



# Why is modelling of extreme rainfall important for urban flood risk estimation?

- ◆ Flow measurements are never enough to support flood modelling using only flow data
- ◆ Particularly, in urban floods the control points are numerous and the flow gauge sites scarce
- ◆ The example of Athens: No flow gauge at all in Kephisos and Ilisos Rivers and in other urban streams
- ◆ The obvious alternative: Use of hydrological models with rainfall input data

# Brief history of the probability of extreme values

- ◆ **1914 (Hazen):** Empirical foundation of hydrological frequency curves known as “duration curves”
- ◆ **1922, 1923 (von Bortkiewicz, von Mises):** theoretical foundation of probabilities of extreme values
- ◆ **1958 (Gumbel):** convergence of empirical and theoretical approaches
- ◆ **Today:** the estimation of hydrological extremes continues to be highly uncertain

*“... the increased mathematisation of hydrological frequency analysis over the past 50 years has not increased the validity of the estimates of frequencies of high extremes and thus has not improved our ability to assess the safety of structures whose design characteristics are based on them. The distribution models used now, though disguised in rigorous mathematical garb, are no more, and quite likely less, valid for estimating the probabilities of rare events than were the extensions ‘by eye’ of duration curves employed 50 years ago.” (Klemeš, 2000)*

# The notion of distribution of maxima

- ◆ Parent variable:  $Y$  (e.g. the hourly or daily rainfall depth)
- ◆ Parent distribution function:  $F(y)$  (with density  $f(y)$ )
- ◆ Variable representing maximum events

$$X := \max \{Y_1, Y_2, \dots, Y_n\}$$

- ◆ Distribution function of maxima:  $H_n(x)$
- ◆ Exact distribution of maxima for constant  $n$ :

$$H_n(x) = [F(x)]^n$$

- ◆ Exact distribution of maxima for randomly varying  $n$ , following a Poisson process

$$H'_\nu(x) = \exp\{-\nu[1 - F(x)]\}$$

# The notion of asymptotic or limiting distribution of maxima

- ◆ Asymptotic or limiting distribution for  $n \rightarrow \infty$  or  $\nu \rightarrow \infty$   
(Generalised extreme value distribution – GEV; Jenkinson, 1955)

$$H(x) = \exp\{-[1 + \kappa(x/\lambda - \psi)]^{-1/\kappa}\} \quad (\kappa x \geq \kappa\lambda(\psi - 1/\kappa))$$

- ◆ In hydrology, an upper bound of  $x$  is not realistic, so  $\kappa \geq 0$
- ◆ If  $\kappa > 0$ ,  $H(x)$  represents the (three-parameter) extreme value distribution of maxima of type II (EV2)
- ◆ In the special case  $\kappa = 0$ ,  $H(x)$  represents the extreme value distribution of maxima of type I (EV1 or Gumbel)

$$H(x) = \exp\{-\exp[-(x/\lambda - \psi)]\} \quad (-\infty < x < +\infty)$$

- ◆ In the special case where the lower bound is zero ( $\kappa \psi = 1$ ),  $H(x)$  is two-parameter EV2 (Fréchet distribution)

$$H(x) = \exp\{-[\lambda/(\kappa x)]^{1/\kappa}\} \quad (x \geq 0)$$

# When and why does the type of extreme value distribution (EV1 or EV2) matter?

- ◆ In typical storm sewer networks, designed on the basis on return periods [ $T = 1 / (1 - H)$ ] of about 5-10 years, the difference of the two distributions may be negligible (in such return periods even interpolation from the empirical distribution would suffice)
- ◆ However, for large  $T$  ( $> 50$  years), for which extrapolation is required, EV1 results in risk (probability of exceedence of a certain value) significantly lower than EV2
- ◆ That is, for large rainfall depths, EV1 yields the lowest possible probability of exceedence (the highest possible  $T$ ) in comparison to those of EV2 for any value of  $\kappa$
- ◆ For  $T > \sim 1000$ , the return period estimated by EV1 could be orders of magnitude higher than that of EV2

# What is the prevailing model in hydrological practice?

- ◆ Definitely, EV1
- ◆ For example, most hydrological textbooks do not mention EV2 at all
- ◆ Also, in most hydrological studies the adoption of EV1 is “automatic” (especially for extreme rainfall)
- ◆ Recently, however, many researchers have expressed scepticism about the appropriateness of EV1

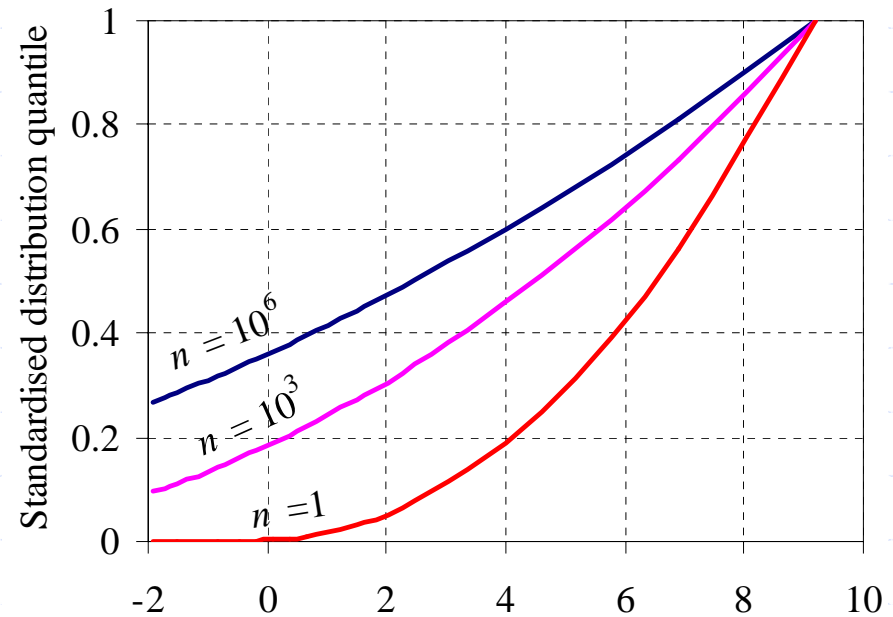
# Are there theoretical reasons favouring EV1 against EV2?

- ◆ Most types of parent distributions functions used in hydrology, such as **exponential, gamma, Weibull, normal** and **lognormal** belong to the domain of attraction of the Gumbel distribution
- ◆ More specifically, rainfall depth at fine time scales (hourly, daily) has been modelled by the gamma or Weibull distributions
- ◆ However, the adoption of these distributions is rather empirical, not based on theoretical reasoning
- ◆ Thus, the above theoretical argumentation is inconsistent



# Assuming that theoretical reasoning supports EV1, what distribution shall I use in my studies?

- ◆ **Intuitive answer:** EV1
- ◆ **Correct answer:** The exact distribution of maxima,  $H_n(x)$  or  $H'_v(x)$
- ◆ The difference of  $H_n(x)$  from EV1 may be large
- ◆ **Practical answer:** EV2 [it yields good approximation of  $H_n(x)$ ]



Gumbel reduced variate

Convergence of distribution of maxima for parent distribution Weibull with shape parameter  $k = 0.5$

Note: The distribution quantiles have been standardised by  $x_{0,9999}$  corresponding to  $z_H = 9.21$

# What could be a basis of developing theoretical arguments for the selection of the type of the parent distribution?

- ◆ The principle of maximum entropy (ME) is a well established mathematical and physical principle that can yield unknown probabilities by theoretical reasoning
- ◆ Entropy is defined on grounds of probability theory
- ◆ The standard or Boltzmann-Gibbs-Shannon or extensive entropy is

$$\phi := E[-\ln f(Y)] = -\int_{-\infty}^{\infty} f(y) \ln f(y) dy$$

- ◆ According to a generalized definition, the Tsallis or nonextensive entropy is

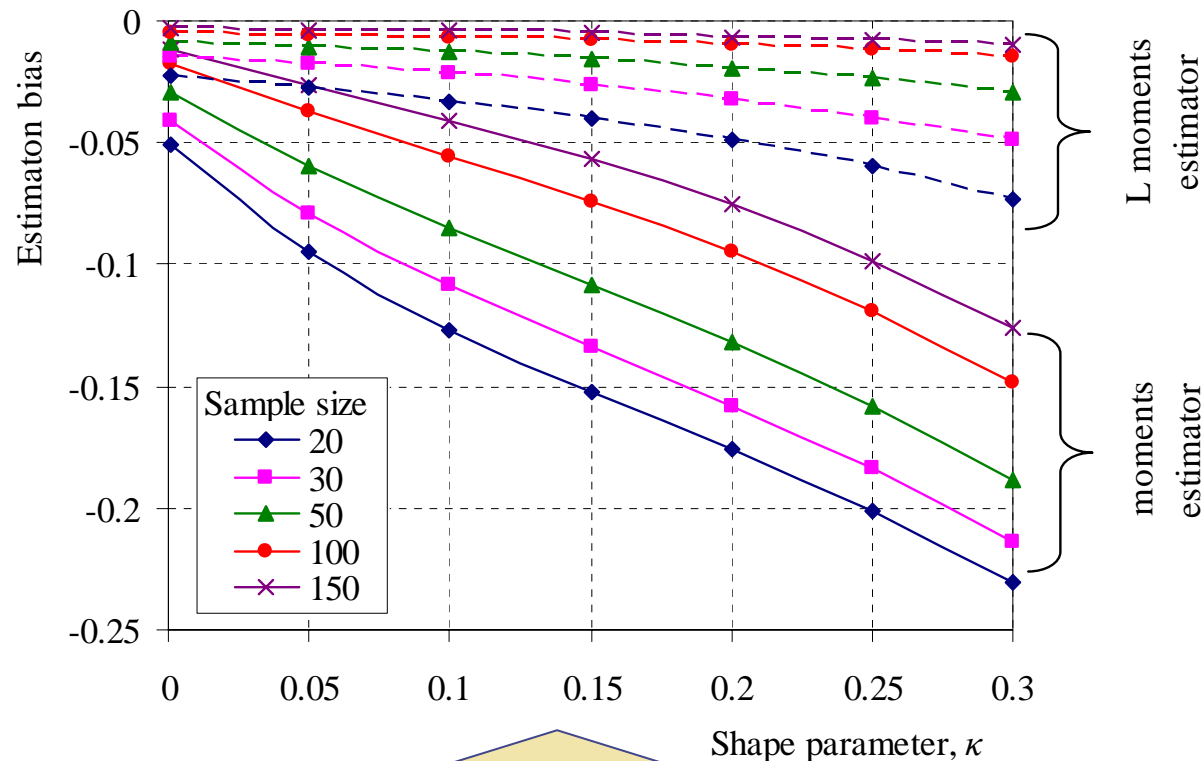
$$\phi_q := \frac{1 - \int_0^{\infty} [f(y)]^q dx}{q - 1}$$

# Application of the ME principle

- ◆ We assume that  $Y$  is continuous and positive, has known mean  $\mu$  and known variation  $\sigma/\mu$ . We can then estimate the distribution function with only this information, applying the ME principle
- ◆ Maximization of the standard entropy is possible when variation is relatively low ( $\sigma/\mu < 1$ ) and yields exponential distribution tails (and distribution from normal to exponential)
- ◆ However, empirical evidence suggests that at small time scales rainfall exhibits high variation ( $\sigma/\mu > 1$ )
- ◆ In this case we should apply Tsallis entropy, which yields power-type (Pareto) distribution

# Why inappropriateness of EV1 has not become evident?

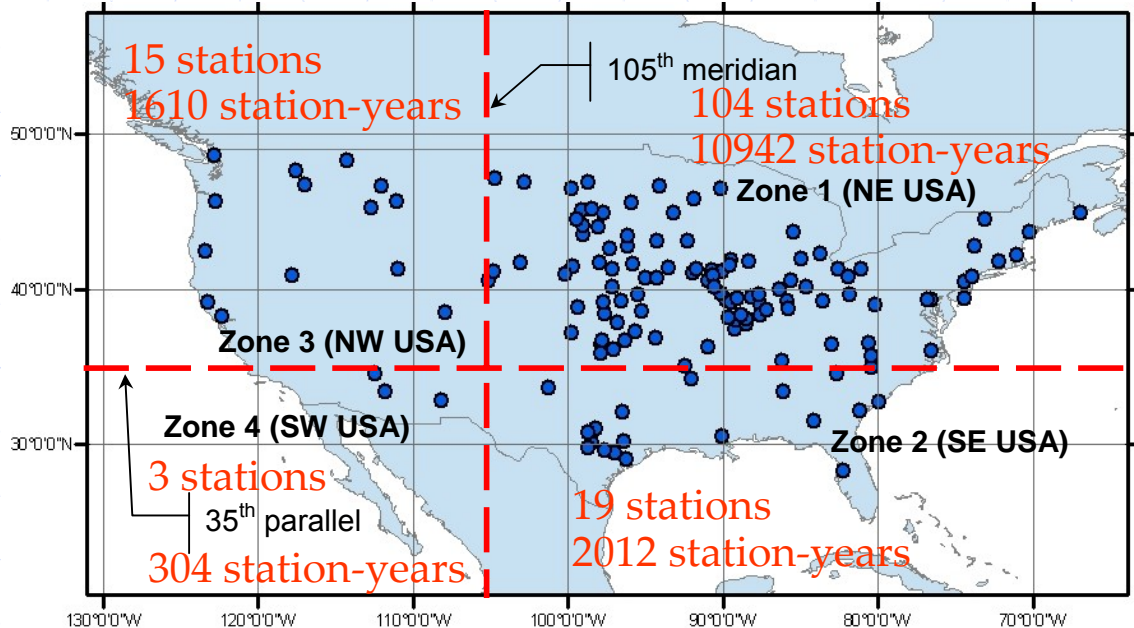
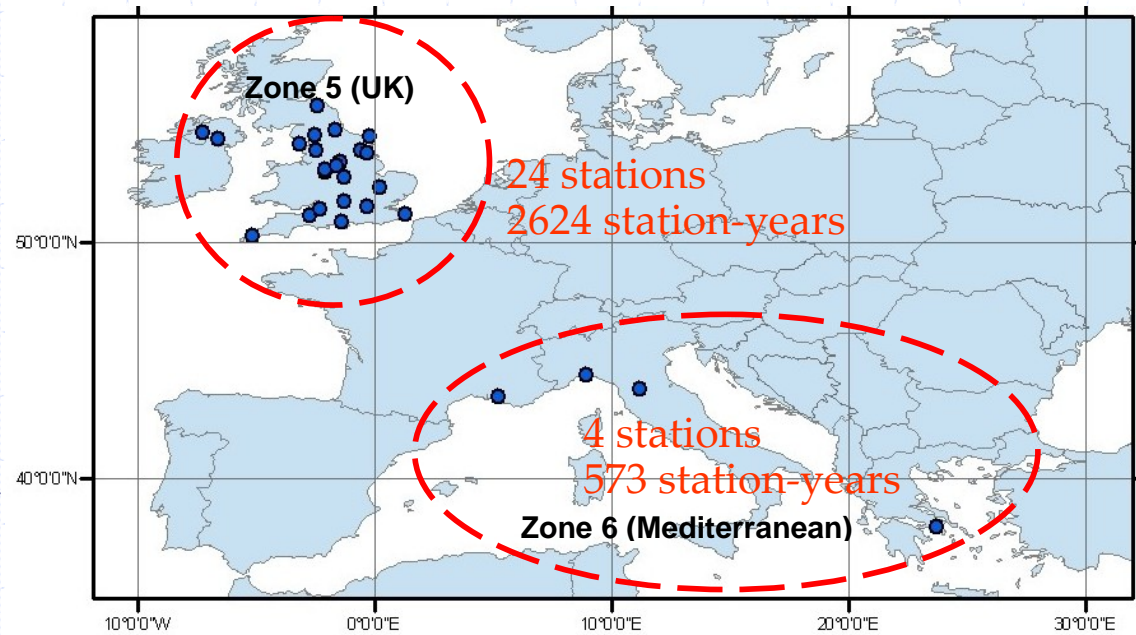
- ◆ Even if one is willing to try EV2 as a potential model, it is very likely that one will reject it due to significant bias of estimators
- ◆ For small samples, the most common method of moments hides completely EV2
- ◆ Even the less biased L-moments method may result in erroneous acceptance of EV1 (e.g. for  $\kappa = 0.15$  and  $m = 20$  the frequency of not rejecting the EV1 distribution is 80%!)



Bias in estimating the shape parameter  $\kappa$  of the GEV distribution  
(Obtained by Monte Carlo simulation)

# Empirical investigation: Data set

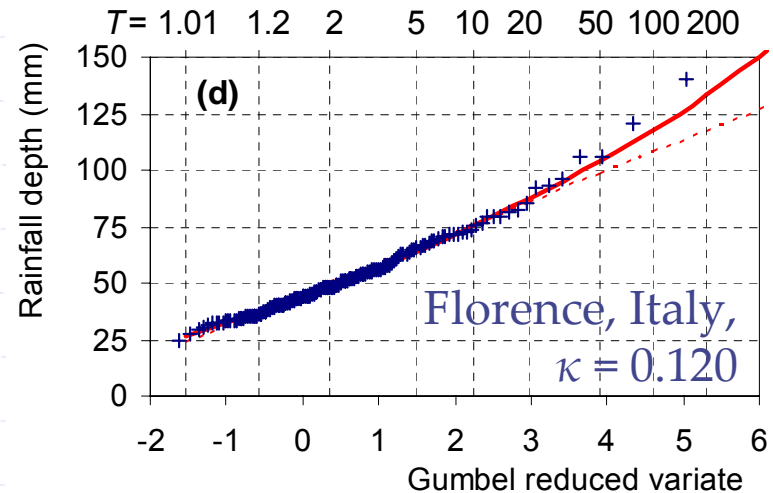
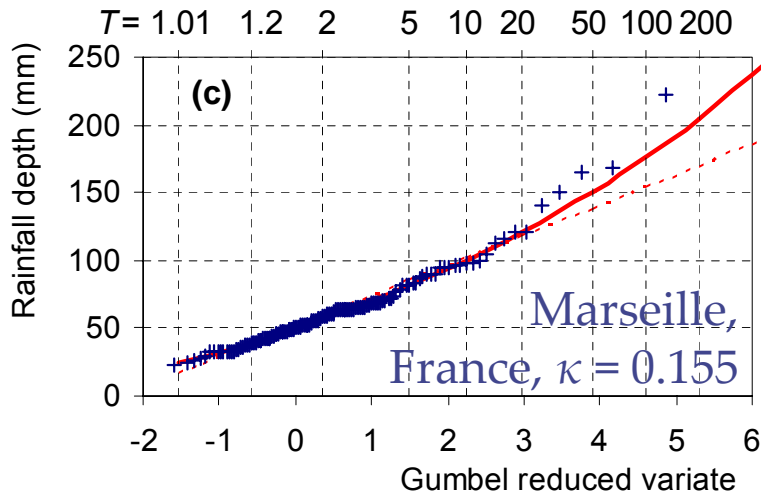
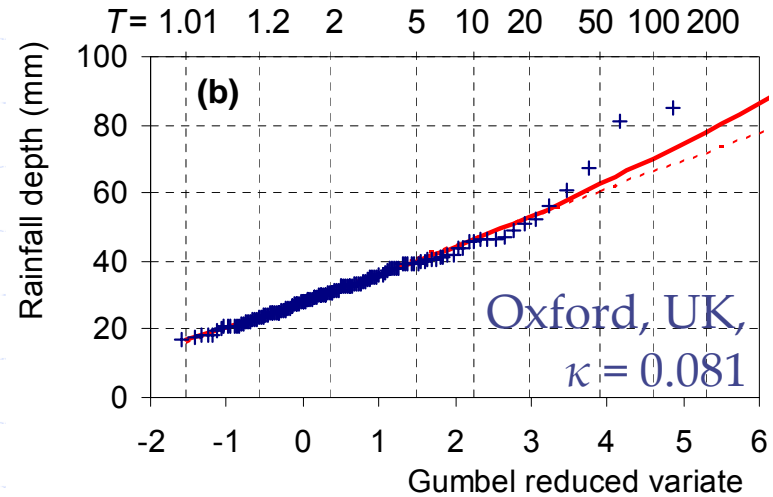
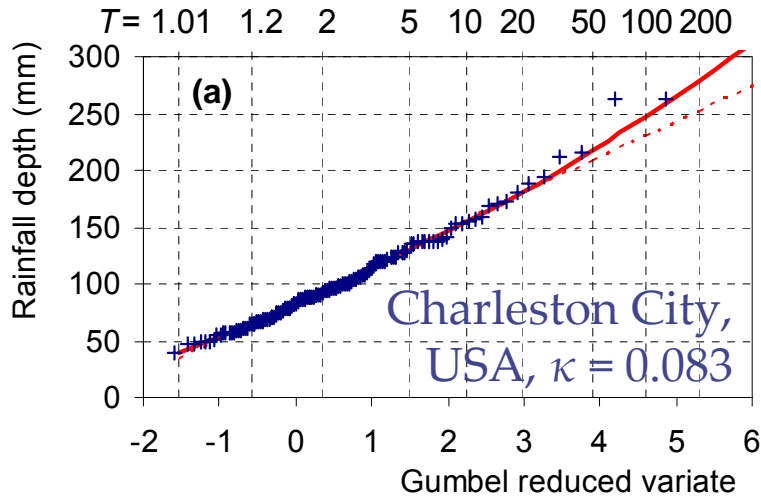
- ◆ 169 stations from Europe and North America
- ◆ Record lengths 100-154 years
- ◆ 18065 station-years in total
- ◆ 6 major climatic zones



# Top ten raingauges (in terms of record length)

Name	Zone /Country /State	Latitude (°N)	Longitude (°)	Elevation (m)	Record length	Start year	End year	Years with missing values
Florence	6/Italy	43.80	11.20	40	154	1822	1979	4
Genoa	6/Italy	44.40	8.90	21	148	1833	1980	
Athens	6/Greece	37.97	23.78	107	143	1860	2002	
Charleston City	2/USA/SC	32.79	-79.94	3	131	1871	2001	
Oxford	5/UK	51.72	-1.29		130	1853	1993	11
Cheyenne	1/USA/WY	41.16	104.82	1867	130	1871	2001	1
Marseille	6/France	43.45	5.20	6	128	1864	1991	
Armagh	5/UK	54.35	-6.65		128	1866	1993	
Savannah	2/USA/GA	32.14	-81.20	14	128	1871	2001	3
Albany	1/USA/NY	42.76	-73.80	84	128	1874	2001	

# Investigation of empirical distributions and comparison with EV2 and EV1 distributions



EV2 and EV1 distributions were fitted by the method of L-moments

# Demonstration of the differences of EV1 and EV2 estimates of quantiles for high return periods

