Muthanna Journal of Engineering and Technology, Vol. (13), Issue (1), Year (2025)

Muthanna Journal of Engineering and Technology

Website: https://muthjet.mu.edu.iq/

Submitted 11 January 2025, Accepted 20 January 2025, Published online 24 January 2025

Best fitting probability distributions of monthly rainfall extremes in Nasiriyah city, southern Iraq

Abaas J. Ismaeel[a], Ahmed A. Dakheel[b] and Basim M. Al-Zaidi[c]

[a]*Civil Engineering Department, College of Engineering, University of Thi-Qar, Thi-Qar, Iraq, E-mail:* a.ismaeel@utq.edu.iq
[b]*University of Thi-Qar, Thi-Qar, Iraq, E-mail:* msc_ahmed@utq.edu.iq
[c]*Civil Engineering Department, College of Engineering, University of Thi-Qar, Thi-Qar, Iraq, E-mail:* basim.m.al-zaidi@utq.edu.iq
**Corresponding author E-mail:* msc_ahmed@utq.edu.iq

DOI:10.52113/3/eng/mjet/2025-13-01-/01-09

Abstract

Analysis of rainfall data is important in the design and planning of water projects in cities. Therefore, in this research, rainfall data recorded at the Nasiriyah station located in the center of Thi-Qar Governorate, southern Iraq, was used for 80 years for the period (1940-2020) to determine the best probability distribution fits this data. All tests and statistical analyzes were carried out using the (HYFRAN-PLUS version 1.2) software, and the method of maximum likelihood was applied to obtain the coefficients of theoretical distributions. Eight distributions were tested: GEV (Generalized Extreme Value), Gumbel, Weibull, Normal, Lognormal, Gamma, Pearson type 3, and Log-Pearson type 3, and adequacy test was conducted by (Chi-Square) test for these distributions, the results showed the GEV, Lognormal, Pearson type 3, and Log-Pearson type 3 distributions are suitable, for describing extreme monthly rainfall in this study area.

Keywords: *Best fitting; Extremes monthly rainfall; Nasiriyah city; Probability analysis; Weibull distribution.*

1. Introduction

The process of designing, building and implementing some projects, such as urban drainage systems, dams, water resource management, reservoirs, and flood risk prevention requires extensive knowledge of extreme events for large return periods. In many cases, return periods exceed the periods of available records, and here it is difficult to extract directly from the previously recorded data, and as an engineering practice, the estimation of peak floods or extreme rainfalls, is according to statistical frequency analysis of the maximum stream flow or the maximum rainfall. Then a suitable distribution can be used in order to estimate events corresponding to return periods less than or greater than the recorded data. Accurate estimation of heavy rainfall helps to reduce the damage that can be caused by floods and storms and also helps to increase the accuracy in the design of hydraulic structures [1]. Heavy rainfall with extreme values has significant impacts and an important role for water resources management. Therefore, the probabilistic analyzes and developments in the statistical theory of extreme values are applied to improve hydrological applications because hydrological processes are often governed by chance and in order to extract information and data and obtain the best mathematical description of these operations must use theories of statistics and probability [2]. Many probability distributions are usually used to calculate rainfall for a particular frequency, such as GEV, lognormal, Gumbel, Weibull, Pearson type III, log-Pearson type III, Gamma, etc., because choosing and relying on a single model is still one of the main problems in engineering practices due to the lack of general agreement on a specific model and its description as being the best for analyzing the frequency of intense rainfall, and in this case a certain range of different probability distributions must be compared. In spite of the great development in probability theories that took place over a long period of time, statistical modeling by models still needs more research [3]. There are many studies in this field presented by many researchers around the world that have reached a set of different distributions of extreme values of rainfall. Alam et al. [4] studied a set of probability distributions for the monthly maximum rainfall data for 35 locations in Bangladesh for a period of 30 years (1984–2013). The results showed that the distributions of both Log-Pearson type 3 and Pearson type 3 are more suitable than others, and the values of return periods for 10 years, 25 years, 50 years and 100 years were also calculated during this study for the data of the monthly maximum values of rainfall and for all sites. Ahmadpari et al. [5] analyzed the monthly rainfall data to get the best probability distributions through the data collected from 6 stations in west Azerbaijan province in Iran, these stations are Pole Miandoab Zarrineh river, Sariqamish, Shahid Kazemi dam, Qareh Papaq, Nezam Abad and Shahin Dezh during 30 years for the period (1989-2018). Data analysis was carried out using the program (SMADA) which depends on the moment's method. Through the results, the probabilistic distributions of the monthly rainfall data for these stations were the Sariqamish and Pole Miandoab Zarrineh river stations with the three parameters log normal distribution. Shahid Kazemi dam and Qareh Papaq stations with a Pearson distribution of the type III, Shahin Dezh station with the Gumble Type I distribution, and Nezam Abad station with the two parameters log normal distribution. Thanh [6] calculated 20 indices of daily average precipitation for a mountainous region in southern Vietnam to get the best probability distributions for this region. Data were collected through four stations for a period of 30 years, and the best appropriate distribution was calculated for each station. The results concluded that the Johnson distribution is the most suitable for the rainfall data over this region. Ximenes et al. [7] analyze the suitability of a set of probability distributions, such as log normal, Gamma, Weibull, Gumbel, Generalized Pareto, and normal for monthly rainfall data, in northeastern Brazil. Data were collected from 293 stations for the period (1988-2017). The results showed that the distributions with two parameters are suitable for describing the studied data, and the distributions of both Gamma and Weibull are the best in general. Barkotulla et al. [8] estimated rainfall status for different return periods in Boalia, Rajshahi, Bangladesh by estimating one day, two days and up to seven consecutive days annual extremes rainfall. Three probability distributions Gamma, log normal, and normal were chosen to determine the best fit for the studied data. The results summarized that the log normal distribution is the best for one day, two days, and seven consecutive days annual extremes rainfall for this region. Bora et al. [9] conducted frequency analyzes of consecutive days maximum rainfall values, data were collected for 23 years through (Indian Meteorological Department), for Jorhat station. A set of probability distributions were used to obtain the most appropriate distribution of rainfall in different return periods, and was concluded that the EV1 distribution is the best among the other distributions, and it was tested with a Chi-square value. Olofintoye et al. [10] tested a set of probability distributions to obtain the characteristics of peak daily rainfall, in southern Nigeria. Daily rainfall data were collected from ten stations for 54-years in order to perform frequency analysis. The results concluded that Pearson type 3 distribution is the best by occupying (50%) of the stations. Kwaku and Duke [11] showed that the log normal distribution is the best probability distribution for one, two, and five consecutive days of the maximum rainfall values for the Accra region located in Ghana. Sukrutha et al. [12] assessmented the best probability distributions of monthly rainfall data for 100-years (1901-2002), for twenty Indian cities. The probability distributions (Lognormal, Normal, Weibull, Fisher, Beta, GEV, Gumbel, Inverse Gaussian, Gamma) were evaluated for all selected cities, and the results showed that both the inverse Gaussian distribution and the generalized extreme value distribution are the best and fit the observed data. Heavy rainfall with extreme values has significant effects and an important role in the management of water resources and in the design of hydraulic structures and sewer networks etc., as previously mentioned, so the topic is still of interest the researchers in many countries of the world. The present work aims to obtain the best models of the probability distributions of the monthly extreme values of rainfall for the Nasiriyah city, southern Iraq for a long period of time from 1940 to 2020 to give a complete perception of the nature of the probability distributions of the extreme values of rainfall. This study is useful and can be used to develop policies and plans in the short and long term to reduce the damage caused by heavy rainfall to many projects under implementation and design, in addition, it is possible to use the resulting best probability distribution to analyze other extreme phenomena for the Nasiriyah city.

2. Description and location of the study area

Nasiriyah is the fourth largest Iraqi city and is the center of Dhi-Qar Governorate. It is located in southern Iraq, about 360 km southeast of the capital Baghdad. It is bordered by Wasit Governorate to the north, Basra to the south, Maysan to the east, and Muthanna and Qadisiyyah to the west. It is located along the banks of the Euphrates River, which divides it into two halves, the first half is called Al-Shamiya and the other half is called Al-Jazeera. Among its famous archaeological landmarks are the ruins of the ancient city of Ur. The area of the Nasiriyah city about (1766) km^2 (see Figure 1). Nasiriyah city is part of the alluvial plain area, which is characterized by flatness in most of its sections, it is also characterized by the presence of a number of depressions and valleys in its western sections, and also distinguished by the marshes in the lowlands, the most famous of which is the Hammar marsh. Nasiriyah located on Latitude ($31^{\circ}\ 2^{'}\ 38^{''}$) North and Longitude ($46^{\circ}\ 15^{'}\ 27^{''}$) East, and it has elevation (7 m) above sea level.

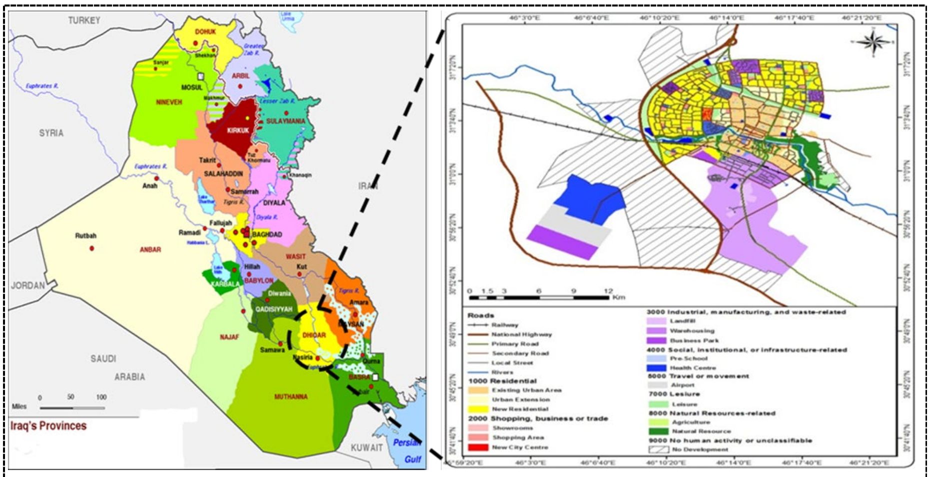

Fig. 1: Location of the study area.

3. Climate conditions of study area

Nasiriyah city is located in southern Iraq, in general southern Iraq is characterized by high temperatures in most months of the year. The summer in Nasiriyah city is characterized by long, hot, dry, windy, and clear while the winter is cold, dry, and mostly clear. With regard to temperatures in Nasiriyah city, and through the temperature data recorded in Nasiriyah station for the period (1941-2020) it was found that the highest annual average temperature was in (2010) and reached (27.5 °C), while the lowest annual average of temperatures was in (1948) and reached (22.9 °C). The highest monthly average was recorded in July (2017), when temperatures reached (40.9 °C) and for the same year and for the same month (July), the highest average for maximum temperature was recorded (49.0 °C), while the lowest monthly average was recorded in January (1964) where the temperature reached (6.2 °C) and the lowest average for minimum temperature was (-0.7 °C) for the same year and for the same month (January) [13]. The annual average of relative humidity recorded in Nasiriyah station reaches for the same period mentioned previously (42.11%), as the air is generally characterized by dryness, and the highest annual average of relative humidity (52.42%) was recorded in a year (1954), while the lowest annual average of relative humidity reached (30.51%) in a year (2017). The prevailing wind speed in Nasiriyah city is irregular, and the average wind speed in Nasiriyah station is 4 m/s, and in general, the highest monthly average recorded in the station was in July of the year (1992) when it reached (9.3) m/s, while the lowest monthly average was in November in the year (2000), when it reached (1.3) m/s. The highest annual average was (5.6) m/s in the years (1989, 1990, 1991), while the lowest annual average was (2.5) m/s in the year (1941) [13,14]. Rainfall in Nasiriyah city is usually fluctuating, and rainfall falls mostly in late winter and in the spring, and rainfall is rare in general from June to October for each year, sometimes large amounts of rainfall fall in November and December. The highest monthly average rainfall was in November in the year (2018) reached (132.1) mm. The years (1964 and 2017) are considered the least years in terms of rainfall, with an annual average of (2.3) mm, while (2006) is the year with the most rainfall, where the annual average is (20.5) mm [13].

4. Description of the rainfall data collected

Extreme changes in climatic phenomena and weather are considered a major source of dangers for all human societies. Defining, diagnosing, and classifying extremist events is very difficult because there is no universal definition of “extreme events” so there is a great and urgent need for more research in such areas and these events [15]-[16]. Usually, extreme events are referred to as the smallest or largest value of meteorological variables that occur among the observed data [17]. All rainfall data were obtained from the hydrological survey report for Iraq [18], and from the records collected by the metrological station in the Nasiriyah city. The rainfall data values used are shown in the Table 1, and represent the highest monthly rainfall during a particular year provided the data, and when there is no precipitation, it is indicated by zero (0.0), while a missing recorded data is indicated by (M). The metrology station is located at latitude (31° 01') and longitude (46° 14'), and has a height (3 m) above sea level [18].

5. Probability distributions for rainfall data

There are many probability distributions used to describe the observed monthly or annual rainfall data and the analysis of these data depends largely on their distribution pattern, and creating a probability distribution for rainfall data is an important topic in meteorology [19]. In this paper, a set of probability distributions such as, GEV, Gumbel, Weibull, Normal, Lognormal, Gamma, Pearson type 3, and Log-Pearson type 3 were used to obtain the best models for the probability distributions of the precipitation data for the Nasiriyah city, southern Iraq.

5.1. GEV distribution

Generalized Extreme Value (GEV), is a family of probability distributions developed within the extreme value theory. From which it is possible to deduce the probabilities of extreme or very rare events. This distribution unites distributions (Frechet, Gumbel, and Weibull) in one family, and these distributions are also known as maximum value distributions of type I, II, and III. The distribution parameters (GEV) are specified using the shape, location, and scale parameter [20]. This distribution can be expressed by the Equation (1) below:

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

Where: $u \in R$, $\alpha > 0$, and $k \in R$ are the location, scale, and shape parameters, respectively. If $k = 0$, the (GEV) distribution coincides with the (Gumbel) distribution, if $k > 0$ it is the reversed (Weibull) distribution, and if $k < 0$ it is the (Frechet) distribution.

Table 1: Monthly meteorological data for rainfall in (mm) of Nasiriyah station for period (1940-2020)

Year	Jan.	Feb.	Mar.	Apr.	May.	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
1940	29.5	27.2	2.9	7.7	0.2	0.0	0.0	0.0	0.0	0.9	3.1	1.2
1941	0.0	31.4	1.0	19.6	0.0	0.0	0.0	0.0	0.0	0.0	0.8	9.5
1942	1.0	16.0	19.2	12.2	0.0	0.0	0.0	0.0	0.0	2.0	3.9	33.5
1943	3.4	2.1	53.3	5.6	0.0	0.0	0.0	0.0	0.0	0.7	0.3	12.5
1944	7.7	0.0	5.9	1.0	5.4	0.0	0.0	0.0	0.0	0.0	60.4	13.6
1945	37.0	16.6	19.6	3.9	46.8	0.0	0.0	0.0	0.0	0.0	31.3	26.6
1946	33.3	3.1	84.7	66.0	3.3	0.0	0.0	0.0	0.0	0.0	5.9	25.8
1947	30.4	5.0	6.8	0.0	7.8	0.0	0.0	0.0	0.0	3.0	34.4	5.9
1948	25.4	8.9	25.9	18.9	2.3	0.0	0.0	0.0	0.0	0.0	1.4	68.9
1949	25.8	2.1	25.0	0.2	7.3	0.0	0.0	0.0	0.0	0.0	0.0	35.4
1950	10.3	4.2	11.1	5.1	6.0	0.0	0.0	0.0	0.0	0.001	53.9	32.6
1951	32.9	13.4	17.2	2.3	0.0	0.0	0.0	0.0	0.0	0.001	5.1	25.1
1952	11.4	12.2	3.7	17.2	0.001	0.0	0.0	0.0	0.0	0.0	59.8	49.4
1953	15.9	19.4	30.6	51.7	M	0.0	0.0	0.0	0.0	10.8	32.4	21.2
1954	2.4	23.6	27.3	5.0	0.7	0.0	0.0	0.0	0.0	8.5	54.7	55.5
1955	14.6	0.4	12.7	7.9	7.5	0.0	0.0	0.0	0.0	0.001	12.3	23.0
1956	12.4	28.1	10.1	4.4	0.0	0.0	0.0	0.0	0.001	0.001	0.001	35.7
1957	19.7	48.5	23.8	59.0	13.8	0.0	0.0	0.0	0.0	0.9	35.1	5.3
1958	39.0	0.001	0.001	0.2	0.7	0.0	0.0	0.0	0.0	0.001	6.4	28.2
1959	17.8	2.6	4.3	2.3	6.1	0.0	0.0	0.0	0.0	0.3	15.3	5.0
1960	9.7	18.3	16.1	8.1	0.001	0.0	0.0	0.0	0.0	0.0	2.9	0.4
1961	41.2	5.4	14.0	13.9	0.001	0.0	0.0	0.0	0.0	0.0	22.6	9.1
1962	7.4	8.1	18.3	40.1	1.2	0.0	0.0	0.0	0.0	0.001	5.3	19.2
1963	7.0	27.9	2.8	49.8	28.5	0.0	0.0	0.0	0.0	0.001	2.8	32.5
1964	0.0	22.4	0.001	0.001	0.001	0.0	0.0	0.0	0.0	0.0	2.5	2.9
1965	50.5	0.001	1.2	4.0	0.001	0.0	0.0	0.0	0.0	26.0	2.8	0.001
1966	4.5	51.3	15.8	6.3	0.001	0.0	0.0	0.0	0.0	4.5	0.001	3.1
1967	23.4	14.6	0.001	0.001	25.6	0.0	0.0	0.0	0.0	7.1	27.2	10.6
1968	0.8	1.8	0.001	52.1	20.2	0.001	0.0	0.0	0.0	1.3	19.4	5.1
1969	64.6	10.1	12.0	39.5	30.5	0.0	0.0	0.0	0.0	0.001	2.4	3.5
1970	36.6	5.5	10.5	2.0	0.0	0.0	0.0	0.0	0.0	0.001	1.5	13.0
1971	7.3	5.8	10.7	33.4	8.0	0.0	0.0	0.0	0.0	0.1	21.0	8.7
1972	66.0	3.4	17.1	22.1	0.6	0.001	0.0	0.001	0.001	0.001	13.8	32.9
1973	0.9	51.7	2.6	0.6	0.6	0.0	0.0	0.0	0.0	0.0	0.001	31.2
1974	20.5	41.3	71.5	5.2	1.7	0.0	0.0	0.0	0.0	0.5	1.7	23.6
1975	76.2	26.3	0.001	9.5	26.2	2.5	0.0	0.0	0.0	0.0	4.5	45.4
1976	31.8	32.6	23.0	27.6	12.9	0.0	0.001	0.0	0.0	0.001	2.1	38.3
1977	12.3	4.6	43.1	4.3	1.7	0.0	0.0	0.0	0.0	29.7	22.9	43.2
1978	7.0	14.4	6.3	0.001	0.3	0.0	0.0	0.0	0.0	0.001	10.0	7.9
1979	29.1	1.9	3.0	2.8	3.9	3.6	0.0	0.0	0.001	7.0	1.7	21.3
1980	13.4	80.9	6.7	1.7	0.1	0.0	0.0	0.0	0.0	0.0	18.0	1.9
1981	18.1	3.8	27.2	1.5	0.0	0.0	0.0	0.0	0.0	3.2	1.4	4.6
1982	92.7	12.3	20.8	8.4	22.2	0.0	0.0	0.0	0.0	32.3	0.0	4.3
1983	11.6	3.3	22.8	10.5	32.1	0.0	0.0	0.0	0.0	0.0	0.3	2.1
1984	17.8	0.2	8.7	5.7	0.0	0.0	0.0	0.0	0.0	0.3	87.5	34.4
1985	41.9	0.0	4.4	10.7	15.1	0.0	0.0	0.0	0.0	0.0	2.0	9.0
1986	39.1	22.5	41.6	4.0	5.6	0.0	0.0	0.0	0.0	0.0	93.0	36.4
1987	0.0	4.4	30.0	0.0	3.7	0.0	0.0	0.0	0.0	17.8	2.9	15.8
1988	24.0	11.5	10.5	44.1	0.0	0.0	0.0	0.0	0.0	0.0	0.2	19.6
1989	0.6	32.6	32.5	0.0	2.1	0.0	0.0	0.0	0.0	0.0	18.4	17.4
1990	14.8	27.1	15.9	4.5	0.0	0.0	0.0	0.0	0.0	0.5	0.0	1.3
1991	66.9	37.0	3.6	27.2	0.0	0.0	0.0	0.0	25.8	50.1	3.0	22.1
1992	11.5	11.9	28.8	0.4	0.0	0.0	0.0	0.0	0.0	0.0	31.0	31.4
1993	30.4	26.5	1.6	32.1	5.7	0.0	0.0	0.0	0.0	9.6	6.5	5.2
1994	26.3	0.0	2.5	11.1	0.6	0.0	0.0	0.0	0.0	4.1	38.2	26.7
1995	30.2	22.8	11.6	17.3	8.6	0.2	0.0	0.0	0.0	0.0	0.0	21.3
1996	72.5	30.8	52.2	6.4	1.2	0.0	0.0	0.0	0.0	0.0	11.5	6.2
1997	27.6	0.0	22.0	10.6	0.0	0.0	0.0	0.0	0.0	14.2	48.3	37.2
1998	34.8	5.7	90.2	21.6	0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.2
1999	21.1	69.8	23.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	14.8	27.6
2000	21.5	4.1	1.5	1.0	1.3	0.0	0.0	0.0	0.0	4.0	7.6	67.0
2001	5.0	3.1	9.1	0.0	0.2	0.0	0.0	0.0	0.0	0.1	3.0	42.4
2002	10.4	7.4	11.4	105.7	1.0	0.0	0.0	0.0	0.0	1.0	8.7	5.4
2003	M	M	M	M	M	M	M	M	M	M	M	M
2004	28.5	0.3	0.8	25.9	0.0	0.0	0.0	0.0	0.0	0.0	26.5	16.6
2005	45.2	0.9	33.7	3.9	0.0	0.0	0.0	0.0	0.0	0.0	0.2	21.8
2006	27.5	59.5	6.1	25.2	1.9	0.0	0.0	0.0	0.0	26.9	17.7	81.0
2007	9.2	0.1	75.8	5.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	21.9
2008	19.4	10.8	0.4	1.4	0.2	0.2	0.0	0.0	0.2	32.2	0.7	0.0
2009	0.3	7.1	18.6	4.6	1.5	0.6	0.0	0.0	0.0	0.2	1.7	22.3
2010	2.6	2.7	0.5	29.2	14.8	0.0	0.0	0.0	0.0	0.1	0.4	7.3
2011	7.5	19.9	13.8	21.2	9.7	0.0	0.0	0.0	0.0	0.0	13.0	0.0
2012	6.2	21.6	1.3	6.7	0.0	0.0	0.0	0.0	0.0	2.6	58.5	19.3
2013	8.8	0.2	0.3	0.0	36.0	0.0	0.0	0.0	0.0	1.8	126.4	1.7
2014	78.3	0.5	83.0	16.1	0.0	0.0	0.0	0.0	0.0	24.6	13.3	3.9
2015	0.9	14.4	7.2	1.0	0.9	0.0	0.0	0.0	0.0	24.4	8.3	36.1
2016	0.4	12.9	10.0	7.6	0.1	0.0	0.0	0.0	0.0	0.0	24.5	2.8
2017	0.0	0.0	13.1	1.8	0.0	0.0	0.0	0.0	0.0	0.0	12.1	0.0
2018	2.5	9.0	0.0	24.4	9.8	0.3	0.0	0.0	0.0	8.2	132.1	40.2
2019	12.8	8.5	31.4	21.0	0.0	0.0	0.0	0.0	0.0	7.5	15.0	7.4
2020	13.7	40.0	10.3	4.7	0.4	0.0	0.0	0.0	0.0	0.0	33.4	59.7

5.2 Gumbel distribution

Gumbel distribution (GUM), usually called the extreme value distribution of the type I, and usually used to represent an extreme process (maximum or minimum), such as; maximum rainfall, minimum stream flow, flood discharge, pollutant concentrations, etc. This distribution can be expressed by the equation below [21]:

$$F(x) = \left(\frac{1}{\alpha}\right) exp\left[-\left(\frac{x-u}{\alpha}\right) - exp\left(-\left(\frac{x-u}{\alpha}\right)\right)\right] \quad (2)$$

$$F(x) = exp\left[-exp\left(-\frac{x-u}{\alpha}\right)\right] \quad (3)$$

$$\alpha = 0.7797\sigma \quad (4)$$

$$u = \mu - 0.5772\alpha \quad (5)$$

Where: α is the scale parameter, and u is the location parameter for the Gumbel distribution.

5.3 Weibull distribution

Weibull distribution, usually called the extreme value distribution of the type III or W2, and usually used to represent minimum or maximum stream flows and relevant fields such as; analysis of maximum rainfall, and this distribution can be expressed by the Equation (6) below [22].

$$F(x) = \left(\frac{c}{\alpha}\right)\left(\frac{x}{\alpha}\right)^{c-1} exp\left[-\left(\frac{x}{\alpha}\right)^{c}\right] \quad (6)$$

Where: α is the scale parameter, c is the shape parameter and also known as the Weibull slope.

5.4 Normal distribution

The normal distribution or N distribution is the most important for symmetrically distributed data and can be applied in runoff analysis and annual precipitation, it is also called the "Gaussian distribution" [23], it can be expressed by the Equation (7) below:

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} exp\left[-(\frac{x-\mu}{2\sigma^2})^2\right] \quad (7)$$

Where: μ is the mean, σ is the variance, and -∞<x<∞.

5.5 Lognormal distribution

Usually, lots of persistent random hydrological variables tend to be asymmetrically distributed. It is mathematically possible to convert this distribution into a normal distribution. In most cases, it is possible to achieve the transformation reasonably by looking at the event logs. If the natural logarithms of variable x are normally distributed, then variable x follows the lognormal probability distribution. The probability density function for this variable F(x) = ln x [23]:

$$F(x) = \frac{1}{x\sigma\sqrt{2\pi}} exp\left[-\left(\frac{\ln x-\mu}{2\sigma}\right)^2\right] \quad for\ 0 \le x \le \infty \quad (8)$$

Where: μ is the mean of (ln x), and σ is the standard deviation of (ln x).

5.6 Gamma distribution

The random variable (x) has a gamma distribution, if the probability density function is given in the equation below, and this distribution is characterized by the presence of usually positive values [23]:

$$F(x) = \frac{\alpha^{\lambda}}{\Gamma_{(\lambda)}} x^{\lambda-1} e^{-\alpha x} \quad for\ 0 \le x \le \infty \quad (9)$$

Where: α is the scale parameter, λ is the shape parameter of the distribution, and the factor Γ(λ) is defined as in Equation (10) below:

$$\Gamma_{(\lambda)} = \int_0^\infty x^{\alpha-1} e^{-x} dx \quad (10)$$

5.7 Pearson type 3 distribution

It is one of the most common and used distributions in hydrology (P3), and is a two parameter Gamma distribution with a third parameter of location, (P3) contains the cumulative distribution function explained as follows [24]:

$$F(x) = \frac{\alpha^{\lambda}}{\Gamma(\lambda)}\left[(x-m)^{\lambda-1}\right]exp[-\alpha(x-m] \quad (11)$$

Where: λ is the shape parameter, α is the scale parameter, m is the location parameter, and Γ(λ) denotes the Gamma function and defined as in Equation (12) below:

$$\Gamma(\lambda) = \int_0^{\infty}(t)^{\beta-1}\, exp(-t)dt \quad (12)$$

5.8 Log-Pearson type 3 distribution

It is another distribution of the Gamma family, also commonly used in hydrology (LP3), where it describes a random variable whose logarithm that follows of the (P3) distribution, and containing the cumulative distribution function shown as follows:

$$F(x) = \frac{\alpha^{\lambda}}{x\Gamma(\lambda)}\left[(ln\, x-m)^{\lambda-1}\right]exp[-\alpha(ln\, x-m] \quad (13)$$

6. Method of research and methodology

To perform the research task and perform the required statistical tests, the (HYFRAN-PLUS version 1.2) software was used, this software is derived from the first letter pairs of the following words (HYdrological FRequency ANalyses PLUS DSS), while (DSS) is derived from (Decision Support System) to help choose the most appropriate class of distributions with respect to extreme values. It was specially designed to allow the fitting of several statistical distributions to a sample of data. It includes a number of flexible, powerful, and easy-to-use mathematical tools, that can be used for the statistical analysis, and especially to extreme value. The program was developed by a chair in Statistical Hydrology (INRS-ETE), at the National Institute of Scientific Research of the University of Québec. The work of the (HYFRAN) program begins as soon as the data is entered, whether through the windows clipboard or entered manually, after which the program automatically calculates the basic statistics values for the entered data as shown in Table 2, in addition, empirical probabilities through many options. After that, the program allows drawing a set of plots such as; the observations data histogram as well as chronological curves, and non-exccedance probability, as shown in Figures 2-3.

Regarding the calculation of the probability of not-exccedance, it was calculated using the Weibull formula shown below:

$$F[x(k)] = (k-a)/(n-2a+1) \quad (14)$$

Where: k is the rank, n is the number of observations, $(0 < a \leq 0.5)$, and for Weibull formula used the value of (a=0), therefore:

$$F[x(k)]= k/(n+1) = P \quad (15)$$

The return period (T) in years is the inverted of (P) and in a mathematical expression:

$$T=1/P \quad (16)$$

For the selected distributions, shown in section (5) to represent the probability analyzes of the extreme rainfall data in this study, the confidence interval (99%) was used, also, the maximum likelihood method was applied, to determine the theoretical distribution coefficients as shown in Figures 4-11 and Table 3. After completing the data analysis process, the adequacy test was achieved by applying the (Chi-Square) test, these values were compared with the tabulated values at a significant level (0.01) as shown in Table 4, according to this test, and after excluding the inappropriate values from the values of (Chi-Square) shown under line which are greater than the tabulated values at the significance level (0.01), found GEV, Lognormal, Pearson type 3, and Log-Pearson type 3 distributions are suitable, for describing extreme monthly rainfall in this study area.

Table 2: Basic statistics of data sets

Basic statistic	Value	Number of data (n)
Minimum	13.1	
Maximum	132	
Average	49.6	
Standard deviation	24.8	
Median	42.8	80
Coefficient of variation (Cv)	0.501	
Skewness coefficient (Cs)	1.11	
Kurtosis coefficient (Ck)	3.96	

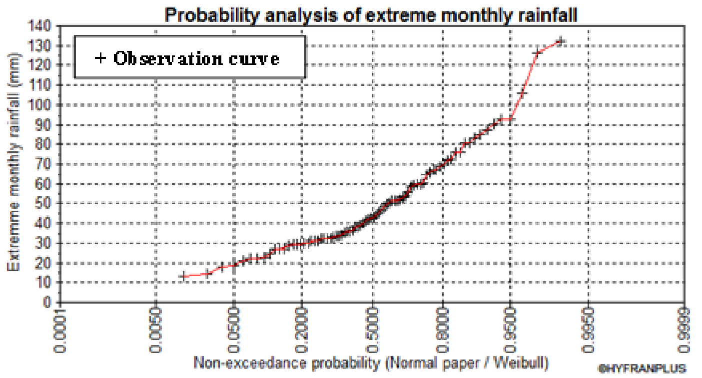

Fig. 2: The observation of data set on probability paper.

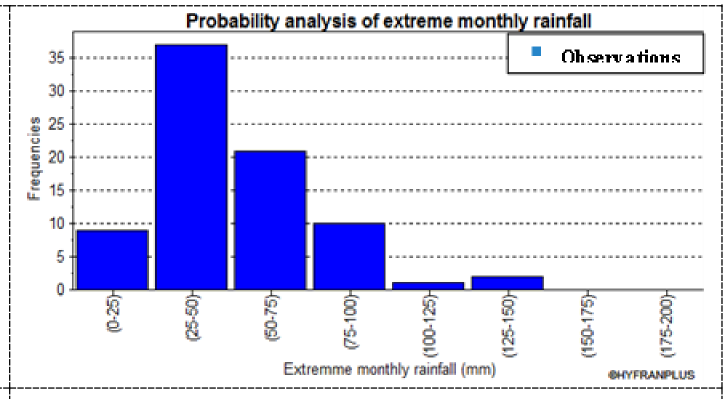

Fig. 3: Histogram of observation data set.

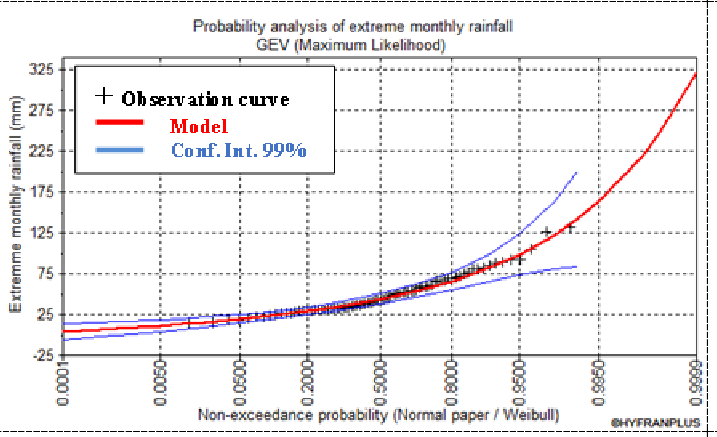

Fig. 4: Theoretical, empirical probabilities and confidence interval against extreme monthly rainfall (GEV distribution).

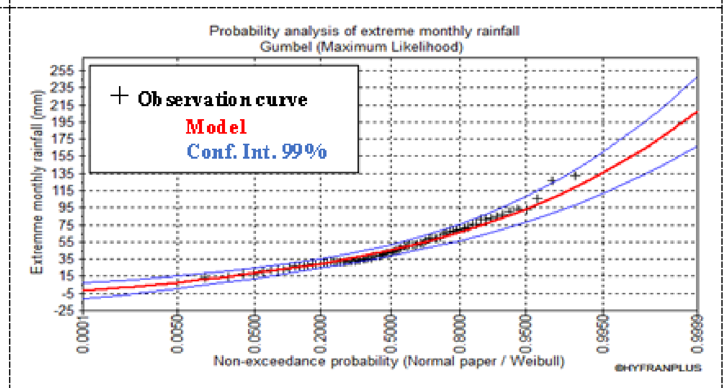

Fig. 5: Theoretical, empirical probabilities and confidence interval against extreme monthly rainfall (Gumble distribution)

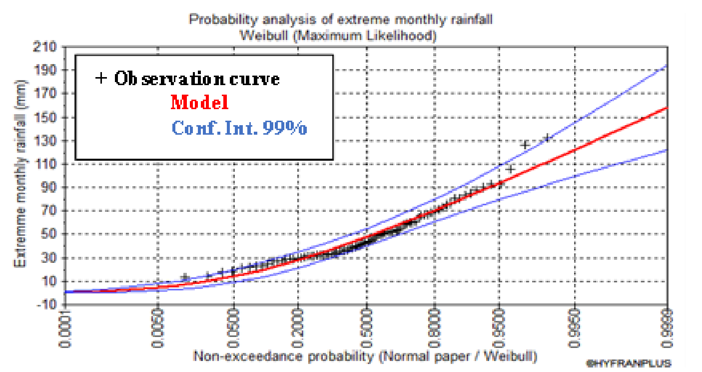

Fig. 6: Theoretical, empirical probabilities and confidence interval against extreme monthly rainfall (Weibull distribution).

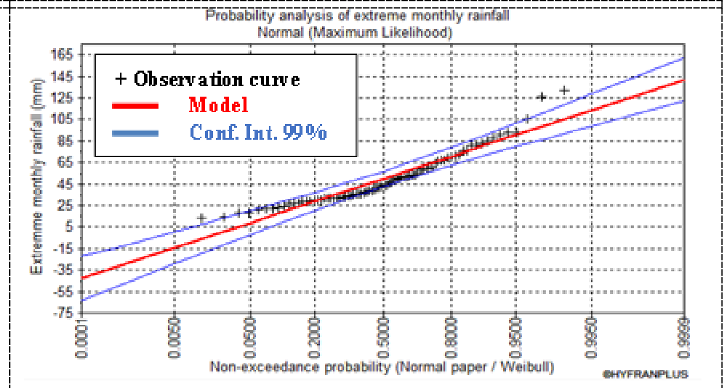

Fig. 7: Theoretical, empirical probabilities and confidence interval against extreme monthly rainfall (Normal distribution).

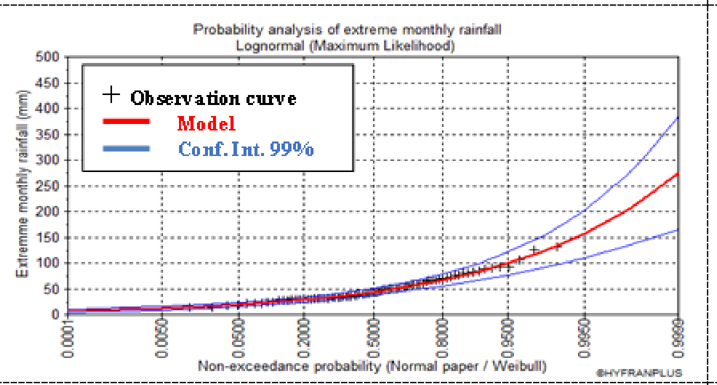

Fig. 8: Theoretical, empirical probabilities and confidence interval against extreme monthly rainfall (Lognormal distribution)

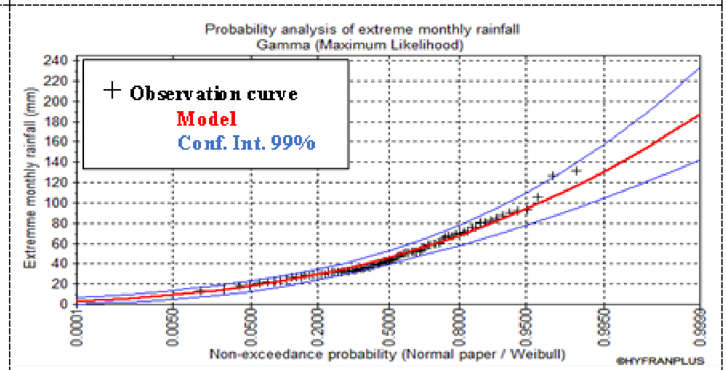

Fig. 9: Theoretical, empirical probabilities and confidence interval against extreme monthly rainfall (Gamma distribution).

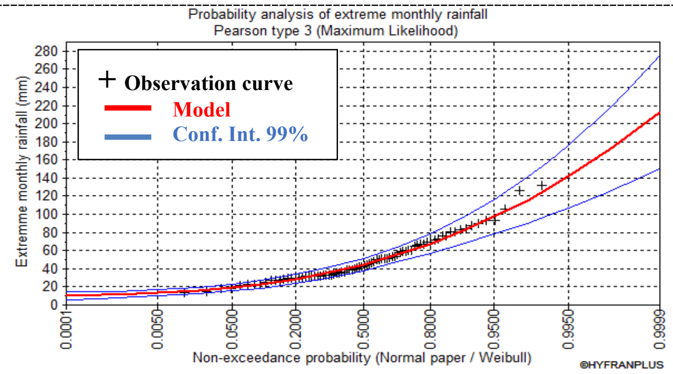

Fig. 10: Theoretical, empirical probabilities and confidence interval against extreme monthly rainfall (Pearson type 3 distribution).

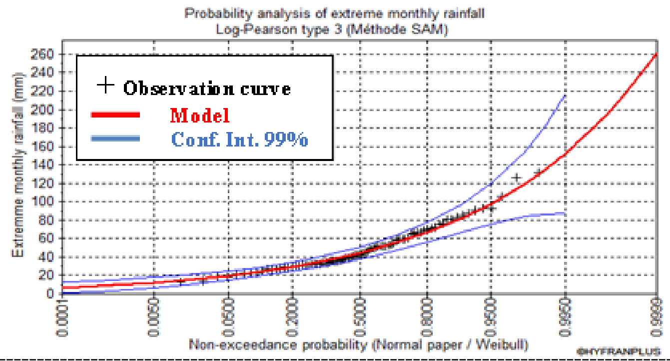

Fig. 11: Theoretical, empirical probabilities and confidence interval against extreme monthly rainfall (Log-Pearson type 3 distribution).

Table 3: Fitting models with the estimated coefficients

No.	Distribution	Estimated coefficients
1	GEV	α =17.2423 k = -0.116208 u = 37.5057
2	Gumbel	u = 38.5234 α = 18.3281
3	Weibull	α = 56.2578 c = 2.1464
4	Normal	μ = 49.6063 σ = 24.8302
5	Lognormal	μ = 3.78637 σ = 0.491929
6	Gamma	α = 0.0888388 λ = 4.40696
7	Pearson type 3	α = 0.0639724 λ = 2.524 m = 10.1517
8	Log-Pearson type 3	α = -212.868 λ = 2027.11 m = 11.1673

Table 4: Adequacy test using Chi-Square

No.	Distribution	Chi-Square test	
		Tabulated value (0.01) significant level	Calculated value
1	GEV	18.475	8.0
2	Gumbel	20.090	11.85
3	Weibull	20.090	18.17
4	Normal	20.090	30.55
5	Lognormal	20.090	8.0
6	Gamma	20.090	13.5
7	Pearson type 3	18.475	5.25
8	Log-Pearson type 3	18.475	7.17

7. Conclusions

The process of analyzing extreme monthly rainfall data is considered important and necessary, which helps in the design, planning, and management of water projects in the Nasiriyah city, the center of Thi-Qar Governorate. Therefore, this data was analyzed using (HYFRAN-PLUS version 1.2) software and eight probability distributions were used to obtain the best one in representing the data of this study area. The maximum likelihood method was applied to these data to obtain the coefficients of the theoretical distributions. The comparison between these probability distributions was done by comparing the values of (Chi-Square) calculated from adequacy test and the tabulated values at a significant level (0.01), and it was concluded that both GEV, Lognormal, Pearson type 3, and Log-Pearson type 3 distributions are suitable, for describing extreme monthly rainfall in this study area.

References

[1] D. Q. Tao, V.T.V. Nguyen, and A. Bourque, ”On selection of probability distributions for representing extreme precipitations in southern quebec” *Annual Conference of the Canadian Society for Civil Engineering*, 2002, pp. 1-8.

[2] F. A. Huff, and J. R. Angel, ”Rainfall frequency atlas of the midwest” Illinois State Water Survey, Champaign, Bulletin 71, MCC Research Report 92-03, 1992.

[3] S. Coles, ”*An introduction to statistical modeling of extreme values*” 1st ed., Springer-Verlag London Berlin Heidelberg, 2001.

[4] MD. A. Alam, K. Emura, C. Farnham, and J. Yuan, ”Best-Fit probability distributions and return periods for maximum monthly rainfall in Bangladesh” *MDPI, Climate*, Vol. 6, No. 9, 2018, pp. 1-16, doi:10.3390/cli6010009.

[5] H. Ahmadpari, E. S. Shokoohi, B. F. Sichani, E. N. Gharghani, and B. R. Ladez, ”Evaluation of the most appropriate statistical distribution for monthly rainfall prediction in the Zarrineh river watershed” *Science Arena Publications, Specialty Journal of Biological Sciences*, Vol. 5, No. 1, pp. 12-23, 2019. [Online]. Available at: www.sciarena.com.

[6] N. T. Thanh, ”Fitting a probability distribution to extreme precipitation for a limited mountain area in Vietnam” Scientific Research Publishing Inc, *Journal of Geoscience and Environment Protection*, Vol. 5, pp. 92-107, 2017. doi: 10.4236/gep.2017.55007.

[7] P. S. M. P. Ximenes, A. S. A. Silva, F. Ashkar, and T. Stosic, ”Best-fit probability distribution models for monthly rainfall of Northeastern Brazil” IWA Publishing, *Water Science & Technolgy*, Vol. 84, No. 6, pp. 1541-1556, 2021. http://iwaponline.com/wst/article-pdf/84/6/1541/942210/wst084061541.pdf.

[8] M. A. B. Barkotulla, M. S. Rahman, and M. M. Rahman, ”Characterization and frequency analysis of consecutive days maximum rainfall at Boalia, Rajshahi and Bangladesh” *Global Science Research Journals*, *Global Journal of Agricultural Economics and Econometrics*, Vol. 7, No. 3, pp. 536-541, 2019.

[9] R. Bora, A. Bhuyan, and B. K. Dutta, ”Analysis of consecutive days maximum rainfall together with dry, wet spell rainfall for crop planning at Jorhat (Assam), India” Communications in Mathematics and Applications, RGN Publications, Vol. 12, No. 4, pp. 835-851, 2021. doi: 10.26713/cma.v12i4.1620.

[10] O. O. Olofintoye, and B. F. Sule, ”Best–fit probability distribution model for peak daily rainfall in southern Nigeria” *1st Annual Civil Engineering Conference*, University of Ilorin, Nigeria, 2009, pp. 227-235.

[11] X. S. Kwaku, and O. Duke, ”Characterization and frequency analysis of one day annual maximum and two to five consecutive days’ maximum rainfall of Accra, Ghana” Asian Research Publishing Network (ARPN), *ARPN Journal of Engineering and Applied Sciences*, Vol. 2, No. 5, 2007, pp. 27-31.

[12] A. Sukrutha, S. R. Dyuthi, and S. Desai, ”Multimodel response assessment for monthly rainfall distribution in some selected Indian cities using best‑fit probability as a tool” *Springer, Applied Water Science*, Vol. 8, No. 145, pp. 1-10, 2018. https://doi.org/10.1007/s13201-018-0789-4.

[13] Iraqi Meteorological Organization and Seismology, Climate department, Directorate of meteorology station, Nasiriya, Iraq, Unpublished data.

[14] M. N. A. Alkinani, ”Climatic change in Thi-Gar governorate” *Journal of Wassit for Human Sciences*, Vol. 10, No. 24, 2014, pp. 271-292.

[15] D. B. Stephenson, ”*Definition, diagnosis, and origin of extreme weather and climate events*” in Diaz, H. F, and Murnane, R. J. ”*Climate extremes and society*” Cambridge University Press, 2009. https://doi.org/10.1017/CBO9780511535840.

[16] M. V. De Pol, S. Jenouvrier, J. H. C. Cornelissen, and M. E. Visser, ”Behavioural, ecological and evolutionary responses to extreme climatic events: challenges and directions” Philosophical Transactions of the Royal Society B: Biological Sciences, pp. 1-16, 2017. http://dx.doi.org/10.1098/rstb.2016.0134.

[17] D. S. Wilks, ”*Statistical methods in the atmospheric sciences*” 2nd ed, Vol. 91 in the International Geophysics Series, Academic Press, Elsevier Inc. 649 pp, 2006.

[18] Summary of monthly precipitation at stations in Iraq for period (1887-1958), Republic of Iraq, Development Board, Baghdad, Iraq, 62 p, 1962.

[19] N. O. S. Alghazali, and D. A. H. Alawadi, ”Fitting statistical distributions of monthly rainfall for some Iraqi stations” *Civil and Environmental Research*, Vol. 6, No. 6, 2014, pp. 40-46.

[20] C. Michele, and F. Avanzi, ”Superstatistical distribution of daily precipitation extremes: A worldwide assessment” Scientific Reports, Vol. 8, pp. 1-11, 2018. doi:10.1038/s41598-018-31838-z.

[21] J. L. Salinas, A. Castellarin, S. Kohnova, and T. R. Kjeldsen, ”Regional parent flood frequency distributions in Europe – Part 2: Climate and scale controls” *Hydrology and Earth System Sciences*, Vol. 18, pp. 4391-4401, 2014. doi:10.5194/hess-18-4391-2014.

[22] J. R. Stedinger, R. M. Vogel, and E. F. Georgia, ”*Frequency analysis of extreme events*” in Maidment, D.R. ”*Handbook of Hydrology*” McGraw Hill, New York, 1993, 1143 pp.

[23] K. Tilahun, ”The characterisation of rainfall in the arid and semi-arid regions of Ethiopia” *Water South Africa*, Vol. 32, No. 3, pp. 429-436, 2006. [Online]. Available at: http://www.wrc.org.za.

[24] W. Sanusi, S. Side, and M. K. Aidid, ”Probability distribution modeling of extremes rainfall series in Makassar city using the L-moments method” *Asian Journal of Applied Sciences*, Vol. 3, No. 5, 2015, pp. 656-663.