Multidimensional context for extreme analysis of daily streamflow, rainfall and accumulated rainfall across USA

European Geosciences Union General Assembly 2020, 4-8 May, 2020, Online HS3.6: Spatio-temporal and/or (geo) statistical analysis of hydrological events, floods, extremes, and related hazards

Maria Nezi[1], Ioannis Tsoukalas[2], Charalampos Ntigkakis[3], and Andreas Efstratiadis[2]

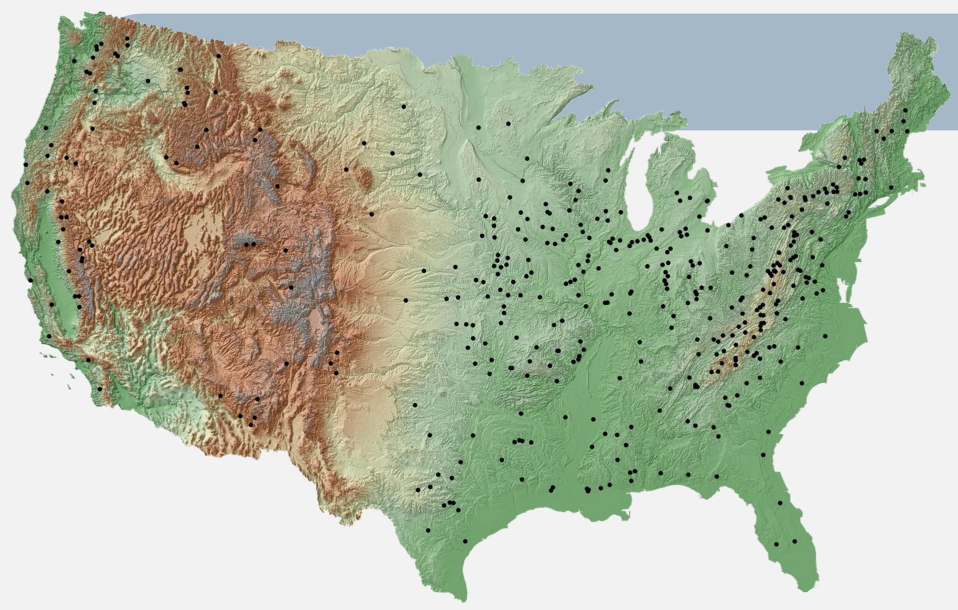
[1]HR Wallingford Ltd, Flood and Water Management Group, Wallingford, United Kingdom (marianezi42@gmail.com), [2]National Technical University of Athens, Department of Water Resources & Environmental Engineering, Athens, Greece, [3]University of Bristol, Faculty of Engineering, Department of Civil Engineering, Bristol, United Kingdom

1. Abstract

Statistical analysis of rainfall and runoff extremes plays a crucial role in hydrological design and flood risk management. Usually this analysis is performed separately for the two processes of interest, thus **ignoring their dependencies**, which appear at multiple temporal scales. Actually, the generation of a flood strongly depends on **soil moisture conditions**, which in turn depends on **past rainfall**. Using daily rainfall and runoff data from about 400 catchments in USA, retrieved from the **MOPEX repository**, we investigate the statistical behaviour of the corresponding annual rainfall and streamflow maxima, also accounting for the **influence of antecedent soil moisture conditions**. The latter are quantified by means of **accumulated daily rainfall at various aggregation scales** (i.e., from 5 up to 30 days) before each extreme rainfall and streamflow event. Analysis of maxima is employed by fitting the **Generalized Extreme Value** (GEV) distribution, using the L-moments method for extracting the associated parameters (shape, scale, location). Significant attention is paid for ensuring statistically consistent estimations of the **shape parameter**, which is empirically adjusted in order to minimize the influence of **sample uncertainty**. Finally, we seek for the possible **correlations** among the derived parameter values and hydroclimatic characteristics of the studied basins, and also depict their **spatial distribution** across USA.

2. Database

The MOPEX project, which initially launched on 1997, aimed to create an extensive database consisted of hydroclimatic observations and catchment's geomorphological characteristics, worldwide. This would allow for better understanding and thus application of hydroclimatic variables leading to a more efficient calibration of the atmospheric and hydrological models. For the purpose of this study we exclusively analyse data for USA.



Why USA database?

- ❑ More than 438 water catchments available
- ❑ Hydroclimatic observations from 1948 to 2003
- ❑ Daily rainfall and runoff observations with average time-series length greater than 30 years
- ❑ Daily observations of maximum and minimum temperature data
- ❑ Geographical data also provided for all catchments

We selected our hydrological data by accounting for two main criteria

- ✓ Time consistency over observations, hence time series without missing data
- ✓ Time series larger or equal to 40 years

Final dataset was formed as follows

- ✓ 423 catchments for the analysis of annual maximum precipitation with associated cumulative time series
- ✓ 299 catchments for the analysis of annual maximum runoff time series

3.Methodology

Steps

Step 1. Extraction of annual extreme rainfall and annual extreme runoff samples

Step 2. Selection of the aggregation scale for cumulative rainfall time series. In this study, the time range starts **five days** before the observed annual maximum continuing up to **one month (30 days) with a 5-days time step**

Step 3. Production of the cumulative rainfall time series for annual maximum rainfall and annual maximum runoff respectively

Step 4. Calculation of Pearson correlation coefficient

Step 5. Fitting the Generalized Extreme Value Distribution using the L-moments method for annual maximum runoff, rainfall and cumulative time series

Step 6. Fixing the GEV shape parameter and calculating the first four moments for every time series

Step 7. Seeking for any possible correlation among the values and depicting their spatial distribution across USA.

For more information, please refer to Nezi (2018)

Generalized Extreme Value Distribution

The generalized extreme value distribution contains three parameters, said to be location, scale and shape parameter. The shape parameter is the most significant over GEV parameters. It resembles the tails "thickness" and shows how fast the tail is approaching zero. Based on its value we identify which asymptotic distribution is the best fit for our data. According to its value, GEV distribution is divined into three asymptotic types: (a) Gumbel (γ=0), (b) Fréchet (γ>0), (c) Weibull (γ<0). Finally, the GEV distribution takes the following form:

$$F(\chi) = \begin{cases} e^{\left(-(1-\gamma*(\frac{\chi-\mu}{\sigma})\right)^{1/\gamma}}, & \gamma \neq 0 \\ e^{\left(-(\chi-\mu)/\sigma\right.}, & \gamma = 0 \end{cases}$$

where: $\mu \in R$ — location, $\sigma > 0$ — scale, $\gamma \in R$ — shape

The method of L-moments has been selected due to its low sensitivity to sample uncertainty and to possible outliers.

GEV shape parameter adjustment

The shape parameter is biased to time series length. Additionally, very negative values of the parameter implies an upper bound that has no physical meaning to hydrological values. In remedy of this, we are 'fixing' the shape parameter using the following equation [2]:

$$\tilde{\gamma}(n) = \frac{\sigma_\gamma}{\sigma_\gamma(n)}\left(\hat{\gamma} - \mu_\gamma(n)\right) + \mu_\gamma$$

where:

a) $\mu_\gamma(n)=\mu_\gamma-0.69n^{(-0.98)}$, $\mu_\gamma \cong 0.114$

b) $\sigma_\gamma(n)=\sigma_\gamma+1.27n^{(-0.70)}$, $\sigma_\gamma \cong 0.045$

The GEV shape parameter equation is derived from the analysis of daily rainfall observations (Papalexiou and Koutsoyiannis, 2013). Here, we assumed that the latter can also be applied to runoff time series.

4.Results

Extreme Value Analysis

- ❑ After producing the annual maximum runoff, rainfall and cumulative samples we fit the Generalized Extreme Value distribution using the L-moments approach;
- ❑ For the case of the annual maximum runoff data we generated two different kinds of cumulative time series, one (day0) that takes into account the day the annual maximum runoff value was observed and one (day1) that starts right before the aforementioned date.

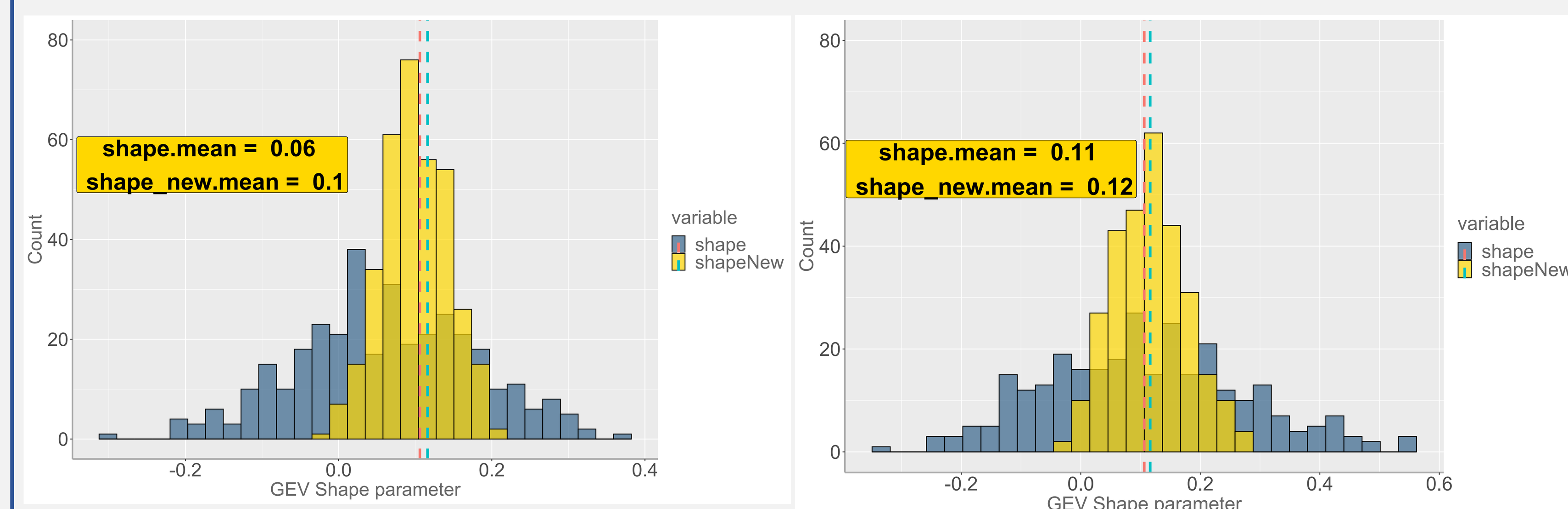


Fig. 1: Histograms showing the changes in GEV shape parameters: (left) annual maximum rainfall time series, (right) annual maximum runoff time series

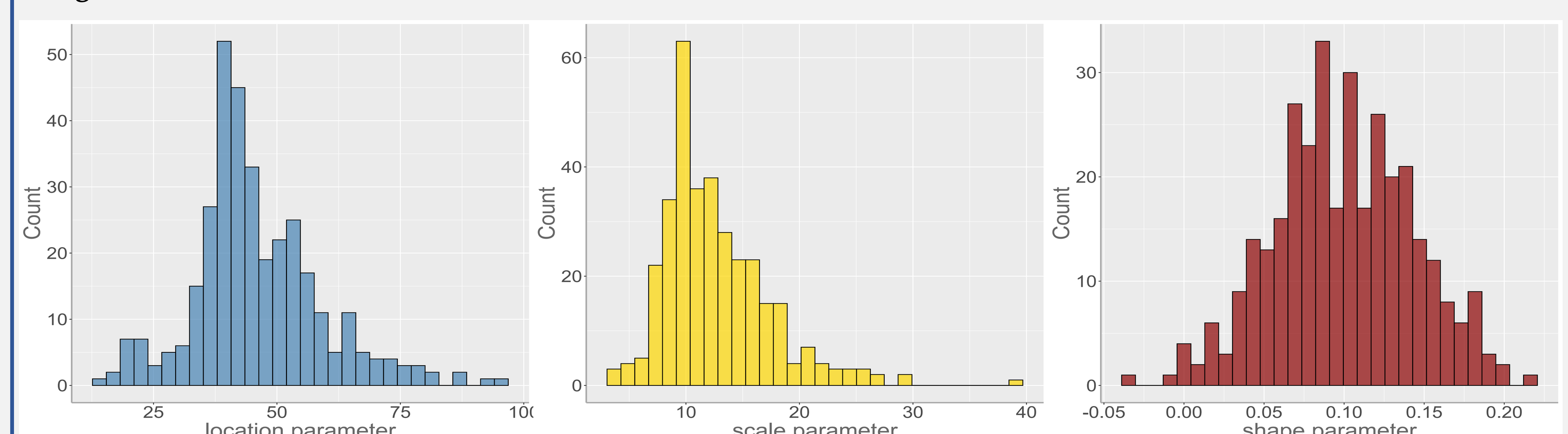


Fig. 2: Histograms illustrating GEV parameters (location, scale, shape) for the annual maximum rainfall time series

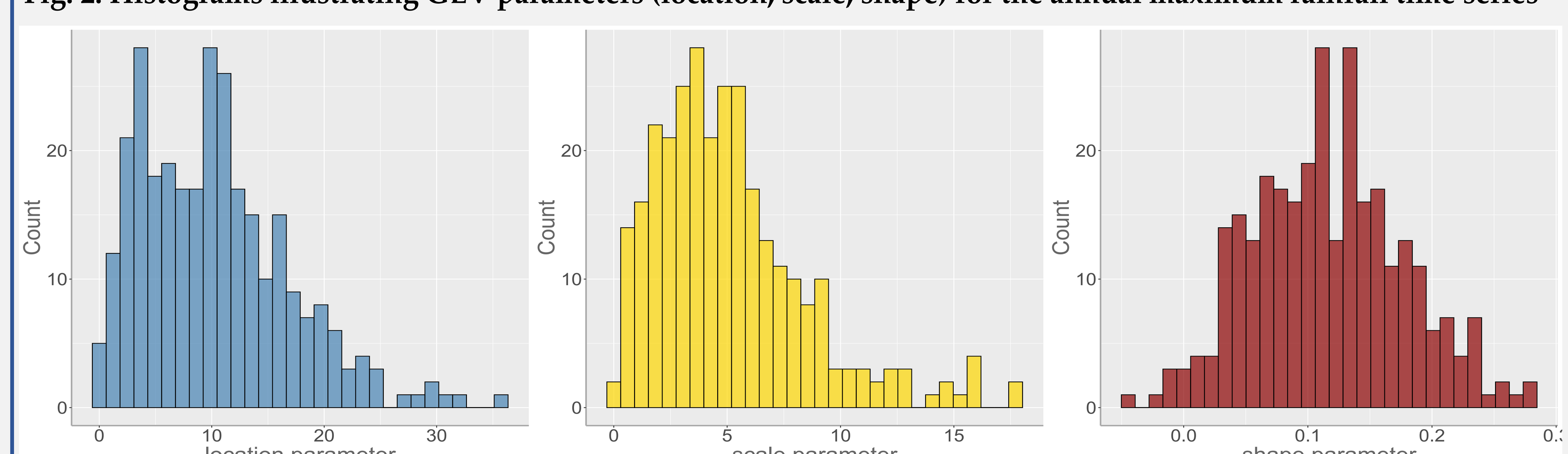


Fig. 3: Histograms illustrating GEV parameters (location, scale, shape) for the annual maximum runoff time series

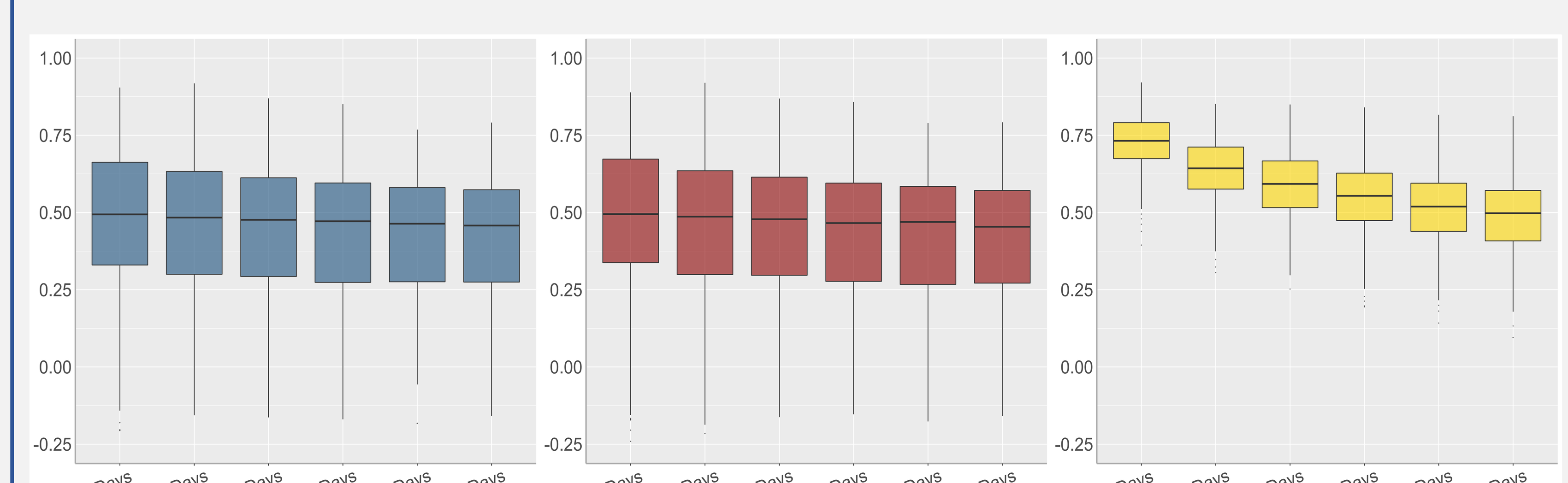


Fig. 4: Box-Plots displaying Pearson correlation coefficient between: (left) annual maximum runoff (day0) and cumulative time series, (centre) annual maximum runoff (day1) and cumulative time series, (right) annual maximum rainfall and cumulative time series

Hydroclimatic Analysis

Finally, we investigate the potential correlation between the hydroclimatic characteristics of the studied basins and the derived parameter value. Following that, we created a so-called "coldness coefficient" which describes the percentage of daily temperature dropping bellow 0C° on annual basis for the whole range of catchments. Our idea was to define any possible statistical relationship between the catchment's climate, hence the effectiveness of coolest days on annual max runoff time series. Consequentially, we depict the maximum correlation for annual maximum runoff and cumulative rainfall against the aforementioned "coldness coefficient", over each catchment.

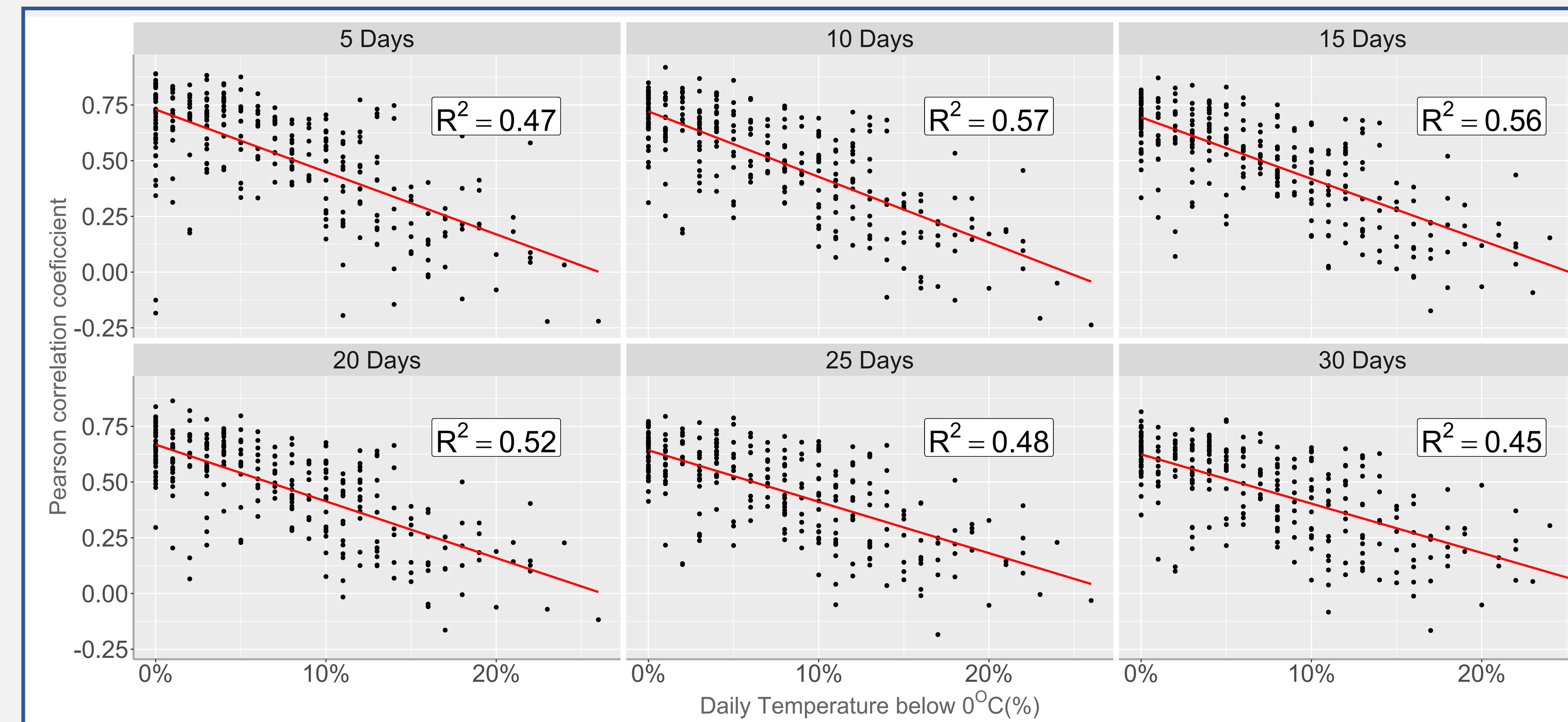


Fig. 5: Scatter plot showing Pearson correlation coefficient between annual maximum runoff time series and "coldness coefficient"

Spatial Distribution Maps

Seeking to analyse the variability of the statistical parameter, we created spatial probabilistic maps using Kriging methodology. The following maps display the fluctuation of GEV shape parameter for annual maximum runoff and rainfall time series across USA. Finally, we depicted the variability of maximum correlation between annual maximum rainfall and runoff with cumulative rainfall respectively.

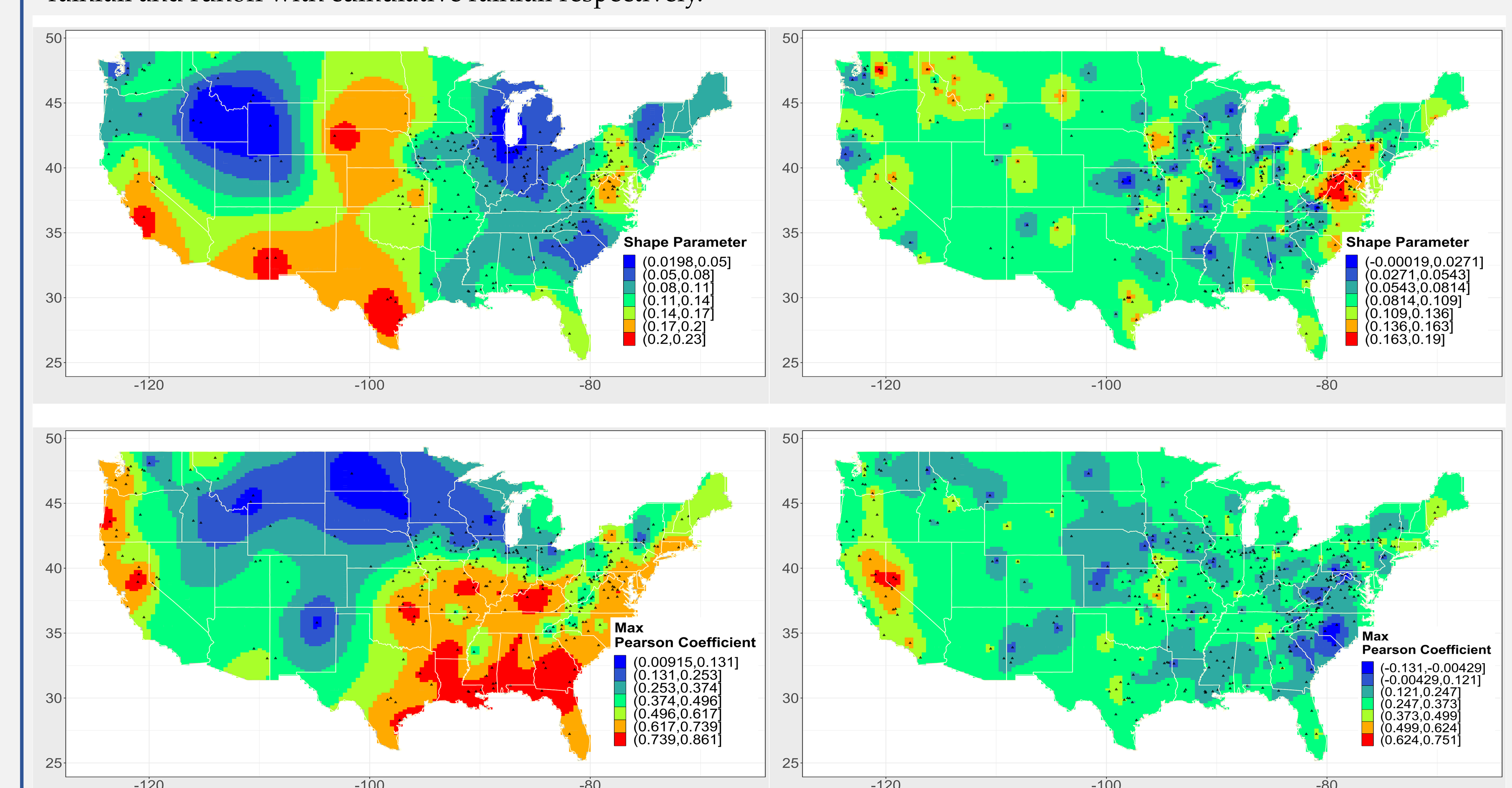


Fig. 6: Spatial distribution maps showing: (top-left) GEV shape parameter of annual maximum runoff time series, (top-right) GEV shape parameter of annual maximum rainfall time series, (bottom-left) Maximum Pearson correlation coefficient between annual maximum runoff and cumulative rainfall time series , (bottom-right) Maximum Pearson correlation coefficient between annual maximum rainfall and cumulative rainfall time series

5.Conclusions

- For annual maximum rainfall time series, the cumulative rainfall of 5 days seems to have the most significant impact among others.
- For runoff time series analysis shows that there is no significant difference between the different scales with the cumulative rainfall of 15 days being the largest by a small margin.
- GEV shape parameter, for annual maximum rainfall time series the parameter approaches on average the value off 0.10 and for annual maximum runoff slightly above 0.11, which indicates that Fréchet distribution is the best fit for both time series.
- For annual maximum runoff the GEV shape parameter takes its lowest values on mountainous locations and the highest ones near the coast line and Pearson correlation takes its lowest values in the north and central USA, with its highest performed in south-east.
- For annual maximum rainfall the shape parameter and Pearson correlation don't change significantly across USA.
- For each ascending condition, the cumulative time series correlation coefficient takes values approximately near 0.50, which reflects the significant role that cold conditions (e.g. snow) play on extreme runoff response (Fig. 5).

References

Nezi, M., Multivariate statistical analysis of extreme rainfall and runoff in a sample of 400 river basins over USA from database MOPEX, Diploma thesis, 103 pages, Department of Water Resources and Environmental Engineering – National Technical University of Athens, Athens, October 2018.

Papalexiou, S.M., and D. Koutsoyiannis, Battle of extreme value distributions: A global survey on extreme daily rainfall, Water Resources Research, 49 (1), 187–201, doi:10.1029/2012WR012557, 2013.