ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΤΜΗΜΑ ΠΟΛΙΤΙΚΩΝ ΜΗΧΑΝΙΚΩΝ
ΤΟΜΕΑΣ ΥΔΑΤΙΚΩΝ ΠΟΡΩΝ
ΔΡ Δ. ΚΟΥΤΣΟΓΙΑΝΝΗΣ, ΕΠ. ΚΑΘΗΓΗΤΗΣ

NATIONAL TECHNICAL UNIVERSITY
FACULTY OF CIVIL ENGINEERING
DEPARTMENT OF WATER RESOURCES
DR D. KOUTSOYIANNIS, ASST. PROFESSOR

15 January 1998

Dr. J. S. G. McCulloch,
Water Resource Systems Research Laboratory
Department of Civil Engineering
University of Newcastle upon Tyne
Newcastle upon Tyne
United Kingdom

Dear Dr. McCulloch,

Please find enclosed four copies of a manuscript entitled *Effect of the sample size on design rainfall inferences: an example using a long record of annual maximum rainfall in Athens, Greece* by D. Koutsoyiannis and G. Baloutsos, to be considered for publication in *Hydrology and Earth System Sciences*.

Although in the focus areas indicated in the leaflet of your Journal the theme of our manuscript is not explicitly included, we do hope that you will find the topic of the manuscript satisfactory for the Journal. May I add that we would very much prefer to publish the manuscript in a European Journal. In case that you decide to process the manuscript I would suggest as possible Scientific Editor Prof. E. Foufoula-Georgiou or Prof. P. E. O'Connell.

I am looking forward to hearing from you at your earliest convenience.

Sincerely,

Demetris Koutsoyiannis

ΗΡΩΩΝ ΠΟΛΥΤΕΧΝΕΙΟΥ 5
157 80 ΖΩΓΡΑΦΟΥ
ΤΗΛ. 772 2831 - ΦΑΞ 772 2853
ΗΛ. ΤΑΧΥΔΡΟΜΕΙΟ dk@hydro.ntua.gr

HEROON POLYTECHNIOU 5
GR-157 80 ZOGRAFOU, ATHENS, GREECE
TEL. +30(1) 772 2831- FAX +30(1) 772 2853
E-MAIL dk@hydro.ntua.gr

Effect of the sample size on design rainfall inferences: an example using a long record of annual maximum rainfall in Athens, Greece

Demetris Koutsoyiannis

Department of Water Resources, Faculty of Civil Engineering,

National Technical University, Athens, Greece

George Baloutsos

Forest Research Institute,

National Agricultural Research Foundation, Athens, Greece

Abstract. An annual series of maximum daily precipitation extending through 1860-1995, i.e., 136 years, was extracted from the daily rainfall measurements of a station at Athens. This is the longest rainfall record available in Greece and its analysis reveals remarkable statistical properties, which are very useful both to predict intense rainfall in Athens, where currently major flood protection works are under way, and to infer some properties of rainfall in other parts of the country where existing records are much shorter. Specifically, the long record allows the comparison of several types of distribution functions such as the Type I (Gumbel) and Type II distributions of maxima and the General Extreme Value (GEV) distribution and the appropriate choice among them. All these distributions might be chosen as appropriate models if less years of measurements were available (i.e., part of this sample were used). Nevertheless, the extrapolation for return periods such as 1000 years or more result in huge differences among the different distributions; for example Type II distribution results in more than twice rainfall in comparison with Type I. The available long record results in rejection of both Type I and II distributions. On the contrary, the General Extreme Value (GEV) distribution appears to be suitable for the examined series and its predictions for large return periods agree with the estimate of probable maximum precipitation obtained by the statistical (Hershfield's) method, when the latter is considered from a probabilistic point of view. Although a 136-year record is still too short to accurately determine the upper tail of

the distribution function of maximum rainfall, it is quite longer than typical samples available for hydrologic applications such as the construction of intensity-duration-frequency relationships. Thus, this record provides a clearer view of such relationships for large return periods, and the fitted GEV distribution is directly utilised for establishing their mathematical expressions.

1. Introduction

In modern times, particularly after World War II, Athens, the capital of Greece, has been continually urbanised, nowadays reaching a population of about four million. Unfortunately, urbanisation has seldom been combined with concurrent infrastructure works, such as natural channel improvement and storm drainage networks. Moreover, there are cases where buildings were illegally constructed over or very close to ephemeral stream beds. Thus, flooding in Athens is probably the most severe among natural hazards in Greece (Koukis and Koutsoyiannis, 1997). Since 1896, at least 179 lives were lost due to floods in Athens. The most catastrophic flood events were those of 14 November 1896, 5-6 November 1961, and 2 November 1977 causing 61, 40, and 38 deaths, respectively (Nicolaidou and Hadjichristou, 1995). These figures are higher than any other part of Greece. Also the number of lives lost due to floods in Athens are greater than those due to any other natural hazards. For example 18 deaths due to earthquakes were reported in the last century in the Attica area that surrounds Athens (Nicolaidou and Hadjichristou, 1995).

Recently, attention has been given to the construction of storm drainage networks and improvement of natural channels in Greater Athens. The main and most threatening river of Attica, Kifissos, with a catchment area of 417 km^2, is also being improved currently by broadening its channel. However, until now the adopted return period for the design of such constructions is traditionally 5-10 years for the storm drainage networks and does not exceed 20-50 years for the main streams including Kifissos. It is estimated that these values of return period do not provide a sufficient protection level and new, more severe, design criteria have to be established (Xanthopoulos et al., 1995) to lower the risk. In doing that, the first step is to estimate the rainfall amount for higher return periods or lower exceedance probabilities.

Rainfall intensity values for exceedance probabilities of the order of 10^{-2} to 10^{-4} as well as those corresponding to the probable maximum precipitation (PMP) must be known not only for the design of new constructions but also for performing simulations of extreme flood events to obtain a sight of the possible impacts of such events.

Numerous studies of maximum rainfall intensities in Athens have been performed. They used empirical or statistical techniques to construct intensity-frequency-duration (IDF) curves for return periods lower than 100 years. A review of such studies is given by Hydrauliki (1980). Typically, these studies used rainfall intensity records of several stations in Greater Athens, whose lengths varied between 26 and 72 years. However, a daily precipitation record can be constructed in Athens for a much longer period, i.e., 136 years, and clearly this can be utilised for a more reliable estimation of rainfall amount for low exceedance probabilities. This is the longest available rainfall record in Greece and its analysis may reveal useful properties of intense rainfall in Athens. From the complete record of daily precipitation measurements extending through 1860-1995 the series of annual maximum daily values can be extracted thus forming a record of 136 years.

Recently, Koutsoyiannis et al. (1998) have shown that records of daily and 2-day maximum rainfall depths can be combined with rain-recording data for lower durations to construct an IDF relationship of the general form

$$i(d, T) = \frac{a(T)}{b(d)} \tag{1}$$

where $i(d, T)$ is the rainfall intensity corresponding to duration d and return period T, and $a(T)$ and $b(d)$ functions of T and d, respectively. Particularly, they have shown that the function $a(T)$ depends on, and can be directly derived from, the distribution function of the rainfall intensity or rainfall depth using data of either recording or non-recording rain gauges. In addition, they concluded that the daily observations of non-recording devices must never be ignored in determining $a(T)$, even in the case of coexistence of recording devices at the same station because autographic devices with their vulnerable mechanisms are more sensible to erroneous recordings, whereas the standard non-recording rain gauges are more reliable due

to their simpler structure. Moreover, non-recording stations typically operate over periods longer than those of recording stations and, therefore, their records lead to a more reliable estimation of $a(T)$. On the other hand, the determination of $b(d)$ apparently needs rain-recording data of lower durations down to some minutes or an hour.

The purpose of this paper is the thorough investigation of the long series of maximum daily rainfall in Athens (described in section 2) in order to reveal interesting properties of the distribution function of the maximum daily rainfall (section 3). Our interest is focused on the upper tail of the distribution and, particularly, we examine the question whether the large record size alters this tail as compared with that obtained by a typical 30 to 40-year record (section 4). In addition, the probable maximum precipitation value is obtained by the typical statistical method, and then is compared with rainfall depth values of low-probability of exceedance obtained by a specific distribution function (section 5). The results of the analyses are combined with rainfall intensity data of lower durations from a nearby station to derive intensity-frequency-duration curves applicable to high return periods (section 6).

2. Brief presentation of the data

Precipitation observations at the city of Athens initiated in 1839; more systematic measurements started in 1958, but a continuous record, free of missing daily values, exists since 1860 (Katsoulis and Kambetzidis, 1989). Since 1890 the location of meteorological station has been fixed at the Nymfon Hill (close to Acropolis) by the National Observatory of Athens (NOA), whose Meteorological Institute became responsible for the observations. Before that year, the station had been located at different sites at distances less than 2 km from its current site at NOA; also, different organisations were responsible for the observations and their processing and publishing; interestingly, during 1884-1890 the observations were published in the Greek Government Paper. The altitude of the various station locations varied between 77.0-124.1 m while that of the current location is 107.1 m. Since 1894 the same type of instrument is used whereas earlier different types of rain gauges were used. This brief history of the station (whose details are given by Katsoulis and Kambetzidis, 1989) indicates that the observation record can be regarded as homogeneous

since 1890's but earlier inconsistencies may appear due to different locations and instruments. However, it is not anticipated that those differences affected the record seriously; for example Mariolopoulos (1938) concluded that the departures due to different instrument types in the collected rainfall amount does not exceed 2%, and Katsoulis and Kambetzidis (1989) using statistical tests concluded that the complete series of precipitation depths can be considered as homogeneous; similar are the results by Zerefos et al. (1977). In any case any suspicion of inhomogeneity does not give reasons for exclusion of the priceless early part of the record (prior to 1890's).

From the continuous record of daily precipitation measurements extending through 1860-1995 the annual maximum daily series was extracted. This work was relatively simple for the years 1936-1995 because the necessary information was included in the Annual Climatological Bulletins published by NOA. Unfortunately, bulletins had not been published before 1936 and so we had to search in the oldest files of the NOA or to contact previous researchers who had used the data for different objectives.

The 136-year annual series of maximum daily values is depicted in Figure 1, along with a smoothed series representing the 21-year moving averages. The most important sample statistics are summarised in Table 1. In Figure 1 we recognise a highly variable random pattern of the annual series (as expected), with the highest value of 150.8 mm being that of year 1899. Also, we perceive a weak falling trend following 1890, which is rather unexpected as we would normally expect a rising trend due to the island effect caused by the intensive urbanisation of the area (such a rising trend was detected in rainfall intensities of low durations in a nearby station by Deas, 1994). However, both the (non-parametric) Kendall's rank correlation test and the (parametric) regression test for linear trend (Kottegoda, 1989, p. 32) result that this falling trend is not statistically significant at a 5% significance level (for a two-tailed test).

As another attempt to detect nonstationarities within the time series, we divided it into four sub-series each corresponding to one quarter of the record length (34 years). Box plots for those sub-series as well as the complete series are shown in Figure 2; they are constructed with the standard rules described by Hirsch et al. (1993, p. 17.10). Some differences appear in

the box plots of the different sub-series (for example, the fourth one appears to have lower values than the others with only one outside value whereas the second has two outside and two far-outside values). However, the application of the (non-parametric) Kruskal-Wallis test (Hirsch et al., 1993, p. 17.25) resulted that the hypothesis that all four sub-series have identical distributions is not rejected at a 5% significance level. All these tests suggest that the complete series may be regarded as stationary, so that the typical statistical analysis of extreme events can be applied to the complete record.

3. Statistical analysis of the record

Three alternative distribution functions of maxima are used to model the annual maximum series of daily rainfall depths of the NOA station in Athens. Namely, these are the Type I (or Gumbel) distribution of maxima

$$F_X(x) = \exp(-e^{-x/\lambda + \psi}) \tag{2}$$

the Type II distribution of maxima

$$F_X(x) = \exp\left\{-\left(\kappa \frac{x}{\lambda}\right)^{-1/\kappa}\right\} \qquad x \geq 0 \tag{3}$$

and the generalised extreme value (GEV) distribution

$$F_X(x) = \exp\left\{-\left[1 + \kappa\left(\frac{x}{\lambda} - \psi\right)\right]^{-1/\kappa}\right\} \qquad x \geq \lambda\,(\psi - 1/\kappa) \tag{4}$$

In all above relationships X and x denote the random variable representing the annual maximum daily rainfall depth and its value, respectively, $F_X(x)$ is the distribution function and κ, λ, and ψ are shape, scale, and location parameters, respectively; κ and ψ are dimensionless whereas λ (> 0) has the same units as x which in our case are mm. Note that both (2) and (3) are two-parameter special cases of the three-parameter (4), resulting when $\kappa = 0$ and $\psi = 1/\kappa$, respectively.

All three distributions fitted by the method of L-moments are shown in Figure 3 on a Gumbel probability paper. The estimated values of parameters are shown in Table 2. The method of L-moments was preferred due to its robustness, i.e., because, unlike other methods, does not overemphasise an occasional extreme event, as it does not involve squaring or cubing of the data. A concise presentation of the method is given by Stedinger et al. (1993). The same text gives the L-moments estimators for the Type I and GEV distributions; the estimators of the Type II distribution are direct consequences of those of the GEV distribution. For the GEV distribution the methods of moments and maximum likelihood were used, as well, for comparison. The equations needed for the application of the method of moments are given by Stedinger et al. (1993, pp. 18.13-18.17). The method of maximum likelihood does not yield simple analytical equations. However, the likelihood function is easy to construct and then its maximisation can be performed using software tools for nonlinear optimisation. In our case we used the embedded Solver utility of the MS-EXCEL spreadsheet software writing no code at all. The parameters estimated by these methods are also shown in Table 2 whereas plots of the distribution functions are shown in Figure 5 (again on Gumbel probability paper).

The empirical distribution function estimated using the Gringorten plotting positions is also shown in Figure 3. Clearly, this figure shows that the Gumbel distribution departs significantly from the empirical distribution as the points corresponding to the empirical distribution do not form a straight line on the Gumbel probability paper. Another indication of the inappropriateness of the Gumbel distribution are the departures of the empirical coefficients of skewness and L-skewness (2.13 and 0.294, respectively; see Table 1) from the theoretical values (1.1396 and 0.1699 respectively). This can be verified by a statistical test. As shown by Hosking et al. (1985; see also Stedinger et al., 1993, p. 18.18) when data are drawn from a Gumbel distribution ($\kappa = 0$) the resultant L-moments estimator of κ has mean 0 and variance

$$\mathrm{Var}\,(\hat{\kappa}) = \frac{0.5633}{n} \tag{5}$$

This allows the construction of a powerful test whether $\kappa = 0$ (i.e. appropriateness of the Gumbel distribution; null hypothesis) or not (alternative hypothesis). Applied to the data of the present study this test results to rejection of the null hypothesis at an attained significance level (i.e., probability of type I error) as low as 0.2%.

Also, Figure 3 displays that the Type II distribution does not fit well the empirical distribution function in its lower tail (this would be more apparent in a Weibull probability plot or, equivalently, if we used logarithmic scale for the vertical axis in Figure 3). By observing that if X has Type II distribution, then the transformed variable $Y = \ln X$ has Type I distribution, we can apply the same test as above to statistically verify whether the Type II distribution is appropriate or not. The parameters of the GEV distribution fitted to the logarithms of our data are shown in Table 2 (last row). The statistical test results that the estimated value of κ differs significantly from zero at an attained level of 1.7% and therefore the Type II distribution should be rejected as well.

On the contrary, in Figure 3 the GEV distribution fits well the empirical distribution. Also the values of L-skewness and L-kurtosis (Table 1) of the available data, if plotted to a generalised diagram of L-kurtosis versus L-skewness (Stedinger et al., 1993, p. 18.9) show that they are consistent with the GEV distribution; moreover, this diagram shows that the GEV distribution is more suitable than other typical distributions such as Gumbel, Pearson II, Log-Normal and Pareto. To test the goodness of fit of the GEV distribution with its three parameters estimated by the method of L-moments we used the χ^2 test. Other tests such as the Kolmogorov-Smirnov and the probability plot correlation test which are studied for the GEV distribution when the location and shape parameters are known (Chowdhury et al., 1991) are not applicable in our case where all three parameters are estimated from the data. The χ^2 test was applied several times with a number of classes varying from 5 to 20 and in all cases the null hypothesis (that the GEV distribution is consistent with the data) was not rejected at the typical 5%-10% (even at a non-typical 40%-50%) significance level.

In conclusion, the above analyses provide evidence that the GEV distribution is a consistent probabilistic model for the annual maximum series of the daily rainfall depth at the NOA station at Athens whereas the Type I and Type II models are inconsistent with the data.

4. Effect of the sample size

In Greece a record length of 136 years is an exception to the rule that sample sizes typically vary between 10 and 50 years. This is because most hydrometeorological stations have been installed after World War II and many of them did not operate continuously. The question arises whether a sample of this typical small size can lead to conclusions similar with those drawn in the previous section, or it delineates a different picture of the distribution function of maximum rainfall.

To study this question, we examined the four sub-series already presented in section 2. Figure 4 depicts a plot of the empirical distribution of the fourth sub-series corresponding to one quarter of the record length (last 34 years) in Gumbel probability paper. We observe that the points corresponding to the empirical distribution form a straight line, which implies the appropriateness of the Gumbel distribution for this sub-series. This Gumbel distribution fitted by the method of L-moments for the fourth sub-series is also shown in Figure 4. For comparison we have plotted on the same figure the empirical, the Gumbel, and the GEV distribution function of the complete series of annual maximum rainfall depths (136 years). We notice the large departure, particularly in the upper tails, of the Gumbel distribution of the 34-year sample from those of both the GEV and the Gumbel distributions of the 136-year sample. Thus, the picture acquired from the sub-series of the last 34 years is deforming: the inappropriate Gumbel distribution appears as appropriate and also shifted towards lower values of rainfall amounts in the upper tail.

Similar are the results for other two sub-series. In summary, the Gumbel distribution tested by the κ-test described in section 3 is not rejected for the three out of four sub-series. Only the second sub-series with the four outside values (out of which two far outside values; see Figure 2) results in a high value of the shape parameter κ, and consequently, a statistically significant departure from the Gumbel distribution. By performing the same analysis for the type II distribution we found that this is not rejected again for the three out of four sub-series, the exception being the third sub-series.

The findings of this investigation have significant meaning from an engineering viewpoint. Typically, engineers start a statistical study regarding storms and floods by

plotting the empirical probabilities on a Gumbel paper and, given that its arrangement is not far from a straight line, they proceed adopting the Gumbel distribution and perform extrapolations for low probabilities of exceedance. As demonstrated above (see also Figure 5), this procedure may underestimate seriously the rainfall amount corresponding to low probabilities of exceedance.

5. Estimation of probable maximum precipitation

The discussion in section 4 showed that a subset of the available 136-year data series may bias seriously our knowledge of the distribution function. Besides, the distribution function obtained by the complete 136-year series, apparently, is not the true population distribution. Thus, the extrapolation of the GEV distribution obtained in section 3 to probabilities of exceedance such as 1/1000, 1/10 000, or even less may lead to inaccurate results. This is a consequence of the so called "Myth of the Tails" (Wileeke, 1980), which reads "Statistical distributions applied to hydrometeorological events that fit through the range of observed data are applicable in the tails" and reminds us that the tails of distributions are highly uncertain (see also Dooge, 1986).

Therefore, as another quantification of an extremely high rainfall magnitude we estimated the daily probable maximum precipitation (PMP) in Athens, based on the available annual series of maximum rainfall depths. More specifically, we used the so called statistical estimation of PMP as developed by Hershfield (1961, 1965) and standardised by the World Meteorological Organization (1986). The method estimates the rainfall depth h_m of the probable maximum precipitation of duration d by the formula

$$h_\mathrm{m} = \overline{h} + k_\mathrm{m}\, s_H \tag{6}$$

where $\overline{h}$ and s_H are the sample mean and standard deviation, respectively, of maximum rainfall depth of duration d, and k_m is a frequency factor given by an empirical nomograph as a function of d and $\overline{h}$. This nomograph can be approximated by the simple analytical equation (Koutsoyiannis and Xanthopoulos, 1997, p.160)

$$k_m = 20 - 8.6 \ln\left(\frac{\bar{h}}{130} + 1\right)\left(\frac{24}{d}\right)^{0.4} \quad (7)$$

The method incorporates some adjustments of mean and standard deviation for sample size and maximum observed event (World Meteorological Organization, 1986, pp. 97-107).

The application of the method to the data of this study (Table 1) with all adjustment factors equal to one, due to the large sample size, results in $k_m = 17.20$ and a PMP value 424.1 mm for daily rainfall. This PMP value when viewed from a probabilistic approach can be assigned a specific return period depending on the particular distribution function. In Table 3 we present the values of return period corresponding to rainfall depth of 424.1 mm for the different distribution functions examined in this study. We observe that the results exhibit a huge variability as the values vary from 4.2 thousand years for the Type II distribution up to 64 billion years for the Gumbel distribution (see also Figure 5). To have an empirical idea about what the true value of the return period might be we recall that the method of Hershfield (1961) was based on the analysis of $m = 95\,000$ station-years of data. That means that the empirical probability corresponding to k_m and consequently h_m is of the order of $p = 1 / m \approx 10^{-5}$ and the return period is of the order of 10^5 years. Indeed, this is the order of magnitude of the return period when this is estimated by the GEV distribution, the estimate with the closest agreement being that resulted by the method of maximum likelihood. An upper confidence limit for p for a confidence coefficient 99% is $p' = p + 2.58\,[p\,(1-p) / m]^{0.5} = 3.7 \times 10^{-5}$ which corresponds to a return period of 27 000 years. This suggests that the estimate of the method of L-moments (28 000 years) is also consistent with the results of the Hershfield's PMP method, when the latter is considered from a probabilistic point of view.

This investigation provides another empirical indication that the GEV distribution performs well in the case study examined whereas the Type I and Type II distributions do not.

6. Estimation of intensity-frequency-duration curves

As already brought up in section 1, the results of the above analyses can be utilised to derive intensity-frequency-duration curves applicable for high return periods. More specifically, these results can be used to determine $a(T)$ in equation (1). At this time we do not have available maximum rainfall data for durations shorter than daily so we have adopted the function $b(d)$ of a nearby rain-recording station, namely Helliniko, that is (Koutsoyiannis et al., 1998)

$$b(d) = (d + 0.189)^{0.796} \quad (d \text{ in h}) \tag{8}$$

We note that this function was derived by making no hypothesis about the distribution function of rainfall depth or intensity using a non-parametric statistical technique (Koutsoyiannis et al., 1998) and the numerical coefficients (0.189 and 0.796) were found to be reasonably constant over a wider geographical area (Koutsoyiannis et al., 1998; Kozonis, 1995).

Solving (4) for $x \equiv h(d, T)$, and replacing F_X by $1 - 1 / T$, we get

$$h(24, T) = \lambda \left\{ \frac{\left[-\ln\left(1 - \frac{1}{T}\right) \right]^{-\kappa} - 1}{\kappa} + \psi \right\} \tag{9}$$

Adopting (for safety) the parameter values of the method of L-moments, which result in higher values of rainfall depth in the upper tail of the distribution (Figure 5) and replacing them in (9) we get

$$h(24, T) = 68.32 \left\{ \left[-\ln\left(1 - \frac{1}{T}\right) \right]^{-0.185} - 0.45 \right\} \tag{10}$$

Combining (1), (8), and (10) and solving for $i(d, T) := h(d, T) / d$ we obtain the idf relationship

$$i(d, T) = \frac{35.95 \left\{ \left[-\ln\left(1 - \frac{1}{T}\right) \right]^{-0.185} - 0.45 \right\}}{(d + 0.189)^{0.796}} \tag{11}$$

For large return periods, e.g., $T \geq 50$ we can write $\ln [1 - (1/T)] = -(1/T) - (1/T)^2 - \cdots \approx -(1/T)$ which simplifies (11). Furthermore, (11) needs an adjustment to account for the fact that the daily rainfall depth is a fixed-interval rainfall amount. Using the adjusting factor of the bibliography (e.g., Linsley et al., 1975, p. 357), which is 1.13, and making the simplification we finally result in

$$i(d, T) = \frac{40.6\,(T^{0.185} - 0.45)}{(d + 0.189)^{0.796}} \quad (i \text{ in mm/h}, t \text{ in h}) \tag{12}$$

The simplified and adjusted equation (12) is also valid for small return periods if we replace the return period T for the annual series with the return period T' for the series over threshold (or partial duration series). Indeed, in that case we have (see, e.g., Raudkivi, 1979, p. 411)

$$-\ln (1 - 1 / T) = 1/T' \tag{13}$$

so that (11) again results in (12) with T' substituted for T.

In the past several idf curves were constructed for Athens using empirical techniques. Among them the most widespread is that by Dallas (1968)

$$i(d, T) = 12.8\, T^{1/3} / d^{2/3} \quad (i \text{ in mm/h}, t \text{ in h}) \tag{14}$$

and a modification due to Memos (1980) which reduces the very high intensities for low durations

$$i = \begin{cases} 12.8\, T^{1/3} / d^{2/3} & i \leq 30 \text{ mm/h} \\ (8.13\, T^{1/2} / d) \exp(0.5 - 0.016\, i) & i > 30 \text{ mm/h} \end{cases} \quad (i \text{ in mm/h}, t \text{ in h}) \tag{15}$$

A plot of idf relationships (12) is given in Figure 6 along with a comparison with the empirical relationships (14) and (15). The large departures among the different sets of curves are apparent in this figure, especially for large return periods T = 1000 - 10 000 years. These departures are explained by the smaller sample sizes and the empirical techniques used to derive the relationships (14) and (15). Due to the longer record used in the present study and

the more thorough study and the refined and consistent methodology, it is suggested that the relationship (12) is more reliable than those developed in the past.

7. Conclusions

The analysis of the longest record of annual maximum daily rainfall in Greece, i.e. the 136-year series of the NOA station in Athens, reveals remarkable statistical properties of the intense rainfall in Athens. This analysis is very useful both to predict intense rainfall in Athens, where currently major flood protection works are under way, and to infer some properties of rainfall in other parts of the country where existing records are much shorter. Specifically, the long record allows the comparison of several types of distribution functions such as the Type I (Gumbel) and Type II distributions of maxima and the General Extreme Value (GEV) distribution and the appropriate choice among them. All these distributions might be chosen as appropriate models if less years of measurements were available (i.e., part of this sample were used). Nevertheless, the extrapolation for return periods such as 1000 years or more result in huge differences among the different distributions; for example Type II distribution results in more than twice rainfall in comparison with Type I. The available long record results in rejection of both Type I and II distributions. On the contrary, the General Extreme Value (GEV) distribution appears to be suitable for the examined series and its predictions for large return periods agree with the estimate of probable maximum precipitation obtained by the statistical (Hershfield's) method, when the latter is considered from a probabilistic point of view. Although a 136-year record is still too short to accurately determine the upper tail of the distribution function of maximum rainfall, it is quite longer than typical samples available for hydrologic applications such as the construction of intensity-duration-frequency relationships. Thus, this record provides a clearer view of such relationships for large return periods, and based on a newly developed methodology, the fitted GEV distribution is directly utilised for establishing their mathematical expressions.

Acknowledgements

The authors are grateful to the National Observatory of Athens and the Academy of Athens - Research Centre of Atmospheric Physics and Climatology for their kind permission of access to some of the oldest files and Annual Climatological Bulletins. They also thank A. Paliatsos of the Technological Educational Institute of Pireas, Greece, for providing useful information for the years 1890-1930. MS-EXCEL is a registered trademark of the Microsoft Corporation.

References

Chowdhury, J. U., J. R. Stedinger, and L.-H. Lu, Goodness-of-fit tests for regional generalized extreme value flood distributions, *Water Resour. Res.*, 27(7), 1765-1776, 1991.

Dallas, S., Study of the rehabilitation of the Kifissos River (in Greek), Sewage Organisation of Athens, 1968.

Deas, N., Geographical distribution on intense rainfall in the Sterea Hellas area, Diploma thesis (in Greek), National Technical University, Athens, 1994.

Dooge, J. C. I., Looking for hydrologic laws, *Water Resour. Res.*, 22(9) pp. 46S-58S, 1986.

Hershfield, D. M., Estimating the probable maximum precipitation, *Proc. ASCE, J. Hydraul. Div.*, 87(HY5), 99-106, 1961

Hershfield, D. M., Method for estimating probable maximum precipitation, *J. American Waterworks Association*, 57, 965-972, 1965.

Hirsch, R. M., D. R. Helsel, T. A. Cohn, and E. J. Gilroy, Statistical analysis of hydrologic data, in *Handbook of Hydrology*, edited by D. R. Maidment, McGraw-Hill, 1993.

Hosking, J. R. M., J. R. Wallis and E. F. Wood, Estimation of the generalized extreme value distribution by the method of probability weighted moments, *Technometrics*, 27(3), 251-261, 1985.

Hydrauliki, Study of the sanitary and storm sewer networks of Acharnes, Ano Liosia, Kamatero and Zefiri, Engineering Report, 204 pp. (in Greek), Sewage Organisation of Athens, 1980.

Katsoulis, B. D., and H. D. Kambetzidis, Analysis of the long-term precipitation series at Athens, Greece, *Climatic change*, 14, 263-290, 1989.

Kottegoda, N. T., *Stochastic Water Resources Technology*, Macmillan Press, London, 1980.

Koukis, G. C., and D. Koutsoyiannis, Greece, in *Geomorphological Hazards of Europe*, edited by C. and C. Embleton, Elsevier, pp. 215-241, Amsterdam, 1997.

Koutsoyiannis, D., D. Kozonis and A. Manetas, A comprehensive study of rainfall intensity-duration-frequency relationship, *J. of Hydrol.*, 1998 (accepted for publication with minor revisions).

Koutsoyiannis, D., and Th. Xanthopoulos, *Engineering Hydrology*, 412 pp. (in Greek), National Technical University, Athens, 1997.

Kozonis, D., Construction of idf curves from incomplete data sets, Application to the Sterea Hellas area, Diploma thesis (in Greek), National Technical University, Athens, 1995.

Linsley, R. K. Jr., M. A. Kohler and J. L. H. Paulus, *Hydrology for Engineers*, McGraw-Hill, Tokyo, 2nd edition, 1975.

Mariolopoulos, H., *The Climate of Greece*, National Observatory of Athens, 280 pp. (in Greek), 1938.

Memos, C., Intensity-duration-frequency curves for large intensities, *Proc. 2nd Greek seminar on hydrology*, 57-65 (in Greek), Ministry of Coordination, Athens, 1980.

Nicolaidou, M. and E. Hadjichristou, Recording and assessment of flood damages in Greece and Cyprus, Diploma thesis (in Greek), National Technical University of Athens,1995.

Xanthopoulos, Th., D. Christoulas, M. Mimikou, D. Koutsoyiannis and M. Aftias, A strategy for the problem of floods in Athens, *Proc. of the Workshop for the Flood Protection of Athens*, Technical Chamber of Greece, Athens, 1995.

Raudkivi, A. J., *Hydrology, An Advanced Introduction to Hydrological Processes and Modelling*, Pergamon Press, Oxford, 1979.

Stedinger, J. R., R. M. Vogel, and E. Foufoula-Georgiou, Frequency analysis of extreme events, Chapter 18 in *Handbook of Hydrology*, edited by D. R. Maidment, McGraw-Hill, 1993.

Willeke, G. E., Myths and uses of hydrometeorology in forecasting, in *Proceedings of March 1979 Engineering Foundation Conference on Improved Hydrological Forecasting – Why and How*, pp. 117-124, American Society of Civil Engineers, New York, 1980.

World Meteorological Organization (WMO), *Manual for Estimation of Probable Maximum Precipitation*, Operational Hydrology Report 1, 2nd edition, Publication 332, World Meteorological Organization, Geneva, 1986.

Zerefos, C. S, G. B. Kosmas, C. C. Repapis, and J. D. Zambakas, Time series analysis of rain at Athens National Observatory during the century 1871-1970, Report (in Greek), National Observatory of Athens, 39 pp., 1977.

List of Tables

Table 1 Statistics of the sample of annual maximum daily rainfall depths (in mm) of the NOA station at Athens.

Table 2 Parameter values of various distribution functions fitted to the sample of annual maximum daily rainfall depths (in mm) of the NOA station at Athens.

Table 3 Return period of the PMP value as estimated using the various distribution functions fitted to the sample of annual maximum daily rainfall depths of the NOA station at Athens.

Table 1 Statistics of the sample of annual maximum daily rainfall depths (in mm) of the NOA station at Athens.

Statistic	Value
Sample size	136
Maximum value	150.8
Minimum value	17.2
Mean	47.9
Median	42.5
Standard deviation	21.7
Interquartile range	19.7
Coefficient of variation	0.454
Coefficient of skewness	2.13
Coefficient of kurtosis	6.30
L-coefficient of variation	0.224
L-coefficient of skewness	0.294
L-coefficient of kurtosis	0.242

Table 2 Parameter values of various distribution functions fitted to the sample of annual maximum daily rainfall depths (in mm) of the NOA station at Athens.

Distribution of variable	Fitting method	Transformation of variable	Parameter values		
			κ	λ	ψ
Type I	Moments	None	(0)	16.89	2.26
Type I	L-moments	None	(0)	15.48	2.51
Type II	L-moments	None	0.292	10.87	(3.43)
GEV	Moments	None	0.118	14.07	2.69
GEV	L-moments	None	0.185	12.64	2.99
GEV	Max. likelihood	None	0.161	12.93	2.94
GEV	L-moments	Logarithmic	–0.136	0.345	10.51

Table 3 Return period of the PMP value as estimated using the various distribution functions fitted to the sample of annual maximum daily rainfall depths of the NOA station at Athens.

Distribution of variable	Type I	Type I	Type II	GEV	GEV	GEV
Fitting Method	Moments	L-moments	L-moments	Moments	L-moments	Max. likelihood
Return period of PMP	8.4×10^9	64×10^9	4 200	210 000	28 000	56 000

List of Figures

Figure 1 Time series of the annual maximum daily rainfall depth at the NOA station.

Figure 2 Box plots of the complete series of annual maximum rainfall depths and the four sub-series each corresponding to one quarter of the record length.

Figure 3 Type I (Gumbel), Type II, and GEV distributions fitted to the sample of annual maximum daily rainfall depths of the NOA station at Athens by the method of L-moments, in comparison with the empirical distribution estimated using Gringorten plotting positions.

Figure 4 Comparison of the distribution function of the complete series of annual maximum rainfall depths (136 years) and the fourth sub-series corresponding to one quarter of the record length (last 34 years).

Figure 5 Plots of several distribution functions in the area of low probabilities of exceedance, and comparison with the estimated PMP value for maximum daily rainfall depths of the NOA station at Athens.

Figure 6 Comparison of intensity-frequency duration curves for Athens. Solid curves represent the idf curves obtained in the present study (eqn. (12)); dashed straight lines represent empirical idf curves given by eqn. (14); dotted curves represent the modification of empirical idf curves given by eqn. (15).

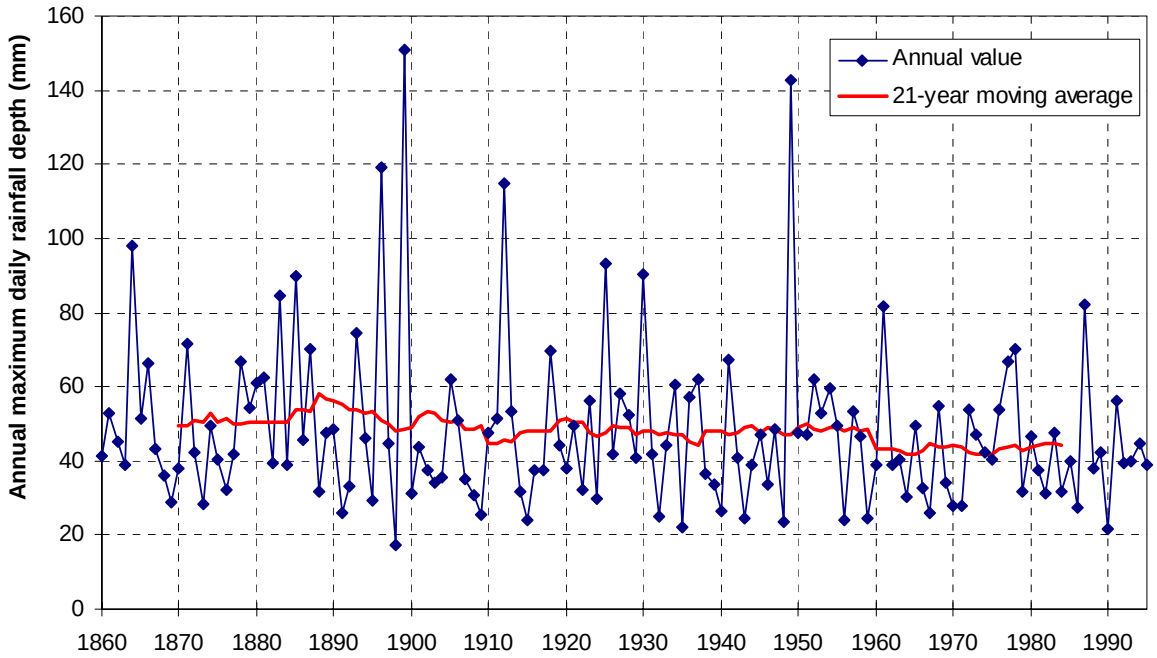

Figure 1 Time series of the annual maximum daily rainfall depth at the NOA station.

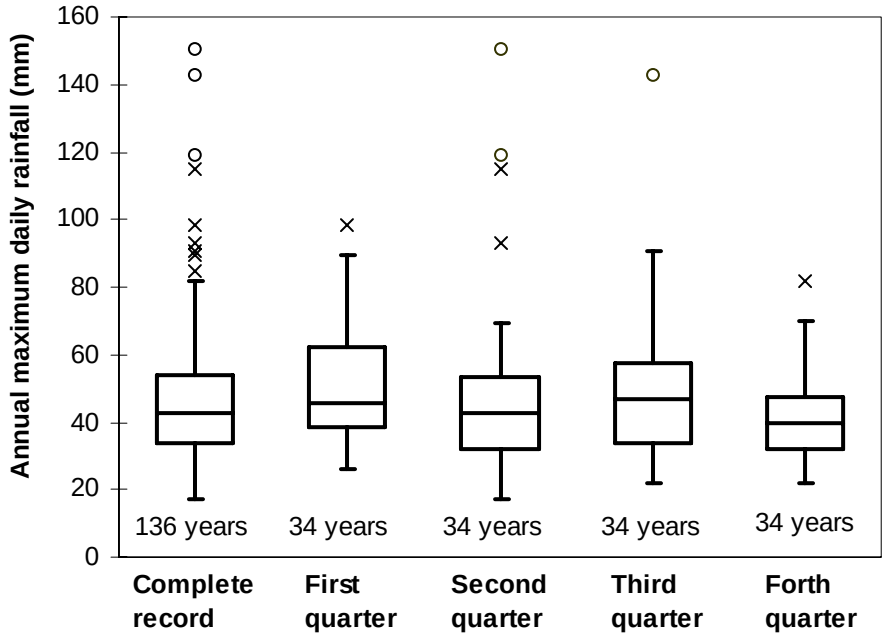

Figure 2 Box plots of the complete series of annual maximum rainfall depths and the four sub-series each corresponding to one quarter of the record length.

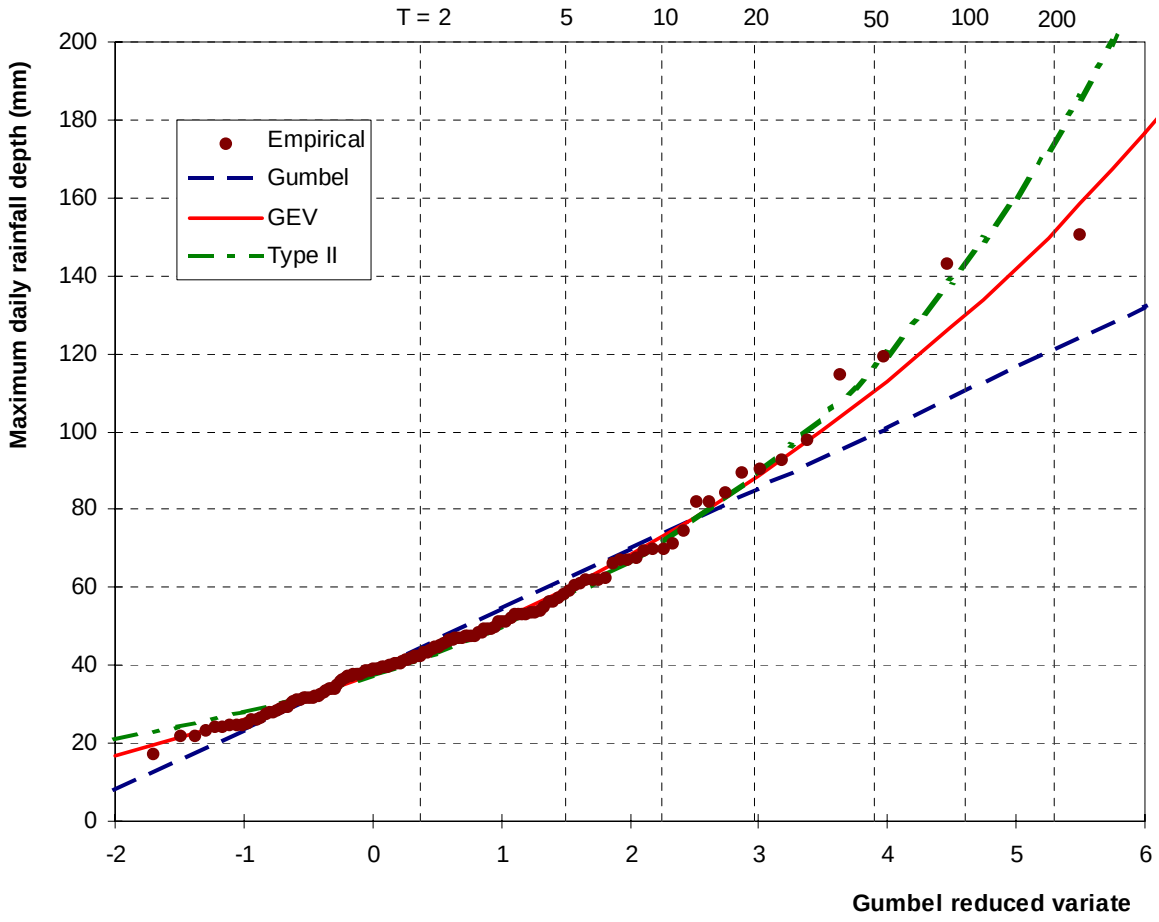

Figure 3 Type I (Gumbel), Type II, and GEV distributions fitted to the sample of annual maximum daily rainfall depths of the NOA station at Athens by the method of L-moments, in comparison with the empirical distribution estimated using Gringorten plotting positions.

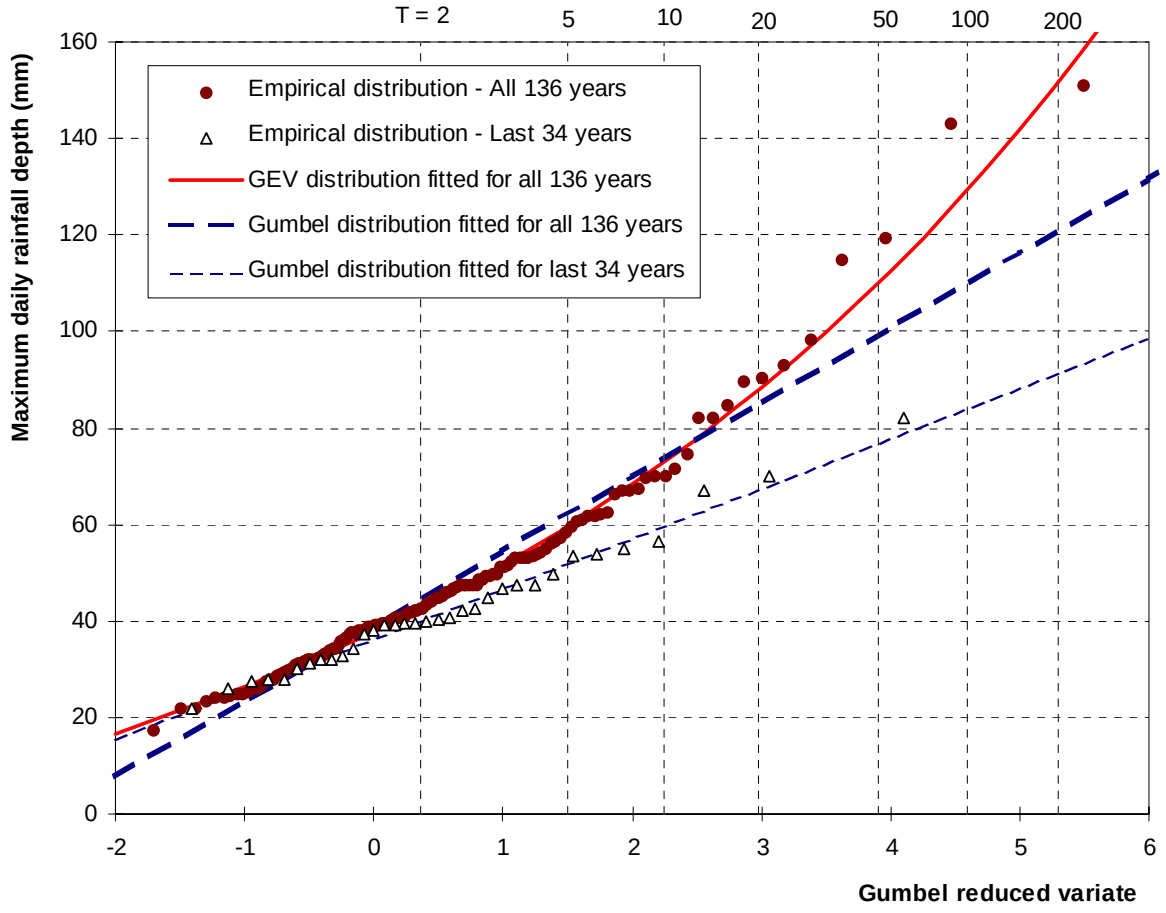

Figure 4 Comparison of the distribution function of the complete series of annual maximum rainfall depths (136 years) and the fourth sub-series corresponding to one quarter of the record length (last 34 years).

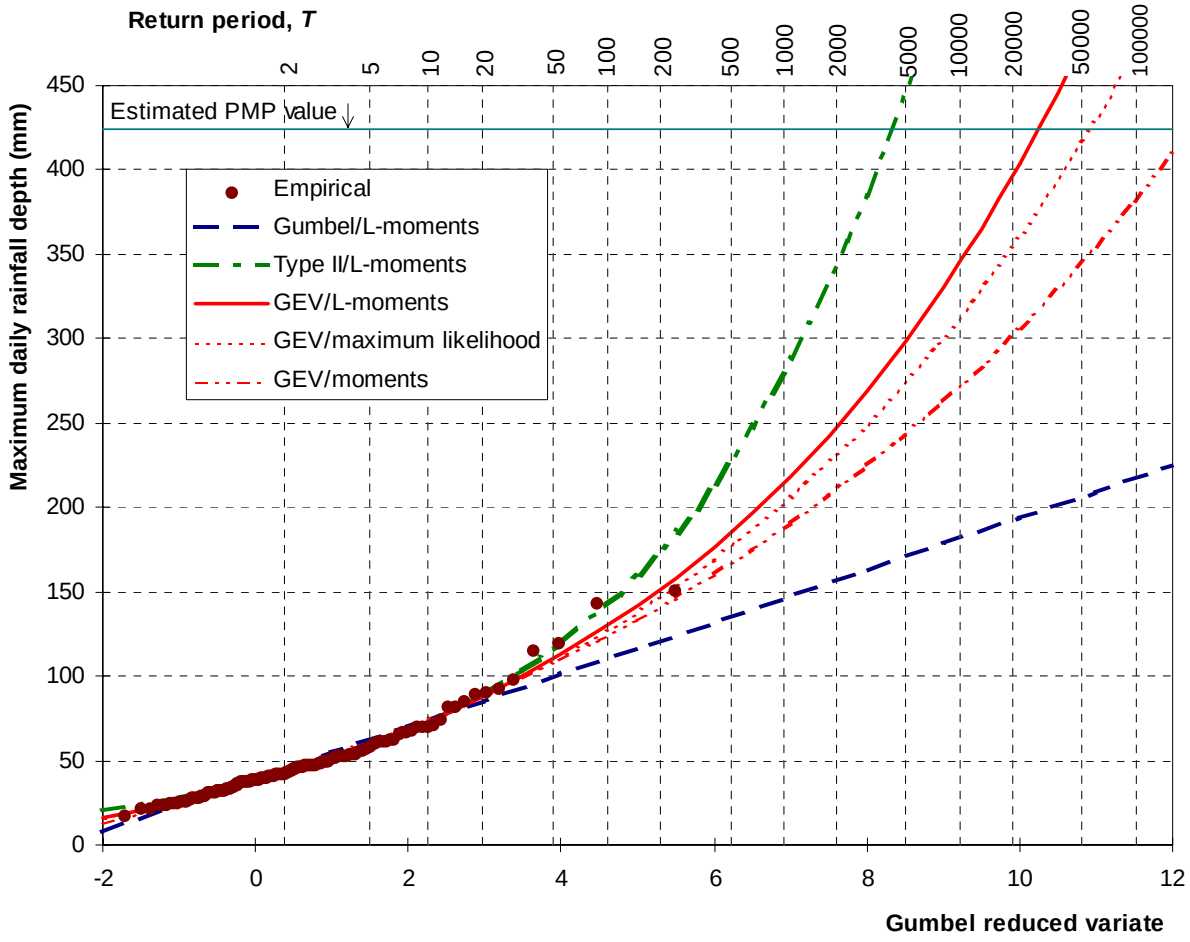

Figure 5 Plots of several distribution functions in the area of low probabilities of exceedance, and comparison with the estimated PMP value for maximum daily rainfall depths of the NOA station at Athens.

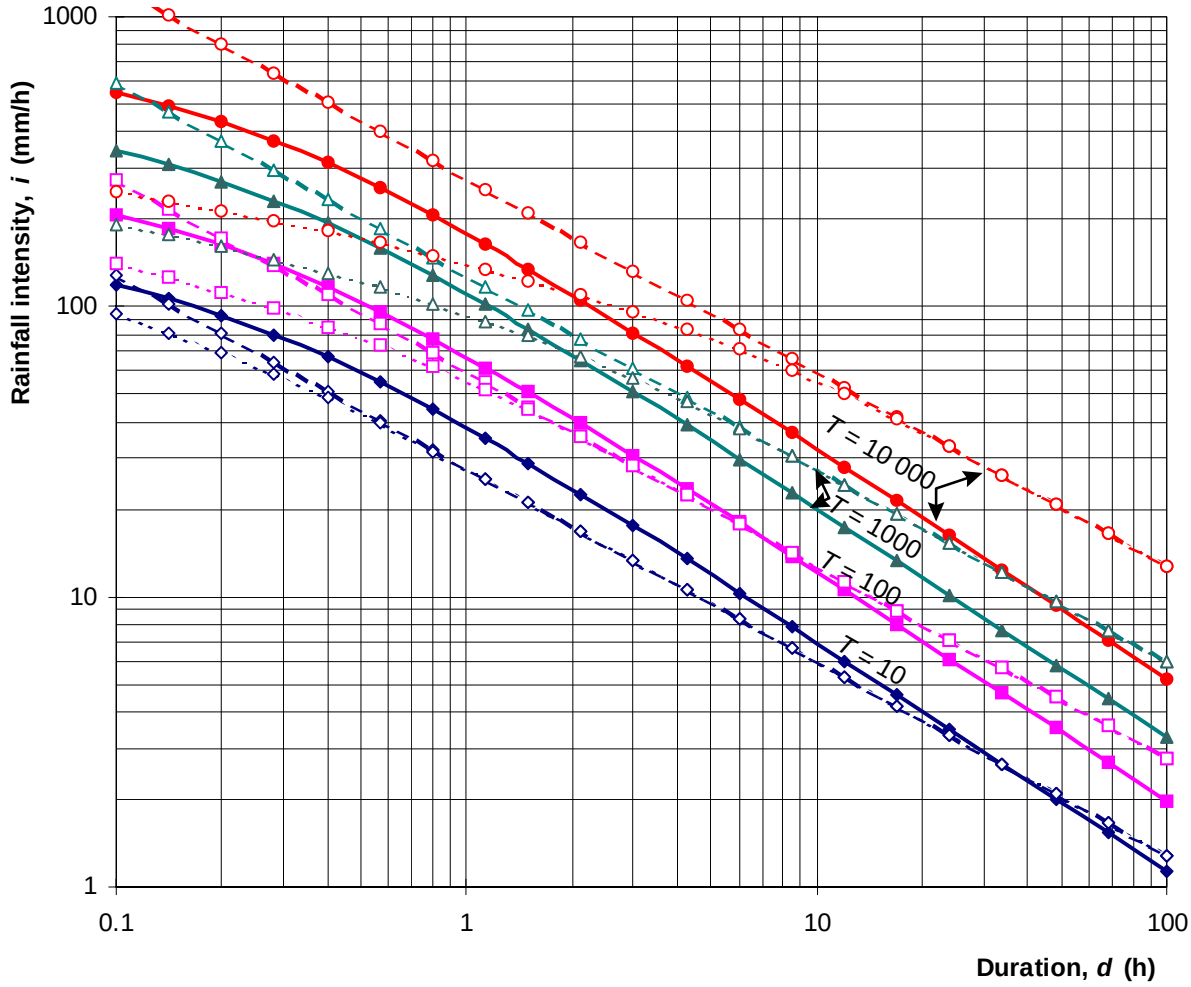

Figure 6 Comparison of intensity-frequency duration curves for Athens. Solid curves represent the idf curves obtained in the present study (eqn. (12)); dashed straight lines represent empirical idf curves given by eqn. (14); dotted curves represent the modification of empirical idf curves given by eqn. (15).