







Risk Management: Extremes of Flood and Drought Europe/China, 28 February 2023

Extreme rainfall modelling for engineering design: a new methodology and its application over the Greek territory





Demetris Koutsoyiannis, Theano Iliopoulou, Antonis Koukouvinos, Nikolaos Malamos, Nikos Mamassis, Panayiotis Dimitriadis, Nikos Tepetidis & David Markantonis

School of Civil Engineering

National Technical University of Athens

Presentation available online: <u>http://www.itia.ntua.gr/2270/</u>

Importance of design rainfall at the regional scale

- Design rainfall estimation at a regional scale is the cornerstone of hydrological design against flooding, particularly essential:
 - for ungauged areas—but even for gauged ones;
 - for hydrological analyses at large areas, i.e., studies of regional flooding and construction of large-scale flood protection works—but even for small spatial scales, e.g., urban stormwater networks.
- Spatial generalization of estimation is essential, as often rainfall data for at-site analysis are missing.
- Design rainfall estimates are conveniently provided in the form of a mathematical relationship linking temporally averaged rainfall intensity to timescale of averaging and return period, usually known by the misnomer 'intensity-duration-frequency' (idf) curves or better named *ombrian curves*.

We aim to revisit design rainfall estimation for Greece:

- benefitting from new advances in the field of regional estimation and ombrian curves (Koutsoyiannis, 2022);
- exploiting new data recorded during the past decade.

For first time we have produced a geographically distributed model for the entire Greek territory.

Ombrian curves: a parsimonious regional approach

Spatial generalization of ombrian curves is particularly complex due to the need to account for spatial dependence together with the increased variability of rainfall extremes in space.

Broadly, the construction of regional ombrian curves can **follow two different approaches**:

- (a) the at-site, independent fitting approach, followed by spatial interpolation methods to map the parameters over the whole region.
- (b) the regional, simultaneous fitting approach, which consists of appropriately pooling the data together and obtaining a single model valid over the entire area, which is, in essence, the inverse approach to (a).

We devise a parsimonious approach to regionalizing the rainfall estimates at the Greek territory without resorting to uncontrolled interpolation.

Overview of theoretical framework (I)

- Koutsoyiannis (2022) developed a new
 framework for ombrian modelling
 that can be applied at any timescale,
 however large or small.
- The example shown is for Bologna, Italy (a station with 206 years of data), for timescales from 1 h to 16 years.
- For large timescales the mathematics are somewhat involved.



□ Under **some simplifying assumptions** the rainfall intensity *x* for **small timescales** *k* (of the order of minutes to a few days) and return period *T* is given by the following relationships, resulting from the full-scale ombrian model with some simplifying assumptions:

o for return period estimated from a full series or of rainfall exceedances over threshold:

$$x = \frac{b(T)}{a(k)} = \lambda \frac{(T/\beta)^{\xi} - 1}{(1 + k/\alpha)^{\eta}}, \qquad \xi > 0$$

• from series of annual maxima (where $\Delta = 1$ year): $x = \lambda \frac{(-(\beta/\Delta) \ln(1 - \Delta/T))^{-\xi} - 1}{(1 + k/\alpha)^{\eta}}, \quad \xi > 0$ Theoretically equivalent for all *T* and for the same parameter values; giving virtually same values for *T* > 10 years

The simplified model parameters are:

- λ a characteristic rainfall intensity (scale parameter) in units of x (e.g., mm/h);
- β a time parameter, related to the mean distance of wet periods, in units of the return period (e.g., years);
- α a timescale parameter in units of timescale (e.g., h) with α > 0;
- η a dimensionless parameter, expressing persistence, with $0 < \eta < 1$;
- $\xi > 0$ the tail index of the process distribution.

5 parameters with physical meaning

Two-step fitting procedure

An attractive feature of this simplification, related to the separable function x = b(T)/a(k), is that it allows the parameters to be estimated by a convenient, two-step procedure.

□ First step:

The timescale parameters (of the expression a(k)) obtained by Koutsoyiannis et al. (1998) optimization procedure.

□ Second step:

The distribution parameters (of the expression b(T)) are obtained by **the newly introduced method of Kmoments** (Koutsoyiannis, 2020), which has the following important properties:

✓ **Intuitive formulation**, as the K-moment of order *p* equals the expected value of the maximum of *p*;

independent stochastic variables identical to \underline{x} , i.e., $K'_p = E[\max(\underline{x}_1, \underline{x}_2, ..., \underline{x}_p)];$

- ✓ **Unbiased** (knowable even for very large orders);
- Can be readily assigned an empirical return period;

✓ Account for the effect of (spatial and temporal) dependence in the estimation of the return period.

Greece's rainfall network



- From the initial set of 940 stations, and after meticulous quality control processing, we compiled a final dataset of 783 stations, comprising:
 - 503 daily rain gauges, 130 of which at locations where there is also a rain recorder;
 - 280 rain gauges (rain recorders) with subdaily resolution.
- □ The stations are distributed over 651 geographical locations.
- □ The longest available record (in Athens) covers the period from 1860 to 2022.

Non-conventional rainfall data



Both data sets (especially the IMERG) underestimate the highest rainfall depths (as seen in the example for the station of Karditsa) and proved not appropriate for the construction of ombrian curves. From satellite-based information, we investigated the usefulness of the IMERG data set (half hourly time step at 0.1° spatial resolution, period 2000-today),
 From the reanalysis information we investigated the usefulness of the ERA5 data set (daily time step at 0.25° spatial resolution; period 1950-today).



On the use of the Hershfield factor:

- In our method, a fixed, rather than a moving, time window is used to extract the maximum for each scale. It is obvious that the maximum extracted from a fixed time window is less than or equal to that extracted from a moving time window, and it is known that the difference between the two is a function of the temporal resolution D of the raw data.
- Application of a correction factor, known as Hershfield factor, distorts the properties of the $\underline{x}_{\tau}^{(k)}$ series, replacing it with the series $\underline{w}_{\tau}^{(k)} := \max_{j} (\underline{x}_{\tau+j}^{(k)}, j = 0, ..., k 1)$. In a consistent stochastic framework, we should do not employ such a factor.

On the exploitation of different sources of rainfall data:

- The fitting of the timescale parameters (of the expression a(k)) is performed using sub-daily or even sub-hourly data, available from tipping-buckets and automated censors.
- The fitting of the distribution parameters (of the expression b(T)) is performed using in priority the daily rainfall records due to:
 - (a) the greater spatial density of the rain gauge network compared to that of rain gauges;
 - (b) the longer duration of rain gauge observations compared to those of rain gauges; and
 - (c) the greater reliability of rainfall measurement during storm events.

Independent at-site procedure

□ First, we perform a spatially-independent fitting of the ombrian curves for each location.

Then, we assess the resulting patterns of variability, and we identify the parameters exhibiting random spatial variation and the ones robust spatial patterns.

Regionalization procedure

- 1. We perform a combined (simultaneous) estimation of the parameters exhibiting random variation in space using the most reliable and relevant data for each case, e.g. we exploit the longest sub-hourly records for the estimation of the α parameter and the longest daily records for the estimation of the tail-index parameter ξ .
- 2. With the common parameters now fixed for all stations, we re-estimate the other parameters and assess their geographical variation.
- 3. In case that systematic patterns are identified, we model their geographic variation using both spatial smoothing and interpolation models, and evaluate their performance based on the accuracy of the fit and cross-validation metrics.
- 4. The best spatial model per parameter is chosen and a map with 5 km resolution is produced with the spatially varying parameters over Greece.

Inverse Distance Weighted (IDW)

The Inverse Distance Weighted (IDW) method is simple with low computational requirements. The IDW estimate for a given point \hat{z}_u is obtained as:

$$\hat{z}_{u} = \sum_{i=1}^{n} w_{i} z_{i}(x_{i'} \ y_{i}) \quad \text{with weights } w_{i} \text{ obtained as:}$$

$$w_{i} = \frac{d_{u \ i}^{-a}}{\sum_{i=1}^{n} d_{u \ i}^{-a}}, \quad \sum_{i=1}^{n} w_{i} = 1$$

The values of α and n are identified as the ones yielding the lowest cross-validation errors.

Bilinear surface smoothing (BSS)

Bilinear Surface Smoothing (BSS) is a flexible spatial interpolation method that adjusts a bilinear surface at known points (x_i, y_i) through linear regression with adjustable weight factors (Malamos and Koutsoyiannis, 2016a,b).

The method is based on compromising two opposing objectives, namely, to minimize the fitting error and the roughness of the fitted bilinear surface.

Additionally, it is possible to integrate, in an objective way, the effect of an explanatory variable available at a spatially denser data set (Bilinear Surface Smoothing with an Explanatory variable-BSSE). In this case, two bilinear surfaces are combined in the same regression model in order to improve the accuracy of the interpolation at the given points.

Regionalization of timescale parameters – α (I)

- \Box We find that the estimation of the parameter α greatly depends on the temporal resolution of the measuring instrument. Specifically, in stations with fine temporal resolution (5 or 10 min) resulting values of the parameter α are small—and vice versa.
- This is interpreted as an artificial statistical effect rather that as representing some physical reality.



Regionalization of timescale parameters – α (II)

 \Box To compensate for the great sensitivity of the α parameter to time resolution of the data, we identify a single value of this parameter for all of Greece, by the following procedure:

- We select the 53 stations with the longest records having temporal resolution 30 min or finer, distributed over all water districts.
- We re-estimate the parameters of the equation a(k), α and η , through optimization in which we set as a constraint that the value of the α parameter is the same among all stations.
- □ As a result of this methodology, the common value of $\alpha = 0.18$ h is obtained, which is used in all further analyses.

Regionalization of η (I)



The map shows point estimates of the η parameter conditional on a common α parameter, $\alpha = 0.18$ h:

- Presence of clusters of low and high values of the η parameter in space.
- Emergence of an inverse relationship η with the altitude (i.e. lower values of the parameter are more likely at high altitudes).

Regionalization of η (II)



We use the BSSE smoothing model with the altitude (derived from SRTM) as an additional explanatory variable.

	Calibration	Leave-one-out-cross- validation, (LOOCV)
Bias (MBE)	0.00	0.00
Mean Absolute Error(M	AE) 0.05	0.06
Root Mean Square Error (I	RMSE) 0.06	0.07
Nash-Sutcliffe efficiency	(EF) 0.57	0.40
Coefficient of Determinati	on(R ²) 0.57	0.40



Regionalization of distribution parameters – ξ (I)



- The parameter ξ (tail index of the distribution) was estimated individually per station and per instrument, and simultaneously with the optimization of the other parameters of the rainfall curves.
- We observe the large spatial variability of the parameter estimates, which reflects both the measurement uncertainty of maximum rainfall as well as the typical absence of systematic patterns in the realization of extreme rainfall.

Regionalization of distribution parameters – ξ (II)



- If we assume that the entire variability of ξ estimates is a statistical effect, then:
 - We can unify (merge) all records at a certain timescale after standardizing with the mean;
 - We can estimate a unique value of ξ from the unified record.
- We have used 61 stations across the Greek territory which have at least 60 years of complete daily timeseries.
- These form a large sample of 299 481 (standardized) nonzero daily rainfall values.
- The resulting ξ is estimated to 0.18 if the different stations are assumed independent (Θ = 0) or larger if dependence is assumed (ξ = 0.23 for Θ = -0.04, where Θ denotes bias; see Koutsoyiannis, 2022 for details).
- The minimal value if $\xi = 0.18$ is finally chosen.

Regionalization of distribution parameters – ξ (III)

Monte Carlo simulation results (70 simulations with Pareto distribution, each corresponding to 70 years of rainfall):

- Show the large variability of the estimated value of ξ (ξ_e), spanning from ~-0.1 to ~0.5, when the true value is ξ = 0.18.
- Verify the consistency of the assumption of a single $\xi = 0.18$ for the entire Greece.



In addition to the direct regionalization of parameters β and λ , we investigated the use of alternative quantities linked to characteristic rainfall intensities, since the statistical behaviour of the latter is more robust and better suitable for regionalization (no boundary issues and better spatial coherence).

Specifically, we express parameters β and λ as functions of either the rainfall intensities x_1 and x_2 corresponding to return periods T_1 = 2 years and T_2 = 100 years, respectively, or equivalently, of x_1 and the ratio $r_x \coloneqq x_2/x_1$, as follows:

$$\beta = \left(\frac{r_x - r_T}{r_x - 1}\right)^{1/\xi} T_1, \qquad r_T \coloneqq (T_2/T_1)^{\xi}, \quad r_x \coloneqq x_2/x_1$$
$$\lambda = b \frac{r_x - r_T}{r_T - 1} x_1, \qquad b \coloneqq (1 + k/\alpha)^{\eta}$$

After examining the correlations between the alternative parameters sets, we chose to use the pair of parameters x_1 and r_x , since they are found uncorrelated with each other and thus the pair's information content is not affected by redundancy.

The intensities x_1 and x_2 are modelled at the 24 h scale (k = 24 h).

Regionalization of x_1



In the case of the x_1 parameter, the best statistical characteristics of the fitting in terms of leave-one-out crossvalidation resulted from the application of the spatial model of the IDW method with the following characteristics:

- Neighboring points included: 5
- Minimum number of points: 1
- 1 sector





Regionalization of r_{χ}



In the case of the r_{χ} parameter, the best statistical characteristics of the fitting in terms of leave-one-out cross-validation resulted from the application of the spatial model of the IDW method with the following characteristics:

- Neighboring points included: 20
- Minimum number of points: 2
- 4 sectors with an angle between the 2 axes of 0°
- Exponent of the distance expression, *a* = 1



Regionalization of distribution parameters – $\theta \& \lambda$



Assessment of regionalization accuracy



Final parameterization

The following generalized form of ombrian curves is derived for rainfall intensity x (mm/h), return period T (years) and temporal scale k (h): $(T/B_{1})^{\xi} = 1$

$$x = \lambda_* \frac{(T/\beta_*)^{\varsigma} - 1}{(1 + k/\alpha)^{\eta_*}}$$

with the following five parameters

- characteristic timescale α = 0.18 h
- tail index $\xi = 0.18$,
- three spatially varying parameters $\eta_*[-]$, β_* (years) and λ_* (mm/h) :



At-site verification: mean % deviations and RMSE



From the analysis of the daily rain gauges, it follows that the median of the average deviation for the 24 h scale is +9.05% while the average value is +8%. The 95% interval of the deviations is [-22.06%, +31.51%], consistent with the uncertainty of precipitation characterized by a high tail index ξ .

The results for the sub-daily rain gauges are also satisfactory, although showing relatively larger ranges of deviations at small scales as rain recorders are fewer and impacted by greater measurement uncertainty.

At-site verification: maximum 24 h depth deviation



We assess the deviations of each record's maximum 24 h rainfall depth to the one obtained for the same return period (assigned through K-moments).

The agreement between the two is very satisfactory ($R^2=0.68$) given the large spatial extent of the analysis and considering that the record's maximum value is a statistical quantity governed by high uncertainty, especially for a large tail index ($\xi = 0.18$).

Impact of regionalization on return period estimates

To inspect the impact of regionalization on design rainfall estimates for various return periods, we compare the deviations between the estimates using regional parameters and the ones obtained using the local (at-site) parameters.



For all scales, the deviations consistently tend to increase towards larger values as the return period increases. This is due to the use of the high single value of the parameter ξ in the regionalization, the influence of which is stronger in large return periods. In the very short return periods (of the order of 2 years), the spatially generalized rainfall model leads to slightly smaller rainfall estimates (for T = 2 years, median -3.55% at 1 h and -5.22% at 24 h). This fact is partly attributed to the non-use of Hershfield factors for the daily rain gauge data which greatly affect the spatial generalization of the distribution parameters.

Mapping characteristic design rainfall depths



Ombrian curves at the catchment scale



The ombrian curves for any region within the Greek territory are derived based on the two constant-value parameters and the three regionally varying parameters which are obtained as a weighted average of the grid points falling within the area.

Inspection of long-term variability – Benchmark series (Bologna)



More than a century long time series in Greece: Athens

- Compared to Bologna, Athens shows climate stability.
- In the last 30 years there has been no remarkable climatic event.
- The largest annual rainfall in history was recorded in the hydrological year 1885-86, and the smallest in 1989-90.
- The all-time high record of rainfall depth, 150.2 mm/d, occurred at the end of the 19th century (1899-90).



More than a century long time series in Greece: Thessaloniki

- Thessaloniki shows climatic stability, similar to Athens.
- In the last thirty years there has been no remarkable climatic event.
- The largest annual rainfall in history was recorded in the hydrological year 1918-19, and the smallest in 1984-85.
- The all-time high record of rainfall depth, 115.9 mm/d, occurred in the hydrological year 1985-86.



Temporal distribution of records of average daily rainfall in 62 stations in the country

- The 1950s and early 1960s were strongly wet.
- About 1/3 of the high records of annual rainfall occurred in a single year, the hydrological year 1962-63.
- The 20-year period centered in 1990 is remarkably dry.
- In particular, about half of the low records of annual rainfall occurred in the 5-year period centered in 1990.
- The other periods, including the current one, are climatically neutral.
- The entire picture suggests the presence of Hurst-Kolmogorov dynamics in time and space.



Temporal distribution of records of maximum daily rainfall in 238 stations in the country

- The distribution is as statistically expected
- An exception is the lack of a record in the three-year period 1982-83 to 1984-85.
- There are no noticeable climatic events.



Climate trends over the last two 30 years: linear trends and differences of two consecutive 30-year climatic periods

- The probability distribution of positive and negative trends is balanced.
- There is an impressive agreement of the empirical variations with the theoretically expected for a stationary process.



- The approach followed incorporates an advanced framework for regional frequency analysis employing knowable (K-) moments that allow:
 - reliable high-order moment estimation; significantly increasing the number of moments that can be justifiably employed in regional analyses of extremes; and
 - ✓ handling of temporal and spatial dependence, which is non-negligible.
- The detailed climatic analysis:
 - did not locate any element that would justify any type of nonstationary analysis;
 - ✓ yet it suggests the presence of changes that can be modelled within a stationary framework of Hurst-Kolmogorov dynamics.
- The final product is a powerful tool, easy to apply for engineering tasks, covering the entire territory of Greece.
- The methodology can be readily applied to other countries or parts thereof.

References

- Iliopoulou, T., Malamos, N., and Koutsoyiannis, D., 2022. Regional ombrian curves: Design rainfall estimation for a spatially diverse rainfall regime, *Hydrology*, 9 (5), 67, doi:10.3390/hydrology9050067.
- Iliopoulou, T., and Koutsoyiannis, D., 2020. Projecting the future of rainfall extremes: better classic than trendy, Journal of Hydrology, 588, doi:10.1016/j.jhydrol.2020.125005.
- Koutsoyiannis, D., 2019. Knowable moments for high-order stochastic characterization and modelling of hydrological processes, Hydrological Sciences Journal, 64 (1), 19–33, doi:10.1080/02626667.2018.1556794.
- Koutsoyiannis, D., 2022. Stochastics of Hydroclimatic Extremes A Cool Look at Risk, Edition 2, ISBN: 978-618-85370-0-2, 346 pages, doi:10.57713/kallipos-1, Kallipos Open Academic Editions, Athens.
- Koutsoyiannis, D., Kozonis, D. and Manetas, A., 1998. A mathematical framework for studying rainfall intensity-duration-frequency relationships, Journal of Hydrology, 206 (1-2), 118–135, doi:10.1016/S0022-1694(98)00097-3.
- Malamos, N. and Koutsoyiannis, D., 2016a. Bilinear surface smoothing for spatial interpolation with optional incorporation of an explanatory variable.
 Part 1:Theory, Hydrological Sciences Journal, 61 (3), 519–526, doi:10.1080/02626667.2015.1051980.
- Malamos, N. and Koutsoyiannis, D., 2016b. Bilinear surface smoothing for spatial interpolation with optional incorporation of an explanatory variable.
 Part 2: Application to synthesized and rainfall data. *Hydrological Sciences Journal*, 61(3), pp.527-540.

Details on the methodology in the book (Edition 2)

Free in open access

https://www.itia.ntua.gr/2000/

Stochastics of Hydroclimatic Extremes is a real monument in stochastics! It is a summary of the lifetime dedication by Demetris Koutsoyiannis to the science of environmental extremes, it is a demonstration of the value of stochastics itself to gain a better understanding of why and how extremes happen. The perspective adopted in the book is that of a scientist who is able to cross and transform disciplines by proposing an innovative synthesis of knowledge. This book is indeed presenting new concepts, new theoretical interpretations and new opportunities for engineering design, for the sake of mitigating the impact of extremes and adapting modern society to environmental variability.

It is fascinating that the book is self-produced and openly available to readers. Like any self-produced creation of the humankind, this book has a unique and independent history that is rooted in the intimate personality of the author. It is a creation that does not require to adhere to any format other than those suggested by the author's vision and creativity. For this reason, its value is incommensurably high, it is a real *Cool Look at Risk* as Demetris says.

I believe time will highlight *Stochastics of Hydroclimatic Extremes* as a transforming masterpiece which will bring illuminating ideas to the reader.

Alberto Montanari Head of the Dept. of Civil, Chemical, Environmental, and Materials Engineering, University of Bologna President of the European Geosciences Union

This is a book that could not only transform your career, but also the entire fields of environmental statistics and stochastic hydrology. This seminal contribution is not like other books you have read which tend to summarize existing knowledge. Rather, it condenses existing knowledge in short order and spends nearly all its time on new knowledge, much of it never before published, communicating effectively both the theoretical and practical aspects of analysis of a wide range of hydroclimatic extremes. The style of presentation itself is novel and compelling, so that I could not resist reading it from cover to cover.

If you think you understand how to apply probability and statistics to predict future extreme events, think again, because very quickly you will be convinced that extremes arise from spatial and temporal stochastic processes, and are neither independent nor identically distributed (iid) events, nor do most of our common probability distributions used for flood and drought frequency analysis capture the type of thick tails which are so convincingly documented in this book.

I predict that many of the novel concepts, examples and techniques introduced here, many for the first time, will find their way into widespread acceptance in hydroclimatology, over time. Foremost, the reader will appreciate the value of viewing extreme events as realizations of stochastic processes rather than a series of iid annual maxima/minima. The climacogram provides a new window into the structure of stochastic processes and may be more fundamental than the correlogram. I can't wait to test out the so-called Pareto-Burr-Feller distribution and the novel knowable moments (K-moments) which appear to have clear advantages over ordinary moments for describing distribution tails.

It is remarkable that after a long career in hydrology, after reading this book, I gained many new insights into common statistical methods as well as new methods documented here for the first time. How I wish my career were just beginning, and thus could have applied all the wonderful ideas and methods in this book during my career. This is literally a treasure for young scholars interested in the probabilistic behaviour of hydroclimatic extremes.

Richard M. Vogel Professor Emeritus and Research Professor, Dept. Civil and Environmental Engineering, Tufts University

Stochastics of Hydroclimatic Extremes A Cool Look at Risk

Demetris Koutsoyiannis National Technichal University of Athens

2nd Edition

Demetris Koutsoyiannis

Extremes

of Hydroclimatic

Stochastics

Kallipos Open Academic Editions Athens 2022

