EGU General Assembly 2020

Session HS7.4
Hydroclimatic change and unchange: exploring the mysteries of variability, nature and human impact (co-sponsored by IAHS, WMO CHy)

Department of Water Resources & Environmental Engineering,
School of Civil Engineering,
National Technical University of Athens,
Greece

#shareEGU20

Georgios T. Manolis
Konstantinos Papoulakos
Theano Iliopoulou
Panayiotis Dimitriadis
Dimosthenis Tsaknias
Demetris Koutsoyiannis

Clustering mechanisms of flood occurrence; modelling and relevance to insurance practices

EGU2020-9357

EGU General Assembly 2020

Presentation Outline

- Abstract
- About Flood Insurance
- Dataset
- Peak Over Threshold Method in insurance
- Methodology
- Evaluating Persistence
- SMA Method - GHK Model
- Results
- Conclusions
- References

The Milky Way is nothing else
but a mass of innumerable stars planted
together in clusters.

Galileo Galilei

Abstract

EGU General Assembly 2020

Population growth, economic development and risk-blind urbanization often increase **exposure to risk**, including that due to floods. While rural flooding may affect much larger areas of land, urban floods are more challenging to manage, since the higher population and asset density in the urban environment **increase** the environmental and social impacts of floods and make the potential flood damages more **costly**.

Therefore, the need for integrated **flood insurance** policy and products on extended parts of the world is pronounced in order to **reduce** the financial consequences of extreme flood events, which endanger in many cases the environmental, social and economic **stability**.

As the assessment of the so-called **collective risk** is a typical issue faced in insurance and reinsurance practices, in this study we investigate the **stochastic dynamics** of daily stream flow series with particular interest to the existence of **clustering mechanisms** in floods, which is known to increase the potential risk.

We analyze collective risk on the US-CAMELS dataset, treating the **stream flow exceedances** over given **thresholds** as proxies for insurance claim amounts (Serinaldi and Kilsby, 2016).

Moreover, we develop **modelling** and **simulation** approaches of extreme flows as a step towards the deeper understanding of the relationship between the **stochastic patterns of flood occurrence** and proxies of insurance claims, paving the way for a more **efficient use** of the available stream flow records.

Acknowledgement: This research is conducted within the frame of the course "Stochastic Methods" of the National Technical University of Athens (NTUA). The School of Civil Engineering of NTUA provided moral support for the participation of the students in the Assembly.

About Flood Insurance

Our research focuses on River Flooding

River flooding can be caused by rainfall, or snowmelt. This type of flood occurs when the river is unable to carry the flow of water, resulting in escape of water from the normal perimeter and submergence of surrounding low-lying land.

The consequences of flood events impact on both individuals and communities, endangering in many cases the social and economic stability.

The severity of these consequences varies greatly depending on location, extent of flooding, vulnerability and the value of the affected natural and constructed environments.

Besides public and individual measures, insurance is an important factor in reducing the financial risk for individuals, enterprises, and even whole societies where extreme flood events are concerned in order to protect the insured from excessive losses that substantially threaten their living or business conditions (Kron, 2005).

Insurance companies should consider clustering mechanisms in their practices to avoid underestimation of the exceedance probability of collective risk.

(Goulianou et. al, EGU 2019)

Dataset

We used and processed the US-CAMELS dataset (Newman et al., 2014), which comprises 671 daily stream flow time series across the major basins and hydrological units in USA.

From this dataset, 360 stream flow time series with the maximum temporal overlap (namely, 35 years from 1980 to 2014) and less than 10% of missing values were selected.

EGU General Assembly 2020

Our Aim

As this dataset spans a very wide range of hydrometeorological processes and conditions across USA, our aim is to develop modelling simulations in order to evaluate the existent clustering mechanisms of extremes and the impact of Hurst-Kolmogorov dynamics on the time series.

Peak Over Threshold Method

in insurance

Peak Over Threshold method (POT) has become one of the most preferable extreme value approaches in insurance.

The threshold should be chosen such that all losses above the threshold could be considered as extreme losses, in the sense of the underlying extreme value analysis.

To characterize the dynamics of extreme stream flow values, we selected four different percentage thresholds (90%, 95%, 98%, and 99%).

EGU General Assembly 2020

Threshold selection

The dominant trend in insurance practices is to select **high percentage** thresholds (99% or greater) in order to analyze exclusively high-impact extreme flood events, which are mainly responsible for the **large amounts** of compensations that companies will have to pay to their clients.

Although this is a **desirable** option, it is not always a **possible** one. The main reason is that, in many cases, the length of the available observed time series is quite **short**.

As a result, in such cases, selecting a high percentage threshold leads to inaccurate conclusions regarding the **statistical behavior** of the mentioned time series.

Methodology

EGU General Assembly 2020

Collective Risk

Collective Risk S is the total claim amount regarding a portfolio of (re)insured properties that produces a random number N of claims in a certain time period. Following Serinaldi and Kilsby (2016), we use POT flows as a proxy for collective risk estimation, defined as:

$$\boldsymbol{S} = \sum_{j=1}^{N} \boldsymbol{Y}_j$$

where Y_j is the jth claim proxy (over-threshold flow fluctuation severity), N is the number of exceedances, and the total claims S=0 if N=0.

Assessing the collective risk S is a typical problem faced in insurance sector.

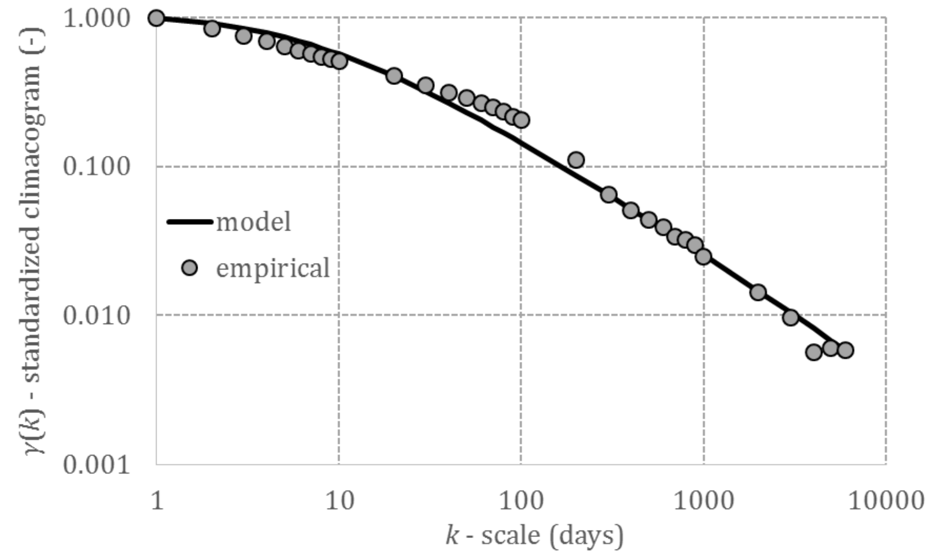
SMA Method - GHK Model

The generalized-HK (GHK) process is a process exhibiting an HK behavior. It is a method that can preserve explicitly (i.e. fully analytical calculations) four marginal moments of a process for any type of dependence structure (Dimitriadis and Koutsoyiannis, 2018).

Evaluating persistence

Based on the mean *climacogram* (Dimitriadis and Koutsoyiannis, 2015) of the 360 empirical stream flow time series of our dataset, the Hurst parameter is estimated as 0.63 indicating a persistent behavior (dependence).

We track the effect of this dependence structure on the behaviors of POT flows in the annual scale and the estimation of the collective risk proxy.



$$\gamma(k) = \frac{\lambda}{(1+k/q)^{2-2H}}$$ (GHK model adjusted for bias)

SMA Method - GHK Model

EGU General Assembly 2020

SMA Method

In this case, we use the symmetric moving average (SMA) method (Koutsoyiannis 2000; 2016), a scheme for approximating the marginal probability function of a process by exactly preserving its first four central moments which is found to be adequate for various distributions commonly applied in geophysical processes (Dimitriadis and Koutsoyiannis, 2018). The SMA method is described by the following equation:

$$\underline{\boldsymbol{x}}_{\boldsymbol{i}} = \sum_{\boldsymbol{l}=-\infty}^{\infty} \boldsymbol{a}_{|\boldsymbol{l}|} \underline{\boldsymbol{v}}_{\boldsymbol{i}+\boldsymbol{l}}$$

where $\underline{x}_i$ is any process with any type of dependence, a_l are coefficients calculated from the autocovariance function and $\underline{v}_i$ is white noise averaged in discrete-time.

GHK Model

For the stochastic simulation of a series with generalized long-range dependence (GHK model) by preserving explicitly, we calculate the first four central moments of the sample series.

In order to produce the synthetic time series from the data of the observed (empirical) one, the model requires as input the following:
mean (*Sm*), variance (*Sv*), skewness and kurtosis coefficients (*Ss* and *Sk*), Hurst parameter of the GHK model (H), scale parameter (q), length of synthetic series (N).

Exploratory Analysis

Phase 1

We choose to develop SMA-GHK modelling simulations of the following stations:

- The USGS **07067000** (Current River) at Van Buren, State of Missouri, USA.
  H = 0.72 and q = 2.71 days
- The USGS **07071500** (Eleven Point River) at Van Buren, State of Missouri, USA.
  H = 0.77 and q = 1.96 days

We developed 1000 synthetic time series for each station.

EGU General Assembly 2020

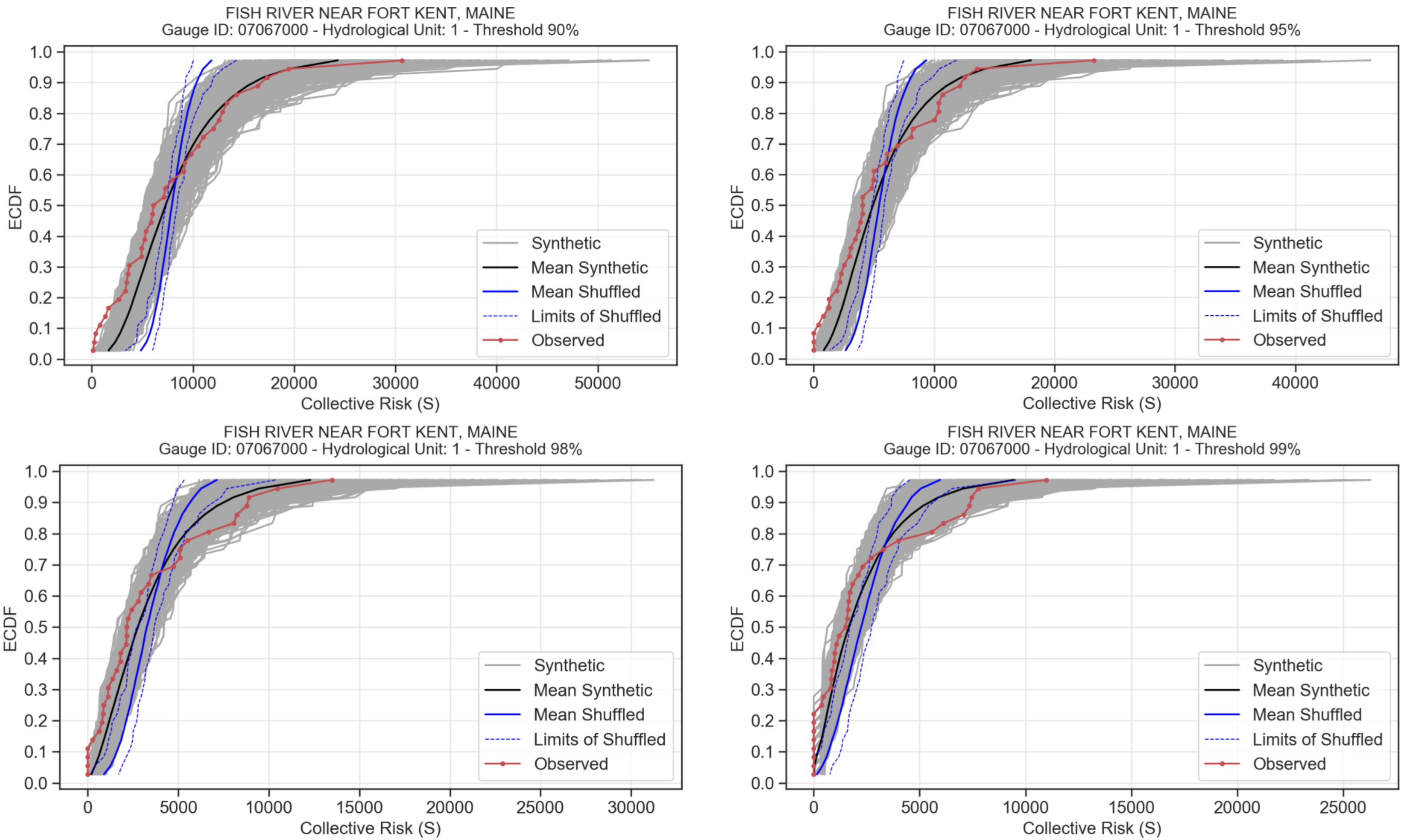
Phase 2

In addition to our modelling simulation, studies that evaluate the clustering mechanisms of extremes on this dataset, based on the observed and shuffled time series in terms of annual collective risk (Papoulakos et al., 2020) are applied.

The diagrams of the empirical cumulative distribution function (ECDF) of collective risk for the four thresholds are extracted, illustrating the following curves: synthetic, mean synthetic, mean shuffled, upper and lower limits of shuffled and the observed.

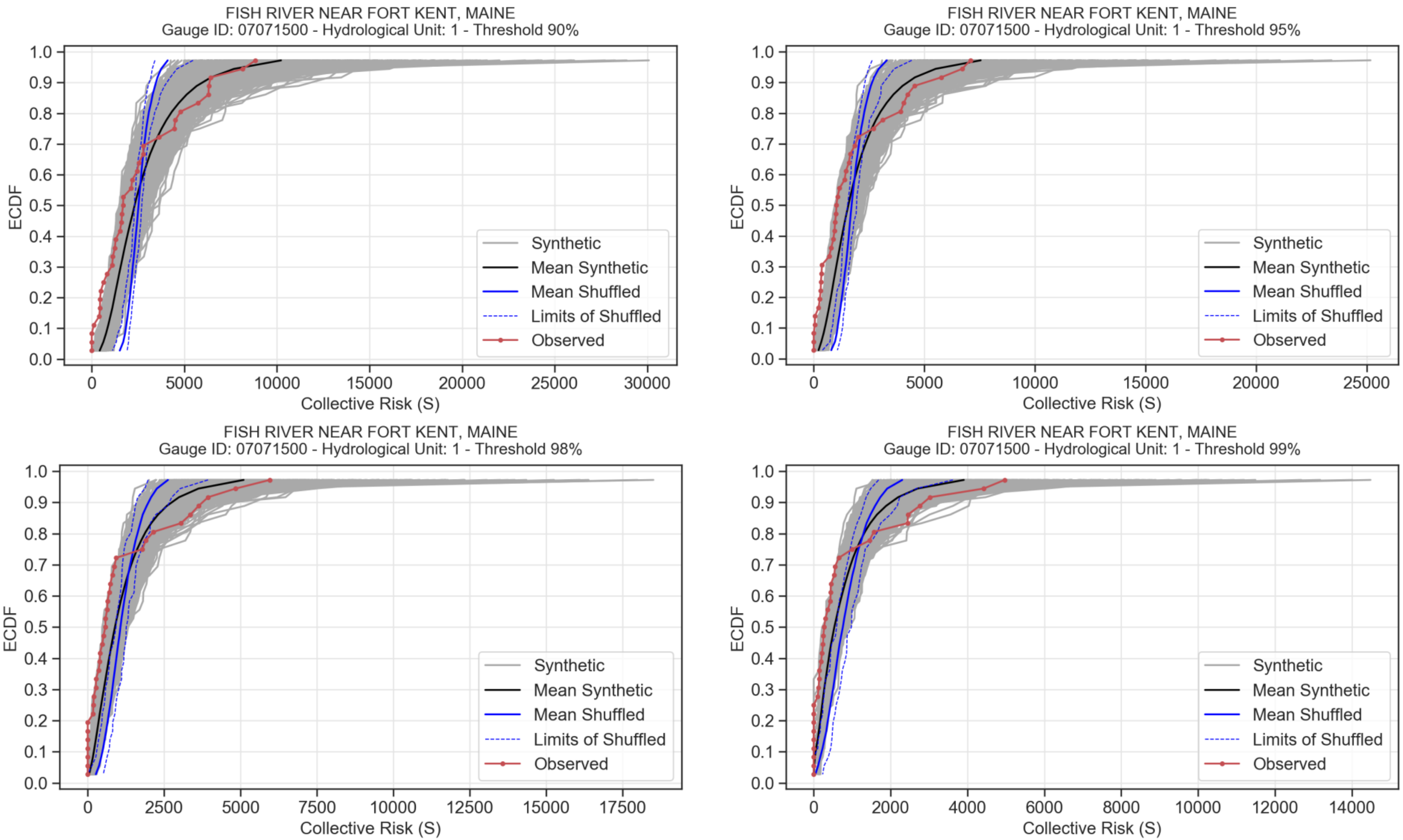
Results

ID 07067000 (Current River) - *H*: 0.72, *q*: 2.71 days



Results

ID 07071500 (Eleven Point River) - *H*: 0.77, *q*: 1.96 days



Conclusions

EGU General Assembly 2020

HK dynamics

The behavior of daily stream flow is found to be consistent with HK dynamics characterized by moderate H parameters (in the range 0.6-0.7), through Monte Carlo simulations.

GHK Model

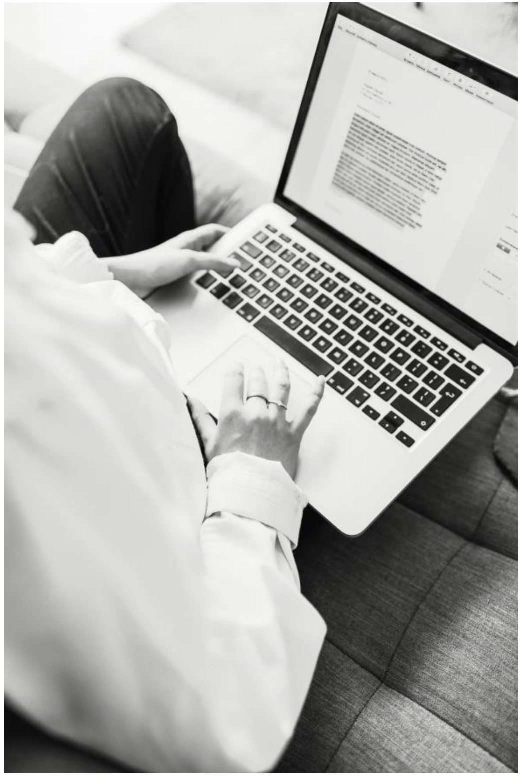
The ECDF curve of the observed collective risk proxy is contained in the Monte Carlo prediction limits by the GHK model, preserving the HK dynamics and the 4 four moments. In contrary, shuffled (randomized) curves have a different behavior, especially in the tails of the distribution.

POT Method

Results encourage the use of the POT method for sampling of extremes.

Threshold Impact

As the threshold increases, the deviation between the observed and the synthetic series increases, too.



Learn More

Additional material will be available at EGU2020.eu until 31st May 2020

Further information about the project are obtainable at our research team's site https://www.itia.ntua.gr/en/

References

- Dimitriadis, P. and Koutsoyiannis, D., 2018. *Stochastic synthesis approximating any process dependence and distribution*. Stochastic Environmental Research and Risk Assessment, 32(6), pp.1493-1515.
- Dimitriadis, P. and Koutsoyiannis, D., 2015. *Climacogram versus autocovariance and power spectrum in stochastic modelling for Markovian and Hurst–Kolmogorov processes.* Stoch. Env. Res. Risk A, 29, 1649–1669.
- Goulianou, T., Papoulakos, K., Iliopoulou, T., Dimitriadis, P. and Koutsoyiannis, D., 2019. *Stochastic characteristics of flood impacts for agricultural insurance practices.* European Geosciences Union General Assembly 2019, Vienna, EGU2019-5891.
- Koutsoyiannis, D., 2000. *A generalized mathematical framework for stochastic simulation and forecast of hydrologic time series.* Water Resources Research, 36 (6), 1519–1533.
- Koutsoyiannis, D., 2016. *Generic and parsimonious stochastic modelling for hydrology and beyond*. Hydrological Sciences Journal, 61 (2), 225–244.
- Kron, W., 2005. *Flood Risk = Hazard • Values • Vulnerability.* Water International, Volume 30, Number 1, 58–68.
- Newman, A., Sampson, K., Clark, M.P., Bock, A., Viger, R.J., and Blodgett, D., 2014. *A large-sample watershed-scale hydrometeorological dataset for the contiguous USA*. Boulder, CO: UCAR/NCAR.
- Papoulakos, K., Iliopoulou, T., Dimitriadis, P., Tsaknias, D. and Koutsoyiannis, D., 2020. *Investigating the impacts of clustering of floods on insurance practices; a spatiotemporal analysis in the USA*. European Geosciences Union General Assembly 2020, Vienna, EGU2020-8667.
- Serinaldi, F., and Kilsby, C.G., 2016. *Understanding Persistence to Avoid Underestimation of Collective Flood Risk*. Water, 2016, 8, 152.