https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

The application of new distribution in determining extreme hydrologic events such as floods

Łukasz Gruss[1], Patrick Willems[2], Paweł Tomczyk[1], Jaroslav Pollert Jr.[3], Jaroslav Pollert Sr.[3], Christoph Märtner[4], Stanisław Czaban[1] and Mirosław Wiatkowski[1]

[1]Institute of Environmental Engineering, Wrocław University of Environmental and Life Sciences, Wrocław, 50-363, Poland
[2]Department of Civil Engineering, KU Leuven, Kasteelpark Arenberg 40, 3001, Leuven, Belgium
[3]Faculty of Civil Engineering, Czech Technical University in Prague, Prague, 16629, Czech Republic
[4]M&S Umweltprojekt GmbH, Pfortenstraße 7, 08527 Plauen, Germany

Correspondence to: Łukasz Gruss (lukasz.gruss@upwr.edu.pl)

Abbreviations and acronyms in alphabetical order

a - micro-catchments
b- meso-catchments
c - macro-catchments
d - large catchments
e - very large catchments
A - catchment area [km2]
Exp – exponential distribution
FFA - Flood frequency analysis
GEV - generalized extreme value distribution
GGEV - Dual Gamma Generalized Extreme Value Distribution
GOF - Goodness of fit test
H – highland
Kurt. - kurtosis
L – lowland
LN2 - 2-parameter log-normal distribution
LN3 - 3-parameter log-normal distribution
M - mountain
MAE - Mean Absolute Error
MAF - mean annual maximum flow
MK - The Mann-Kendall trend test
MLE - maximum likelihood estimation
N - sample size
NMT - no trend
NT - negative trend
P3 - Pearson III type distribution
PT - positive trend
Qp - Peak flow
RDA - Redundancy Analysis
RMSE - Root mean square error
SD - standard deviation

https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

Skew. - empirical skewness
Var - variation
VIF - Variance Inflation Factor

Abstract. Climate change has already impacted global water resources, and it is expected to have even more severe consequences in the future. Advancing climate change will necessitate the use of new distributions that are more flexible in adapting to changes in stationarity or the presence of trends in the sample. In this work, we compare the best fit of three-parameter distributions such as lognormal, Generalized Extreme Value (GEV), Pearson type III, and a new extension of GEV - Dual Gamma Generalized Extreme Value Distribution (GGEV) under different trends in the time series and by adding criteria such as catchment area and peak flow magnitude. The research pertains to catchments in the temperate climate zone of Poland, covering 678 water gauges in 340 rivers. Based on a trend criterion, the GGEV distribution compared to the analyzed three-parameter distributions, and the GEV distribution compared to the other three-parameter distributions, were the best fit for most samples. Based on the trend criterion and catchment size it was found that the GEV distribution is best suited for micro- and meso-catchments, while the GGEV distribution is ideal for macro- to large-catchments where the series exhibits a trend, either positive or negative. The major benefit of the GGEV distribution is its flexibility when the data are influenced by temporal non-stationarities. The additional shape parameter compensates for the limitations of the other shape parameter in distributions with lighter tails. Analysis of the dependence relationships between environmental indicators such as geographic, physiographic and hydrological indicators and the distribution parameters is less conclusive. In order to test the risk of overparameterization and overfitting for the distributions with more parameters, Kolmogorov-Smirnov tests and K-Fold cross validation shows that the GEV and GGEV distributions perform better compared to the exponential and two-parameter lognormal distributions. As an overall conclusion, the study showed that for the analyzed samples in the temperate climate zone in the era of climate change, distributions that better respond to trends, like GGEV, are more likely to be applied.

1 Introduction

Climate change has already impacted global water resources, and it is expected to have even more severe consequences in the future (Dakhlaoui et al., 2019; Pokhrel et al., 2021; Połomski and Wiatkowski, 2023; Tomczyk et al., 2023; Willems, 2013). The significance of climate change lies in the substantial impacts it brings, including the increased occurrence of floods (Gruss et al., 2023; Tabari et al., 2021b). In modeling extreme hydrological events, such as floods, stochastic modeling is commonly used. This approach relies on historical data and employs probability distributions (Gruss et al., 2022; Młyński et al., 2020) to account for the uncertainty and variability of these phenomena (Szulczewski and Jakubowski, 2018). Such methods include the at-site flood frequency analysis [FFA] (Cassalho et al., 2018). The choice of probability distribution should be verified through the assumptions of stationarity and independence, as any deviation can lead to biased distributions and potentially catastrophic consequences, such as inappropriate designs that jeopardize property and human life (Ologhadien, 2021). However, the assumption of stationarity has faced increasing challenges due to the intensification of climate change and human activities (Gruss et al., 2022; Jiang and Kang, 2019; Milly et al., 2008). Many studies present series consisting of annual maximums where, for some water gauges, the assumption of stationarity, randomness, or non-monotonic trend (NMT) is not met (Cassalho et al., 2018; Szulczewski and Jakubowski, 2018). Advancing climate change will necessitate the use of new distributions that are more flexible in adapting to changes in stationarity or the presence of trends in the sample.

In many countries, two- and three-parameter distributions are used to estimate the magnitude and frequency of annual maximum streamflow (AMS) (e.g. Valentini et al., 2024 ; Gruss et al., 2022; Młyński et al., 2018; Pitlick, 1994; Rutkowska et al., 2015; Rutkowska et al., 2015; Bezak et al., 2014; Morlot et al., 2019; Šraj et al., 2016; Ul Hassan et al., 2019; Berton and Rahmani, 2024). The Pearson Type III distribution provides the best fit to both annual minimum and annual average streamflows, assuming the series is stationary but with a linear trend (Vogel and Wilson, 1996). There are also many studies

https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

among them on Pearson type 3 (P3) (Cassalho et al., 2018) log Pearson Type 3 (LP3) (Berton and Rahmani, 2024; Morlot et al., 2019) and log-normal distributions in at-site FFA (Cassalho et al., 2018). In recent decades, a significant amount of research has been dedicated to the GEV distribution. Extreme events are often better modeled using distributions with heavy tails (Karczewski et al., 2022; Karczewski and Michalski, 2022), a characteristic of the GEV distribution (Cassalho et al., 2018; Morlot et al., 2019; Otiniano et al., 2019; Rutkowska et al., 2015).

However, some extreme event data do not follow the GEV distribution because they require a more asymmetric distribution or one with a heavier tail. As a result, new classes of probability distributions have been developed that extend beyond the GEV distribution, such as the Dual Gamma GEV distribution (GGEV) (Otiniano et al., 2019). GGEV distribution is regarded as highly flexible for several reasons: 1. It introduces an additional parameter that adjusts tail weight and skewness, making it more adaptable to diverse datasets. 2. This added flexibility allows the GGEV distribution to capture the nuances of empirical data more effectively than the standard GEV distribution. 3. As a result, the GGEV distribution is often preferred in practical applications where accurate modeling of complex data is essential (Nascimento et al., 2016). The additional shape parameter enables the GGEV distribution to adapt to various data characteristics, especially in terms of tail behavior. Notably, when this parameter is less than 1, the GGEV exhibits a heavier tail than the GEV, making it more effective for modeling extreme events that may occur more frequently than lighter-tailed distributions would predict (E Silva and Do Nascimento, 2022).

Next to the influence of non-stationarities, it is well-known that various environmental factors, including land use, may significantly influence the tail of flood frequency distributions, although this depends on the region. Pitlick (1994) found that the mean annual flood is most closely correlated with watershed area, but did not find an influence of other measures of catchment physiography on the differences in flood frequency distributions. In contrast, research by Ahilan et al. (2012) confirms that the type of landscape influences the GEV distribution. Other research by Sampaio and Costa (2021) and Tyralis et al. (2019) has shown that morphological catchment characteristics correlate with these distributions. Also Kusumastuti (2007) highlighted the role of environmental factors in influencing flood frequency and the occurrence of flood events. Although single factors may not always correlate well with the distribution parameters, it may be the combined influence of multiple factors that explain the differences in flood quantiles (Allamano et al., 2009). Understanding this influence may offer valuable insights for regionalization (He et al., 2015) and reduce uncertainties in inferences made using regional FFA frameworks (Hu et al., 2020; Tyralis et al., 2019). In this study, the assumption is made that if environmental factors have an influence on distribution parameters, one can expect dependence relationships between the parameters when different distributions are calibrated to the flood data.

The aim of the study is to analyze the fit of the GGEV distribution versus three-parameter distributions (GEV, LN3, P3) to empirical data for river catchments in Poland. The study also aims to analyze the consistency of patterns exhibited by environmental factors concerning the parameters of the examined distributions and to conduct tests for overparameterization and overfitting of the analyzed distributions.

2 Study area

The research area spans 678 water gauge situated within the drainage basins of the Dniester, Dunajec, Neman, Oder, Pregoła, Vistula, and other rivers that flow into the Baltic Sea and, covering the territory of Poland in Central Europe (Figure 1). The area of Poland is located within a temperate climate zone.

https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

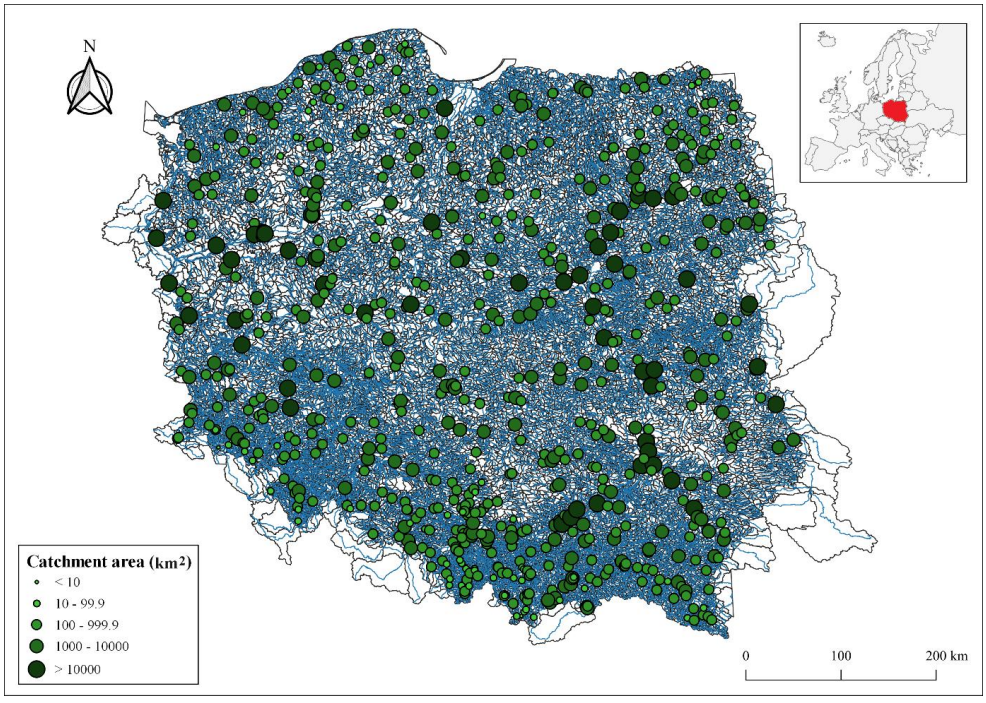

Figure 1. Location of analyzed 678 water gauges (Source: hydrographic map of Poland).

Depending on the size of the catchment area in the studied area, micro-catchments ($A < 10$ km^2), meso-catchments ($10 \leq A < 100$ km^2), macro-catchments ($100 \leq A < 1{,}000$ km^2), large catchments ($1{,}000 \leq A < 10{,}000$ km^2), and very large catchments ($A > 10{,}000$ km^2) were distinguished. The presented division criterion was adopted based on (Bertola et al., 2020). The least number of micro-catchments was recorded (for 2 stream gauge profiles), and the highest number of macro-catchments (for 388 stream gauge profiles). In between were very large catchments, meso-catchments, and large catchments, in amounts of 50, 68, and 170 stream gauge profiles, respectively (Fig. 1).

https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

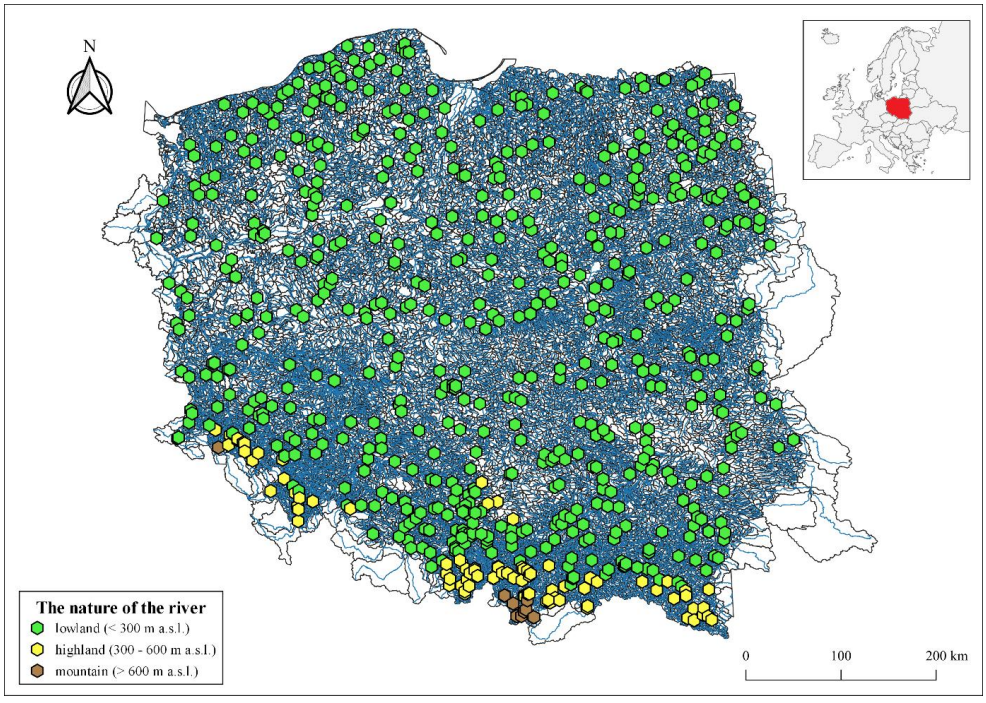

Figure 2. The terrain characteristics of the analyzed 678 stream gauge profiles consist of lowlands, uplands, and mountains (Source: hydrographic map of Poland).

The terrain of the studied area is uneven. The majority of stream gauge profiles (as many as 582 stream gauge profiles) are located in lowland areas (Fig. 2). These are catchments situated within the provinces of the Central Polish Lowland, Eastern Baltic-Belarusian Lowland, and the Czech Massif, on the Polish Uplands. A smaller quantity, specifically 86 stream gauge profiles, were located in highland areas. They are located within the provinces of the Polish Uplands, Czech Massif, and the Western Carpathians with the Western and Northern Podkarpacie, as well as the Eastern Carpathians with the Eastern Podkarpacie. Stream gauge profiles located in mountainous areas, totalling 10, are situated within the provinces of the Czech Massif and the Western Carpathians with the Western and Northern Podkarpacie, as well as the Eastern Carpathians with the Eastern Podkarpacie.

3 Methods

3.1 Data collection and extraction of flow extremes

For 678 gauge stations located in the basins of the Vistula, Oder, coastal rivers, Pregoła, and Neman, Dniester, Dunajec, maximum annual flows were collected. The source of the data (flows) is the Institute of Meteorology and Water Management - National Research Institute (IMGW-PIB). These data have been processed. The gauge stations for which only series equal to or longer than 30 years could be collected were retained (Gruss et al., 2022; Tabari et al., 2021a). The data periods used for analysis varied across stations, from 30 to 70. In this way, maximum annual flows were collected for 678 stations. The data are compiled in the hydrological year, which for Poland begins in November and ends in October. For each hydrological year, the annual maximum flow was extracted. These are hereafter referred to as peak flows (Qp) and are often associated with floods or extreme hydrological events (Gruss et al., 2022; Langridge et al., 2020; Northrop, 2004). Qp help in understanding the maximum capacity of rivers or streams to handle water, which is essential for infrastructure planning, floodplain

https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

management, and disaster mitigation efforts (Langridge et al., 2020). The Qp were utilized in this study for calibrating and evaluating the probability distributions.

Per station, the mean annual flood or mean annual maximum flow (MAF) represents the average of the Qp over the period of record (Nyeko-Ogiramoi et al., 2012; Pastor et al., 2014). Hydrologists often analyze mean annual maximum flows to understand the long-term characteristics of river systems, including flood frequency, river behavior, and water resource management (Merz and Blöschl, 2009; Nyeko-Ogiramoi et al., 2012; Pastor et al., 2014). The MAF was utilized in this study for the redundancy analysis.

3.2 Trend detection

For all analyzed time series, a test was conducted to ascertain the presence of a trend. The Mann-Kendall test was utilized for this purpose. This allowed grouping the obtained distributions into three categories: without a trend, with a positive trend, and with a negative trend.

The Mann-Kendall test (MK) is frequently used to detect monotonic trend in long time series of hydrological data (Cassalho et al., 2018; Gruss et al., 2022, 2023; Rutkowska, 2015; Svensson et al., 2005).

The null hypothesis is that the data are identically distributed, the alternative hypothesis is that the data follow a monotonic trend. A two sided test was performed and the significance level was set at 5%.

3.3 Extreme value distributions

This study considered the following type of extreme value distributions: the four-parameter GGEV distribution and the three-parameter distributions GEV, LN3, and P3.

The GGEV probability density function (PDF) proposed by (Nascimento et al., 2016), is given by:

$$f(x;\ \mu;\ \sigma;\ \xi;\ \delta) = \begin{cases} \frac{\sigma^{-1}}{\Gamma(\delta)}[1+\frac{\xi(x-\mu)}{\sigma}]^{-\left(\frac{\delta}{\xi}\right)-1}\exp\left\{-[1+\frac{\xi(x-\mu)}{\sigma}]^{-\frac{1}{\xi}}\right\}, \xi \neq 0 \\ \frac{\sigma^{-1}}{\Gamma(\delta)} exp\{-\delta[(x-\mu)/\sigma]\} exp\left\{-exp\left\{[-\frac{x-\mu}{\sigma}]\right\}, \xi \to 0\right. \end{cases}, \quad (1)$$

where:

μ – location parameter

σ – scale parameter

ξ – shape parameter

δ – shape parameter of GGEV extension.

This GGEV is a four-parameter extension of the (GEV) distribution with an additional shape parameter (δ). The additional parameter allows for varying tail weights and skewness, making it more adaptable to different types of data (Nascimento et al., 2016). A Bayesian Monte Carlo Markov Chain (MCMC) approach is used to estimate the posterior parameters of the GGEV distribution. . For the additional shape parameter δ, the optimal value is assessed after using the Akaike Information Criterion (AIC) (E Silva and Do Nascimento, 2022; Nascimento et al., 2016). The estimation of the parameters and fitting of probability distribution was done using the following R package: 'MCMC4Extremes'.

The Generalized Extreme Value Distribution (GEV) was used in many studies (Abida and Ellouze (2008), Bezak, Brilly and Šraj (2014), Cassalho et al. (2018), Kidson and Richards (2005), Szulczewski and Jakubowski (2018)). The GEV PDF function is given in equation (2):

$$f(x) = \exp[-\{1+\frac{s(x-\alpha)}{b}\}^{-1/s}], \quad (2)$$

where:

α, b, s are location, scale and shape parameters, respectively.

https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

for 1+s(x-α)/b>0, where b>0.

The parameters of this distribution were estimated by the maximum likelihood method (MLE), as described by Smith (1985). The estimation of the parameters and fitting of GEV distribution was done using the following R packages: 'evd' and 'fExtremes'.

The LN3 distribution function is given by the formula (3):

$$f(x) = \frac{1}{(x-\alpha)\sigma_y\sqrt{2\pi}} \exp\{-\frac{1}{2\sigma_y^2}[\log(x-\alpha) - \mu_y]^2\}, \quad (3)$$

where:

μ_y, σ_y^2 , α are shape, scale and location, parameters, respectively.

The LN3 is similar to the two-parameter LN2 distribution, except that x is subtracted by a value α in the former, which represents the lower bound (Cassalho et al., 2018). The parameters of this distribution were estimated by MLE as shown by Meeker and Escobar (1998). The estimation of the parameters and fitting of probability distribution was done using the following R packages: 'EnvStats' and 'weibulltools'. For the MLE method used to estimate the distribution parameters a confidence level of 0.95 was assumed.

The PDF of the P3 distribution is given by (4):

$$f(x) = \frac{1}{|s|^\alpha \Gamma(\alpha)} |x-\lambda|^{\alpha-1} e^{-\frac{x-\lambda}{s}}, \quad (4)$$

for s≠0, a>0 and $\frac{x-\lambda}{s} \geq 0$.

Where:

α, s, λ are shape, scale and location parameters, respectively.

The MLE was used to estimate the parameters for the P3 distribution. In the gamma distribution developed by Becker and Klößner (2017), this function allows negative scale parameters to allow for negative skewness. The estimation of the parameters and fitting of probability distribution was done using the R package 'PearsonDS'.

3.4 Accuracy measures

The goodness of fit of the four probability distributions to the empirical data was evaluated based on the accuracy measures Mean Absolute Error (MAE) and Root mean square error (RMSE). The MAE is recommended for leptokurtic distributions (MAE) and RMSE is preferred for platykurtic distributions (Karunasingha, 2022). Among the 678 samples, the kurtosis value exceeded 3 for 560 samples (leptokurtic distributions), while kurtosis less than 3 was observed in 118 samples (platykurtic distributions).

3.5 Redundancy analysis

Redundancy analysis (RDA) was applied as a canonical technique, to investigate the influence of environmental variables and sample characteristics on the parameters of the extreme value distributions. It aims to identify common patterns and key factors affecting the distribution parameters.

The environmental factors examined included the watershed area, categorised by catchment type, and the nature of the watercourse (Lowlands, Highlands, Mountains) (Bertola et al., 2020; Han et al., 2023; Tyralis et al., 2019). Sample characteristics considered included the highest Qp, MAF, sample size, empirical moments of standard deviation (SD), variance (Var), skewness, kurtosis, third-moment center, fourth-moment center (which measures the intensity of the distribution tails) of the Qp, and trend measures.

https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

RDA was performed separately for each distribution. The final RDA model was selected by evaluating environmental variables and sample characteristics using the Variance Inflation Factor (VIF). RDA, standardized by response variables (center and standardize) and environmental variables (center), was performed using Canoco 5.12.

3.6. Assessing overparameterization and overfitting

A distribution with more parameter does not necessarily lead to a more accurate distribution. It obviously will lead to a better goodness-of-fit because due to the higher flexibility during the calibration, but more parameters will lead to higher uncertainty in parameter calibration. So, the three- and more parameter distribution may lead to "overparameterization" and "overfitting". With more parameters, there is also a risk that distribution extrapolations may be more erroneous (Alsadat et al., 2023).
In order to evaluate whether the increased complexity of multi-parameter distributions offers a substantial improvement in fit or merely results in overfitting, the procedure shown in Figure 3 was applied.
The one- and two-parameter distributions of the exponential (Exp) and 2-lognormal (LN2) distributions were designated to serve as a reference point for evaluating the overparameterization and overfitting in the three- and four-parameter distributions GEV and GGEV.

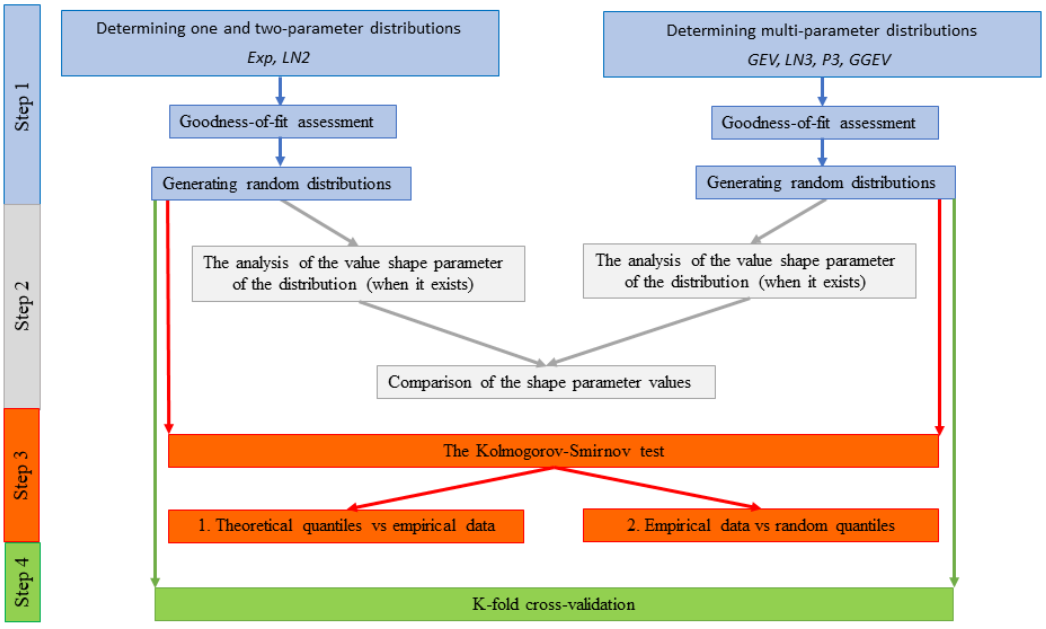

Figure 3. Workflow for evaluating overparameterization and overfitting in multi-parameter probability distributions.

The first analysis focused on examining whether the theoretical GEV and GGEV distributions significantly alter the shape parameter compared to the LN2 distribution (Fig. 3: Step 2) (Raynal-Villasenor and Raynal-Gutierrez, 2014). This investigation aimed to determine whether the GEV and GGEV distributions are more complex (overparameterized) than necessary and whether fitting these distributions improves their ability to accurately predict extreme values, particularly for very high return periods.

https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

In the second analysis, the Kolmogorov-Smirnov test was used (Kim et al., 2017) in two testing variants: 1. theoretical quantiles with empirical data and 2. empirical data with random quantiles (Fig.3: Step 3). This was done for the GEV, GGEV, Exp, and LN2 distributions. The hypothesis was that a p-value less than 0.05 would suggest rejecting the hypothesis that the samples come from the same distribution. The KS test was used to determine whether multi-parameter distributions (such as GEV and GGEV) might provide slightly better fits in some cases (variant 1) and whether they could be more prone to overfitting (variant 2) (Ozonur et al., 2021).

In the third analysis, the K-fold cross-validation (split sample test) was used to validate the distribution's performance (Fig. 3: Step 4) (Kim et al., 2017; Xu and Goodacre, 2018). In this study, we employed the k-fold cross-validation technique, specifically dividing the data series into 5 equal folds (also called 5-folds) (Rohani et al., 2018; Yadav and Shukla, 2016). The distribution is trained on k−1 subsets and tested on the remaining subset. This process is repeated k times until each subset has been used as the test set (Prusty et al., 2022). K-fold cross-validation is often used for comparing and selecting the best distribution for a given predictive problem. This method allows for evaluating which distribution generalizes best to new data (Brunner et al., 2018; Jaiswal et al., 2022). Cross-validation was performed for GEV and GGEV distributions. To check the results, the following measures were used: MAE (for leptokurtic distributions), RMSE (for platykurtic distributions) (Karunasingha, 2022). In response to the question of how these analyses would be conducted for distributions with fewer than three parameters, two additional distributions – Exp and LN2 – were selected for testing. Finally, a comparison of cross-validation results between GEV, GGEV and Exp, LN2 distributions was conducted. For the GEV and GGEV, only the distribution with the best fit following the MAE and RMSE, was considered in that analysis. That means that the total number of tested samples was 678 for each of the Exp and LN2 distributions. In contrast, there were 172 samples for the GEV distribution and 281 for the GGEV distribution.

The methods for determining the Exp and LN2 distributions and their goodness-of-fit assessment are presented in Supplementary Material S1 and S2. The generation of random samples for the Exp, LN2, GEV and GGEV distributions are described in the Supplementary Material S3.

4 Results and discussion

4.1 Goodness of fit results in relation to the trend category

Among the 678 samples, a no trend (NMT) was observed in the highest number of cases (446). Conversely, a negative trend (NT) was identified in 200 samples, while the least number of samples exhibited a positive trend (PT) (32 samples) (Figure 4, Table S1).

https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

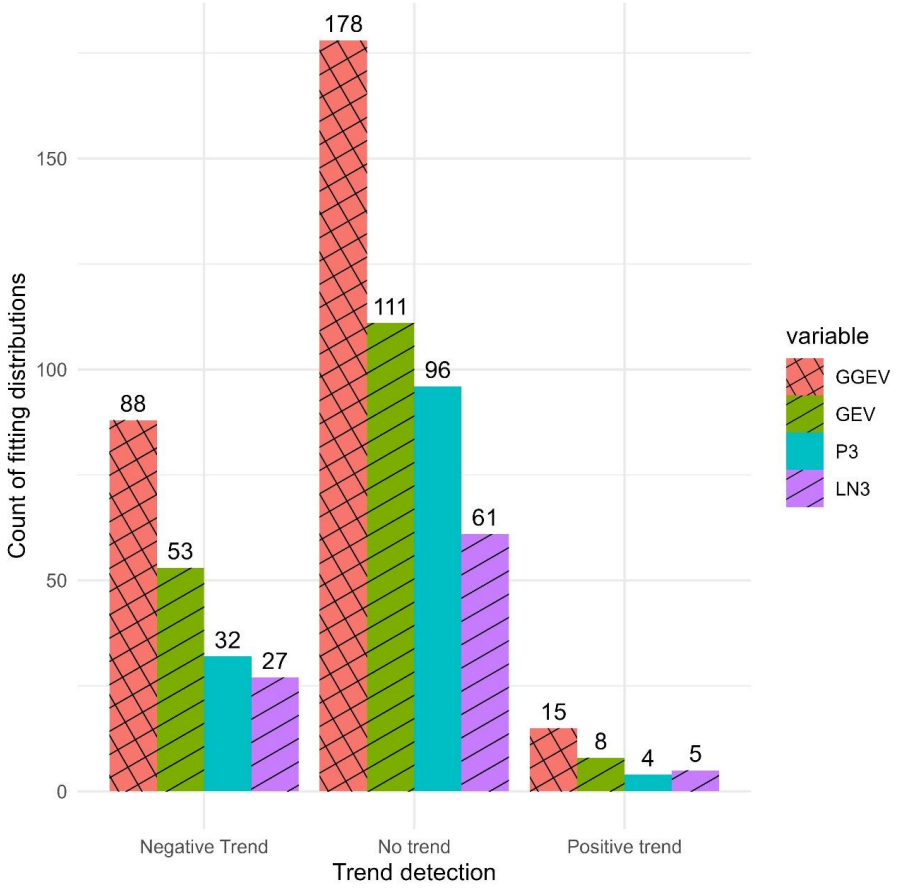

Figure. 4. Count of four fitting distributions (GEV, GGEV, LN3, P3) with marked trend categories.

In the case of no trend (NMT), the accuracy measures show a good fit for the LN3 distribution to 61 samples, the Pearson III type distribution to 96 samples, the GEV distribution to 111 samples, and the GGEV distribution to 178 samples (Figure 4). In instances of a negative trend (NT), the accuracy measures led to fitting the LN3 distribution to 27 samples, the Pearson III type distribution to 32 samples, the GEV distribution to 53 samples, and the GGEV distribution to 88 samples (Figure 4). In turn, for a positive trend (PT), the accuracy measures resulted in fitting the Pearson III type distribution to 4 samples, the LN3 distribution to 5 samples, the GEV distribution to 8 samples, and the GGEV distribution to 15 samples (Figure 4). In NMT samples, the GGEV distribution was the most frequently identified and the LN3 distribution was the least common. Similarly, for NT and PT samples, the GGEV distribution was most frequently observed, while the LN3 distribution was least frequently encountered in NT samples, and the P3 distribution was the least frequently observed in PT samples (Figure 4). Among the four examined distributions, the GGEV distribution predominates in general. The GGEV distribution prevails in terms of count for all trend categories (Figure 4, Table S1) when compared to the three-parameter distributions. This is consistent with the findings by Nascimento et al., (2016), who found for maximum monthly flow data that the best model was the generalized GGEV model rather than the GEV model.

Focusing solely on the three-parameter distributions (P3, LN3, and GEV), it is evident that the GEV distribution is most frequently fitted best, followed by the P3 and LN3 distributions (for 172, 132, and 93 samples, respectively). This applies to both NMT and NT samples. In contrast, for the PT samples, , the GEV distribution has the highest number of best-fit samples among the three-parameter distributions, followed by the LN3 distribution and, finally, the P3 distribution (Figure 4, Table

https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

S1). This is consistent with Kumar et al. (2003), who argue that the GEV distribution in terms of the L-moments was the best fit compared to P3. Bezak et al. (2014) obtained a completely different result, indicating that the best results were obtained with the P3 distribution in terms of the MLE.

4.2. Goodness of fit results in relation to the trend and catchment size categories

It was checked whether a similar pattern of results is obtained when considering catchment area size ranges (Figure 5).

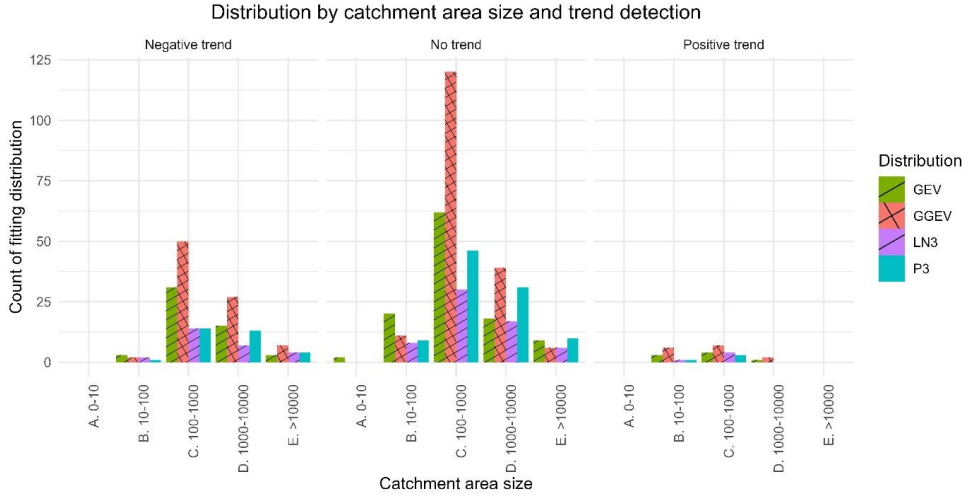

Figure 5. Count of four fitting distributions (P3, LN3, GEV and GGEV) with marked trend and catchment size categories.

For the NMT samples, the best-fitted distribution is the GEV distribution for samples where the catchment area is less than 10 km^2, in the range 10 – 100 km^2, and P3 above 10,000 km^2. Meanwhile, the GGEV distribution is the best fit for samples with catchment areas in the ranges 100-1000 km^2 and 1000 – 10,000 km^2 (Figure 5, table S2). Comparing this results for three-parameter distributions utilizing the MLE estimation method, Gruss et al. (2022) obtained findings for six data series with no trend. Among these, the Weibull, GEV, LN3, and P3 distributions were best fitted to the empirical data from sub-catchments with areas ranging from 100 to 1000 km^2. Moreover, as reported by Gruss et al. (2022) the GEV distribution was fitted best for two catchment areas ranging from 1000 to 10000 km^2. In this study, in the context of NMT samples, the least fitted distributions to the empirical data were: LN3 (samples with catchment areas in the range of 10-100 km^2, 100-1000 km^2, 1000-10,000 km^2) and additionally LN3 and GGEV for areas larger than 10,000 km^2.

There are no samples with a NT trend for catchments smaller than 10 km^2. The GEV distribution best fits to empirical data from catchments with areas in the range 10-100 km^2. Conversely, the GGEV distribution has the best fit for catchments in the ranges of 100-1000 km^2, 1000-10,000 km^2, and above 10,000 km^2 (Figure 5, table S2). This is consistent with Silva and Nascimento (2022) for catchments with areas greater than 10,000 km^2 like the Gurguéia Rive catchment in Brazil. As they reported, the GGEV distribution has a better fit than the GEV distribution. Gruss et al. (2022) concluded for Czech Republic and Poland that the Weibull distribution fits best for catchment areas ranging from 100 to 1000 km^2, the Weibull and P3 distributions for catchment areas from 1000 to 10,000 km^2, and the GEV distribution for catchment areas above 10,000 km^2. In this study, in the context of NT samples, the least fitted distributions to the empirical data are P3 (samples with catchment areas in the range of 10-100 km^2), P3 and LN3 (samples with catchment areas in the range of 100-1000 km^2), LN3 (samples with catchment areas in the range of 1000-10,000 km^2), and additionally GEV for >10,000 km^2 (Figure 5, table S2).

The fewest samples exhibit a PT trend, occurring only in catchments within the ranges of 10-100 km^2, 100-1000 km^2, and 1000-10,000 km^2. The GGEV distribution fits best for these samples (Figure 5, table S2). In the context of PT samples, the least fitted distributions to the empirical data are P3 (samples with catchment areas in the range of 10-100 km^2), and P3

https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

(samples with catchment areas in the range of 100-1000 km²). and additionally GEV (samples with catchment areas in the range of 1000-10,000 km²) (Figure 5, table S2).

For NMT samples, GEV fits best for catchments under 100 km², GGEV for 100-10,000 km², P3 above 10,000 km²; for NT, GEV suits 10-100 km², GGEV for 100-10,000 km² and over 10,000 km²; and for PT, GGEV fits for 10-10,000 km².

4.3. Goodness of fit results in relation to catchment size and peak flow

Next, it was checked whether the relationship between catchment area (A) and registered maximum peak flow (Q_p) could influence the choice of distribution.

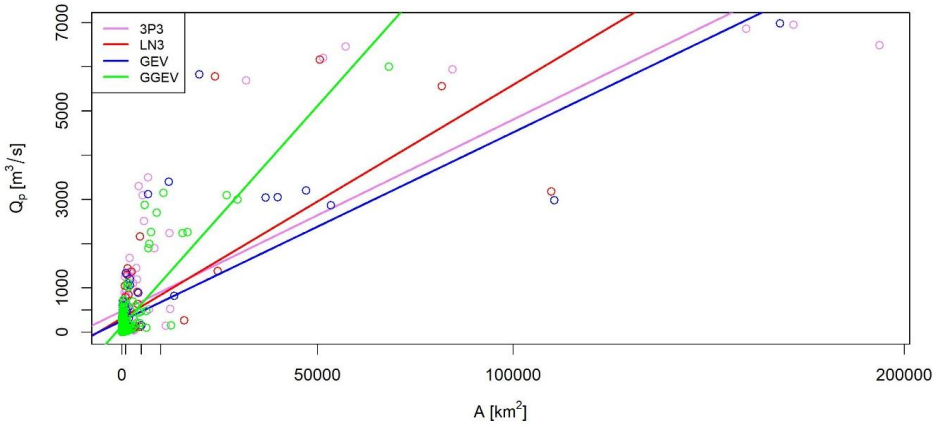

Figure 6. Relationship between catchment area (A) and peak flow magnitude (Qp) for observational series with a no trend and for four fitting probability distributions (P3, LN3, GEV and GGEV).

Probability distributions determined for NMT samples show a relationship between catchment area size (A) and peak flow magnitude (Qp), represented by a simple regression line (Figure 6, Table S3). The widest range of A is characterized by the samples for which the P3 distribution (30 - 20,000 km²) and the GEV distribution (3.5 – 170,000 km²) best fits, while LN3 (35 – 110,000 km²) and the GGEV (50 – 70,000 km²) fit more limited ranges. Moreover, the widest range of Qp is characterized by the samples for which the P3 distributions (1.9 - 7,000 m³/s) and GEV (1.6 - 7,000 m³/s) fits best, suggesting that these distributions are the most flexible in modeling extreme flows for different catchment sizes, while LN3 (8.5 to 6.500 m³/s) and GGEV (2 to 6,000 m³/s). The P3 and GEV distributions are typical distributions for FFA. These distributions also show the widest range of applicability in this study. In contrast, the LN3 and GGEV distributions show a more limited applicability (Figure 6, Table S3). Because this appears conflicting with the results in previous sections, similar analysis is done hereafter but for each trend category separately.

https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

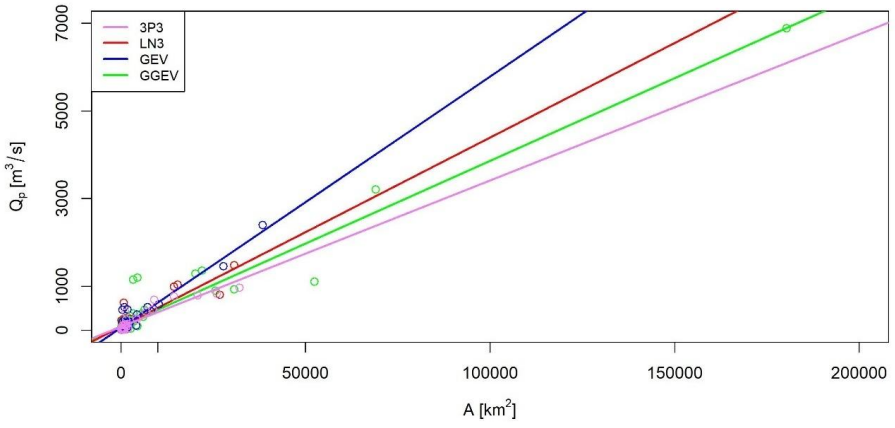

Figure 7. Relationship between catchment area (A) and peak flow magnitude (Qp) for observational series with a negative trend and for four fitting probability distributions (P3, LN3, GEV and GGEV).

345 When we focus on the observational series with a negative trend (Figure 7, Table S4), the widest range of area A is characterized by samples fitted to the distribution GGEV (80 - 180,000 km²), whereas other distributions (P3, LN3, GEV) have more limited ranges. The widest range of Qp is characterized by samples fitted to the GGEV distribution (5 – 7,000 m³/s), indicating its flexibility in modeling extreme flows for samples with a detected NT. The P3 and LN3 distributions have narrower ranges, making them less flexible for samples with a detected NT. This suggests that the GGEV distribution is
350 particularly well-suited for extreme flow events with negative trend. Moreover, the GEV distribution fits much better than the other three-parameter distributions for samples with NT.

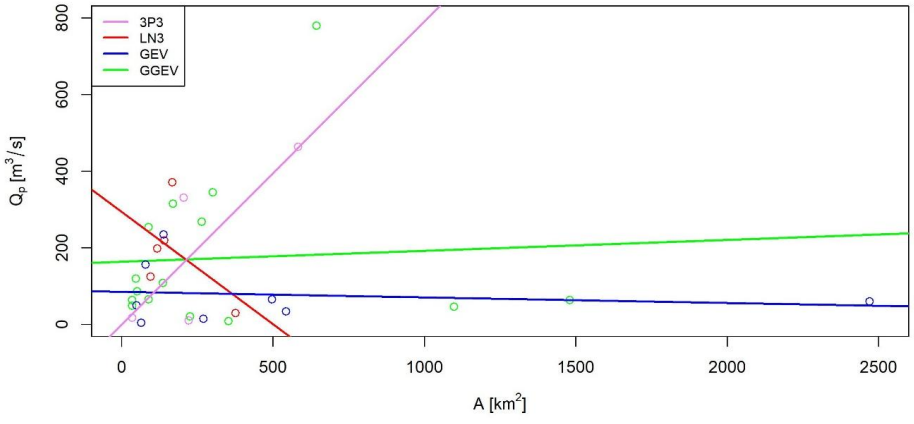

Figure 8. Relationship between catchment area (A) and peak flow magnitude (Qp) for observational series with a positive trend and for four fitting probability distributions (P3, LN3, GEV and GGEV).

355 For the positive trend samples (Figure 8, Table S5), the widest range of area A is characterized by samples fitted to the GGEV distribution (35 - 1,500 km²), while other distributions (P3, LN3, GEV) have more limited ranges. The widest range of Qp is also characterized by samples fitted to the GGEV distribution (8 - 750 m³/s), indicating its flexibility in modeling extreme flows. The P3 and LN3 distributions have narrower ranges, making them less flexible for samples with a detected positive trend. This also confirms that the GGEV distribution has the best fit for data with detected positive trends. In another study on
360 the evaluation of the GEV and LN3 distributions with L-moments estimation, Kousar et al. (2020) concluded for the Sewden River two locations with catchment areas ranging from 1000-10,000 km², and above 10,000 km² exhibit a platykurtic

https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

distribution and fit best to the GEV. In turn, the LN3 distribution was the best fit for three other locations that exhibit a leptokurtic distribution for areas ranging from 100-1000 km^2, 1000-10,000 km^2, and above 10,000 km^2 (Kousar et al., 2020).

4.4 Influence of environmental factors on probability distribution parameters

Redundancy analysis (RDA) was performed separately for each distribution to examine whether environmental factors have a similar impact on the parameters of the distributions.

The independent variables are catchment area ranges: a - micro-catchments, b - meso-catchments, c – macro-catchments, d – large catchments, e - very large catchments; flow peak (Qp); trend (NMT is no trend, PT is positive trend and NT is negative trend); the nature of the watercourse (L - lowlands, H - highlands, M - mountains); empirical skewness (Skew.) and 4th center moment (4thMoment); empirical kurtosis (Kurt.), sample size (N) (Fig. 9a-9d). The response variables are add. shape (Fig. 9), location, scale, shape (Fig. 9-12, Table S3),

Since the mean, standard deviation, variance, third moment, and 4thMoment are interrelated, it is essential to carefully select the set of explanatory variables. It was also confirmed that multicollinearity exists between skewness and kurtosis as explanatory variables (Variance Inflation Factor, VIF > 10). Collinearity between skewness and kurtosis may result from the fact that both of these measures are defined using the standard deviation (SD). Therefore, RDA was conducted separately for kurtosis (Fig. 9a, 10a, 11a, 12a) and skewness (Fig. 9b, 10b, 11b, 12b). Additionally, RDA was performed with the inclusion of catchment area ranges and Kurt. was replaced with the 4thMoment and Skew. (see Figures 9c, 10c, 11c, and 12c).

The decision to replace Kurt. with the 4th Moment was made because both Skew. and Kurt. are functions of the standard deviation, making them potentially collinear. The use of the Skew. and the 4thMoment can allow for capturing more detailed aspects of the data distribution. The 4thMoment measures the overall Kurt., which is the tail heaviness of the distribution, while Kurt. is the normalized version of this moment. Following the initial RDA, subsequent analyses considered only the changes that were not identified in the first analysis. A detailed description of the RDA analyses for the distributions is provided in the Supplementary Materials Section S4.

The use of topography in modeling Qp helps to uncover the runoff mechanism prevailing in the catchment (Valeo, 2013).

4.4.1 GGEV distribution

The first two axes (RDA 1 and RDA 2) explain 54.00% of the variance (45.76% and 8.24%, respectively) (Fig. 9a).The Qp and A are strongly correlated with RDA 1 and Kurt. with RDA 2. According to the response variables, scale and location are related to RDA 1. Shape and add. shape are related to RDA 2. Scale and Location are strongly positive correlated with Qp (score 0.95) and A (score 0.94). The shape and add. shape are positive related to Kurt. (score 0.90) The add. shape and shape are inversely proportional to N (score -0.14). The H (score -0.68), L (score 0.70), M (score -0.14), PT (score -0.33) and NT (score 0.25) are correlated with RDA 3. Samples with NMT (score 0.14) do not affect the distribution parameters for add. shape and shape.

a)	b)	c)

https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

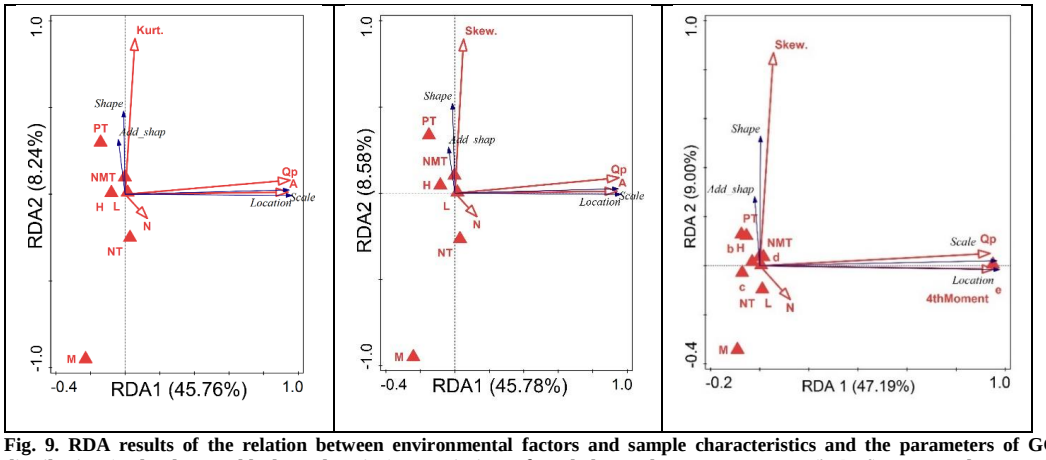

Fig. 9. RDA results of the relation between environmental factors and sample characteristics and the parameters of GGEV distribution (scale, shape, add. shape, location). Descriptions of symbols: catchment area ranges (in km²): a - < 10, b - 10-100, c – 100 – 1000, d – 1000-10000, e - > 10000; flow peak (Q_p); trend (NMT is no trend, PT is positive trend and NT is negative trend); nature of the watercourse (L - lowlands, H - highlands, M - mountains), parameters of the empirical sample (Skew. is empirical skewness, Kurt. is empirical kurtosis and 4thMoment is a 4th center moment), sample size (N).

In the second RDA the first two axes (RDA 1 and RDA 2) explain 54.36% of the variance (45.78% and 8.58%, respectively) (Fig. 9b). Compared to the previous RDA, more significant changes, as described below, are observed. The Qp and A are strongly correlated with RDA 1 and Skew. with RDA 2. The shape and add. shape are positive related to Skew. (0.89). The H (score -0.69), L (score 0.73) and PT (score -0.29) are correlated with RDA 3. In turn, the M (score 0.30) is correlated with RDA 4. Samples with NMT (score 0.14) and NT (score -0.18) do not affect the distribution parameters for add. shape and shape.

Both Skew. and Kurt. (Fig. 9a-b), are correlated with the shape and add. shape parameters, which may support the additional parameter mechanism described by Nascimento et al. (2016) for this distribution.

In the third RDA the first two axes (RDA 1 and RDA 2) explain 56.19% of the variance (47.19% and 9%, respectively) (Fig. 9c). Thus, the explained variance is slightly higher compared to the previous two canonical analyses for the GGEV distribution. Compared to the previous RDA, more significant changes, as described below, are observed. The Qp (e) and 4thMoment are strongly correlated with RDA 1 and Skew. with RDA 2. Scale is strongly positive correlated with Qp and location is strongly positive correlated with (e) and 4thMoment (Fig. 9c). This means that higher Qp values correspond to a larger scale parameter in the GGEV distribution. Larger catchment areas (e) lead to an increase in the location parameter, which shifts the central point of the distribution. Observing the biplot (Fig. 9a), it is noted that the Kurt. parameter affects the shape and add. shape, while the 4thMoment influences the location. Meanwhile, the third moment strongly correlates with the scale (not shown on the graph). An increase in the 4thMoment, which measures the concentration of values around the mean and is also related to Kurt., indicates an increase in the value of the location parameter. The location parameter determines where the center of the distribution is located on the number line. The greater the 4thMoment, the higher the location parameter in a heavy-tailed distribution. This means that, where more extreme values occur, the central tendency of the distribution (measured by the location parameter) shifts towards these higher values to better reflect the influence of extremes on the distribution. This means that higher values of the 4thMoment cause the central value or location of these extreme values to shift towards higher values. If, with an increase in the 4thMoment, the location parameter increases, it means that the center of the distribution shifts to the right on the number line. In turn, the 3rd central moment correlates with the distribution parameter known as scale because the scale affects the magnitude of deviations from the mean, and the 3rd central moment measures precisely these deviations.

The shape is positive related to Skew. The add. shape and shape are inversely proportional to N(Fig. 9c). In practice, this might suggest that with larger N, the distribution becomes less extreme or lighter. The shape parameters likely adjust to reflect a more stable and less variable distribution as the amount of data increases. Shape and add. shape are negative correlated with

https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

NT (Fig. 9c). This may indicate that in situations where there is a downward trend in the data, the distribution becomes less varied or more flattened. A weak correlation could suggest that with a NT, the values of add. shape and shape may slightly decrease.

We observe that add. shape is not as strongly correlated with Skew. as shape is. Add. shape serves as a supplementary parameter, and the canonical analysis shown in Fig. 9a-b indicated that add. shape has similar properties to shape. Since the 4thMoment is not associated with RDA2, it will not directly influence shape and add. shape, or its impact will be limited for the samples examined (Fig. 9a). However, the 4thMoment used to determine Kurt. will cause Kurt. to strongly correlate with shape parameters (Fig. 9c).

Scale and location are negative correlated with (c) (Fig. 9c). Moreover, (b), (d), H, M, L and PT do not have influence to the distribution parameters (Fig. 9c). This suggests that terrain topography does not have a direct impact on the parameters of the GGEV distribution. Additionally, it can be suggested that not all types of catchments influence the shaping of distribution parameters. Very large catchments (e) have a strong positive impact, while macro-catchments (c) have a weak influence, and there is no effect on the parameters of meso-catchments (b) and large catchments (d). This may be because distribution parameters affecting larger areas may not have as strong an impact on smaller catchments, where local effects dominate over the effects associated with distribution parameters (Arnaud et al., 2011; Roodsari and Chandler, 2017). An additional advantage of the distribution is its weak sensitivity to trends. Only a NT affects the shape parameter and the additional shape parameter. Shape is correlated with Skew. and Kurt. of the empirical data. This means that the shape parameter influences the asymmetry and tail distribution of empirical flow data, which is consistent with the description by (Nascimento et al., 2016).

This also means that temporal trends such as NMT and NT do not affect the parameters of the GGEV distribution.

In summary, distribution parameters are more closely related to the hydrological characteristics of flows than to geographic or temporal features.

The RDA analysis indicates that the GGEV distribution is anticipated to be the least sensitive to landscape forms and N.

4.4.2 GEV distribution

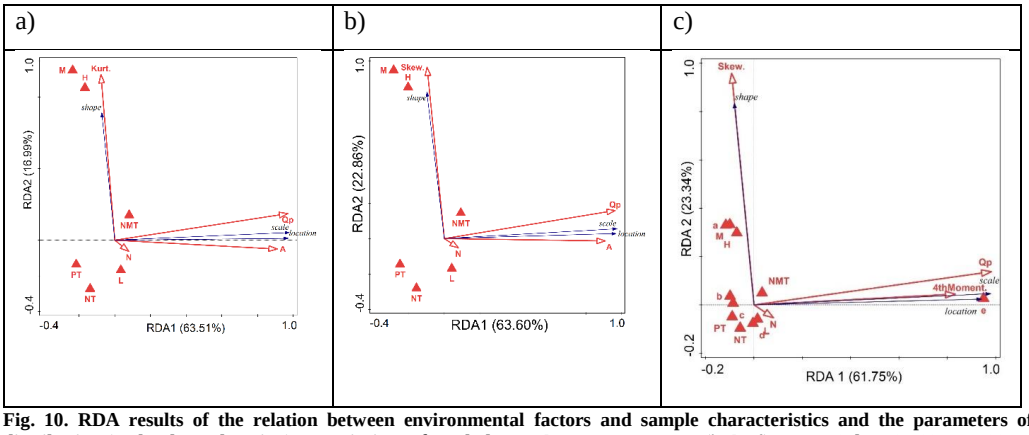

Fig. 10. RDA results of the relation between environmental factors and sample characteristics and the parameters of GEV distribution (scale, shape, location). Descriptions of symbols: catchment area ranges (in km²): a - < 10, b - 10-100, c – 100 – 1000, d – 1000-10000, e - > 10000; flow peak (Q_P); trend (NMT is no trend, PT is positive trend and NT is negative trend); nature of the watercourse (L - lowlands, H - highlands, M - mountains), parameters of the empirical sample (Skew. is empirical skewness, Kurt. is empirical kurtosis and 4thMoment is a 4th center moment), sample size (N).

The first two axes (RDA 1 and RDA 2) explain 80.50% of the variance (63.51% and 16.99%, respectively) (Fig. 10a). The Qp and A are strongly correlated with RDA 1 and Kurt. with RDA 2. According to the response variables, scale and location are related to RDA 1, and shape is related to RDA 2. It exhibits a strong correlation with the scale and location parameters, whereas this relationship is not observed for the shape parameter (Tabari et al., 2021b). Scale and location are strongly positive correlated with Qp (score 0.96) and A (score 0.91) (Fig. 10a). This is consistent with the findings of (Villarini and Smith,

https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

2010). What is particularly noteworthy is that hydrological signatures related to flow magnitude, such as the location and scale parameters, are primarily dependent on A, which significantly influences their values, while other attributes have a lesser impact on the response variable. This is consistent with previous findings by He et al. (2015), Northrop (2004), and Tyralis et al. (2019).

The shape is positive related to Kurt. (score 0.93), H (score 0.41), M (score 0.18), and NMT (score 0.22) (Fig. 10a). The shape is inversely proportional to L (score -0.45) and NT (score -0.21) (Fig. 10a). However, the shape parameter is more likely linked to hydrological processes and meteorological conditions than to catchment area (He et al., 2015). The N (score 0.14) is correlated with RDA 3 (Fig. 10a).

In the second RDA, the first two axes (RDA 1 and RDA 2) explain 54.36% of the variance (63.60% and 22.86%, respectively) (Fig. 10b). Compared to the previous RDA, more significant changes, as described below, are observed. The Qp and A are strongly correlated with RDA 1 and Skew. with RDA 2. The shape is positive related to Skew. (score 0.97), NMT (score 0.20), H (score 0.35), and M (score 0.15) (Fig. 10b). Scale and Location are inverse proportional to PT (score -0.06) (Fig. 10b).

In the third RDA, the first two axes (RDA 1 and RDA 2) explain 85.09% of the variance (61.75% and 23.34%, respectively) (Fig. 10c). Compared to the previous RDA, more significant changes, as described below, are observed. The Qp, e and 4thMoment are strongly correlated with RDA 1 and Skew. with RDA 2. Scale is strongly positive correlated with Qp (score 0.98) and location is strongly positive correlated with e (score 0.76) and 4thMoment (score 0.83) (Fig. 10c). This is consistent with (Tabari et al., 2021b), which report that the scale parameter of the GEV distribution, representing the deviation around the mean and serving as an indicator of the variance (Tabari et al., 2021b). The location parameter indicates the center of the distribution, acting as an indicator of the mean (Tabari et al., 2021b).

The shape is positive related to Skew. (score 0.96) (Fig. 10c). The shape parameter determines the tail behavior of the distribution (He et al., 2015). Specifically, higher values of the shape parameter lead to heavier tails (Tabari et al., 2021b; Tyralis et al., 2019; Villarini and Smith, 2010). Shape is negative correlated with L (score -0.39), NT (score -0.19), and d (score -0.11) (Fig. 10c). Scale and location are negative correlated with PT (score -0.05), c (-0.29). The NMT (score 0.19) had weak effect on shape. In turn, H (score 0.35) and M (score 015) had a weak relation to the shape (Fig. 10c). The shape parameter of the GEV distribution is correlated with nature (terrain elevation) (Sampaio and Costa, 2021; Tyralis et al., 2019). However, morphologic characteristics of the catchments in the regression model for the GEV shape parameter is small (Sampaio and Costa, 2021). Capturing the spatial variation of the GEV shape parameter by means of covariates, such as terrain elevation, remains a challenging task (Sampaio and Costa, 2021). On the other hand, (Ahilan et al., 2012) research confirms that the type of landscape affects the distribution of GEV. The shape parameter of the GEV distribution determines the behavior of the upper tail. Specifically, higher values of the shape parameter lead to heavier tails. The shape parameter dependency is mainly influenced by climatic indices, while other catchment characteristics are less significant (Tyralis et al., 2019). This is consistent with He et al. (2015) which found no relationship between the shape parameter and catchment area, suggesting that hydrological heterogeneity is implicitly captured by the shape parameter.

Moreover, N (score 0.293), b (score -0.16) is related to RDA 3 (Fig. 10c).

RDA analysis confirms that the shape of the distribution is strongly dependent on the Skew. of the empirical sample (Fig. 10b-c). The location parameter of the GEV distribution is positively correlated with the weight and the tail distribution in the empirical data.

For the impact of catchment size on GEV parameters, Villarini and Smith (Villarini and Smith, 2010) found that scale and location are positively correlated, while shape is negatively correlated. The magnitude of the shape parameter of the GEV distribution depends on the location of the gauge, whether it is in lowlands, highlands, or mountains (Villarini and Smith, 2010).

In sample (e), which is a very large catchment, the 4thMoment affects the location parameter and is strongly positively correlated with the location. In sample (c), the 4thMoment is weakly negatively correlated with the location and with PT.

https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

Landscape forms and trends, as well as catchment types (with the exception of micro-, meso-, and very large catchments), have a weak influence on the parameters of this distribution.

The scale parameter of the GEV distribution in FFA is strongly correlated with Qp and MAF (Villarini and Smith, 2010). This relationship is further supported by the use of scale-invariant statistics, which show good correlations with historical flood-frequency records (Turcotte 1993). However, it is important to note that the scale parameter can vary over time, as demonstrated by the application of a non-stationary GEV model to account for changing streamflow series (Jiang 2019).

The shape parameter of the GEV distribution in FFA is a critical factor, but its determinants have been elusive (Tyralis 2019). While it is known to influence the upper tail of the distribution, its relationship with catchment attributes is not well understood. Morrison and Smith (2002) found that the shape parameter is not dependent on catchment morphological parameters or land cover properties, suggesting that other factors may be at play. Sampaio (2021) and Kumar (2003) both highlight the importance of the GEV distribution in regional FFA, but do not specifically address the relationship between the shape parameter and highlands area. Indeed, in their work (Northrop, 2004), they analyzed the relationship between the location, scale, and shape parameters of the GEV distribution, among other factors, of annual maxima and catchment descriptors like area and base flow index. In this study, the MLE method was also used for estimating distribution parameters. The work indicates a linear relationship between the location and scale parameters, which means that as the catchment area increases, so do these parameters. Current research confirms this trend (Fig. 10a-b). On the other hand, regarding the shape parameter, Northrop (2004) states that there is no trend. Current studies show a negative trend, which is explained by Tyralis et al. (2019). According to Tyralis et al. (2019), the shape parameter exhibits a negative linear correlation with the catchment mean elevation. As elevation increases, the value of the shape parameter slightly decreases (Fig. 10a-c).

4.4.3 LN3 distribution

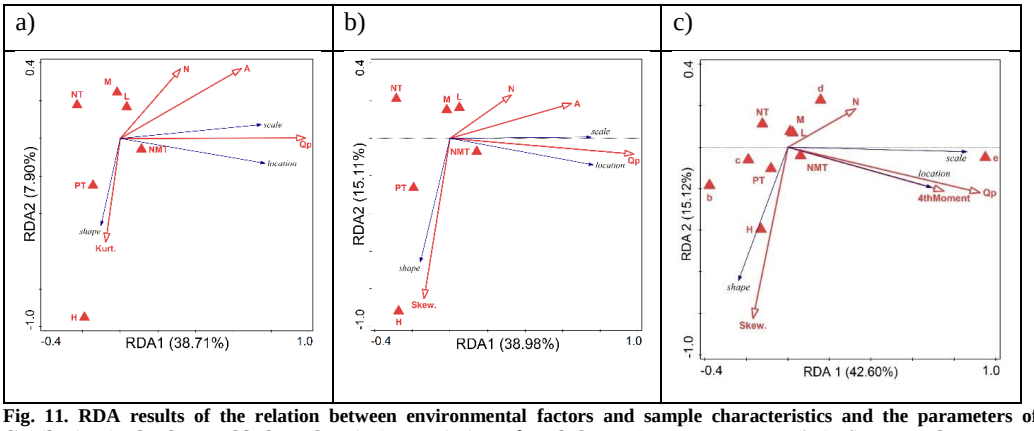

Fig. 11. RDA results of the relation between environmental factors and sample characteristics and the parameters of LN3 distribution (scale, shape, add. shape, location). Descriptions of symbols: catchment area ranges (in km²): a - <10, b - 10-100, c – 100 – 1000, d – 1000-10000, e - > 10000; flow peak (Q_p); trend (NMT is no trend, PT is positive trend and NT is negative trend); nature of the watercourse (L - lowlands, H - highlands, M - mountains), parameters of the empirical sample (Skew. is empirical skewness, Kurt. is empirical kurtosis and 4thMoment is a 4th center moment), sample size (N).

The first two axes (RDA 1 and RDA 2) explain 46.61% of the variance (38.71% and 7.90%, respectively) (Fig. 11a).The Qp and A are strongly correlated with RDA 1 and Kurt. with RDA 2. According to the response variables, scale and location are related to RDA 1, and shape is related to RDA 2. Scale and location are strongly positive correlated with Qp (score 0.98) and A (score 0.64) (Fig. 11a). The shape parameter of the LN3 has found to be correlated with extreme flows and catchment area (Haktanir and Horlacher, 1993; Smith, 1989).

The shape is negative related to Kurt. (score -0.55) and H (score -0.51) (Fig. 11a). The shape is inversely proportional to N (score 0.37) (Fig. 11a). Location is weakly correlation to NMT (score 0.2) (Fig. 11a). The shape is weakly correlated to N

https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

(score 0.37) and inverse proportional to H (score -0.51). The PT (score 0.21), NT (score -0.28), M (score 0.2), and L (score -0.56) are correlated with RDA 3 (Fig. 11a).

In the second RDA, the first two axes (RDA 1 and RDA 2) explain 54.09% of the variance (38.98% and 15.11%, respectively) (Fig. 11b). Compared to the previous RDA, more significant changes, as described below, are observed. The Qp and A are strongly correlated with RDA 1 and Skew. with RDA 2. The shape is negative related to Skew. (score -0.85) and H (score -0.39) (Fig. 11b). The location is weakly inversely proportional to N (scores 0.33) (Fig. 11b). Location is weakly correlated to NMT (score 0.2) (Fig. 11b). The PT (score 0.18), NT (score -0.22), M (score 0.2), and L (score -0.39) are correlated with RDA 3 (Fig. 11b).

In the third RDA, the first two axes (RDA 1 and RDA 2) explain 57.72% of the variance (42.60% and 15.12%, respectively) (Fig. 11c). Compared to the previous RDA, more significant changes, as described below, were observed. The Qp, (e) and 4thMoment are strongly correlated with RDA 1 and Skew. with RDA 2 (Fig. 11c). Scale is strongly positive correlated with (e) and location is strongly positive correlated with Qp and 4thMoment (Fig. 11c). The shape is negative related to Skew. and weakly correlated with H, while positive with (d). Location is negative correlated with NT (Fig. 11c). Scale is negative correlated with (b) and (c), while positive correlated with N (Fig. 11c). As reported (Kamal et al., 2017) the larger the N, the better the result for the LN3. Moreover, M, L, NMT and PT do not have influence on the distribution parameters (Fig. 11c).

4.4.4 P3 distribution

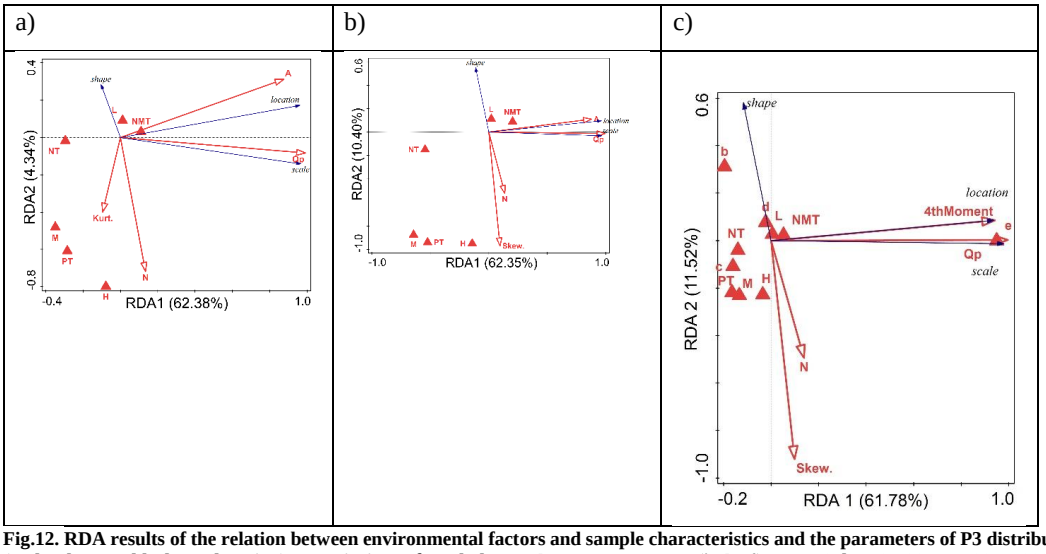

Fig.12. RDA results of the relation between environmental factors and sample characteristics and the parameters of P3 distribution (scale, shape, add. shape, location). Descriptions of symbols: catchment area ranges (in km²): a - <10, b - 10-100, c – 100 – 1000, d – 1000-10000, e - > 10000; flow peak (Q_p); trend (NMT is no trend, PT is positive trend and NT is negative trend); nature of the watercourse (L - lowlands, H - highlands, M - mountains), parameters of the empirical sample (Skew. is empirical skewness, Kurt. is empirical kurtosis and 4thMoment is a 4th center moment), sample size (N).

The first two axes (RDA 1 and RDA 2) explain 66.72% of the variance (62.38% and 4.34%, respectively) (Fig. 12a).The Qp and A are strongly correlated with RDA 1 and N with RDA 2. According to the response variables, scale and location are related to RDA 1, and shape is related to RDA 2. Scale is strongly positive correlated with Qp (score 0.99) and Location is strongly positive correlated with A (score 0.88) (Fig. 12a). This contrasts with the findings of Hebson and Wood (1986); Hu et al. (2020), Farooq et al. (2018), Flynn et al. (2006); Ribeiro-Correa and Rousselle (1993) who concluded that the scale parameter of the P3 is typically strongly correlated with increasing catchment area.

https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

The shape is strongly negative correlated with N (score -0.72) (Fig. 12a). In turn, the shape is weakly negative correlated with Kurt. (score -0.4) and H (score -0.26). The shape is weakly positive correlated with L (score 0.27) (Fig. 12a). The M (score 0.08), NT (score -0.17), PT (score 0.12), and NMT (score -0.22) are correlated with RDA 3 (Fig. 12a).

In the second RDA, the first two axes (RDA 1 and RDA 2) explain 66.72% of the variance (62.35% and 10.40%, respectively) (Fig. 12b). Compared to the previous RDA, more significant changes, as described below, are observed. The Qp and A are strongly correlated with RDA 1, and Skew. and N with RDA 2 (Fig. 12a). The shape is strongly negatively correlated with Skew. (score -0.97) and N (score -0.52) (Fig. 12a). The shape is weakly negative correlated with H (score -0.17), M (score -0.04) and PT (score -0.09). The shape is weakly positive correlated with L (score 0.17) (Fig. 12a). The NMT (score 0.18) and NT (score -0.17) are weakly correlated with location (Fig. 12a).

In the third RDA, the first two axes (RDA 1 and RDA 2) explain 73.30% of the variance (61.78% and 11.52%, respectively) (Fig. 12c). Compared to the previous RDA, more significant changes, as described below, are observed. The Qp, 4thMoment and (e) are strongly correlated with RDA 1 and Skew. with RDA 2. Scale is strongly positive correlated with Qp, and location is strongly positive correlated with (e) and 4thMoment (Fig. 12c). In contrast to the results shown in Fig. 12a-b Hebson and Wood, 1986; Hu et al., 2020, Farooq et al., 2018; Flynn et al., 2006; Ribeiro-Correa and Rousselle, 1993 came to the same conclusions, stating that the scale parameter of the P3 is indeed strongly correlated with increasing catchment area in FFA.

The larger the Qp, the greater the scale, and the larger the 4thMoment – especially for the largest catchments (e) – the greater the location (Fig. 12c). In the current analysis, Qp and MAF are strongly correlated with the scale (Fig. 13-14).

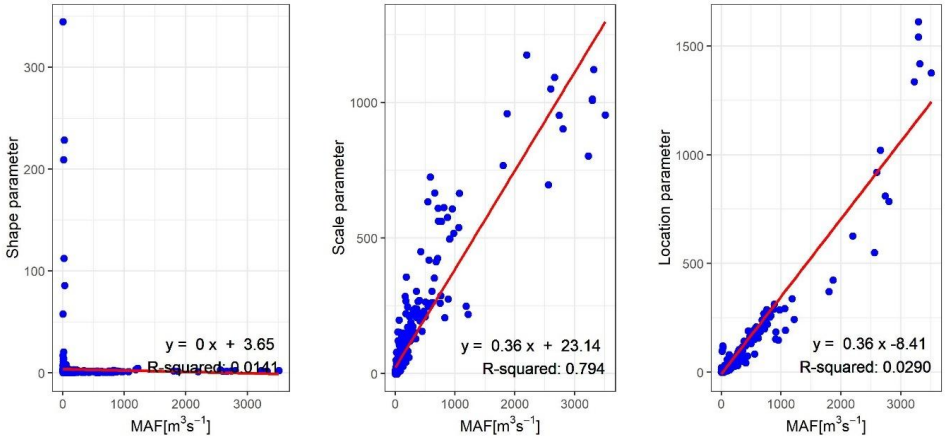

Fig. 13. Scatter plots of the P3 distribution parameters versus predictor variable MAF.

https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

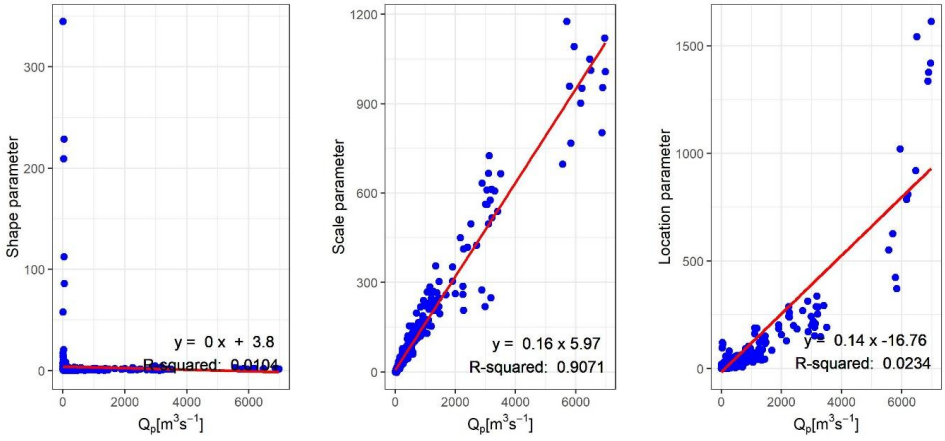

Fig. 14. Scatter plots of the P3 distribution parameters versus predictor variable Qp.

In previous work by (Hu et al., 2020), the location parameter of the log-P3 distribution was not found to be correlated significantly with flow, type of trend, landscape, or sample size, which is consistent with current findings. However, as noted by Hu et al. (2020), the location parameter of the P3 distribution is also not significantly correlated with catchment area. Our findings confirm this but also show that it is positively and significantly correlated only for large catchments (A>10,000 km^2). The shape parameter is inversely proportional to Skew. and N (Fig. 12c). This means that with smaller sample sizes and lower Skew., the shape parameter is larger. As reported by Hu et al. (2020), Skew. in the log-P3 distribution is very sensitive to N. On the other hand, Jia et al. (2023) found that a trimmed L-moments method allowed for good P3 parameter estimation even for samples with small sizes and skewness greater than twice the coefficient of variation.

The parameters (b) and (d) have a weak positive relation to the shape parameter (Fig. 12c). In catchments (b) and (d), the shape parameter increases. The location parameter is weakly negatively correlated with (c) and PT (Fig. 12c). Catchments (c) and detected PT lead to a decrease in the location parameter (Fig. 12c). In contrast, H, L, M, NT, and NMT do not influence the distribution parameters (Fig. 12c). Thus, distribution parameters are not sensitive to landscape forms or the absence of trends or negative trends. This contrasts with the findings of Farooq et al. (2018), Jain and Singh (1987), and Vogel and McMartin (1991), who concluded that the shape parameter of the P3 distribution is strongly correlated with the NMT. Moreover, Valeo and Rasmussen (2000), who investigated the log-P3 distribution, states that not only the catchment area as an independent variable determines the Qp rate, but also the topographical distribution.

Konrad (2021) highlighted the importance of understanding trends in annual peak streamflow, which can be influenced by factors such as reservoir operation and urban development (Konrad and Restivo, 2021).

4.4.5 Key points on the influence of environment factors

The following points summarize key findings regarding the relationships between environmental factors and the parameters of the studied probability distributions:

- GGEV and P3 distributions tend to have the parameters H, M, L located outside RDA1.
- GGEV and P3 share a common feature of a negative correlation between N and shape, while GEV and LN3 exhibit more complex correlation patterns.
- The GGEV, GEV, and LN3 distributions show similar correlations between the parameters A, c, and e in the context of RDA1, whereas P3 differs in this respect.
- GGEV and GEV share a pattern where NMT appears in RDA2. LN3 shows a broader presence in RDA1 and RDA2, while P3 has a different configuration.

https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

- The pattern for GGEV differs from the other distributions because, in GEV, LN3, and P3, skewness is distinctly concentrated only in RDA2.

Based on the above comparison, the GGEV distribution shows some similarities with other distributions regarding the occurrence and correlation of the distribution parameters. However, there are differences in certain aspects, such as the distribution of parameters in the principal components and parameter correlations, which indicates unique characteristics of GGEV compared to GEV, LN3, and P3. GGEV often differs from other distributions in how its parameters spread within the principal component space, which may be significant when modeling and interpreting extreme flow data analysis results.

4.5. Overparameterization check

In order to evaluate the overparameterization or overfitting problem, results are hereafter summarized for the 4 steps of the methodology outlined in Figure 3.

Step 1

Out of the two distributions (Exp and LN2), only LN2 demonstrates the best fit across all 678 profiles based on accuracy measures. Nevertheless, we performed the tests for both distributions. Additionally, these distributions were evaluated against three- and four-parameter distributions (GEV and GGEV) using the same criteria.

Step 2

The GEV distribution, analyzed for 172 profiles, exhibited shape parameter values consistently near zero, remaining below one. In turn, the fitted LN2 distribution has a shape parameter value greater than 1 for only seven stations. In contrast, the GGEV distribution has a shape parameter value greater than 1 for a smaller number of four stations. It is worth noting that the additional shape parameter reached a value greater than 1 for 45 profiles out of the 281 analyzed. However, as shown by the RDA analysis, the contribution of the add. shape parameter relative to the shape parameter is smaller. This may suggest that the add. shape parameter primarily compensates for the limitations of the shape parameter in distributions with lighter tails. The details are in Supplementary Materials Tables S3-S5. This is confirmed by the research of Nascimento et al. (2016), who state that when the additional shape parameter is less than 1, the GGEV exhibits a heavier tail than the GEV, making it more effective for modeling extreme events that may occur more frequently than lighter-tailed distributions would predict.

Step 3

The Kolmogorov-Smirnov test comparing theoretical quantiles (Exp distribution) with empirical data found no significant differences (p-value > 0.05) for 372 out of 678 profiles, indicating agreement between the distributions. Similarly, when comparing empirical quantiles to random samples (Exp distribution), 279 profiles showed no significant differences, suggesting a comparable nature of the empirical and random distributions in these cases. This corresponds to a fit rate of 74.9% (279/372).

A better fit was obtained for the LN2 distribution. The Kolmogorov-Smirnov test comparing theoretical quantiles (LN2 distribution) with empirical data found no significant differences (p-value > 0.05) for 678 out of 678 profiles, indicating agreement between the distributions. Similarly, when comparing empirical quantiles to random samples (LN2 distribution), 661 profiles showed no significant differences, suggesting a comparable nature of the empirical and random distributions in these cases. This corresponds to a fit rate of 97.5% (678/661).

A much better fit was obtained for the GEV and GGEV distributions. The Kolmogorov-Smirnov test comparing theoretical quantiles (GEV distribution) with empirical data found no significant differences (p-value > 0.05) for 172 out of 172 profiles, indicating agreement between the distributions. Similarly, when comparing empirical quantiles to random samples (GEV distribution), 171 profiles showed no significant differences, suggesting a comparable nature of the empirical and random distributions in these cases. This corresponds to a fit rate of 99.4% (171/172). The Kolmogorov-Smirnov test comparing theoretical quantiles (GGEV distribution) with empirical data found no significant differences (p-value > 0.05) for 281 out of 281 profiles, indicating agreement between the distributions. Similarly, when comparing empirical quantiles to random

https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

samples (GGEV distribution), 281 profiles showed no significant differences, suggesting a comparable nature of the empirical and random distributions in these cases. This corresponds to a fit rate of 100% (281/281).

The results showed that distributions with higher parameters (three or more, such as GEV and GGEV) not only provided slightly better fit in some cases (in the empirical data vs. theoretical quantiles scenario) but were also less prone to overfitting (in the empirical data vs. random quantiles scenario).

Step4

The GEV and GGEV distributions were subjected to K-Fold cross-validation alongside the LN2 and Exp distributions.

The K-Fold cross-validation results are presented as the percentage distribution of outcomes across individual intervals relative to the total number of results (Table 1).

Table 1. Percentage distribution of K-Fold cross-validation results across individual intervals relative to the total number of outcomes.

Distribution	MAE (0 - 100)	MAE (101 - 1000)	RMSE (0 - 100)	RMSE (101 - 1000)	RMSE (1001 - 10 000)
2LN	94.8%	5.22%	92.37%	7.63%	0%
Exp	94.8%	5.22%	92.37%	4.24%	3.39%
GEV	95.09%	4.91%	100%	0%	0%
GGEV	94.9%	5.12%	98.48%	1.51%	0%

Explanation: MAE – mean absolute error, RMSE - root mean square error, 0–100 – best-fitting model, 101–1000 – well-fitting model, 1001–10,000 – poorest-fitting model.

MAE and RMSE values vary significantly, ranging from very low (close to 0) to much higher values (e.g., 1000). High values suggest that the model predicts river flows less accurately for certain rivers. The intervals represent the quality of model fit, with 0–100 indicating the best fit, 101–1000 well fitting model, and 1001–10,000 the poorest fit. The GEV distribution achieved the highest percentage of best-fitting models (95.09% for MAE and 100% for RMSE), indicating superior performance compared to the other distributions. The GGEV distribution also showed strong results, with 94.9% of models falling in the best-fit category for MAE and 98.48% for RMSE. The LN2 and Exp distributions performed similarly, with over 94% of results in the best-fit category for both MAE and RMSE. However, the Exp distribution showed a small proportion (3.39%) of poorest-fitting models in the RMSE category, which was not observed for the other distributions. Overall, the GEV and GGEV distributions demonstrated the most consistent performance across both error metrics.

To summarize, Step 4 showed that the GGEV and GEV distributions have excellent predictive efficiency (better than distributions with fewer parameters), demonstrating that in most cases analyzed, they are quite robust to overparameterization and overfitting.

Although the study used observational series of 30 years or more, the number of profiles analyzed in highland and mountainous areas was considerably lower than those in lowland areas. Furthermore, the number of observational series exhibiting a positive trend in the analyzed region was limited.

5 Conclusions

The main findings of this research can be summarized as follows:

1. Based on the trend criterion, the GGEV distribution, compared to the analyzed three-parameter distributions, and the GEV distribution compared to the other three-parameter distributions were the best fit for most samples.
2. Based on the trend criterion and catchment size it was found that the GEV distribution is best suited for micro- and meso-catchments, while the GGEV distribution is ideal for macro- to large-catchments where the series exhibits a trend (either negative trend or no trend). The P3 distribution is preferred for very large catchments but only when the

https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

sample has no trend. In contrast, for samples with a positive trend, the GGEV distribution performs best across meso- to very large catchments.

3. The GGEV distribution, in comparison with the analyzed P3, LN3, and GEV distributions, was not flexible regarding the parameters A and Qp for samples with no trend, but it was flexible for samples with detected positive or negative trends.
4. Catchment size types influence the distribution parameters, with the most types affecting the parameters of GEV and LN3, and the least types affecting the parameters of GGEV and P3.
5. Additionally, the parameters of P3 and GGEV distributions are not correlated with landscape forms.
6. Our findings showed that some patterns between the P3 and GEV distributions were identical to those with GGEV. A similar pattern between GGEV and P3 was observed: highland, lowland and midland topographies were outside RDA1, indicating no influence of these factors on the scale and location parameters, and N was negatively correlated with the shape parameters. The GGEV distribution had a similar pattern to the GEV distribution, where the no trend samples appeared in RDA2 for both distributions.
7. It was found that adding the shape parameter of the GGEV distribution primarily compensates for the limitations of the shape parameter in distributions with lighter tails.
8. Using the Kolmogorov-Smirnov test it was found that the GEV and GGEV not only provided slightly better fit in some cases (in the empirical data vs. theoretical quantiles scenario) but were also less prone to overfitting (in the empirical data vs. random quantiles scenario) in comparison to Exp and LN2. Furthermore, the robustness of GEV and GGEV distributions to overparameterization and overfitting is confirmed by K-Fold cross validation.
9. Based on the above, in the era of climate change, distributions like GGEV are expected to be better suited under the presence of trends, having a clear performance benefit.

The results of this study highlight several promising avenues for future research. One potential direction is the further exploration of the GGEV distribution in the context of various hydrological and meteorological phenomena. Given its superior performance in fitting most samples and its sensitivity to trends, especially under non-stationary conditions like climate change, future studies could examine its applicability across different geographical regions and climatic conditions.

The findings on the influence of catchment types on distribution parameters indicate that more research is needed to refine our understanding of how landscape characteristics interact with hydrological distributions. A deeper exploration into the relationship between catchment area characteristics, especially in varied topographies and land-use patterns, could yield more universal insights. Expanding the range of predictor variables used in modeling, beyond trend detection, nature of catchment, catchment area, and the hydrological characteristics, might also improve the accuracy and flexibility of distribution selection.

Author contributions

LG: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Writing – original draft preparation, Writing – review & editing; PW: Conceptualization, Methodology, Supervision, Validation, Writing – review & editing; PT: Investigation, Resources, Visualization, Writing – original draft preparation, Writing – review & editing; CM: Data curation, Writing – review & editing; MW: Supervision, Validation; JPJr: Conceptualization, Supervision; JPSen: Supervision; Scz: Supervision, Validation.

Competing interests

The authors declare that they have no conflict of interest.

https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

Acknowledgments
The authors would like to express their sincere gratitude to the MSc Eng. Karol Misdziol for for help with coding in R and the Institute of Meteorology and Water Management—National Research Institute in Warsaw for the release of the flow data.

References:

Ahilan, S., O'Sullivan, J. J., and Bruen, M.: Influences on flood frequency distributions in Irish river catchments, Hydrol. Earth Syst. Sci., 16, 1137–1150, https://doi.org/10.5194/hess-16-1137-2012, 2012.

Allamano, P., Claps, P., and Laio, F.: An analytical model of the effects of catchment elevation on the flood frequency distribution, Water Resources Research, 45, 2007WR006658, https://doi.org/10.1029/2007WR006658, 2009.

Alsadat, N., Elgarhy, M., Hassan, A. S., Ahmad, H., and El-Hamid Eisa, A.: A new extension of linear failure rate distribution with estimation, simulation, and applications, AIP Advances, 13, 105019, https://doi.org/10.1063/5.0170297, 2023.

Arnaud, P., Lavabre, J., Fouchier, C., Diss, S., and Javelle, P.: Sensitivity of hydrological models to uncertainty in rainfall input, Hydrological Sciences Journal, 56, 397–410, https://doi.org/10.1080/02626667.2011.563742, 2011.

Bertola, M., Viglione, A., Lun, D., Hall, J., and Blöschl, G.: Flood trends in Europe: are changes in small and big floods different?, Hydrol. Earth Syst. Sci., 24, 1805–1822, https://doi.org/10.5194/hess-24-1805-2020, 2020.

Berton, R. and Rahmani, V.: Differences in Flood Quantiles Estimate of Disturbed and Undisturbed Watersheds in the United States, Hydroecology and Engineering, 1, 10002–10002, https://doi.org/10.35534/hee.2024.10002, 2024.

Bezak, N., Brilly, M., and Šraj, M.: Comparison between the peaks-over-threshold method and the annual maximum method for flood frequency analysis, Hydrological Sciences Journal, 59, 959–977, https://doi.org/10.1080/02626667.2013.831174, 2014.

Cassalho, F., Beskow, S., De Mello, C. R., De Moura, M. M., Kerstner, L., and Ávila, L. F.: At-Site Flood Frequency Analysis Coupled with Multiparameter Probability Distributions, Water Resour Manage, 32, 285–300, https://doi.org/10.1007/s11269-017-1810-7, 2018.

Chen, X., Ye, C., Zhang, J., Xu, C., Zhang, L., and Tang, Y.: Selection of an Optimal Distribution Curve for Non-Stationary Flood Series, Atmosphere, 10, 31, https://doi.org/10.3390/atmos10010031, 2019.

Dakhlaoui, H., Ruelland, D., and Tramblay, Y.: A bootstrap-based differential split-sample test to assess the transferability of conceptual rainfall-runoff models under past and future climate variability, Journal of Hydrology, 575, 470–486, https://doi.org/10.1016/j.jhydrol.2019.05.056, 2019.

E Silva, W. V. M. and Do Nascimento, F. F.: MCMC4Extremes: an R package for Bayesian inference for extremes and its extensions, Communications in Statistics - Simulation and Computation, 51, 432–442, https://doi.org/10.1080/03610918.2019.1653914, 2022.

Farooq, M., Shafique, M., and Khattak, M. S.: Flood frequency analysis of river swat using Log Pearson type 3, Generalized Extreme Value, Normal, and Gumbel Max distribution methods, Arab J Geosci, 11, 216, https://doi.org/10.1007/s12517-018-3553-z, 2018.

Flynn, K. M., William H. Kirby, and Paul R. Hummel: User's Manual for Program PeakFQ, Annual Flood-Frequency Analysis Using Bulletin 17B Guidelines, 2006.

Gruss, Ł., Wiatkowski, M., Tomczyk, P., Pollert, J., and Pollert, J.: Comparison of Three-Parameter Distributions in Controlled Catchments for a Stationary and Non-Stationary Data Series, Water, 14, 293, https://doi.org/10.3390/w14030293, 2022.

https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

Gruss, Ł., Wiatkowski, M., Połomski, M., Szewczyk, Ł., and Tomczyk, P.: Analysis of Changes in Water Flow after Passing through the Planned Dam Reservoir Using a Mixture Distribution in the Face of Climate Change: A Case Study of the Nysa Kłodzka River, Poland, Hydrology, 10, 226, https://doi.org/10.3390/hydrology10120226, 2023.

Haktanir, T. and Horlacher, H. B.: Evaluation of various distributions for flood frequency analysis, Hydrological Sciences Journal, 38, 15–32, https://doi.org/10.1080/02626669309492637, 1993.

Han, H., Yan, X., Xie, H., Qiu, J., Li, X., Zhao, D., Li, X., Yan, X., and Xia, Y.: Incorporating a new landscape intensity indicator into landscape metrics to better understand controls of water quality and optimal width of riparian buffer zone, Journal of Hydrology, 625, 130088, https://doi.org/10.1016/j.jhydrol.2023.130088, 2023.

He, J., Anderson, A., and Valeo, C.: Bias compensation in flood frequency analysis, Hydrological Sciences Journal, 60, 381–401, https://doi.org/10.1080/02626667.2014.885651, 2015.

Hebson, C. S. and Wood, E. F.: A Study of Scale Effects in Flood Frequency Response, in: Scale Problems in Hydrology, vol. 6, edited by: Gupta, V. K., Rodríguez-Iturbe, I., and Wood, E. F., Springer Netherlands, Dordrecht, 133–158, https://doi.org/10.1007/978-94-009-4678-1_7, 1986.

Hu, L., Nikolopoulos, E. I., Marra, F., and Anagnostou, E. N.: Sensitivity of flood frequency analysis to data record, statistical model, and parameter estimation methods: An evaluation over the contiguous United States, J Flood Risk Management, 13, e12580, https://doi.org/10.1111/jfr3.12580, 2020.

Iliopoulou, T., Koutsoyiannis, D., and Montanari, A.: Characterizing and Modeling Seasonality in Extreme Rainfall, Water Resources Research, 54, 6242–6258, https://doi.org/10.1029/2018WR023360, 2018.

Jain, D. and Singh, V. P.: Comparison of Some Flood Frequency Distributions Using Empirical Data, in: Hydrologic Frequency Modeling, edited by: Singh, V. P., Springer Netherlands, Dordrecht, 467–485, https://doi.org/10.1007/978-94-009-3953-0_33, 1987.

Jiang, S. and Kang, L.: Flood frequency analysis for annual maximum streamflow using a non-stationary GEV model, E3S Web Conf., 79, 03022, https://doi.org/10.1051/e3sconf/20197903022, 2019.

Kamal, V., Mukherjee, S., Singh, P., Sen, R., Vishwakarma, C. A., Sajadi, P., Asthana, H., and Rena, V.: Flood frequency analysis of Ganga river at Haridwar and Garhmukteshwar, Appl Water Sci, 7, 1979–1986, https://doi.org/10.1007/s13201-016-0378-3, 2017.

Karczewski, M. and Michalski, A.: A data-driven kernel estimator of the density function, Journal of Statistical Computation and Simulation, 92, 3529–3541, https://doi.org/10.1080/00949655.2022.2072503, 2022.

Karczewski, M., Kaźmierczak, B., Michalski, A., and Kuchar, L.: Probability Function Estimation for the Maximum Precipitation Model Using Kernel Estimators, Rocznik Ochrona Środowiska, 24, 260–275, https://doi.org/10.54740/ros.2022.019, 2022.

Karunasingha, D. S. K.: Root mean square error or mean absolute error? Use their ratio as well, Information Sciences, 585, 609–629, https://doi.org/10.1016/j.ins.2021.11.036, 2022.

Kasiviswanathan, K. S., He, J., and Tay, J.-H.: Flood frequency analysis using multi-objective optimization based interval estimation approach, Journal of Hydrology, 545, 251–262, https://doi.org/10.1016/j.jhydrol.2016.12.025, 2017.

Kim, H., Kim, S., Shin, H., and Heo, J.-H.: Appropriate model selection methods for nonstationary generalized extreme value models, Journal of Hydrology, 547, 557–574, https://doi.org/10.1016/j.jhydrol.2017.02.005, 2017.

Konrad, C. and Restivo, D.: Assessment and significance of the frequency domain for trends in annual peak streamflow, J Flood Risk Management, 14, e12761, https://doi.org/10.1111/jfr3.12761, 2021.

https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

Kousar, S., Khan, A. R., Ul Hassan, M., Noreen, Z., and Bhatti, S. H.: Some best-fit probability distributions for at-site flood frequency analysis of the Ume River, J Flood Risk Management, 13, e12640, https://doi.org/10.1111/jfr3.12640, 2020.

Kumar, R., Chatterjee, C., Kumar, S., Lohani, A. K., and Singh, R. D.: Development of Regional Flood Frequency Relationships Using L-moments for Middle Ganga Plains Subzone 1(f) of India., Water Resources Management, 17, 243–257, https://doi.org/10.1023/A:1024770124523, 2003.

Kusumastuti, D. I., Struthers, I., Sivapalan, M., and Reynolds, D. A.: Threshold effects in catchment storm response and the occurrence and magnitude of flood events: implications for flood frequency, Hydrol. Earth Syst. Sci., 11, 1515–1528, https://doi.org/10.5194/hess-11-1515-2007, 2007.

Langridge, M., Gharabaghi, B., McBean, E., Bonakdari, H., and Walton, R.: Understanding the dynamic nature of Time-to-Peak in UK streams, Journal of Hydrology, 583, 124630, https://doi.org/10.1016/j.jhydrol.2020.124630, 2020.

Merz, R. and Blöschl, G.: Process controls on the statistical flood moments - a data based analysis, Hydrological Processes, 23, 675–696, https://doi.org/10.1002/hyp.7168, 2009.

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., and Stouffer, R. J.: Stationarity Is Dead: Whither Water Management?, Science, 319, 573–574, https://doi.org/10.1126/science.1151915, 2008.

Młyński, D., Petroselli, A., and Wałęga, A.: Flood frequency analysis by an event-based rainfall-runoff model in selected catchments of southern Poland, Soil Water Res., 13, 170–176, https://doi.org/10.17221/153/2017-SWR, 2018.

Młyński, D., Wałęga, A., Ozga-Zielinski, B., Ciupak, M., and Petroselli, A.: New approach for determining the quantiles of maximum annual flows in ungauged catchments using the EBA4SUB model, Journal of Hydrology, 589, 125198, https://doi.org/10.1016/j.jhydrol.2020.125198, 2020.

Morlot, M., Brilly, M., and Šraj, M.: Characterisation of the floods in the Danube River basin through flood frequency and seasonality analysis, Acta hydrotechnica, 73–89, https://doi.org/10.15292/acta.hydro.2019.06, 2019.

Morrison, J. E. and Smith, J. A.: Stochastic modeling of flood peaks using the generalized extreme value distribution, Water Resources Research, 38, https://doi.org/10.1029/2001WR000502, 2002.

Nascimento, F., Bourguignon, M., and Leao, J.: Extended generalized extreme value distribution with applications in environmental data, HJMS, 46, 1–1, https://doi.org/10.15672/HJMS.20159514081, 2016.

Northrop, P. J.: Likelihood-based approaches to flood frequency estimation, Journal of Hydrology, 292, 96–113, https://doi.org/10.1016/j.jhydrol.2003.12.031, 2004.

Nyeko-Ogiramoi, P., Willems, P., Mutua, F. M., and Moges, S. A.: An elusive search for regional flood frequency estimates in the River Nile basin, Hydrol. Earth Syst. Sci., 16, 3149–3163, https://doi.org/10.5194/hess-16-3149-2012, 2012.

Ologhadien, I.: Selection of Probabilistic Model of Extreme Floods in Benue River Basin, Nigeria, EJENG, 6, 7–18, https://doi.org/10.24018/ejeng.2021.6.1.2300, 2021.

Otiniano, C. E. G., Paiva, B. S. D., and Martins Netob, D. S. B.: The transmuted GEV distribution: properties and application, CSAM, 26, 239–259, https://doi.org/10.29220/CSAM.2019.26.3.239, 2019.

Ozonur, D., Pobocikova, I., and De Souza, A.: Statistical analysis of monthly rainfall in Central West Brazil using probability distributions, Model. Earth Syst. Environ., 7, 1979–1989, https://doi.org/10.1007/s40808-020-00954-z, 2021.

https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

Pastor, A. V., Ludwig, F., Biemans, H., Hoff, H., and Kabat, P.: Accounting for environmental flow requirements in global water assessments, Hydrol. Earth Syst. Sci., 18, 5041–5059, https://doi.org/10.5194/hess-18-5041-2014, 2014.

Pitlick, J.: Relation between peak flows, precipitation, and physiography for five mountainous regions in the western USA, Journal of Hydrology, 158, 219–240, https://doi.org/10.1016/0022-1694(94)90055-8, 1994.

Pokhrel, Y., Felfelani, F., Satoh, Y., Boulange, J., Burek, P., Gädeke, A., Gerten, D., Gosling, S. N., Grillakis, M., Gudmundsson, L., Hanasaki, N., Kim, H., Koutroulis, A., Liu, J., Papadimitriou, L., Schewe, J., Müller Schmied, H., Stacke, T., Telteu, C.-E., Thiery, W., Veldkamp, T., Zhao, F., and Wada, Y.: Global terrestrial water storage and drought severity under climate change, Nat. Clim. Chang., 11, 226–233, https://doi.org/10.1038/s41558-020-00972-w, 2021.

Połomski, M. and Wiatkowski, M.: Impounding Reservoirs, Benefits and Risks: A Review of Environmental and Technical Aspects of Construction and Operation, Sustainability, 15, 16020, https://doi.org/10.3390/su152216020, 2023.

Prusty, S., Patnaik, S., and Dash, S. K.: SKCV: Stratified K-fold cross-validation on ML classifiers for predicting cervical cancer, Front. Nanotechnol., 4, 972421, https://doi.org/10.3389/fnano.2022.972421, 2022.

Raynal-Villasenor, J. A. and Raynal-Gutierrez, M. E.: Estimation procedures for the GEV distribution for the minima, Journal of Hydrology, 519, 512–522, https://doi.org/10.1016/j.jhydrol.2014.07.045, 2014.

Ribeiro-Correa, J. and Rousselle, J.: A hierarchical and empirical Bayes Approach for the regional Pearson type III distribution, Water Resources Research, 29, 435–444, https://doi.org/10.1029/92WR02086, 1993.

Rohani, A., Taki, M., and Abdollahpour, M.: A novel soft computing model (Gaussian process regression with K-fold cross validation) for daily and monthly solar radiation forecasting (Part: I), Renewable Energy, 115, 411–422, https://doi.org/10.1016/j.renene.2017.08.061, 2018.

Roodsari, B. K. and Chandler, D. G.: Distribution of surface imperviousness in small urban catchments predicts runoff peak flows and stream flashiness, Hydrological Processes, 31, 2990–3002, https://doi.org/10.1002/hyp.11230, 2017.

Rutkowska, A.: Properties of the Cox–Stuart Test for Trend in Application to Hydrological Series: The Simulation Study, Communications in Statistics - Simulation and Computation, 44, 565–579, https://doi.org/10.1080/03610918.2013.784988, 2015.

Rutkowska, A., Kohnová, S., Banasik, K., Szolgay, J., and Karabová, B.: Probabilistic properties of a curve number: A case study for small Polish and Slovak Carpathian Basins, J. Mt. Sci., 12, 533–548, https://doi.org/10.1007/s11629-014-3123-0, 2015.

Sampaio, J. and Costa, V.: Bayesian regional flood frequency analysis with GEV hierarchical models under spatial dependency structures, Hydrological Sciences Journal, 66, 422–433, https://doi.org/10.1080/02626667.2021.1873997, 2021.

Smith, J. A.: Regional flood frequency analysis using extreme order statistics of the annual peak record, Water Resources Research, 25, 311–317, https://doi.org/10.1029/WR025i002p00311, 1989.

Šraj, M., Viglione, A., Parajka, J., and Blöschl, G.: The influence of non-stationarity in extreme hydrological events on flood frequency estimation, Journal of Hydrology and Hydromechanics, 64, 426–437, https://doi.org/10.1515/johh-2016-0032, 2016.

Svensson, C., Kundzewicz, W. Z., and Maurer, T.: Trend detection in river flow series: 2. Flood and low-flow index series / Détection de tendance dans des séries de débit fluvial: 2. Séries d'indices de crue et d'étiage, Hydrological Sciences Journal, 50, 6, https://doi.org/10.1623/hysj.2005.50.5.811, 2005.

https://doi.org/10.5194/egusphere-2025-860
Preprint. Discussion started: 10 April 2025
© Author(s) 2025. CC BY 4.0 License.

Szulczewski, W. and Jakubowski, W.: The Application of Mixture Distribution for the Estimation of Extreme Floods in Controlled Catchment Basins, Water Resour Manage, 32, 3519–3534, https://doi.org/10.1007/s11269-018-2005-6, 2018.

Tabari, H., Hosseinzadehtalaei, P., Thiery, W., and Willems, P.: Amplified Drought and Flood Risk Under Future Socioeconomic and Climatic Change, Earth's Future, 9, e2021EF002295, https://doi.org/10.1029/2021EF002295, 2021a.

Tabari, H., Moghtaderi Asr, N., and Willems, P.: Developing a framework for attribution analysis of urban pluvial flooding to human-induced climate impacts, Journal of Hydrology, 598, 126352, https://doi.org/10.1016/j.jhydrol.2021.126352, 2021b.

Tomczyk, P., Wierzchowski, P. S., Dobrzyński, J., Kulkova, I., Wróbel, B., Wiatkowski, M., Kuriqi, A., Skorulski, W., Kabat, T., Prycik, M., Gruss, Ł., and Drobnik, J.: Effective microorganism water treatment method for rapid eutrophic reservoir restoration, Environ Sci Pollut Res, 31, 2377–2393, https://doi.org/10.1007/s11356-023-31354-2, 2023.

Tyralis, H., Papacharalampous, G., and Tantanee, S.: How to explain and predict the shape parameter of the generalized extreme value distribution of streamflow extremes using a big dataset, Journal of Hydrology, 574, 628–645, https://doi.org/10.1016/j.jhydrol.2019.04.070, 2019.

Ul Hassan, M., Hayat, O., and Noreen, Z.: Selecting the best probability distribution for at-site flood frequency analysis; a study of Torne River, SN Appl. Sci., 1, 1629, https://doi.org/10.1007/s42452-019-1584-z, 2019.

Valentini, M. H. K., Beskow, S., Beskow, T. L. C., De Mello, C. R., Cassalho, F., and Da Silva, M. E. S.: At-site flood frequency analysis in Brazil, Nat Hazards, 120, 601–618, https://doi.org/10.1007/s11069-023-06231-3, 2024.

Valeo, C. and Rasmussen, P.: Topographic Influences on Flood Frequency Analyses, Canadian Water Resources Journal, 25, 387–406, https://doi.org/10.4296/cwrj2504387, 2000.

Villarini, G. and Smith, J. A.: Flood peak distributions for the eastern United States, Water Resources Research, 46, 2009WR008395, https://doi.org/10.1029/2009WR008395, 2010.

Vogel, R. M. and Wilson, I.: Probability Distribution of Annual Maximum, Mean, and Minimum Streamflows in the United States, J. Hydrol. Eng., 1, 69–76, https://doi.org/10.1061/(ASCE)1084-0699(1996)1:2(69), 1996.

Vogel, R. W. and McMartin, D. E.: Probability Plot Goodness-of-Fit and Skewness Estimation Procedures for the Pearson Type 3 Distribution, Water Resources Research, 27, 3149–3158, https://doi.org/10.1029/91WR02116, 1991.

Willems, P.: Revision of urban drainage design rules after assessment of climate change impacts on precipitation extremes at Uccle, Belgium, Journal of Hydrology, 496, 166–177, https://doi.org/10.1016/j.jhydrol.2013.05.037, 2013.

Willmott, C. and Matsuura, K.: Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance, Clim. Res., 30, 79–82, https://doi.org/10.3354/cr030079, 2005.

Xu, Y. and Goodacre, R.: On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning, J. Anal. Test., 2, 249–262, https://doi.org/10.1007/s41664-018-0068-2, 2018.

Yadav, S. and Shukla, S.: Analysis of k-Fold Cross-Validation over Hold-Out Validation on Colossal Datasets for Quality Classification, in: 2016 IEEE 6th International Conference on Advanced Computing (IACC), 2016 IEEE 6th International Conference on Advanced Computing (IACC), Bhimavaram, India, 78–83, https://doi.org/10.1109/IACC.2016.25, 2016.