Scientific Review Engineering and Environmental Sciences (2026), 35 (1), 20–39

Sci. Rev. Eng. Env. Sci. (2026), 35 (1) https://srees.sggw.edu.pl

ISSN 1732-9353 (suspended) eISSN 2543-7496 DOI 10.22630/srees.10847

Received: 03.10.2025
Accepted: 30.12.2025

Yerlan MUKHANBET[1]✉
Jarosław CHORMAŃSKI[2]
Dana TUNGATAR[1]
Mariusz Paweł BARSZCZ[2]
Ainura ALDIYAROVA[1]

[1] Kazakh National Agrarian Research University, Faculty of Water Resources and Information Technology, Kazakhstan
[2] Warsaw University of Life Sciences – SGGW, Institute of Environmental Engineering, Poland

Probabilistic assessment of annual maximum precipitation in Almaty, Kazakhstan

Keywords: urban flooding, extreme rainfall, Gumbel distribution, generalized extreme value, GEV, return period analysis

Introduction

Extreme hydrometeorological events have become a growing global concern due to their direct socio-economic and environmental impacts. Numerous studies have shown that intense rainfall events and flash floods are increasing in both frequency and severity due to climate change. These phenomena not only damage infrastructure and property but also pose significant risks to human safety, particularly in rapidly urbanizing areas with insufficient drainage systems (D'iya et al., 2014).

✉mukhanbet.yerlan@inbox.ru

Globally, urban flooding has become one of the most critical consequences of climate-induced hydrological extremes. For example, severe rainfall events in Europe, such as the 2021 floods in Germany, Belgium, and Poland, resulted in extensive infrastructural damage and loss of life (Kreienkamp et al., 2021; Pietras & Pyrc, 2025). Similar patterns are observed in South Asia, where countries like India, Pakistan, and Bangladesh experience increasingly frequent monsoon-driven floods (Latif et al., 2017; Muhammad Iskandar et al., 2025). In China and Japan, rapid urbanization combined with changing rainfall regimes has intensified flood hazards, prompting the development of advanced stormwater management systems and statistical models for rainfall prediction (Jayawardane et al., 2024; Ling et al., 2025). These international examples highlight that the challenge of managing extreme precipitation and urban flooding is not unique to a single region but represents a widespread, global issue requiring localized yet methodologically comparable approaches.

In Kazakhstan, Almaty city has recently witnessed a marked increase in the frequency and intensity of extreme precipitation events. Several episodes of intense rainfall and sudden surface runoff have caused localized flooding, especially in low-lying districts such as Almalinskiy and Bostandyk. The city's mountainous surroundings, coupled with rapid urban expansion and the loss of permeable green spaces, accelerate surface runoff and exacerbate the overload of the existing drainage infrastructure (BES Media, 2025). Furthermore, the absence of a comprehensive stormwater management system across many parts of the city leaves Almaty particularly vulnerable to the adverse effects of hydrological extremes.

These challenges are not unique to Almaty city. Similar situations have been observed in other urban centers of Kazakhstan, such as Astana city, where episodes of heavy rainfall and rapid snowmelt have overwhelmed drainage systems, resulting in urban flooding (Nakispekova, 2024). Almaty city has typical characteristics of the Central Asia urbanization processes and, in the authors' opinion, can explain similar processes in this region, where rapidly developing cities are dominant. Almaty city faces the dual challenge of adapting to intensifying climate extremes while managing the consequences of rapid urbanization. Traditional stormwater infrastructure and planning approaches, based on outdated climatic assumptions, are increasingly inadequate for modern conditions (Dong et al., 2020). This underscores the urgent need for more adaptive and data-driven methods in urban flood risk assessment and management.

International experience demonstrates that statistical modeling of extreme rainfall using probability distribution functions has become a widely accepted approach to understanding and predicting hydrological extremes. The generalized extreme value (GEV) distribution remains the most commonly applied model in Europe, China, and Australia for estimating design rainfall and return periods of floods due to its flexibility in capturing both short-term and long-term extremes (Coles, 2001; Katz et al., 2002; Gentilucci et al., 2023).

Recent studies have shown that lognormal and gamma distributions continue to be preferred in regions where rainfall intensity exhibits moderate variability and observational records are limited, owing to their computational simplicity and stable fitting performance (Stedinger et al., 1993; Cho et al., 2004; Montes-Pajuelo et al., 2024). However, each model presents specific limitations: for instance, the GEV may overestimate rare events when sample sizes are small, while lognormal and gamma distributions tend to underestimate the upper tail of extreme rainfall (Rima et al., 2025).

Therefore, the combined application and intercomparison of multiple statistical models provide a more robust and reliable basis for assessing precipitation extremes under non-stationary climate conditions, as emphasized in recent hydrological analyses (Gentilucci et al., 2023; Rima et al., 2025).

Between 2015 and 2025, Almaty city experienced several instances of localized flooding and mudflows, primarily triggered by heavy rainfall, rapid snowmelt, and natural runoff from the surrounding mountainous terrain. Due to its geographic location at the foothills of the Trans-Ili Alatau (a subrange of the Tian Shan mountain belt), Almaty is particularly vulnerable to sudden water surges from upstream rivers and glacial sources. Each spring and summer, rising temperatures accelerate glacial melt, leading to swollen rivers and, occasionally, the formation and bursting of moraine-dammed glacial lakes (Kyrgyzbay et al., 2023; Choudhary & Kumar, 2025).

Although none of these incidents matched the scale of the 2015 glacial lake outburst flood, they collectively posed a significant challenge to the urban infrastructure. The city's peripheral districts – such as Alatau, Turksib, and Bostandyk – were the most frequently affected, often facing obstructed roads, flooded basements, and temporary evacuations (Duskayev et al., 2023; Mirlas et al., 2024). Rapid urban expansion has increased vulnerability: changes to the hydrographic network due to infrastructural development, land use patterns, and inadequate drainage have amplified local flood risks. Human activity, particularly the construction of new neighborhoods and modification of watercourses, has further complicated natural water flow, leading to more severe and recurrent urban flood events (Duskayev et al., 2023).

Despite ongoing efforts by the city and national authorities to improve early warning systems and build protective infrastructures such as mudflow barriers and improved river channels, these events highlight the central role of sustainable urban planning and adaptation to climate change. The continued increase in average temperatures and more frequent episodes of intense precipitation intensify the risk of flash floods and mudflows in Almaty's metropolitan area (Duskayev et al., 2023; Kyrgyzbay et al., 2023; Choudhary & Kumar, 2025).

This study provides a preliminary analysis of rainfall data collected from meteorological stations in Almaty city prior to their application in urban flood modeling.

The aim of this study is to analyze trends and spatial variability of extreme precipitation and runoff potential in Almaty using ground-based meteorological data (2000–2023) through statistical analysis.

Material and methods

Study area

Almaty city is located at the foothills of the Zailiyskiy Alatau Mountains (also known as the Trans-Ili Alatau), in southeastern Kazakhstan. It serves as the administrative center of Almaty city and lies at elevations ranging from approx. 700 m to 900 m a.s.l. The city includes both lowland urbanized areas and mountainous catchments, which form the headwaters of several small rivers, such as Ulken Almaty, Kishi Almaty, Esentai, Aksay, and Kargaly (Kyrgyzbay et al., 2023).

Over the past century, Almaty has undergone rapid urban expansion, with the city's built-up area increasing from approx. 3 km² in the 19th century to over 700 km² in recent decades. This urban growth has resulted in significant transformations of the natural hydrographic network, including the channelization of streams, construction of stormwater infrastructure, and creation of artificial drainage pathways (Choudhary & Kumar, 2025). These changes have substantially altered the hydrological behavior of the area, increasing the likelihood of urban flooding. The area of Almaty city is now approx. 682 km².

The city is highly susceptible to natural hazards, including heavy rainfall, flash floods, and debris flows. These risks are exacerbated by the combination of steep mountainous topography, convective precipitation events, and reduced infiltration capacity in urbanized zones (Mirlas et al., 2024). Historical flood events, including the catastrophic mudflows of 1921 and the urban floods in 2015, have caused significant damage to infrastructure and property.

Furthermore, recent groundwater modeling studies in the northern parts of the city, particularly in the Akbulak Micro District, have identified recurring groundwater flooding events caused by poor drainage and excessive surface runoff. These issues are compounded by the obstruction of natural subsurface flows and the degradation of the historical Karasu stream system (Mirlas et al., 2024).

The combination of climate variability, increasing rainfall intensity, land-use change, and hydrological modification has created a complex and dynamic urban flood regime. Therefore, Almaty represents a critical case study for assessing the impacts of climate-induced extreme rainfall and evaluating probabilistic models for sustainable flood risk management.

Data

The rainfall data used in this study were obtained from the official website of Kazhydromet (Republic State Enterprise Kazhydromet, n.d.) covering a continuous 23–24-year period from 2000 to 2023 (from 2001 to 2023). The dataset used in the analysis in that paper consists of the total annual precipitation sums recorded at five meteorological stations within Almaty city: Almaty, Shymbulak, Mynzhylyk, Bolshoe Almatinskoe Ozero (BAL), and Kemen. The names and geographical coordinates of these stations, together with key descriptive statistics, are presented in Table 1. The 24-year data series (2000–2023) can be considered relatively short for long-term climatological analysis, when the normal analytical period for hydrometeorological characteristics should not be shorter than 30 years (World Meteorological Organization [WMO], 1989). However, this period was selected based on the availability of continuous, high-quality digital data provided by the Republican State Enterprise "Kazhydromet" Kazakhstan's official hydrometeorological service. Earlier data (prior to 2000) are mainly stored in paper archives and contain numerous gaps and inconsistencies, which reduce their homogeneity and reliability. Therefore, the 2000–2023 dataset obtained from the official Kazhydromet portal was considered optimal for ensuring completeness, calibration comparability, and statistical reliability.

To evaluate the variability of precipitation, several statistical indicators were calculated. The standard deviation (*SD*), which represents the dispersion of precipitation around the mean, ranged from 135 mm at Almaty station to 196 mm at Shymbulak, indicating moderate interannual variability. The coefficient of variation (*CV*), computed as the ratio of *SD* to the mean (*M*), varied between 0.176 and 0.202, corresponding to moderate variability (10–20%), which suggests relative stability of precipitation across the studied period.

To further describe the statistical characteristics of rainfall, skewness and kurtosis were analyzed. Positive skewness values ranging from 0.63 to 1.34 indicate a right-skewed distribution, meaning that most rainfall events are moderate, but rare extreme events occasionally occur. Kurtosis coefficients (γ_2) between 2.87 and 4.19 show that certain stations (e.g., Shymbulak) exhibit more pronounced peaks and heavier tails compared to a normal distribution (kurtosis = 3), which is typical for regions exposed to intense orographic precipitation. All descriptive parameters were calculated using the formulas presented in Wilks (2011).

TABLE 1. Geographic characteristics of meteorological stations and descriptive statistics of annual precipitation sum data series observed in the period of 2000–2023

Geographical coordinate	Meteorological station				
	Almaty	Shymbulak	Mynzhylky	BAL	Kemen
Latitude (N)	43°24′	43°12′	43°08′	43°05′	43°18′
Longitude (E)	76°93′	77°08′	77°07′	76°98′	76°96′
Descriptive statistics of annual precipitation sum					
Minimum [mm]	489	687	668	636	615
Maximum [mm]	1,010	1,480	1,240	1,320	1,360
Mean (*M*) [mm]	678	968	875	866	904
Standard deviation (*SD*) [mm]	135	196	154	165	179
Coefficient of variation (*CV*)	0.198	0.202	0.176	0.190	0.198
Skewness (γ_1)	0.995	1.34	0.739	1.12	0.632
Kurtosis coefficient (γ_2)	3.24	4.19	2.87	3.60	2.88
Data length [year]	2000–2023 (24 years)	2001–2023 (23 years)	2000–2023 (24 years)	2000–2023 (24 years)	2000–2023 (24 years)

Source: own work.

To verify the reliability and temporal consistency of the precipitation dataset, three non-parametric statistical tests were conducted for all five meteorological stations (Table 2). The Wald–Wolfowitz runs test (*U*) was used to evaluate the independence of the data series, ensuring that precipitation values are not serially correlated. The Kruskal–Wallis test (*K*) was applied to assess stationarity, confirming that there are no systematic changes or trends within the time series. The Pettitt test (*W*) was employed to check homogeneity, detecting possible abrupt shifts or inhomogeneities in the data over time.

The calculated test statistics fall within the following ranges:
– independence (*U*): from 1.30 to 1.83,
– stationarity (*K*): from 0.74 to 1.76,
– homogeneity (*W*): from 0.22 to 1.01.

All values are below the critical threshold of ±1.96 at a 95% confidence level, indicating that none of the null hypotheses were rejected. Therefore, the precipitation series for all stations can be considered independent, stationary, and homogeneous, providing a reliable basis for further hydrological and probabilistic modeling.

TABLE 2. Statistical characteristics of annual precipitation sum data series observed in the period of 2000–2023 by stations

Statistical test value	Meteorological station				
	Almaty	Shymbulak	Mynzhylky	BAL	Kemen
Independence	$\lvert U \rvert = 1.5$	$\lvert U \rvert = 0.83$	$\lvert U \rvert = 0.38$	$\lvert U \rvert = 0.24$	$\lvert U \rvert = 1.23$
Stationarity	$\lvert K \rvert = 1.76$	$\lvert K \rvert = 1.32$	$\lvert K \rvert = 0.74$	$\lvert K \rvert = 0.86$	$\lvert K \rvert = 0.82$
Homogeneity	$\lvert W \rvert = 1.01$	$\lvert W \rvert = 0.22$	$\lvert W \rvert = 1.01$	$\lvert W \rvert = 0.72$	$\lvert W \rvert = 0.43$

Source: own work.

The magnitude of the test statistics varies among stations, reflecting spatial differences in precipitation behavior and local influences. The Almaty station exhibits higher $|U|$ and $|K|$ values (1.5 and 1.76), indicating a slight upward trend likely associated with urbanization and impervious surfaces that affect local convection and moisture dynamics. In contrast, the mountain stations Shymbulak, Mynzhylky, BAL, and Kemen exhibit lower values of $|U|$ and $|K|$ (< 1), confirming stable and homogeneous precipitation patterns typical of natural high-altitude environments with minimal anthropogenic influence. The homogeneity test ($|W| < 1.1$) for all stations confirms the absence of structural breaks in the data. Overall, the precipitation records are statistically consistent, although the elevated values at Almaty suggest emerging urban–climate interactions influencing rainfall variability.

The topographic complexity of the research area requires the inclusion of meteorological data collected across the entire area. As shown in Figure 1, the network of meteorological stations covers both the urbanized areas and the surrounding mountainous part of Almaty city. Located in the southeastern part of Kazakhstan, within the Almaty Region, the city lies at the foothills of the Zailiyskiy Alatau mountains. This spatial distribution enables a more accurate representation of precipitation variability and supports the identification of potential flash-flood risk zones in Almaty city.

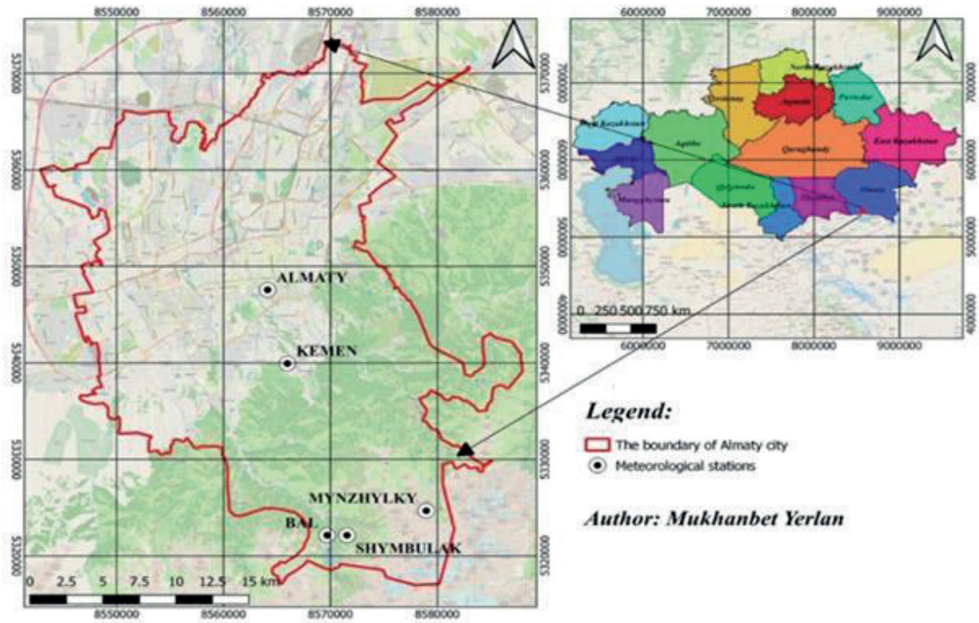

FIGURE 1. Location map of the meteorological stations in Almaty city, with a general map of Kazakhstan
Source: own work.

As illustrated in Figure 2, the spatial distribution of annual precipitation across Almaty city demonstrates a clear north-south gradient. Precipitation amounts increase significantly from the northern plains toward the foothills and mountainous zones of the Zailiyskiy Alatau. The highest annual totals, exceeding 900 mm, are observed at high-altitude stations such as Shymbulak and Mynzhylyk, while the lowest values, below 700 mm, occur in the northern and central lowland areas. This orographic influence results in substantial spatial variability of rainfall, which plays a crucial role in shaping surface runoff and the occurrence of flash floods in the city.

Given the high spatial and seasonal variability of climatic parameters, it is crucial to determine the probabilistic distributions of atmospheric precipitation. This enables:

– Assessment of the risk of extreme weather events (e.g., rare but high-impact rainfall).
– Hydrological modeling of stormwater runoff based on the statistical characteristics of precipitation.
– Development of adaptation scenarios and sustainable stormwater management measures under climate variability.

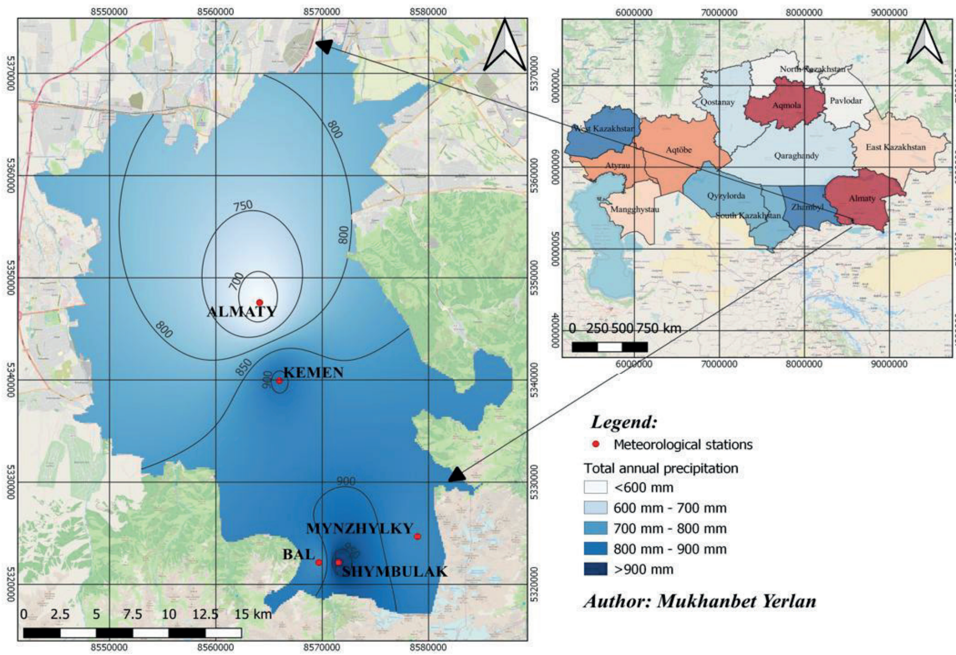

FIGURE 2. Map of annual precipitation distribution in Almaty city
Source: own work.

The application of probabilistic approaches, such as the five tested probability distributions, allows for a shift from descriptive analysis to quantitative risk assessment of pluvial flooding.

In recent decades, Almaty city has regularly faced challenges related to stormwater flooding, caused by a combination of intense precipitation events, rapid urbanization, and insufficient capacity of the existing drainage infrastructure. This problem is particularly acute during spring and summer, when short-duration but high-intensity rainfall generates substantial surface runoff that exceeds the capacity of current stormwater systems.

According to the Almaty Department for Emergency Situations, more than 50 incidents of localized flooding affecting streets, residential basements, and public infrastructure were recorded between 2010 and 2022. One of the most significant events occurred in July 2015, when over 50 mm of rainfall fell within a few hours, flooding central areas of the city. According to municipal authorities, total material damage exceeded 1.2 billion tenge, including damage to vehicles, underground utilities, and commercial facilities. One can see one of the flooding outcomes in the following link (BES Media, 2025).

Methodology

This section presents the methodology used to identify the most suitable probability distribution function for estimating design values of annual maximum precipitation. Such a model can be applied by engineers and water resource specialists for a priori assessments in frequency analysis.

During the first stage, five different probability distribution functions (PDFs) were tested on the series of annual maximum precipitation values to determine which best fit the data. These were selected due to their widespread use in contemporary hydrometeorological studies (Kundzewicz et al., 2013; Dong et al., 2020). The annual maximum precipitation series, obtained from various meteorological stations, were fitted using the HYFRAN software package (Duskayev et al., 2023), a practical tool for evaluating probability distributions. The list of applied distribution functions found in the software is presented in Table 3.

TABLE 3. Probability density functions, equations, and descriptions

Distribution	Probability formula	Typical use	Advantages	Limitations
Gumbel	$f\left(x\right)\frac{1}{\beta}\exp\left[-\frac{x-\mu}{\beta}-\exp\left(-\frac{x-\mu}{\beta}\right)\right]$	Maxima (annual peak precipitation)	Simple, widely used in hydrology	May under-estimate tail behavior (underpredict extremes)
Generalized extreme value (GEV)	$f\left(x\right)=\frac{1}{\sigma}\left(1+\xi\frac{x-\mu}{\sigma}\right)^{-\frac{1}{\xi}-1}\exp\left[-\left(1+\xi\frac{x-\mu}{\sigma}\right)^{-\frac{1}{\xi}}\right]$	Flexible for maxima with different tail behavior	Accounts for skewness and heavy tails	More complex parameter estimation
Log-Pearson Type III	log (*X*) ~ *Pearson Type III*	Flood and rainfall extremes	Recommended by the U.S. Water Resources Council	Sensitive to sample size and outliers
Pearson Type III	$f\left(x\right)=\frac{\beta^{\alpha}\ (x-x_0)^{\alpha-1}e^{-\beta\left(x-x_0\right)}}{\Gamma\left(\alpha\right)}$, where $x>x_0$	General hydrologic use	Good fit for varied data with skewness	Assumes continuous positive data
Lognormal	$f\left(x\right)=\frac{1}{x\sigma\sqrt{2\pi}}\exp\left(-\frac{(\ln x-\mu)^2}{2\sigma^2}\right),x>0$	Precipitation intensities	Suitable for moderately skewed data	Sensitive to zero and near-zero values
Normal	$f\left(x\right)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$	Not recommended for extremes	Easy to interpret	Poor fit for skewed and extreme event data

Source: own work.

To fit the parameters of each PDF, maximum likelihood estimation (MLE) and method of moments (MoM) approaches were used. These methods enabled the estimation of distribution parameters based on a series of annual maximum precipitation values.

In addition, it was necessary to determine the empirical exceedance probability of the observed annual maximum precipitation values. For this purpose, the following formula was used:

$$F\left(x_k\right) = \frac{k-\alpha}{n-2\alpha+1}, 0 \le \alpha \le 0.5,$$

where: $F(x_k)$ is empirical probability corresponding to the value x_k, k is rank of the value in the ascending ordered series, n is total number of observations, α is offset parameter depending on the selected formula (cf. Table 4).

TABLE 4. Values of parameter α for estimating the empirical exceedance probability

α	Comment	Reference
0.4	Recommended for hydrometeorological data; provides a balanced estimation for both small and large sample sizes	Cunnane (1978)
0.5	Universal, but performs poorly at the tails (rare events). Frequently used in engineering calculations, but may overestimate probabilities at the distribution tails	Hazen (1914)
0	Simple and widely used, but may introduce bias	Weibull (1939)
0.3	Based on Soviet tradition, but often inadequate in mountainous regions	Chegodayev (1955)
0.375	Used in statistics; closer to the normal distribution	Blom (1958)
≈0.333	Sometimes applied in scientific publications	Tukey (1962)

Source: own work.

To calculate the empirical exceedance probability of annual maximum precipitation in Almaty city, the Cunnane formula (Cunnane, 1978) was selected. The choice of this formula is justified by the region's topographic and climatic characteristics. Almaty is situated in the foothills of the Zailiyskiy Alatau, where orographic influences result in high spatial and temporal variability in precipitation. This is particularly evident in the case of intense convective rainfall events occurring during the warm season, often associated with moisture-laden mountain air and strong atmospheric instability.

Under such conditions, the Cunnane formula provides the most balanced probability estimates, especially when dealing with a limited number of observations ($n \leq 25$ years), which is typical for meteorological stations in the city.

The Cunnane formula minimizes systematic errors at the distribution tails, offering a more accurate representation of both frequent and rare (extreme) events. This makes it particularly useful for comparing empirical data with theoretical distributions, such as the Gumbel, GEV, or log-Pearson Type III distributions, which are widely used in hydrological and climate-related studies (Kundzewicz et al., 2013; Dong et al., 2020). Moreover, the World Meteorological Organization recommends the use of the Cunnane formula for analyzing extreme hydrometeorological data in mountainous regions (WMO, 2009, 2017).

Thus, the use of the Cunnane formula is both scientifically and methodologically justified for analyzing extreme precipitation in the mountainous conditions of Almaty city. It provides a reliable foundation for estimating return periods and assessing the risks of pluvial flooding.

Results and discussion

In this study, the 24-year annual precipitation records (2000–2023) from five meteorological stations near Almaty city were analyzed. For hydrological planning and design purposes, the best available long-term precipitation data were used. The annual maximum precipitation values (in mm) were statistically processed to estimate design rainfall intensities for return periods of 10 (10%), 50 (2%), and 100 (1%) years using several PDFs, namely the exponential, GEV, normal, lognormal, and gamma distributions.

The goodness of fit of each distribution was evaluated using multiple criteria: the chi-square (χ^2) test, Akaike information criterion (AIC), Bayesian information criterion (BIC), and posterior model probability $P(M_i|x)$. Table 5 presents the comparison of empirical and theoretical precipitation quantiles (in mm) for each station.

Figures 3 and 4 show the results of fitting the GEV and lognormal distributions to the observed values of annual precipitation maxima at the Almaty weather station. The red line shows the theoretical model, the black crosses show the empirical data, and the blue lines show the 95% confidence interval. These two graphs are provided as illustrative examples.

TABLE 5. Design rainfall intensities for return periods of 1%, 2% and 10% calculated by selected empirical probability functions applied to precipitation annual sum time series in meteorological stations of Almaty city

Meteorological station	Design rainfall for return periods calculated by selected probability distribution functions [mm]														
	exp.			GEV			normal			lognormal			gamma		
	10%	2%	1%	10%	2%	1%	10%	2%	1%	10%	2%	1%	10%	2%	1%
Almaty	936	1,250	1,390	851	1,020	1,090	851	955	991	849	983	1030	845	964	1,010
Shymbulak	1,350	1,820	2,030	1,210	1,480	1,590	1,220	1,370	1,420	1,210	1,400	1,470	1,210	1,380	1,440
Mynzhylky	1,160	1,500	1,650	1,070	1,260	1,330	1,070	1,190	1,230	1,070	1,230	1,280	1,070	1,200	1,250
BAL	1,180	1,570	1,730	1,080	1,310	1,410	1,080	1,200	1,250	1,070	1,230	1,290	1,070	1,210	1,270
Kemen	1,300	1,780	1,990	1,140	1,330	1,400	1,130	1,270	1,320	1,140	1,330	1,400	1,130	1,290	1,350

exp. – exponential distribution, GEV – generalized extreme value distribution, normal – normal distribution, lognormal – lognormal distribution, gamma – gamma distribution.

Source: own work.

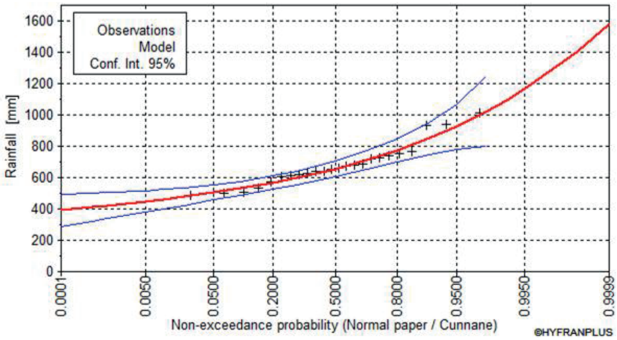

FIGURE 3. Annual maximum precipitation distribution fitted with the generalized extreme value distribution using the maximum likelihood estimation method for the Almaty meteorological station

Source: own work.

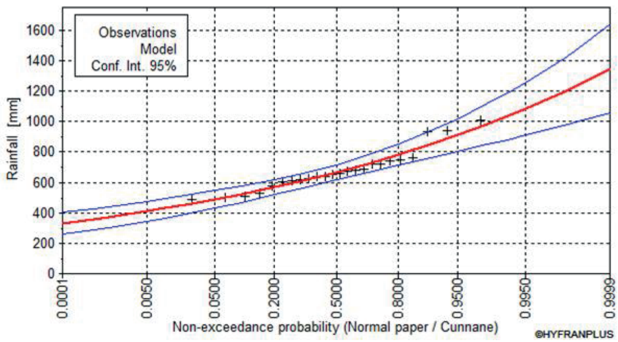

FIGURE 4. Annual maximum precipitation distribution fitted with the lognormal distribution using the maximum likelihood estimation method for the Almaty meteorological station

Source: own work.

The analysis of AIC, BIC, and $P(M_i|x)$ for the Almaty meteorological station, which serves as a representative location for an urban area in this study, shows that the lognormal distribution demonstrates superior performance, yielding the highest posterior probability $P(M_i|x) = 39.94$ along with the lowest AIC (303.2) and BIC (305.6) values (Table 6).

TABLE 6. Comparison of the analysis for the meteorological station based on AIC, BIC, $P(M_i|x)$, χ^2 test, and p-value criteria

Meteorological station	Parameter	Probability distribution type				
		exp.	GEV	normal	lognormal	gamma
Almaty	AIC	305.7	304.2	306.4	303.2	304.0
	BIC	308.1	307.7	308.7	305.6	306.3
	$P(M_i\|x)$	11.2	13.4	8.2	39.9	27.1
	χ^2	9.2	2.8	7.5	5.1	5.1
	p	0.05	0.41	0.11	0.27	0.27
Shymbulak	AIC	311.4	307.2	311.0	306.9	309.2
	BIC	313.7	310.6	313.3	309.2	310.2
	$P(M_i\|x)$	4.5	21.1	5.5	43.4	25.3
	χ^2	7.1	2.8	4.7	1.6	3.4
	p	0.13	0.41	0.32	0.80	0.48
Mynzhylky	AIC	310.0	312.2	312.8	310.8	311.2
	BIC	312.4	315.7	315.2	313.1	313.6
	$P(M_i\|x)$	37.4	7.1	9.2	25.8	20.4
	χ^2	8.6	4.0	2.8	4.5	4.5
	p	0.07	0.26	0.58	0.33	0.33
BAL	AIC	315.1	312.8	316.0	312.5	313.4
	BIC	317.4	316.3	318.4	314.9	315.8
	$P(M_i\|x)$	10.9	18.8	6.7	38.9	24.5
	χ^2	12.7	5.7	1.6	5.1	5.1
	p	0.01	0.12	0.79	0.27	0.27
Kemen	AIC	326.0	320.3	320.2	318.5	318.5
	BIC	328.3	323.9	322.5	320.8	320.8
	$P(M_i\|x)$	0.9	8.6	16.9	39.3	34.1
	χ^2	11.0	6.3	4.00	8.0	6.3
	p	0.02	0.09	0.40	0.08	0.17

exp. – exponential distribution, GEV – general extreme value distribution, normal – normal distribution, lognormal – lognormal distribution, gamma – gamma distribution.

Source: own work.

Although the GEV showed the best agreement with empirical data according to the χ^2 test, the lognormal distribution proved more favorable when accounting for model simplicity and goodness-of-fit through information-theoretic metrics. The gamma distribution also performed relatively well, with a posterior probability of 27.1 and competitive AIC (304.0) and BIC (306.3) values (Table 6).

The variation in the best-fit distribution across different stations and criteria underscores the complexity of hydrological modeling in a region with complex climatic conditions, such as those in Almaty city. The discrepancy between the χ^2 test (favoring the GEV) and information criteria/Bayesian probabilities (favoring lognormal/gamma) highlights the importance of using multiple goodness-of-fit measures for robust model selection. The GEV, while showing good agreement with empirical data via the χ^2 test, might be penalized by AIC and BIC if its additional parameter does not significantly improve the fit relative to its complexity.

Estimating design rainfall with different return periods (e.g., 1%, 2%, 5%, and 10%) is crucial for infrastructure planning and for mitigating risks from extreme precipitation events, such as flash floods and urban inundations, which are intensified by ongoing urbanization and climate variability in Almaty. The selection of the most appropriate PDF directly impacts the accuracy of these design values.

Similar approaches have been widely applied in other countries, where various probability distributions have been tested to model extreme rainfall events. For instance, the GEV is the most commonly used model in Europe, China, and Australia for estimating design rainfall and flood return periods due to its flexibility in capturing both short-term and long-term extremes (Coles, 2001; Burnham & Anderson, 2002; Katz et al., 2002; Gentilucci et al., 2023). In contrast, lognormal and gamma distributions are frequently applied in regions with moderate rainfall variability and limited data availability, such as parts of South Asia and Africa, owing to their computational simplicity and stable fitting performance (Stedinger et al., 1993; Cho et al., 2004; Montes-Pajuelo et al., 2024). However, these models differ in their ability to represent the upper tail of extreme precipitation: the GEV tends to perform better for rare events, whereas the lognormal and gamma distributions may underestimate extremes. Therefore, comparing multiple statistical distributions, as presented in this study, provides a more robust and comprehensive understanding of precipitation extremes under non-stationary climate conditions.

Conclusions

This study presents a comprehensive probabilistic assessment of extreme precipitation in the city of Almaty, based on 24 years of ground-based meteorological observations (2000–2023) from five stations: Almaty, Shymbulak, Mynzhylky, BAL, and Kemen. The analysis used annual maximum precipitation values and applied five theoretical probability distribution functions – exp., GEV, normal, lognormal, and gamma – to model extreme rainfall behavior.

All rainfall series were confirmed to be independent, stationary, and homogeneous, as verified by statistical tests (Wald–Wolfowitz, Kruskal–Wallis, and Pettitt). The results of the MLE showed that the GEV and lognormal distributions provided the best fit across most stations.

Quantitatively, the maximum annual precipitation during the study period ranged from 615 mm (Kemen) to 1,010 mm (Shymbulak), while the mean annual precipitation varied between 678 mm (Almaty) and 866 mm (BAL). The *CV* ranged from 0.17 to 0.20, indicating moderate interannual variability.

Based on the best-fitting probability models, the estimated rainfall quantiles showed a clear increase with longer return periods. On average across all stations, the 10-year event (10% exceedance probability) corresponded to rainfall amounts of approximately 950–1,050 mm, the 50-year event (2% exceedance) to around 1,100–1,200 mm, and the 100-year event (1% exceedance) to 1,200–1,300 mm. These values represent the expected magnitudes of extreme precipitation for the respective return periods and can serve as reference thresholds for hydrological design and risk assessment in Almaty city. The lognormal distribution demonstrated the lowest AIC and BIC values and the highest posterior probability $P(M_i|x) > 0.35$ for the Almaty and Shymbulak meteorological stations, while the GEV was slightly better for BAL and Kemen meteorological stations. The gamma model also performed reasonably well (AIC differences < 5), suggesting its potential suitability for areas with complex mountain–urban topography.

These quantitative results confirm that extreme rainfall events have intensified in Almaty city, especially during the warm season (May–September). The observed upward trend in the magnitude of extreme precipitation highlights growing risks of urban flash floods and surface runoff overload.

From a practical standpoint, the derived rainfall quantiles can be directly applied to the design of stormwater drainage networks, flood risk assessments, and climate-resilient infrastructure planning in Almaty. For instance, a 100-year design rainfall of approx. 1,250 mm can serve as a baseline parameter for the dimensioning of retention basins and flood control structures.

References

BES Media. (2025, 2 March). *Ocherednoy liven zatopil ulitsy Almaty: livnevyye kanalizatsii, kak obychno, ne spravlyayutsya s potokom vody*. BES.media. https://bes.media/news/ocherednoy-liven-zatopil-ulitsi-almati-livnyovki-kak-vsegda-ne-spravlyayutsya-696088/

Blom, G. (1958). *Statistical estimates and transformed beta-variables*. Wiley.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach*. Springer. https://doi.org/10.1007/b97636

Chegodayev, L. (1955). *Raspolozheniye elementov na grafike dlya raboty po teorii veroyatnostey*. Gidrometeoizdat.

Cho, H. K., Bowman, K. P., & North, G. R. (2004). A comparison of gamma and lognormal distributions for characterizing satellite rain rates from the tropical rainfall measuring mission. *Journal of Applied Meteorology*, *43*(11), 1586–1597. https://doi.org/10.1175/JAM2165.1

Choudhary, K., & Kumar, R. (2025). An analysis of stormwater management with the Internet of Things (IoT). *Nature Environment and Pollution Technology*, *24*(3), B4271. https://doi.org/10.46488/NEPT.2025.v24i03.B4271

Coles, S. (2001). *An introduction to statistical modeling of extreme values.* Springer. https://doi.org/10.1007/978-1-4471-3675-0

Cunnane, C. (1978). Unbiased plotting positions – A review. *Journal of Hydrology*, *37*(3–4), 205–222. https://doi.org/10.1016/0022-1694(78)90017-3

D'iya, S. G., Gasim, M. B., Toriman, M. E., & Abdullahi, M. G. (2014). Floods in Malaysia: Historical reviews, causes, effects and mitigations approach. *International Journal of Interdisciplinary Research and Innovations*, *2*(4), 59–65. https://www.researchpublish.com/upload/book/FLOODS%20IN%20MALAYSIA-760.pdf

Dong, X., Jiang, L., Zeng, S. H., Guo, R., & Zeng, Y. (2020). Vulnerability of urban water infrastructures to climate change at city level. *Resources, Conservation and Recycling*, *161*, 104918. https://doi.org/10.1016/j.resconrec.2020.104918

Duskayev, K., Mussina, A., Rodrigo-Ilarri, J., Zhanabayeva, Z., Tursyngali, M., & Rodrigo--Clavero, M. E. (2023). Study of temporal changes in the hydrographic network of small mountain rivers in the Ile Alatau, Kazakhstan. *Hydrology Research*, *54*(11), 1420–1431. https://doi.org/10.2166/nh.2023.305

Gentilucci, M., Rossi, A., Pelagagge, N., Aringoli, D., Barbieri, M., & Pambianchi, G. (2023). GEV analysis of extreme rainfall: Comparing different time intervals to analyse model response in terms of return levels in the study area of central Italy. *Sustainability*, *15*(15), 11656. https://doi.org/10.3390/su151511656

Hazen, A. (1914). Storage to be provided in impounding reservoirs for municipal water supply. *Transactions of the American Society of Civil Engineers*, *77*(1), 1539–1640. https://doi.org/10.1061/taceat.0002563

Jayawardane, J. M. P. M., Rajapakse, R. L. H. L., & Siriwardana, C. S. A. (2024). Urban flood assessment targeting flood risk mitigation: A case study focusing on changing environments. In H. R. Pasindu, H. Damruwan, P. Weerasinghe, L. Fernando & C. Rajapakse (Eds.), *Proceedings of Civil Engineering Research Symposium 2024* (pp. 13–14). Department of Civil Engineering, University of Moratuwa. https://doi.org/10.31705/CERS.2024.7

Katz, R. W., Parlange, M. B., & Naveau, P. (2002). Statistics of extremes in hydrology. *Advances in Water Resources*, *25*(8–12), 1287–1304. https://doi.org/10.1016/S0309-1708(02)00056-8

Kreienkamp, F., Philip, S. Y., Tradowsky, J. S., Kew, S. F., Lorenz, P., Arrighi, J., Belleflamme, A., Bettmann, T., Caluwaerts, S., Chan, S. C., Ciavarella, A., De Cruz, L., Demuth, N., Ferrone, A., Fischer, E. M., Fowler, H. J., Goergen, K., Heinrich, D., Henrichs, Y., … Wanders, N. (2021). *Rapid attribution of heavy rainfall events leading to the severe flooding in Western Europe during July 2021*. World Weather Attribution Project. https://www.worldweatherattribution.org/wp-content/uploads/Scientific-report-Western-Europe-floods-2021-attribution.pdf

Kundzewicz, Z. W., Kanae, S., Seneviratne, S. I., Handmer, J., Nicholls, N., Peduzzi, P., Mechler, R., Bouwer, L. M., Arnell, N., Mach, K., Muir-Wood, R., Brakenridge, G. R., Kron, W., Benito, G., Honda, Y., Takahashi, K., & Sherstyukov, B. (2013). Flood risk and climate change: Global and regional perspectives. *Hydrological Sciences Journal*, *59*(1), 1–28. https://doi.org/10.1080/02626667.2013.857411

Kyrgyzbay, K., Kakimzhanov, Y., & Sagin, J. (2023). Climate data verification for assessing climate change in the Almaty region of the Republic of Kazakhstan. *Climate Services*, *32*, 100423. https://doi.org/10.1016/j.cliser.2023.100423

Latif, M., Syed, F. S., & Hannachi, A. (2017). Rainfall trends in the South Asian summer monsoon and its related large-scale dynamics with focus over Pakistan. *Climate Dynamics*, *48*(11), 3565–3581. https://doi.org/10.1007/s00382-016-3284-3

Ling, Z., Jing, Y., Aqeel, M., Siddiqa, F., Yan, L., & Wenlong, M. (2025). Urban Flood Modeling and Mitigation Strategies Using Remote Sensing and GIS. Scholars *Journal of Engineering and Technology*, *13*(7), 535–551. https://doi.org/10.36347/sjet.2025.v13i07.007

Mirlas, V., Zhakyp, A., Auelkhan, Y., & Anker, Y. (2024). Assessment of urbanization-related groundwater flooding process via Visual MODFLOW modeling: A case study for the northern part of Almaty city, Kazakhstan. *Journal of Flood Risk Management*, *18*(1), e13029. https://doi.org/10.1111/jfr3.13029

MontesPajuelo, R., RodríguezPérez, Á. M., López, R., & Rodríguez, C. A. (2024). Analysis of probability distributions for modelling extreme rainfall events and detecting climate change: Insights from Mathematical and Statistical Methods. *Mathematics*, *12*(7), *1093.* https://doi.org/10.3390/math12071093

Muhammad Iskandar, M. F., Ibrahim, A., Mohtar, Z. A., Pakir Mohamed Latif, M. F., & EM Yahaya, N. K. (2025). Assessing rainfall trends and variability in a climate change. *ESTEEM Academic Journal*, *21*, 91–105. https://doi.org/10.24191/esteem.v21iMarch.4892.g3087

Nakispekova, A. (2024). *Kazakhstan unites in response to flood crisis*. The Astana Times. Bringing Kazakhstan to the World. https://astanatimes.com/2024/04/kazakhstan-unites-in-response-to-flood-crisis/

Pietras, B., & Pyrc, R. (2025). Extreme short-duration rainfall and urban flood hazard: Case studies of convective events in Warsaw and Zamość, Poland. *Water*, *17*(18), 2671. https://doi.org/10.3390/w17182671

Republic State Enterprise Kazhydromet. (2025, October 2). *Weather forecast.* https://www.kazhydromet.kz/ru/

Rima, L., Haddad, K., & Rahman, A. (2025). Generalised Additive Model-Based Regional Flood Frequency Analysis: Parameter Regression Technique Using Generalised Extreme Value Distribution. *Water*, *17*(2), 206. https://doi.org/10.3390/w17020206

Stedinger, J. R., Vogel, R. M., & Foufoula-Georgiou, E. (1993). Frequency analysis of extreme events. In D. R. Maidment (Ed.), *Handbook of Hydrology* (pp. 18.1–18.66). McGraw-Hill.

Tukey, J. W. (1962). The future of data analysis. In *Breakthroughs in Statistics: Methodology and Distribution* (pp. 408–452). Springer. https://doi.org/10.1007/978-1-4612-4380-9_31

Weibull, W. (1939). *A statistical theory of the strength of materials.* Generalstabens Litografiska Anstalts Förlag.

Wilks, D. S. (Ed.) (2011). *Statistical methods in the atmospheric sciences*. Academic Press.

World Meteorological Organization [WMO]. (1989). *Calculation of Monthly and Annual 30-year Standard Normals (WMO/TD-No. 341)*. World Meteorological Organization. https://posmet.ufv.br/wp-content/uploads/2017/04/MET-481-WMO-341.pdf

World Meteorological Organization [WMO]. (2009). *Manual on estimation of Probable Maximum Precipitation (PMP).* World Meteorological Organization. https://library.wmo.int/records/item/35708-manual-on-estimation-of-probable-maximum-precipitation-pmp?offset=2

World Meteorological Organization [WMO]. (2017). *WMO Guidelines on the Calculation of Climate Normals (WMO-No. 1203).* World Meteorological Organization. https://www.agroorbi.pt/livroagrometeorologia/DocsProg/Temas&Exerc%C3%ADciosExtraPor-Cap%C3%ADtulo/Cap1_Introdu%C3%A7%C3%A3o/Docs/WMO%20Guidelines%20on%20the%20Calculation%20of%20Climate%20Normals_en.pdf

Summary

Probabilistic assessment of annual maximum precipitation in Almaty, Kazakhstan. Accurate selection of a best-fit probability distribution function for rainfall data is crucial in hydrological studies and plays a fundamental role in the planning and design of infrastructure for the city of Almaty. This study presents a comprehensive statistical and probabilistic assessment of extreme precipitation in the city of Almaty, Kazakhstan, based on annual maximum precipitation data from five meteorological stations for the period 2000–2023. Given the complex mountainous terrain and distinct seasonal precipitation regimes, selecting an appropriate distribution is particularly critical for modeling design rainfall and flood risks. The reliability of the rainfall data was verified through tests for independence and stationarity. Five theoretical probability distributions – exponential, generalized extreme value, normal, lognormal, and gamma – were evaluated using the maximum likelihood estimation method. The best-fit distribution was determined using the chi-square goodness-of-fit test. The results indicate that the generalized extreme value distribution provides the best fit for most stations, followed by the lognormal and gamma

distributions, confirming its robustness in representing extreme precipitation in mountainous urban environments such as Almaty. Furthermore, spatial variability and increasing intensity of extreme rainfall events were observed, especially during the warm season. Design rainfall estimates were calculated for various exceedance probabilities (e.g., 1%, 2%, and 10%), corresponding to return periods of 100, 50, and 10 years, respectively. These findings are critical for flood risk assessment and the development of climate-resilient urban drainage systems, highlighting the broader applicability of this distribution-fitting methodology in regions exposed to hydrological extremes.