days 120 and 150, i.e. between the end of April and the middle of May. However, a few flood data were also observed at different periods of the year, which shows that the data does not show 100$\%$ seasonal intensity but presents a general trend.

In this context, floods are significantly concentrated in spring at station 2333, which shows that seasonal predictions are feasible from a hydrological perspective. Moderate homogeneity supports the reliability of the predictions.

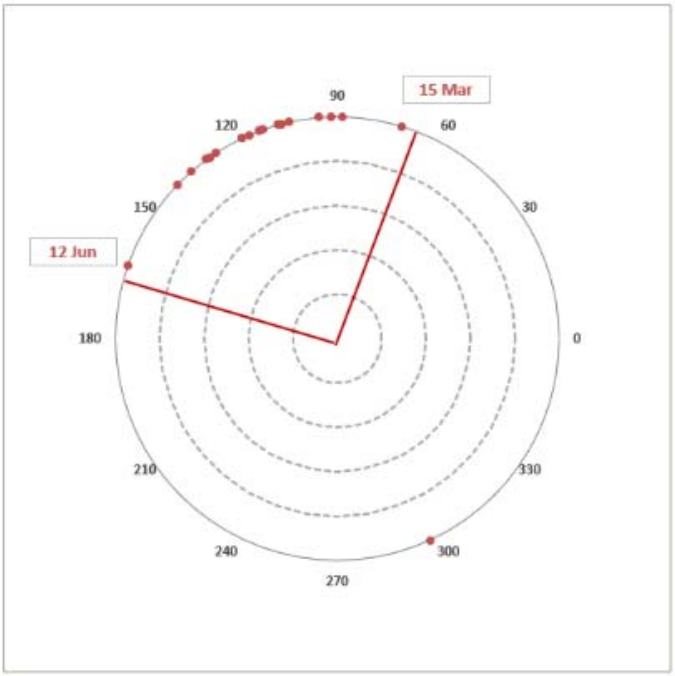


Figure 3.69 Unit Circle of the Station 2334

The average occurrence day of flood events for station number 2334 is the 115th calendar day, which corresponds to April 25. This time period coincides with the middle of spring and represents a period when snowmelt and spring precipitation are effective in the region. This seasonal effect plays an important role in the hydrological cycle of the region.

The r value calculated for the station was found to be 0.86. This high value shows that floods are concentrated in a very specific period during the year. In other words, flood events at this station show a high seasonal regularity in terms of time. As can be seen in the unit circle graph, flood observations are largely collected between mid-April and early May. However, there are rare flood observations from other periods of the year.

In the light of these data, it can be said that the majority of the floods observed at station number 2334 occurred in the spring season, especially during the transition period when snowmelt is effective. Due to the high degree of homogeneity, the reliability of seasonal based flood Estimations at this station is quite high.

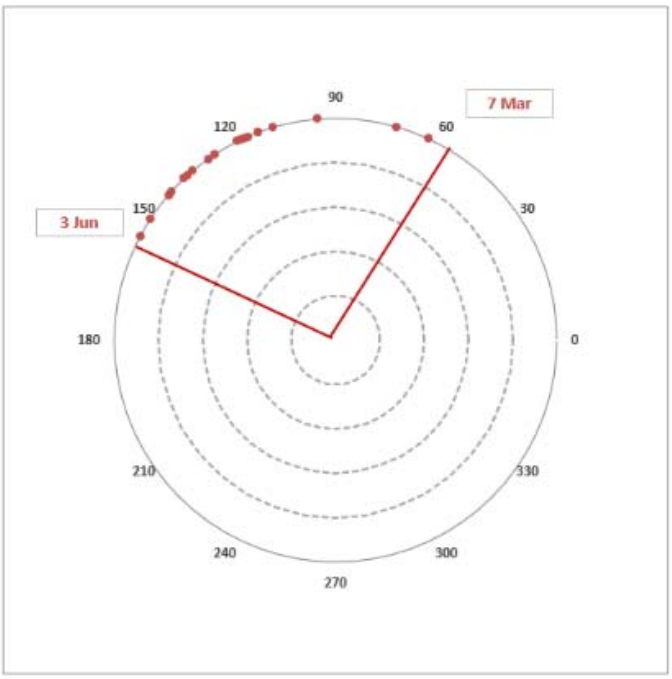


Figure 3.70 Unit Circle of the Station 2335

According to the angular seasonality analysis conducted at station 2335, the average occurrence day of flood events during the year was calculated as the 120th calendar day. This date corresponds to April 30, which coincides with the middle of the spring season. This period is known as the time period when snowmelt is completed and spring precipitation intensifies.

The r value of this station is 0.93, which is quite high. This indicates that flood events are clustered in a certain period of the year and show a distinct seasonality in terms of time. Flood days on the unit circle are largely concentrated between the end of April and the beginning of May, which reveals that seasonal regularity is high.

This regularity observed at station 2335 can be associated with the hydrological regime of the region. Located at the intersection of the Black Sea and Eastern Anatolian climate characteristics, this region remains at risk of flooding due to the combined effects of snowmelt and precipitation in the spring period. Therefore, the station's flood observations increase the reliability of seasonal Estimations and support the effectiveness of early warning systems.

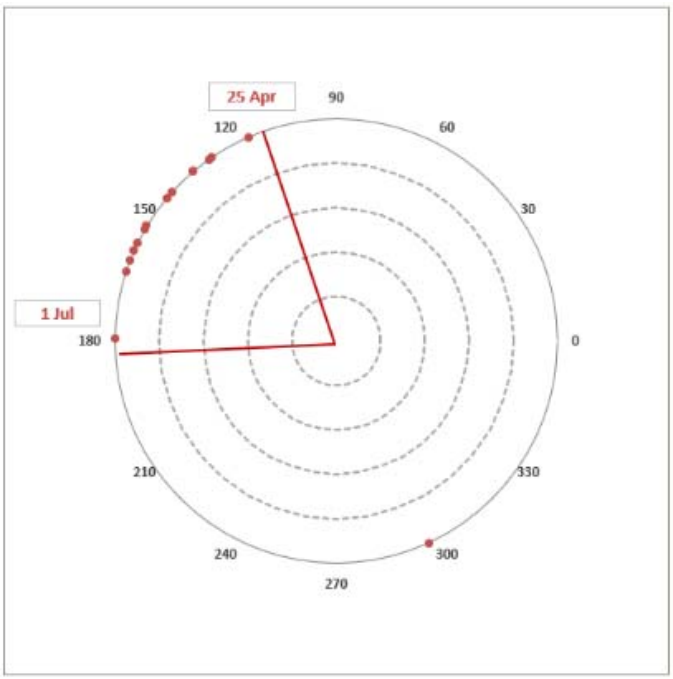


Figure 3.71 Unit Circle of the Station 2336

According to the results of the angular analysis performed at station 2336, the average occurrence day of flood events during the year was calculated as the 148th calendar day, i.e. May 28. This date corresponds to the end of spring and the beginning of summer. It indicates a period when snowmelt accelerates with the increase in temperatures in the region and increases the risk of floods by combining with precipitation.

The r value of the station is 0.83, which indicates that floods are concentrated in a certain time period during the year and have a fairly homogeneous distribution in terms of timing. The clustering observed in the polar graph shows that the data is largely collected in a certain sector and contains only a few extreme examples. This supports the fact that floods occur seasonally consistently.

According to the analysis results, the main cause of floods for station 2336 is the precipitation that is effective together with snowmelt in the late spring and early summer periods. The fact that most of the flood observation dates are concentrated in the last week of May reveals that this period is critical in terms of flood risk. The high homogeneity value obtained by the station is an important indicator for the reliability of flood Estimations.

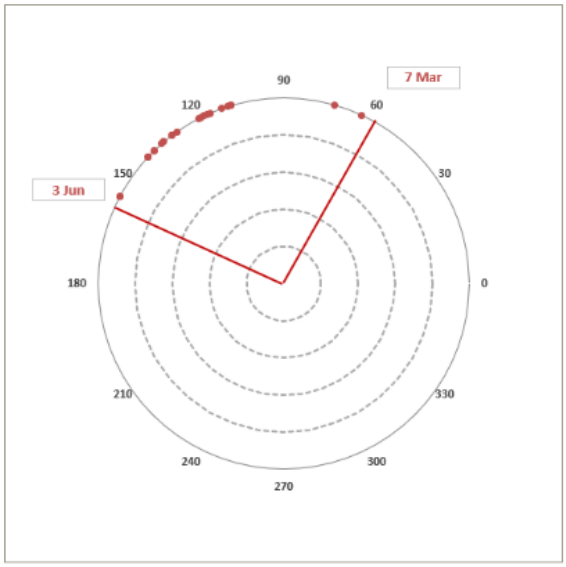


Figure 3.72 Unit Circle of the Station 2337

According to the results of the angular analysis conducted at station 2337, the average occurrence day of flood events during the year was determined as the 122nd calendar day, i.e. May 2. This date coincides with the middle of spring and reflects the period when the effects of snowmelt and spring precipitation in the region intensify.

The r value of the station indicates a very high level of homogeneity with 0.94. This value shows that the flood observation dates are concentrated in a certain time interval during the year and that there is a seasonal pattern far from randomness. It is also clearly seen in the polar graph that the data are mostly clustered in the same sector, therefore the floods are seasonally predictable.

These results reveal that the hydrological behavior in the region where station 2337 is located has a strong seasonal structure. Especially the snowmelt that starts with the increase in temperature in the spring months and the simultaneous precipitation can be considered as the main factors triggering these floods. This station offers a highly reliable analysis opportunity in terms of predicting and managing flood risk.

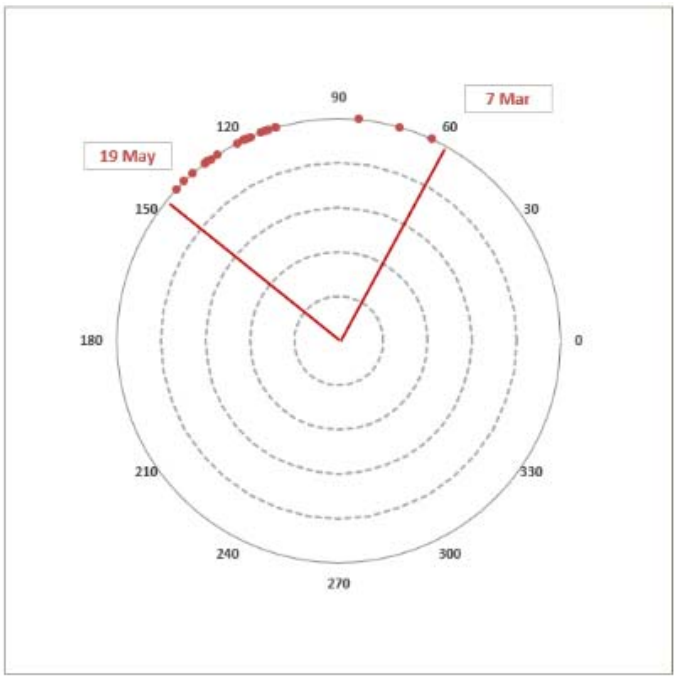


Figure 3.73 Unit Circle of the Station 2338

According to the angular seasonality analysis conducted at station 2338, the average occurrence day of flood events was determined as the 115th calendar day, namely April 25. This date coincides with the middle of the spring season and shows that snowmelt in the region as well as increased spring precipitation are effective in flood formation.

The r value of the station is 0.95, indicating a very high degree of homogeneity. This situation shows that floods occur in a very specific period during the year and are densely clustered in time. The collection of data in the polar graph in a narrow angular range reveals that the seasonality effect is very clear for this station.

According to the findings, floods were observed for station 2338 mainly in the spring months, especially in the late April and early May periods. Floods in this period may have occurred due to the combination of snowmelt and increased precipitation with the seasonal transition. The high r value supports the usability of the data obtained from this station in analyses and their reliability in flood timing estimates.

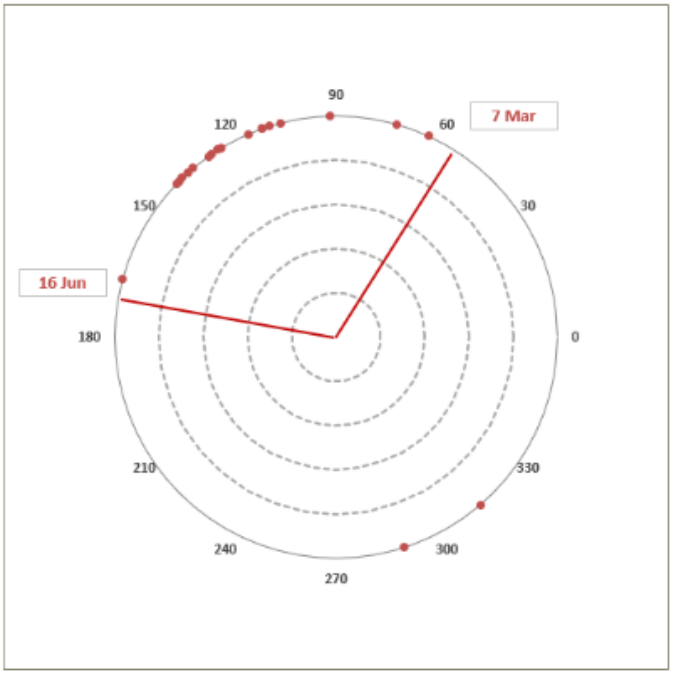


Figure 3.74 Unit Circle of the Station 2340

According to the angular seasonality analysis conducted at station 2340, the average time of occurrence of floods corresponds to the 118th calendar day, approximately April 28. This date coincides with the middle period of the spring season in the region and indicates a period when floods are caused by precipitation that increases with snowmelt.

The r value of this station was calculated as 0.73. This value shows that floods occur at a certain time period during the year, but have a lower temporal density than some other stations. In other words, it is understood that flood events are focused on a certain season, but some floods are observed at other times of the year.

When the polar graph is examined, it is seen that although the majority of the data is clustered within the spring period, a few data points are scattered at the end of summer and the beginning of autumn (for example, around days 300 and 330). This situation shows that although the floods observed at the station are mostly spring-related, floods can also occur at other times of the year. This diversity can perhaps be related to the microclimatic characteristics of the location of the station or the local precipitation regime.

As a result, it has been shown, supported by the r value, that floods at station number 2340 occur predominantly in spring seasonally, but low-intensity events can also occur at other times of the year. This necessitates careful evaluation in terms of seasonal prediction.

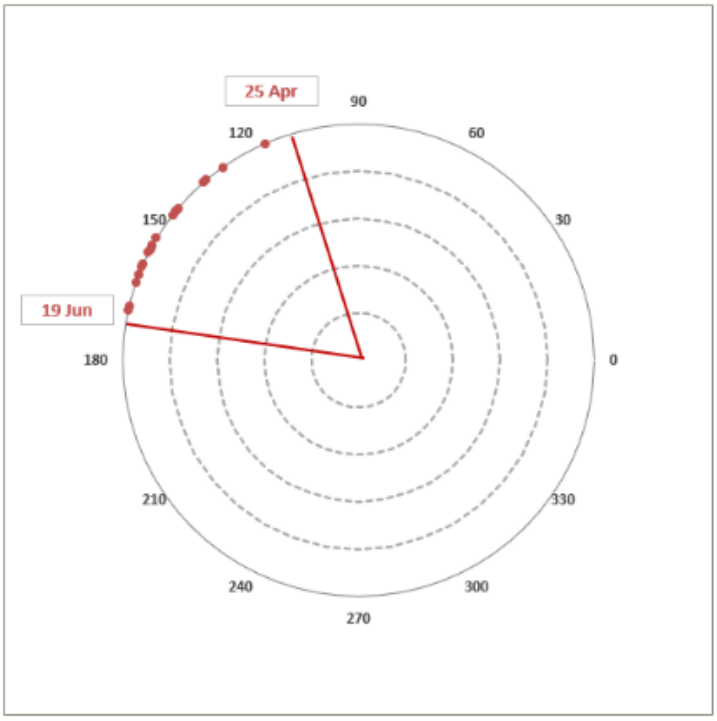


Figure 3.75 Unit Circle of the Station 2342

The angular seasonality analysis for station 2342 was conducted to determine the temporal distribution of flood events throughout the year. As a result of the calculations, the average occurrence day of floods for this station was found to be the 149th calendar day. This date corresponds approximately to May 29 and shows that floods are concentrated in the late spring period.

The r value for this station was calculated as a very high value of 0.97. This situation indicates that floods occur in a very specific and narrow time interval during the year. In other words, the timing of flood events at station 2342 is highly homogeneous and largely depends on a specific season.

When the polar graph is examined, it is clearly seen that almost all of the data are clustered in a time interval close to the end of spring and the beginning of summer. The concentration of the points on a small angular arc also supports that the floods at the station largely coincide with the same period.

This situation shows that the hydrological cycle in the region has a high potential to produce floods, especially during a period when snowmelt and

seasonal precipitation coincide. At the same time, this seasonal repetition can be evaluated in a way that allows flood management and early warning systems to focus on certain time periods.

As a result, the majority of the floods observed at station number 2342 occur on approximately the 150th day of the year, and this tendency has been shown to show a strong seasonality supported by the high r value (0.97). This homogeneity provides high reliability in seasonal estimating and planning studies at the station.

Unit Circles of Stations Together:

As seen in the Figure 80, the average day of annual maximum flood flows in the vast majority of the stations examined is concentrated between the 110th and 150th days of the year. This shows that flood events generally occur between the end of April and the end of May. In addition, the fact that the points are located close to the circle reveals that floods are concentrated in a certain period during the year and that seasonality is evident. This supports the fact that the hydrological regime in the region depends on snowmelt and spring precipitation.

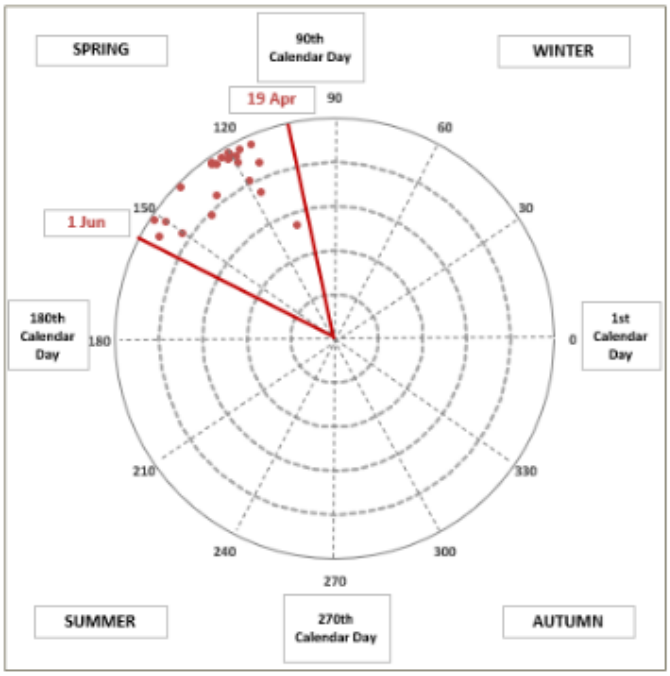


Figure 3.76 Flood Seasonality Representation on Unit Circle

3.8.3 Seasonal Similarities of Stations Calculated with Euclidean Distances

In this study, in order to analyze the similarities of 23 stations according to their seasonal characteristics, Euclidean distance was calculated for each station pair. In the calculations, two-dimensional Euclidean distance was used in accordance with the formula (Eq. 2.109) based on the $\bar{x}$ and $\bar{y}$ coordinates obtained from the angular seasonality analysis. These calculated distances are presented in Table 38 in the form of a symmetric matrix. These similarity values vary between 0 and 1, and when the value approaches 0, it is stated that the stations are more similar to each other seasonally, and when it approaches 1, the similarity decreases.

In this study, in order to analyze the similarities of 23 stations according to their seasonal characteristics, Euclidean distance was calculated for each station pair. In these calculations, distances were obtained in accordance with the two-dimensional Euclidean distance formula (Eq. 2.109) based on the $\bar{x}$ and $\bar{y}$ coordinates obtained from the angular seasonality analysis. The distance values for each station pair are presented in Table 38 in a symmetric matrix structure. These similarity values vary between 0 and 1, and the value approaching 0 indicates a high level of similarity in seasonal behavior between stations, while the value approaching 1 indicates a difference.

As a result of the calculations, the highest similarity value was found between stations 2331 and 2342 with 0.65; the lowest value was obtained between stations 2304 and 2316 with 0.01. Based on these values, a threshold value was determined to classify similarities between stations. This threshold, determined based on the arithmetic average of the maximum and minimum values, was accepted as 0.33. Accordingly, all station pairs with $d_{ij} < 0.33$ were considered “similar” in terms of season, while values above the threshold value were evaluated as pairs with significant differences.

This classification between station pairs is visualized in Table 39 with blue (+) and red (×) colors. Blue cells represent similar station pairs below the threshold value, and red cells represent different pairs above the threshold value.

According to the analysis results, a very low distance of 0.02 was obtained between stations 2305 and 2327. Similarly, low distances were recorded in station pairs such as 2316 and 2320; 2320 and 2305; 2322 and 2327; 2323 and 2325, and it was revealed that these stations were quite compatible in terms of seasonal flood behavior. On the other hand, it is seen that station 2331 has a high similarity value with many stations (for example, 0.65 with 2342 and 0.59 with 2330), and therefore it is different from other stations in terms of the timing of floods during the year.

In this context, in terms of seasonality, stations such as 2305, 2316, 2320, 2322, 2323, 2335 and 2337 are included in groups that exhibit high levels of seasonal coherence due to their low-distance connections with many other stations. On the other hand, stations 2331, 2342, 2321 and 2330 are separated by their relatively high distances and show more isolated flood behaviors.

As a result, thanks to this similarity study supported by angular seasonality analysis, clusters with high levels of similarity between the stations were determined. These clusters obtained will be an important reference for future regional flood risk estimates and hydrometeorological models and will contribute to the determination of homogeneous sub-regions.

Table 3.38 Similarity Values Calculated with Euclidean Distances of the Stations

d_s^{ij}	2302	2304	2305	2316	2320	2321	2322	2323	2325	2326	2327	2328	2329	2330	2331	2333	2334	2335	2336	2337	2338	2340
2304	0,27																					
2305	0,19	0,08																				
2316	0,27	0,01	0,08																			
2320	0,25	0,03	0,06	0,03																		
2321	0,24	0,49	0,42	0,49	0,47																	
2322	0,18	0,09	0,02	0,09	0,08	0,41																
2323	0,23	0,04	0,04	0,03	0,03	0,46	0,06															
2325	0,28	0,05	0,10	0,06	0,05	0,49	0,12	0,08														
2326	0,17	0,19	0,14	0,20	0,17	0,32	0,15	0,17	0,17													
2327	0,18	0,10	0,02	0,09	0,08	0,41	0,01	0,06	0,12	0,14												
2328	0,31	0,14	0,17	0,16	0,13	0,48	0,19	0,16	0,09	0,16	0,19											
2329	0,32	0,05	0,12	0,05	0,07	0,53	0,14	0,08	0,06	0,23	0,14	0,14										
2330	0,17	0,42	0,35	0,42	0,40	0,07	0,34	0,39	0,42	0,26	0,33	0,42	0,46									
2331	0,55	0,44	0,46	0,45	0,43	0,63	0,48	0,46	0,39	0,39	0,47	0,29	0,43	0,59								
2333	0,18	0,28	0,23	0,29	0,26	0,25	0,24	0,27	0,26	0,09	0,23	0,24	0,32	0,20	0,39							
2334	0,37	0,14	0,20	0,15	0,14	0,56	0,22	0,17	0,10	0,24	0,22	0,09	0,11	0,50	0,33	0,32						
2335	0,29	0,02	0,10	0,04	0,04	0,50	0,11	0,06	0,03	0,19	0,12	0,13	0,03	0,43	0,42	0,29	0,11					
2336	0,21	0,42	0,35	0,42	0,39	0,10	0,35	0,39	0,41	0,24	0,34	0,39	0,46	0,08	0,53	0,16	0,48	0,43				
2337	0,26	0,01	0,07	0,02	0,01	0,48	0,08	0,03	0,05	0,18	0,09	0,14	0,06	0,41	0,43	0,27	0,14	0,03	0,41			
2338	0,37	0,10	0,18	0,11	0,12	0,59	0,20	0,14	0,10	0,27	0,20	0,16	0,06	0,52	0,42	0,37	0,09	0,09	0,51	0,11		
2340	0,36	0,22	0,24	0,23	0,21	0,50	0,26	0,24	0,17	0,20	0,26	0,07	0,21	0,45	0,22	0,25	0,13	0,20	0,40	0,21	0,22	
2342	0,19	0,45	0,38	0,45	0,43	0,08	0,37	0,42	0,46	0,31	0,36	0,47	0,50	0,06	0,65	0,26	0,54	0,47	0,14	0,44	0,55	0,50

Table 3.39 Similarity Groups Calculated with Euclidean Distances of the Stations

d_s^{ij}	2302	2304	2305	2316	2320	2321	2322	2323	2325	2326	2327	2328	2329	2330	2331	2333	2334	2335	2336	2337	2338	2340
2304	+																					
2305	+	+																				
2316	+	+	+																			
2320	+	+	+	+																		
2321	+	x	x	x	x																	
2322	+	+	+	+	+	x																
2323	+	+	+	+	+	x	+															
2325	+	+	+	+	+	x	+	+														
2326	+	+	+	+	+	+	+	+	+													
2327	+	+	+	+	+	x	+	+	+	+												
2328	+	+	+	+	+	x	+	+	+	+	+											
2329	+	+	+	+	+	x	+	+	+	+	+	+										
2330	+	x	x	x	x	+	x	x	x	+	x	x	x									
2331	x	x	x	x	x	x	x	x	x	x	x	+	x	x								
2333	+	+	+	+	+	+	+	+	+	+	+	+	+	+	x							
2334	x	+	+	+	+	x	+	+	+	+	+	+	+	x	+	+						
2335	+	+	+	+	+	x	+	+	+	+	+	+	+	x	x	+	+					
2336	+	x	x	x	x	+	x	x	x	+	x	x	x	+	x	+	x	x				
2337	+	+	+	+	+	x	+	+	+	+	+	+	+	x	x	+	+	+	x			
2338	x	+	+	+	+	x	+	+	+	+	+	+	+	x	x	x	+	+	x	+		
2340	x	+	+	+	+	x	+	+	+	+	+	+	+	x	+	+	+	+	x	+	+	
2342	+	x	x	x	x	+	x	x	x	+	x	x	x	+	x	+	x	x	+	x	x	x

3.8.4 Result

In seasonality part, angular seasonality analysis was performed to evaluate the timing of flood events during the year. The calendar days when floods occurred were determined for each station, these values were converted to radians and the $\bar{x}$ and $\bar{y}$ coordinates on the unit circle were calculated. The averages of the obtained coordinates were taken to determine the average direction of floods during the year and the periods when they intensified.

In line with the angular data of the stations, the homogeneity degrees according to the distribution of flood days during the year were evaluated with the r coefficient and polar graphs were created for each station to visualize the temporal patterns. As a result of the analysis, it was determined that floods intensified in the spring months at many stations.

In addition, the Euclidean distance was calculated for each station pair to evaluate the seasonal similarity between stations. In these calculations,

$\bar{x}$ and $\bar{y}$ averages were taken as basis and similarity matrices between stations were created. Based on the determined threshold value of 0.33, stations with similar seasonal behaviors were grouped and visually classified. The findings revealed that some station groups showed high similarity in terms of

intra-year flood timing, while some stations were clearly separated. These results constitute an important basis for regional flood predictions and determination of homogeneous regions in hydrological models.

3.8.5 Regional Seasonality Analysis: Comparison of Flood Occurrence Times

The timing of flood events plays an important role in determining regional hydrometeorological similarities. In this context, the average flood calendar days (Flood Calendar Day) of the stations in each region were evaluated and their seasonal similarities were analyzed.

Region 1: In this region consisting of stations 2305, 2316, 2320, 2323, 2326, 2327, 2328, 2331, 2334, 2335, 2337 and 2338, floods were generally observed between the end of April and mid-May. For example, in 2334, the flood was observed on April 25, in 2338 on April 25 and in 2305 on May 6. This distribution indicates that the floods occurred in a relatively homogeneous time period throughout the region. However, relatively late floods are observed at a few stations (e.g. 2321: 1 June). This may indicate the effect of microclimatic or topographic differences within the region.

Region 2: In this region, which includes stations 2302, 2329, 2333, 2304, 2325 and 2330, floods are generally concentrated between the end of April and the beginning of May. The flood days for stations such as 2329 and 2330 are 28 April and 28 May, respectively, and for 2302 it is 18 May. This shows a slightly wider temporal distribution compared to Region 1. However, the floods still appear to be clustered in a certain period and show a seasonally consistent structure.

Region 3: When the flood days in this region are examined at stations 2321, 2342, 2336, 2340 and 2322, a similar homogeneity is observed. For 2340, 2322 and 2336, the floods occur between April 28 and May 29. This distribution reveals that the flood season of the region is concentrated in May and that the stations in the region show high seasonal coordination.

General Assessment: The analysis conducted for the three regions shows that the times of occurrence of floods during the year are similar between the regions, but there are also intra-regional variations. This situation shows that it is appropriate to base regional analyses on seasonality consistency between stations. In particular, while the stations in Region 3 have the highest time consistency, some stations in Region 1 (e.g. 2321) deviate from the general trend.

This assessment demonstrates the importance of zoning studies based on flood seasonality and will provide a solid basis for future hydraulic modelling.

CONCLUSION AND SUGGESTION

Statistical Values of the Stations:

In this section, basic statistical values for flood flows were calculated for 23 stations. The variation between stations was examined using mean values $\bar{x}$, standard deviations (σ) and coefficients of variation (C_{vx}). The highest coefficient of variation was observed at station 2335 with a value of 0.722, indicating that the flood magnitude fluctuated significantly at this station. The lowest value of C_{vx} was found at station 2321 with a value of 0.2353, indicating that this station had a relatively more stable flood regime. These analyses formed the basis of regional homogeneity studies and prepared the ground for the comparisons made in the following sections.

Goodness of Fit Tests for Distributions:

In this study, Chi-Square and Kolmogorov-Smirnov (K-S) goodness-of-fit tests were performed with EasyFit software for flood frequency analysis at 23 different stations. The test statistics calculated for each distribution were compared with the critical values and the goodness-of-fit of the distributions was evaluated according to the statistical significance level.

According to the K-S test results, all tested distributions were found to be acceptable at almost all stations. Only at station 2333, the normal distribution did not fit the data, while other distributions gave significant results at all stations including this station.

In terms of the Chi-Square test, the normal distribution did not fit the data at stations 2304, 2320 and 2333. On the other hand, Log-Normal (LN), Log-Normal 3 (LN3) and Generalized Extreme Value (GEV) distributions were successfully applied to all stations. The Gumbel distribution could not fit the data at stations 2304 and 2326.

According to the findings, GEV, LN and LN3 distributions stand out as more reliable and generally accepted models in flood analyses in the region. Especially the fact that GEV distribution gives successful results at every station

supports the preference of this distribution. The fact that the normal distribution cannot fit the data at some stations requires the careful use of this distribution.

Parameters of Probability Distributions for Station Data:

For each station, parameters belonging to Normal, Log-Normal (LN), Three Parameter Log-Normal (LN3), Gumbel and Generalized Extreme Value (GEV) distributions were estimated. L-moment method was used in estimating distribution parameters. While LN3 distribution is generally suitable for data with high skewness, similar parameter estimates were obtained with symmetric distributions (e.g. Normal) at some stations. This diversity reflects the difference in hydrometeorological conditions of the stations in the region.

Estimation of Flood Flow Rates at Various Return Intervals:

For all stations, 2, 5, 10, 25, 50, 100 and 500-year flood magnitudes were estimated. For example, for station No. 2302, the flood value corresponding to a 100-year return period was found to be approximately 342.8 m^3/s in the LN3 distribution and 92.03 m^3/s in the GEV distribution. This difference clearly reveals the effect of the distribution selection on the estimation results. While the LN3 distribution estimated the extreme values larger, the GEV and Gumbel distributions gave more conservative results. Therefore, it is of great importance to determine which distribution is used in critical infrastructure designs.

Trend Analysis:

In the study, the Mann-Kendall test was applied to 23 hydrological stations on annual maximum flood flow rates; only four stations (2328, 2330, 2340, 2342) had statistically significant trends ($|Z|>1.96$). In order to evaluate the trend directions and explanatory power at these stations, regression analyses were performed on dimensionless flood flow rate $(Q/\bar{Q})$ and R^2 values were calculated.

A negative trend was observed at all stations, and R^2 values ranged between 0.17 and 0.24. This shows that although the trends are statistically significant, their linear explanatory power is limited.

As a result, it was concluded that flood trends in the region are generally weak and non-linear, and therefore it would be appropriate to evaluate them by

taking into account longer time periods and climatic factors. These results should be taken as reference only.

Regional Flood Analysis:

Initially it was assumed that all stations were a single homogeneous region, but this hypothesis was rejected according to the Wiltshire homogeneity test results. Then, two regions were defined according to the coefficients of variation and these regions were tested again. It was determined that these two regions were not homogeneous. It was observed that the stations with high coefficients of variation were excluded in the second region, while the stations with low coefficients disrupted the homogeneity in the first region. In line with this observation, the stations with low variation were grouped as a new third region. All three regions created met the homogeneity condition. This regional grouping is a critical step in increasing statistical reliability in flood predictions.

Seasonality Analysis:

The distribution of flood events for each station during the year is expressed in angular values. The average flood day is concentrated around May 18, and flood intensity is observed between the end of spring and the beginning of summer in most of the stations. The highest homogeneity is detected at station 2342 with $r = 0.97$.

Flood dates for each station are shown on the unit circle. These graphical representations have been an important tool for visualizing the temporal distribution of floods throughout the year. Clustered points on the circle indicate homogeneity of flood timing, while scattered structures indicate irregularity.

The similarity between stations according to angular parameters was calculated with Euclidean distance. The lowest distance was calculated as 0.01 (2304–2316) and the highest distance was calculated as 0.65 (2331–2342). The threshold value was taken as 0.33 and accordingly the stations were divided into three similarity clusters. This clustering can be used in timing-based flood risk management.

REFERENCES

Akyer, M.K. (1995). *Regional flood frequency analysis of the Büyük Menderes river basin.* DEU Graduate School of Natural and Applied Sciences, M.Sc. Thesis in Civil Engineering. (Advisor: E. Benzeden).

Bayazıt, M. (1996). *İnşaat Mühendisliğinde Olasılık Yöntemleri*. İTÜ İnşaat Fakültesi Matbaası.

Bayazıt, M. (1998). *Hidrolojik Modeller*. İTÜ İnşaat Fakültesi Matbaası.

Bayazıt, M., Cığızoğlu, H.K. and Önöz, B. (2002). Türkiye akarsularında trend analizi. *Türkiye Mühendislik Haberleri*, 420-422, 8-10.

Bayazıt, M. and Önöz, B. (2004). Sampling variances of regional flood quantiles affected by intersite correlation. *Journal of Hydrology*, 291, 42-51.

Bayazıt, M. and Yeğen Oğuz, E.B. (2005). *Mühendisler için İstatistik*. Birsen Yayınevi.

Bayazıt, M. and Önöz, B. (2007). To prewhiten or not to prewhiten in trend analysis? *Hydrological Sciences Journal*, 52 (4), 611-624.

Bayazıt, M. and Önöz, B. (2008). *Taşkın ve Kuraklık Hidrolojisi*. Nobel Yayın Dağıtım.

Beard, L.R. (1974). Flood flow frequency techniques. *Center for Research in Water Resources*, University of Texas.

Benson, M.A. (1968). Uniform flood-frequency estimating methods for federal agencies. *Water Resources Research*, 4(5), 891-908.

Burn, D.H. (2000). The formation of groups for regional flood frequency analysis. *Journal of Hydrology*, 229, 138-147.

Chowdhury, J.U., Stedinger, J.R. and Lu, L.H. (1991). Goodness-of-fit tests for regional generalized extreme value flood distributions. *Water Resources Research*, 27 (7), 1765-1776.

Cicioni, G., Guiliano, G. and Spaziani, F.M. (1973). Best fitting of Probability functions to a set of data for flood studies. In: *Floods and Droughts*, Proc. of the Second Int. Symp. in Hydrology, Water Resources Publications, Fort Collins, pp. 304-314.

Crutcher, H.L. (1975). A note on the use of the Kolmogorov--Smirnov test. *Journal of Applied Meteorology*, 14, 1600-1601.

Dinpashoh, Y., Jhajharia, D., Fakheri-Fard, A., Singh, P. and Kahya, E. (2011). Trends in reference crop evapotranspiration over Iran. *Journal of Hydrology*, 399, 422-433.

Gedikli, D. (1994). *Analysis of Floods in the Scope of GAP*. DEU Graduate School of Natural and Applied Sciences, M.Sc. Thesis in Civil Engineering. (Advisor: E. Benzeden).

Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., Kim, H. and Kanae, S. (2013). Global flood risk under climate change. *Nature Climate Change*, 3, 816-821. \url{https://doi.org/10.1038/nclimate1911}

Hirsch, R.M. and Slack, J.R. (1984). A non-parametric trend test for seasonal data with serial dependence. *Water Resources Research*, 20, 727-732.

Intergovernmental Panel on Climate Change (IPCC). (2021). *Climate Change 2021: The Physical Science Basis*. Contribution of Working Group I to the Sixth Assessment Report of the IPCC, Cambridge University Press. \url{https://doi.org/10.1017/9781009157896}

Kahya, E. and Kalaycı, S. (2002). Trend analysis of streamflow in Turkey. *Journal of Hydrology*, 289(1-4), 128-144.

Kahya, E., Kalaycı, S. and Karabörk, M.C. (2004). Temporal and spatial variability of precipitation over Turkey. *Hydrological Processes*, 18 (5), 847-866.

Lettenmaier, D.P. and Potter, K.W. (1985). Testing flood frequency estimation methods using a regional flood generation model. *Water Resources Research*, 21(2), 1903-1914.

McCuen, R.H. and Beighley, R.E. (2003). Seasonal flow frequency analysis. *Journal of Hydrology*, 279, 43-56.

Sudheer, K.P. and Nayak, P.C. (2013). Successive-station monthly streamflow prediction using neuro-wavelet technique. *Environmental Modelling & Software*, 41, 56-65.

Saf, B. (1995). *Regional flood frequency analysis of west Mediterranean river basins*. DEU Graduate School of Natural and Applied Sciences, M.Sc. Thesis in Civil Engineering. (Advisor: E. Benzeden).

Şen, Z. (2009). *Taşkın Afet ve Modern Hesaplama Yöntemleri*. Su Vakfı Yayınları.

Şenocak, S. ve Taşçı, S. (2020). Çoruh Havzası Taban Akışının İngiliz Hidroloji Enstitüsü Yuvarlatılmış Minimumlar Yöntemi ile Belirlenmesi. *Artvin Çoruh Üniversitesi Orman Fakültesi Dergisi*, 21(1), 73-82.

Van Belle, G. and Hughes, J.P. (1984). Non-parametric test for trend in water quality. *Water Resources Research*, 20, 127-136.

Vogel, R.W. and McMartin, D.E. (1991). Probability plot goodness-of-fit and skewness estimation procedures for the Pearson type 3 distribution. *Water Resources Research*, 27(12), 3149-3158.

Vogel, R.M., McMahon, T.A. and Chiew, F.H.S. (1993). Flood flow frequency model selection in Australia. *Journal of Hydrology*, 146, 421-449.

Wiltshire, S.W. (1986). Identification of homogeneous regions for flood frequency analysis. *Journal of Hydrology*, 84, 287-302.

Yıldız, M., Saraç, M., Malkoç, Y. and Uçar, İ. (2004). Türkiye akarsularındaki akımların trendleri ve bu trendlerin hidroelektrik enerji üretimine etkileri. *IV. Ulusal Hidroloji Kongresi*, İTÜ, 23--25 Haziran, pp. 59-70.

APPENDICES

Table A.1 Δ_a Values of K-S Test

N	0,2	0,1	0,05	0,01
5	0,45	0,51	0,56	0,57
10	0,32	0,37	0,41	0,49
15	0,27	0,30	0,34	0,4
20	0,24	0,26	0,29	0,36
25	0,21	0,24	0,27	0,32
30	0,20	0,22	0,24	0,29
35	0,18	0,21	0,23	0,27
40	0,17	0,19	0,21	0,25
45	0,16	0,18	0,2	0,24
50	0,15	0,17	0,19	0,23
> 50	$\frac{1,07}{\sqrt{n}}$	$\frac{1,22}{\sqrt{n}}$	$\frac{1,36}{\sqrt{n}}$	$\frac{1,63}{\sqrt{n}}$

Table A.2 Δ_a values of K-S test for Normal and Gumbel Distribution (Crutcher, 1975)

Distribution	n	α				
		0,2	0,15	0,1	0,05	0,01
Normal	25	0,142	0,147	0,158	0,173	0,2
	30	0,131	0,136	0,144	0,161	0,187
	> 30	$\frac{0,736}{\sqrt{n}}$	$\frac{0,768}{\sqrt{n}}$	$\frac{0,805}{\sqrt{n}}$	$\frac{0,886}{\sqrt{n}}$	$\frac{1,031}{\sqrt{n}}$
Gumbel	25	0,152	0,157	0,17	0,183	0,209
	30	0,134	0,14	0,149	0,164	0,15
	> 30	$\frac{0,738}{\sqrt{n}}$	$\frac{0,769}{\sqrt{n}}$	$\frac{0,816}{\sqrt{n}}$	$\frac{0,888}{\sqrt{n}}$	$\frac{1,041}{\sqrt{n}}$

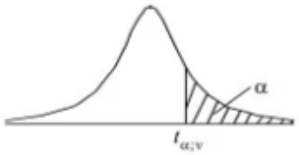
Table A.3 Z Table

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
+0	.50000	.50399	.50798	.51197	.51595	.51994	.52392	.52790	.53188	.53586
+0.1	.53983	.54380	.54776	.55172	.55567	.55966	.56360	.56749	.57142	.57535
+0.2	.57926	.58317	.58706	.59095	.59483	.59871	.60257	.60642	.61026	.61409
+0.3	.61791	.62172	.62552	.62930	.63307	.63683	.64058	.64431	.64803	.65173
+0.4	.65542	.65910	.66276	.66640	.67003	.67364	.67724	.68082	.68439	.68793
+0.5	.69146	.69497	.69847	.70194	.70540	.70884	.71226	.71566	.71904	.72240
+0.6	.72575	.72907	.73237	.73565	.73891	.74215	.74537	.74857	.75175	.75490
+0.7	.75804	.76115	.76424	.76730	.77035	.77337	.77637	.77935	.78230	.78524
+0.8	.78814	.79103	.79389	.79673	.79955	.80234	.80511	.80785	.81057	.81327
+0.9	.81594	.81859	.82121	.82381	.82639	.82894	.83147	.83398	.83646	.83891
+1	.84134	.84375	.84614	.84849	.85083	.85314	.85543	.85769	.85993	.86214
+1.1	.86433	.86650	.86864	.87076	.87286	.87493	.87698	.87900	.88100	.88298
+1.2	.88493	.88686	.88877	.89065	.89251	.89435	.89617	.89796	.89973	.90147
+1.3	.90320	.90490	.90658	.90824	.90988	.91149	.91308	.91466	.91621	.91774
+1.4	.91924	.92073	.92220	.92364	.92507	.92647	.92785	.92922	.93056	.93189
+1.5	.93319	.93448	.93574	.93699	.93822	.93943	.94062	.94179	.94295	.94408
+1.6	.94520	.94630	.94738	.94845	.94950	.95053	.95154	.95254	.95352	.95449
+1.7	.95543	.95637	.95728	.95818	.95907	.95994	.96080	.96164	.96246	.96327
+1.8	.96407	.96485	.96562	.96638	.96712	.96784	.96856	.96926	.96995	.97062
+1.9	.97128	.97193	.97257	.97320	.97381	.97441	.97500	.97558	.97615	.97670
+2	.97725	.97778	.97831	.97882	.97932	.97982	.98030	.98077	.98124	.98169
+2.1	.98214	.98257	.98300	.98341	.98382	.98422	.98461	.98500	.98537	.98574
+2.2	.98610	.98645	.98679	.98713	.98745	.98778	.98809	.98840	.98870	.98899
+2.3	.98928	.98956	.98983	.99010	.99036	.99061	.99086	.99111	.99134	.99158
+2.4	.99180	.99202	.99224	.99245	.99266	.99286	.99305	.99324	.99343	.99361
+2.5	.99379	.99396	.99413	.99430	.99446	.99461	.99477	.99492	.99506	.99520
+2.6	.99534	.99547	.99560	.99573	.99585	.99598	.99609	.99621	.99632	.99643
+2.7	.99653	.99664	.99674	.99683	.99693	.99702	.99711	.99720	.99728	.99736
+2.8	.99744	.99752	.99760	.99767	.99774	.99781	.99788	.99795	.99801	.99807
+2.9	.99813	.99819	.99825	.99831	.99836	.99841	.99846	.99851	.99856	.99861
+3	.99865	.99869	.99874	.99878	.99882	.99886	.99889	.99893	.99896	.99900
+3.1	.99903	.99906	.99910	.99913	.99916	.99918	.99921	.99924	.99926	.99929
+3.2	.99931	.99934	.99936	.99938	.99940	.99942	.99944	.99946	.99948	.99950
+3.3	.99952	.99953	.99955	.99957	.99958	.99960	.99961	.99962	.99964	.99965
+3.4	.99966	.99968	.99969	.99970	.99971	.99972	.99973	.99974	.99975	.99976
+3.5	.99977	.99978	.99978	.99979	.99980	.99981	.99981	.99982	.99983	.99983
+3.6	.99984	.99985	.99985	.99986	.99986	.99987	.99987	.99988	.99988	.99989
+3.7	.99989	.99990	.99990	.99990	.99991	.99991	.99992	.99992	.99992	.99992
+3.8	.99993	.99993	.99993	.99994	.99994	.99994	.99994	.99995	.99995	.99995
+3.9	.99995	.99995	.99996	.99996	.99996	.99996	.99996	.99996	.99997	.99997
+4	.99997	.99997	.99997	.99997	.99997	.99997	.99998	.99998	.99998	.99998

Table A.4 t-Students Table

Table of the Student's *t*-distribution

The table gives the values of $t_{\alpha;\nu}$ where
$\Pr(T_\nu > t_{\alpha,\nu}) = \alpha$, with ν degrees of freedom



ν \ α	0.1	0.05	0.025	0.01	0.005	0.001	0.0005
1	3.078	6.314	12.076	31.821	63.657	318.310	636.620
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.767
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.090	3.291

CURRICULUM VITAE