

# Potential distribution of extreme rainfall in the Basque Country

R. Moncho<sup>1,2</sup> and V. Caselles<sup>2</sup>

<sup>1</sup>AZTI-Tecnia, Unit of Marine Research, Txatxarramendi ugarteia z/g, 48395 Sukarrieta

<sup>2</sup>Department of Earth Physics and Thermodynamics, Universitat de València, C/ Doctor Moliner, 50, 46100, Burjassot

Received: 20-I-2010 – Accepted: 10-IX-2010 – **Translated version**

Correspondence to: robert@temps.cat

## Abstract

*The Potential Distribution model can describe at once the probability distribution and temporal distribution of rainfall. It also allows the incorporation of the dependence of the probability with the number of independent stations. This paper analyzes the probability distribution and temporal distribution of extreme rainfall in the Autonomous Community of the Basque Country using daily data from 43 stations, with an average of 31 active stations between 1961 and 2000. It was found that this model provided a good adjustment (NMAE = 1.4%) for rainfall over 115 mm in 24 h, and is consistent with other analyses of extremes. Finally, the article proposes a mathematical relationship to estimate the maximum expected rainfall for a return period equal to or more than 10 years, with a duration longer than 1 minute, and for a given set of independent stations. This relationship depends on two exponents, one for the return period ( $m = 0.23 \pm 0.02$ ) and another for the duration ( $n = 0.63 \pm 0.06$ ). In addition, it also depends on a scale factor, which takes values between  $P_o = 42 \pm 2$  mm to the south of the region and  $P_o = 71 \pm 5$  mm to the north, with an average value for the Basque Country equal to  $58 \pm 2$  mm.*

**Key words:** extreme rainfall, potential distribution, extreme climatology

## 1 Introduction

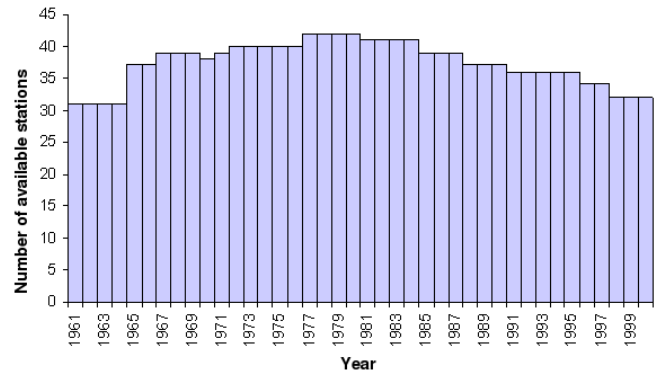
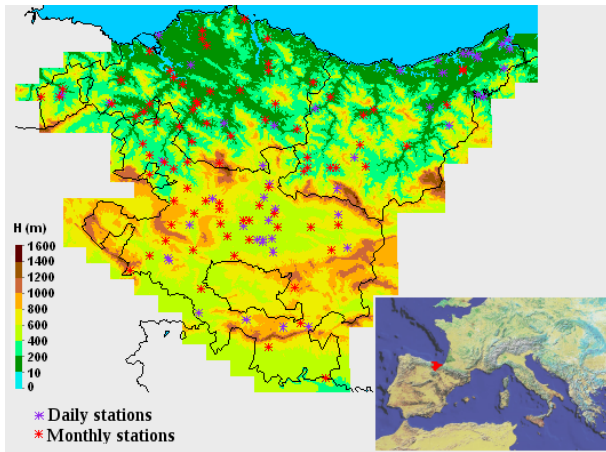
One of the variables that have the greatest impact on human resources is heavy rainfall. Flood risk depends more on the vulnerability of the infrastructure and especially on excessive exposure to the phenomenon, usually due to poor land use planning with respect to “flood zones” (Marco, 1999). On the other hand, when speaking of “climate risk” there is reference to the probability of occurrence of a climatic extreme, such as the frequency of heavy rain episodes, as well as of prolonged dry episodes.

The area of greatest risk of torrential rain in Spain is the Mediterranean coast (Font-Tullot, 2000; Llasat et al., 1996); but several episodes of significant flooding have also been recorded in the Basque Country, one of them out of the ordinary, such as the case of the floods in Bilbao in 1983 (Ugarte and González, 1984). Due to the large natural variability of precipitation, it is necessary to have long series of records to estimate both the recurrence of extreme rainfall and the pos-

sible climatic trend (Kundzewicz and Robson, 2000; Lana et al., 2009).

Normally the frequency of occurrence of a heavy rainfall episode is designated by the term “return period”, which refers to the average wait time needed to pass above a certain threshold at a given point (Témez, 1978). Therefore, this precise value does not adequately represent the episodes of a large area, which are responsible for major flooding. In fact, the repetition for “any point of an area” will always be higher than the repetition for a single point. In addition, due to the lack of data inherent to situations that are out of the ordinary that can be used as references, or due to the nature of extreme precipitation, probabilistic models often cannot successfully reproduce the associated return periods (Ministerio de Fomento, 1999; de Salas and Fernández, 2006; Soro et al., 2010). Therefore, this paper attempts to focus the statistics in the curve of extreme rainfall using a solid methodology valid for large areas of the Basque Country.





**Figure 1.** Stations of the Spanish Meteorological Agency (AEMET) available for the study. (a) (left) Geographical distribution, where  $H$  represents the altitude in meters regarding sea level. (b) (right) Time availability for the period 1961-2000.

## 2 Material and methodology

### 2.1 Rainfall data

For analysis of heavy rainfall in the Basque Country, there is daily data available from 43 meteorological stations of the network of the Spanish Meteorological Agency (Figure 1a). However, the availability of these stations depends on time. There is a peak period registered between 1977 and 1981 (Figure 1b).

Many stations have no continuity even between the starting year and end of activity year, but they can be activated and deactivated on an irregular basis, showing interruptions of months or years (Ruíz-Urrestarazu, 1983). In fact, of the 43 stations, there are only 498 092 daily records, which represent 79% of the maximum availability for the 1961-2000 series and only 19 stations fully complete it. The maximum number of active stations in one day was 36 and the minimum 19, while the average is 31.

To complete the study with the temporal distribution, there is hourly detailed information available of 22 of the 30 episodes (1950-2008) in which floods of diverse impact were registered in any given area of the Basque Country, according to various documents of the Spanish Meteorological Agency (AEMET).

### 2.2 Distribution of daily rainfall probability

Daily precipitation data do not follow a normal distribution due to the presence of a large number of zero values and no values under zero. Therefore, it is necessary to consider other probability distributions such as Gamma Distribution, Generalized Pareto Distribution or the SQRT-ET<sub>max</sub>, among others (de Salas and Fernández, 2006; Ministerio de Fomento, 1999). The following software has been used for the statistical treatment of data: IDL, R, STATGRAPHICS and EASYFIT.

The highest rainfall for each day was considered from all the available stations. The frequency of each value of rainfall,  $P$ , was calculated from these records, as well as the cumulative probability,  $\pi(p \geq P)$ , defined as the probability that a group of  $N$  stations records a maximum daily rainfall equal to or higher than  $P$ . Therefore, the empirical return period  $\rho(P)$  (in years) associated with this cumulative probability  $\pi(p \geq P)$  is given by the following relationship:

$$\rho(P) \equiv \frac{1}{\Pi(p \geq P)} \approx \frac{1}{365.25 \cdot \pi(p \geq P)} \quad (1)$$

where  $\Pi(p \geq P)$  is the probability that a maximum rainfall equal to or higher than  $P$  with a certain duration and in a certain group of stations, happens in a year.

On the other hand, according to Moncho et al. (2009), the occasional maximum accumulation  $P_{max}$  of heavy rainfall, with a duration  $t$  and a timely return period  $\rho$  between 1 and 50 years, in the Iberian Peninsula can be approximated to the shape of a curve MAI-IDF (Maximum Average Intensity - Intensity-Duration-Frequency) as:

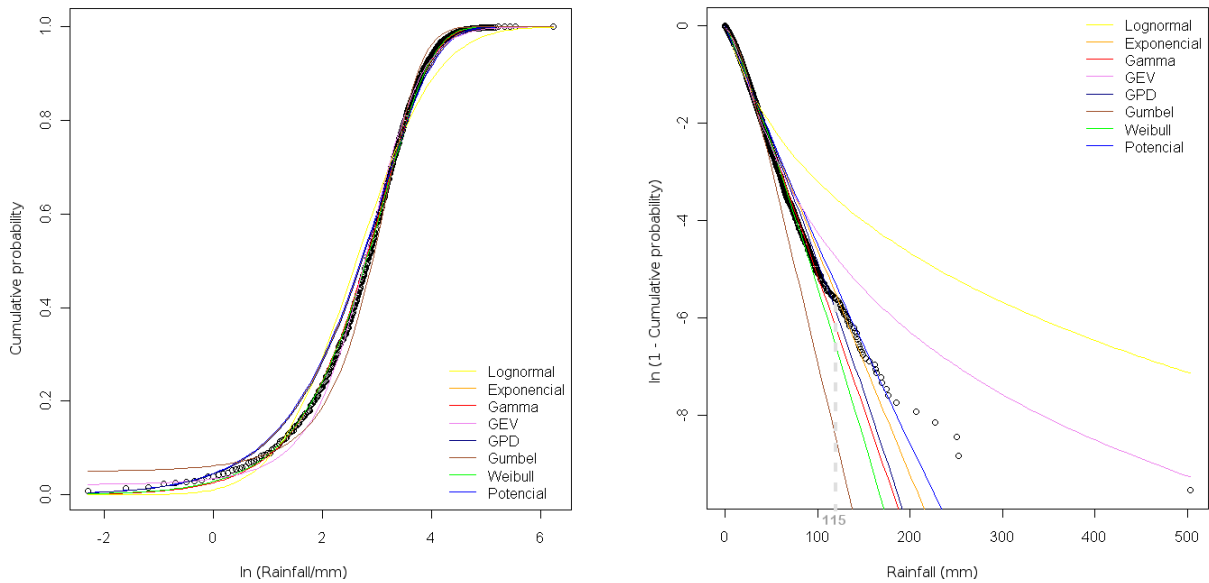
$$P_{max}(\rho, t) \approx P_o(\rho_o, t_o) \left( \frac{\rho}{\rho_o} \right)^m \left( \frac{t}{t_o} \right)^{1-n} \quad (2)$$

where  $P_o(\rho_o, t_o)$  is the reference accumulation for a reference period  $\rho_o$ , and a duration of  $t_o$  (which was chosen as a day),  $m$  is a dimensionless parameter that for the main stations of the Iberian Peninsula has been estimated as  $0.24 \pm 0.06$  (Moncho et al., 2009), and where  $n$  is an index of precipitation (Appendix A).

Given Equation 1, it can be proven that Equation 2 implies a Potential Probability Distribution or Pareto distribution (Annex B), according to the following equation:

$$\Pi(p \geq P) = \left( \frac{P_1}{P} \right)^{\frac{1}{m}} \quad (3)$$

where  $P_1$  and  $m$  are adjustable constants, while  $P$  is the considered precipitation threshold, which must be larger than



**Figure 2.** (a) (left) Adjustment of different distribution models to the accumulated probability observed in the maximum daily rainfall. (b) (right) Detail of the logarithm of the cumulative probability in the extreme part of the curve.

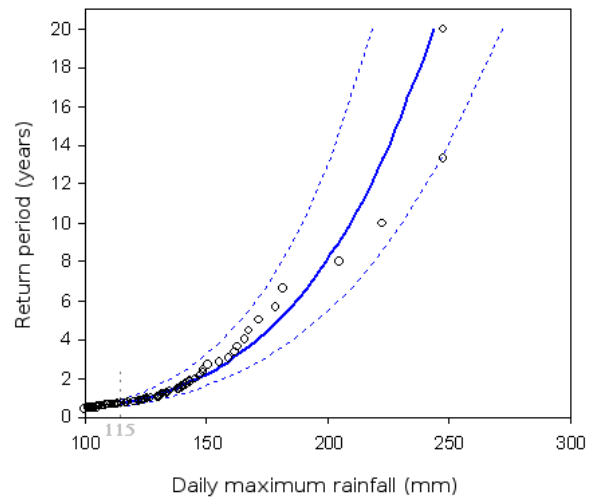
$P_1$ . In this work the MAI-IDF model (Equation 2) was chosen because it only needs three parameters adjusted ( $n$ ,  $m$  and  $P_o(\rho_o, t_o)$ ), regardless of duration  $t$ , thus it is more robust than conventional distributions, such as distribution SQRT-ET $_{max}$  (Etoh et al., 1986; Ferrer, 1996), and even the Generalized Pareto Distribution (GPD) using PDS (Ben-Zvi, 2009), as these models use different parameters for each duration  $t$ .

The maximum daily precipitation ( $P$ ) recorded by a group of stations depends largely on the number of stations ( $N$ ) of the group and the number of days ( $D$ ) in which the group of stations is active. The larger the number of days and mutually independent stations in each group, the better the statistics. In addition, the analyzed groups must be climatically similar, so it can be assumed that the probability is not conditioned to a type of group (depending on climatic parameters). With all these assumptions, the highest recorded rainfall  $P$  can be expected to be a function of the product between the number of days of activity and the number of independent stations ( $N \cdot D$ ), as this is equivalent to having only one station with the total number of days from all stations ( $D_{total}$ ).

Therefore, the maximum daily precipitation expected in the total number of days available ( $D_{total}$ ) is given by Equation 2, taking that time as period of return,  $\rho = D_{total}$ , and taking  $t = t_o = 1$  day as duration, that is to say:

$$P(D, N) \approx P_o(D_o, N_o) \left( \frac{DN}{D_o N_o} \right)^m \quad (4)$$

On the other hand, the cumulative probability  $\pi(P)$  is estimated assuming that all days present the same probability  $\Pi(P)$  that the group of stations records a value equal



**Figure 3.** Adjustment of the return period curve (Equation 2) for the maximum daily rainfall of the Basque Country. The adjusted parameters are:  $m = 0.23 \pm 0.02$  and  $P_o(t_o) = 125 \pm 4$  mm (error interval of 95%).

to or higher than  $P$ . Since the maximum recorded rainfall is expected to be larger as the number of stations is larger, it may be necessary to correct the expected contribution to each daily register, according to the number of active stations, based on Equation 4.

### 2.3 Curve of extreme precipitation

To study the frequency of extreme rainfall events in the Basque Country, a precipitation threshold high enough so that it is only exceeded by a few days can be considered, for example 0.05% (Begueria, 2005). The threshold is defined as the maximum rainfall,  $P_U$ , accumulated during a period of time  $t_U$ . Meanwhile, the return period  $\rho(P_U)$  can be defined as the expected time between two events that exceed the threshold  $P_U$ . To simplify the equations, a theoretical return period is considered,  $\rho_i(P_{Teo})$ , from the average of the occurrences of an episode and the next,  $\rho_{ij}(P_i)$ , ie:

$$\rho_i(P_{Teo}(P_i)) \equiv \frac{1}{N} \sum_{j=1}^N \rho_{ij}(P_i) \quad (5)$$

where the subscript  $i$  enumerates the maximum accumulation episodes equal to  $P_i$  so that  $P_i > P_U$ ,  $t_{ef} < t_U$ , according to Equation 3. On the other hand  $P_{Teo}$  is the theoretical accumulation for the same set, taken according to the following proposed equation:

$$P_{Teo}(P_i) \equiv \frac{(N-2)P_{Min} + \langle P_i \rangle}{N-1} \quad (6)$$

where  $N$  is the number of episodes considered equal to or larger than 3, and  $P_{Min}$  is the lowest maximum accumulation of all the episodes. This way, the representativeness of the minimum intensity is adjusted with the number of available data and considering the average of intensities. Finally, the theoretical curve for the rainfall extremes in the Basque Country is considered, according to Equation 2, designed from the average of the three parameters above:

$$P_o \equiv \langle P_{Teo}(P_i) \rangle \quad (7)$$

$$\rho_o(P_o) \equiv \langle \rho_i(P_{Teo}) \rangle \quad (8)$$

$$n_o \equiv \langle n_i(P_i) \rangle \quad (9)$$

where Equations 7 and 8 refer to the average values obtained by the above Equations 5 and 6, while  $n_i(P_i)$  is the value of the index  $n$  adjusted for each of the episodes- $i$  of extreme precipitation.

## 3 Results

### 3.1 Probabilistic distribution of daily rainfall

Once estimated the cumulative probability of maximum daily precipitation for the entire period 1961-2000, different probabilistic models were applied (Generalized Extreme Value, Generalized Pareto Gamma, Gumbel and Weibull Distribution) and it was found that none of them fit properly for the entire rainfall, according to various non-parametric tests. For example, in a confidence interval of 95% and for 878 classes, the Kolmogorov-Smirnov parameter is larger than

0.1 and the parameter Anderson and Darling (1952) is larger than 39 in all of them (Siegel, 1986; Corder and Foreman, 2009). One of the models that best fits the data is the Gamma distribution with  $\alpha = 19.7 \pm 0.7$  and  $\beta = 0.97 \pm 0.02$  (Figure 2), which has a mean absolute error (NMAE) of 8.6% in the estimation of the precipitation values, being 4.1% for values between 14 and 115 mm in 24 h, and 12% for values between 115 and 250 mm.

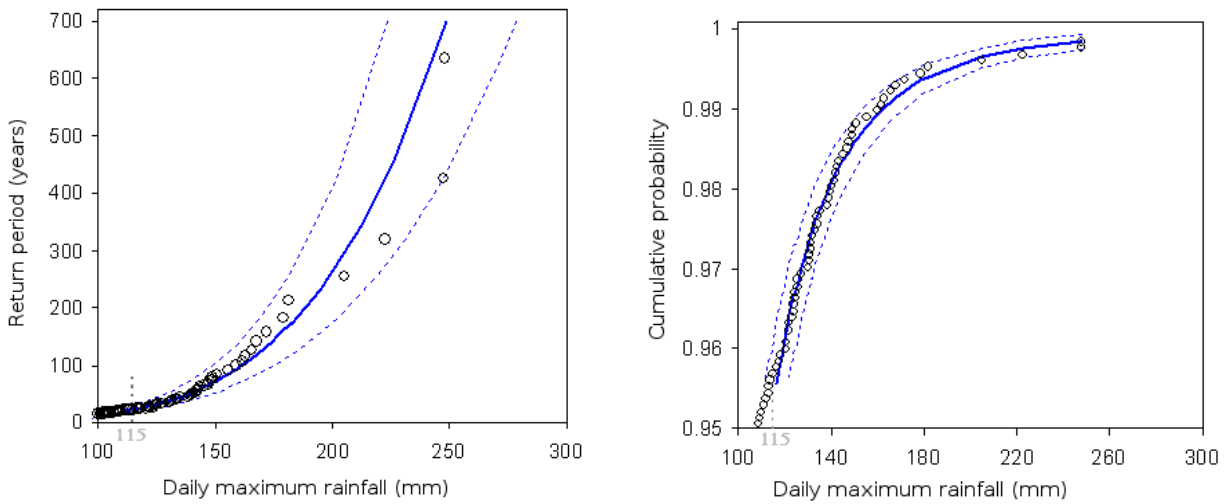
A better setting of the Gamma distribution was found for values between 14 and 115 mm in 24 h, with  $\alpha = 17.0 \pm 0.2$  and  $\beta = 1.20 \pm 0.05$ , showing a NMAE of 1.6%. However, for values between 115 and 250 mm the error is larger than in the previous setting (17%). The interval between 115 to 250 mm was also tested with the same distributions and it was observed that the best distributions were: the Potential Distribution (p-value < 0.0001), the Exponential Distribution (p-value = 0.0001) the Generalized Pareto distribution (p-value = 0.0003) and Gamma Distribution (p-value = 0.0005). In particular, for Potential Distribution (Equation 2) a NMAE of 1.4% was obtained, with the following adjusted parameters:  $m = 0.23 \pm 0.02$  and  $P_o(t_o) = 125 \pm 4$  mm (Figure 3).

The probability of recording a daily maximum (larger than a threshold  $P$ ) may vary over the study period (1961-2000) according to the number of active stations. Therefore, the maximum precipitation threshold was corrected according to Equation 4, so that the final values represent a constant set of 31 stations. However, it was noted that the correction does not significantly change the probability distribution, as similar parameters are obtained:  $m = 0.22 \pm 0.02$  and  $P_o(t_o) = 126 \pm 4$  mm (compared to  $m = 0.23 \pm 0.02$  and  $P_o(t_o) = 125 \pm 4$  mm, before correcting).

On the other hand, considering that the setting obtained refers to the average group of 31 stations in the Basque Country with daily availability, it is possible to estimate the return period equivalent to a single station (average return period). To do this, Equation 4 was again applied, as it relates maximum daily precipitation and the number of records available, ie the product of the set of  $N$  stations and the expected return period in years. As a first approximation it was considered that the  $N$  stations have an independent probability of recording a maximum rainfall record. With all this, it was found that the adjusted parameters with Equation 4 for one station are:  $m = 0.22 \pm 0.02$  and  $P_o(t_o) = 58 \pm 2$  mm (Figure 4).

### 3.2 Curve of extreme precipitation

To design a curve of extreme precipitation a threshold exceeded on 0.05% of the days during 1950-2008 was considered. Specifically, the threshold was set at 150 mm in 24 h, so 14 days in which a station exceeded the threshold were found. These 14 entries correspond to 11 stations with high average rainfall, in the north and northeast of the Basque Country. Therefore, to estimate the return periods the 14 cases were analyzed together. However, to estimate the aver-



**Figure 4.** (a) (left) Register time or total period (expressed in number of stations for the period of return) for each daily maximum rainfall and (b) (right) cumulative probability associated to that measuring period. The adjusted parameters are:  $m = 0.22 \pm 0.02$  and  $P_o(t_o) = 58 \pm 2$  mm (error interval of 95%), with a reference return period  $\rho_o = 1$  year.

**Table 1.** Some of the episodes between 1950 and 2008, whose precipitation in one day ( $P$ ) exceeds the threshold of 150 mm, and there is information available about their temporal distribution ( $n$ ).

Episode (day of the maximum)	Station of the maximum	$P$ (mm)	$n$
11-15 October 1953 (14)	Oyarzun (Arditurri)	313.5	$0.65 \pm 0.12$
24-25 September 1959 (24)	Fuenterrabia (Aeropuerto)	214	$0.71 \pm 0.15$
26 August 1983	Larraskitu	503	$0.60 \pm 0.09$
1 June 1997	San Sebastián (Igeldo)	251	$0.53 \pm 0.07$
25 August 2002	San Sebastián (Ategorrieta)	224	$0.64 \pm 0.06$

age temporal distribution only the index  $n$  is available for 5 of the 14 extreme episodes (Table 1).

There are enough cases available (14) to adjust the empirical return periods (Equation 5) to theoretical accumulations (Equation 6) and get the value of the parameter  $m$  according to Equation 2, in which case it results in Equation 10. However, it is also possible to treat the 14 cases together to get an average value of the maximum expected rainfall and return period, thus it is possible to use the same equation, but assuming that the value of  $m$  is equal to  $0.24 \pm 0.02$  (average value for the whole Iberian Peninsula, according to Moncho et al., 2009). In this case the result is Equation 11. For both estimates it is considered that the value of index  $n$  is the average of the episodes in Table 1, ie  $n = 0.63 \pm 0.06$ .

$$P_{max} \approx 117 \text{ mm} \left( \frac{\rho}{\rho_o} \right)^{0.28} \left( \frac{t}{t_o} \right)^{1-0.63} \quad (10)$$

$$P_{max} \approx 125 \text{ mm} \left( \frac{\rho}{\rho_o} \right)^{0.24} \left( \frac{t}{t_o} \right)^{1-0.63} \quad (11)$$

where  $P_{max}$  is the maximum accumulation for a return period  $\rho$  and a duration  $t$ , while a  $\rho_o$  reference period of one

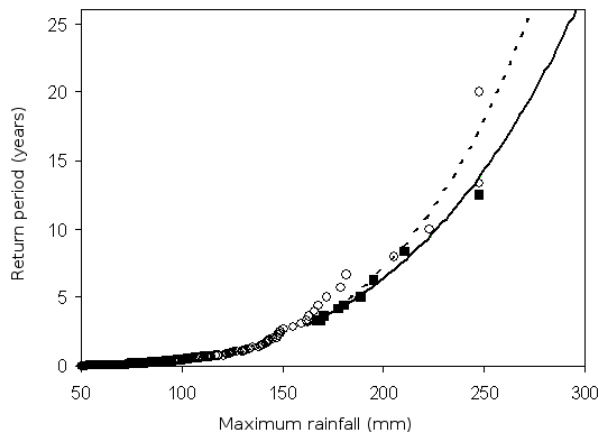
year and a duration  $t_o$  of one day have been chosen. It is important to highlight that the resulting curve from adjusting the value of  $m$  (Equation 10) shows more extreme precipitation values than the average curve (Equation 11), as shown in Figure 5.

Calculations were repeated for the 10 stations with registers of lower extreme rainfall records, and it was found that the maximum daily rainfall with a return period of 1 year is  $P_o(t_o) = 71 \pm 5$  mm with parameter  $m = 0.23 \pm 0.05$ . That is, using Equation 2, the mean for a station is  $P_o(t_o) = 42 \pm 2$  mm in a day, with a return of one year. The location of the 10 stations is distributed in the southern province of Alava, which is also consistent with the records of mean annual precipitation (Martín-Vide, 2004).

## 4 Discussion

### 4.1 Choice of methodology

The precipitation stations with daily availability of the Basque Country in general have a short series of records,



**Figure 5.** Comparison among the maximum daily rainfall (circles) of the set of 31 stations (period 1961-2000) and the extreme rainfall (squares) obtained from Equations 5 and 6, for the set of 11 stations that have exceeded the considered threshold (150 mm in 24 h) in the period 1950-2008. The solid line is the adjustment of Equation 10 to extreme rainfall, with the parameters  $m = 0.28 \pm 0.06$  and  $P_o(t_o) = 117 \pm 10$  mm, while the broken line is the result of adjusting Equation 11, with parameters  $m = 0.24 \pm 0.06$  and  $P_o(t_o) = 125 \pm 9$  mm (error interval of 95%).

making it very difficult to estimate the return period of extreme rainfall. Therefore, in this paper we have chosen to study the available stations together, both analyzing all the stations and trying separately the most and least extreme sets of stations. This way the number of calculable measurements in a given statistics is increased, which means having much larger series (number of independent stations per number of years). However, due to the nature of extreme precipitation, probabilistic models often present difficulties to reproduce satisfactorily the higher return periods. In this paper we chose the Potential or Pareto Distribution as it is among those which provide the best results and mathematically it is the simplest. This facilitates the interpretation of results by mathematically explicit IDF curves that require few parameters. In other models such a GEV, SQRT-ET<sub>max</sub>, Gumbel, etc., 2 or 3 parameters are needed for each different duration (Ferrer, 1996; Pereyra-Diaz et al., 2004), and therefore a large number of parameters is required to write an overview of the representative IDF curves of any length.

The potential shape of the IDF curve needs only three parameters ( $P_o$ ,  $m$ ,  $n$ ) to represent all the possible temporal distributions of maximum rainfall, for a given set of  $N$  stations, and for a return period  $\rho$ . On the other hand, the reason to choose a reference period  $\rho_o$  of one year responds to reasons of mathematical simplicity (the equations are more synthetic), but usually the equations are applied for return periods over 10 years.

## 4.2 Effect of the number of stations

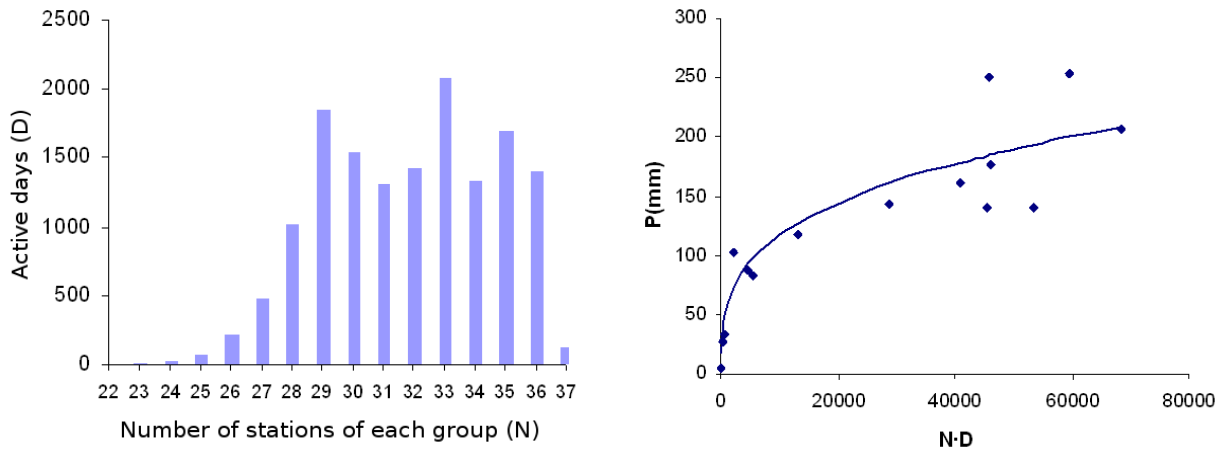
For the combined statistics of the whole Basque Country, a correction of the cumulative probability according to the variable number of stations has been applied, estimating probable maximum daily precipitation that would have been registered if the number of stations had been constant in 31 (in accordance with Equation 4). This is to avoid overestimation or underestimation of the probability of a register conducted in a given group of stations different than the average.

The relationship of probability based on the number of active stations can be seen in Figure 6. Given the number of days,  $D$ , there are  $N$  active stations available (Figure 6a), the maximum precipitation obtained depends on the total number of records (the product of the  $N$  stations by the  $D$  days when they are active), according to Equation 4, with  $m = 0.30 \pm 0.05$  and  $P_o = 7 \pm 1$  mm for a 1 *day-station*, which corresponds to  $44 \pm 6$  mm for a 1 *year-station*. If compared with the curve in Figure 4a, we can see that the value is very close to the reference value 58 mm for a 1 *year-station*. The difference between both values may be that in Figure 6b it was considered that all groups are equally likely to register a maximum, which a priori is not true. Therefore, it appears that the initial assumption that the probability depends on the number of stations is justified.

According to the extreme curves corresponding to the group of 11 stations that recorded the rainfall records (Equations 10 and 11), it is possible to compare them with the estimated IDF curves for the stations of Bilbao-Sondica and San Sebastian-Igueldo (Table 2). In this comparison we can appreciate that the index  $n$  of these stations is consistent with that obtained for the records of the Basque Country ( $0.63 \pm 0.06$ ), especially that of the San Sebastian-Igueldo station, which has the largest maximum accumulation in 1 hour. Specifically, the accumulation corresponding to one day for the same station is 147 mm for a return period of 25 years (Equation 2). For a return period of 1 year, this value corresponds to 70 mm, assuming  $m = 0.23$ , which is almost half the 125 mm value of Equation 11.

The difference is due to the fact that Equation 11 corresponds to a statistic of 11 stations, and not just one station, so that the joint probability of registering a maximum is significantly increased. Again, applying Equation 4, it is possible to predict that the value of 125 mm corresponds to  $N/N_o = 11 \pm 2$  regarding Igueldo. In other words, it is possible to interpret that the effective number of independent stations in the group of 11 stations is about 11, or very close, regarding Igueldo. If the same relationship is applied with the statistical average of all stations ( $58 \pm 2$  mm in 1 day per year), we get that a set of  $30 \pm 4$  seasons with the average statistics is necessary to match the statistics of the group of 11 stations compatible with that of Igueldo. This is logical because the statistics for all stations (31) include records from the records at the 11 stations.





**Figure 6.** (a) (left) Number of days,  $D$ , when there are  $N$  active stations. (b) (right) Maximum rainfall in one day registered in each group, and comparison with the product  $N \cdot D$  (number of stations and duration of each one of that group). The adjusted curve corresponds to the model of Equation 4, where  $m = 0.30 \pm 0.05$  and  $P_o = 7.4$  mm for a  $N_o \cdot D_o$  of reference equal to one station per one day.

**Table 2.** Values of the adjustments to the IDF curves for the stations of Bilbao-Sondica and San Sebastián-Igueldo, according to Moncho et al. (2009).  $n_{med}$  is the mean value of the index  $n$ , while the reference times are 60 minutes for the duration ( $t_o$ ) and 25 years for the return period ( $\rho_o$ ).

Station	$P(t_o, \rho_o)'$	$m$	$\varepsilon(m)$	$n_{med}$	$\varepsilon(n_{med})$
1024E San Sebastián Igueldo	45.5	0.23	0.02	0.63	0.03
1082 Bilbao Sondica	35.8	0.24	0.02	0.574	0.002

Therefore, one might think that the 11 stations are enough to reproduce most of the statistical maximum of the whole group of daily stations. However, it would be worth examining in detail the daily series of such stations as well as the rest of the Basque Country to estimate a function of the effective number of independent stations in relation to the total number of stations in each random group ( $N$ ).

As an example of the application of the results obtained, the return period of the 503 mm rainfall recorded in 24 h during August 26, 1983 in Larraskitu (Ugarte and González, 1984) was estimated, and it resulted in  $300 \pm 150$  years (95% reliability) for all the 11 stations with the most extreme rainfall. Taking into account that on January 9, 1875, 495 mm were registered in 24 h (AEMet, 2010) at the main station of Bilbao (1859-1920), it is confirmed that the repetition of 500 mm in 24 h the Basque Country is quite low.

### 4.3 Spatial distribution

Regarding the spatial distribution of maximum precipitation in the Basque Country, this work is limited to estimating the overall range of values without going into great detail, because it is considered that the individual length of the analyzed series (40 years) is insufficient, at least with this methodology, to adequately estimate the spatial variabil-

ity referred to the same period of study and a real climatic representation. Therefore, as discussed above, an alternative is to simulate a longer time by combining the statistics from several stations, or it is also possible to estimate the IDF curves for a few stations and try to interpolate the spatial distribution with the help of orographic predictor variables. For example, using the series of the rain gauges of Igueldo, Bilbao and others from autonomous communities nearby (Moncho et al., 2009), it was estimated that in the Basque Country the maximum precipitation oscillates between 30 and 50 mm in an hour for a return period of 25 years (Annex B), which corresponds to values between 40 and 70 mm in 24 h with a return period of 1 year, ie, an average of about 55 mm in 24 h. This is consistent with the value obtained in this work of  $P_o = 58 \pm 2$  mm, referred to one station, using that  $m = 0.24 \pm 0.02$ .

The same value of  $P_o = 58 \pm 2$  mm (with a return equal to one year) implies that for a return period of 10 years, the maximum expected precipitation for one station is about 100 mm, on average. This is consistent with Pérez-Cueva (1983), who gets a maximum precipitation for the same return period of 70 mm in the south of lava to 150 mm in Guipuzcoa. Besides the match on the average, it is worth mentioning that the range of values is almost coincident in the three papers, that is, hopefully, in the Basque Country the maximum rainfall in a day is between 40 and 70 mm in a

return period of 1 year, or what is the same, between 70 and 150 over a period of 10 years.

## 5 Conclusions

From the records of the 43 stations available daily (498 092 data), the daily maximum rainfall from 1961 to 2000 was analyzed and the associated daily probability distribution was estimated. This has shown that the usual models of probability distribution do not adequately conform to the curve (Kolmogorov-Smirnov is larger than 0.1 in all of them, with a reliability of 95% and 878 classes considered). However, analyzing the series from a certain extreme threshold (115 mm in 24 h), the models show a remarkable improvement, especially in the case of the Potential Distribution Model (mean error 1.4%).

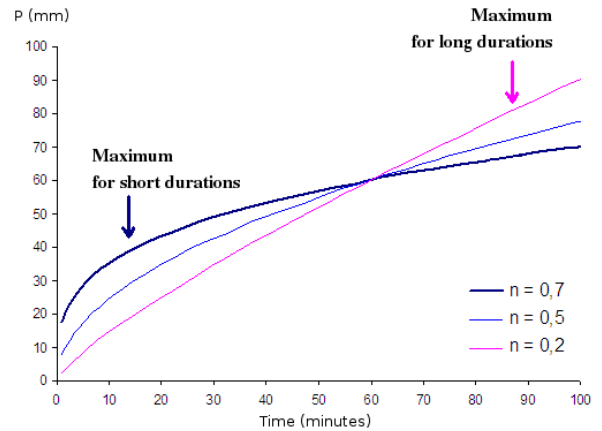
It was observed that the maximum rainfall recorded in a group of stations depends on the number of stations belonging to the group. In a first approximation we can consider that in the Basque Country stations with daily availability have virtually independent chances of registering an extreme of precipitation. However, it was observed that for the average probability, the number of stations was fairly constant over time, since the correction of probability did not incorporate significant changes in the parameters set.

Finally, a mathematical expression is proposed to estimate the maximum expected precipitation for a return period  $\rho$  (in years) equal to or larger than 10 years, with a duration  $t$  (in days) over 1 minute, and for a set of  $N$  independent stations, according to:

$$P_{max} \approx P_o (N\rho)^m t^{1-n} \quad (12)$$

where  $m = 0.23 \pm 0.02$ ,  $n = 0.63 \pm 0.06$ , while the scale factor,  $P_o$ , depends on whether we consider the set of all stations in the Basque Country ( $P_o = 58 \pm 2$  mm) or a particular area, whose rainfall significantly differs from the average: for the rainiest area (north and northeast of the Basque Country) the result is that  $P_o = 71 \pm 5$  mm, while for the less rainy area (South Álava) the result is  $P_o = 42 \pm 2$  mm. This model has important advantages over the others because it incorporates the dependence with time and the number of stations using only 3 adjustable parameters.

**Acknowledgements.** Roberto Moncho has been a beneficiary of the Iñaki Goenaga scholarship from the Technology Center Foundation, without which this work would not have been possible. In addition, we appreciate the cooperation of the Spanish Meteorological Agency (AEMET) and especially that of Margarita Martín, Director of the Provincial Delegation of the AEMET in the Basque Country, who provided us the meteorological data needed to develop this study. Likewise, we acknowledge the input from Guillem Chust (AZTI). This work is part of a project funded by the grant Ertortek of the Basque Government (*K-Egokitzen II*).



**Figure A1.** Comparison of three kinds of time distributions of rainfall, according to three different values for index  $n$ , taking as a reference the accumulation of 60 mm in one hour that, according to the Spanish Meteorological Agency, represents the torrentiality threshold in Spain.

## Appendix A Temporal distribution of rainfall

The most commonly used pluviometric variable in hydrology is the maximum average intensity, especially on Intensity-Duration-Frequency curves (Témez, 1978). The maximum average intensity is defined as the quotient between the maximum accumulation for each time interval and that time interval, ie:

$$I(t) \equiv \frac{P(t)}{t} \quad (A1)$$

where  $P(t)$  is the maximum accumulation in a time  $t$ , and  $I(t)$  is the maximum average intensity in a time  $t$ . According to Chow (1962), the maximum average intensity can be adjusted as a triparametric function:

$$I(t) = \frac{a}{t^n + b} \quad (A2)$$

where  $a$ ,  $b$  and  $n$  are the adjustable parameters for each case. And this equation in turn can be rewritten as:

$$\frac{\frac{I_{max}(0)}{I(t)} - 1}{\frac{I_{max}(0)}{I_o(t_o)} - 1} = \left(\frac{t}{t_o}\right)^n \quad (A3)$$

where  $I_{max}(0)$  is the absolute maximum average intensity (at time zero as reference),  $I_o(t_o)$  is the maximum average intensity of arbitrary length  $t_o$ , and  $I(t)$  is the maximum average intensity in the duration of interest  $t$ . Since the moment of maximum intensity is in general  $I_{max}(0) \gg I(t)$ , then Equation A3 can be approximated to the mathematical relationship:

$$I \approx I_o(t_o) \left(\frac{t_o}{t}\right)^n \quad (A4)$$



This approach is valid for any episode of precipitation and for any time scale equal to or more than a minute (Moncho, 2008). Furthermore, by Equation A1, Equation A4 can be rewritten for the maximum accumulation  $P(t)$ , in a time  $t$ , as:

$$P(t) \approx P_o(t_o) \left( \frac{t}{t_o} \right)^{1-n} \quad (\text{A5})$$

where  $P(t_o)$  is the maximum accumulation in a reference time  $t_o$ . Note that the parameter  $n$  provides information on the type of precipitation (Moncho et al., 2009). For instance, low values of  $n$  are more associated with a greater persistence of the intensity than high values of  $n$  (see Figure A1).

## Appendix B Probabilistic distribution of rainfall

Defining  $\Pi(p \geq P)$  as the probability that a maximum precipitation  $p$  (of a certain duration) equal to or greater than  $P$  happens in a year, the Potential or Pareto Distribution is:

$$\begin{aligned} \Pi(p \geq P) &= 1 & \text{si } P < P_1 \\ \Pi(p \geq P) &= \left( \frac{P_1}{P} \right)^{\frac{1}{m}} & \text{si } P \geq P_1 \end{aligned} \quad (\text{B1})$$

where  $m$  is a positive parameter and  $P_1$  is the maximum expected rainfall for a year, with the same length as  $P$ .

The return period exceeding one year,  $\rho(P)$ , is defined as the inverse of the probability  $\Pi(p \geq P)$ . Therefore, dividing the return period  $\rho_o(P_o)$  of the reference accumulation  $P_o$ , Equation B1 can be written as:

$$\frac{\rho_o(P_o)}{\rho(P)} \approx \left( \frac{P_o}{P} \right)^{\frac{1}{m}} \quad (\text{B2})$$

Reversing Equation B2, the expected precipitation for each return period,  $P(\rho)$ , can be written as:

$$P(\rho) \approx P_o(\rho) \left( \frac{\rho}{\rho_o} \right)^m \quad (\text{B3})$$

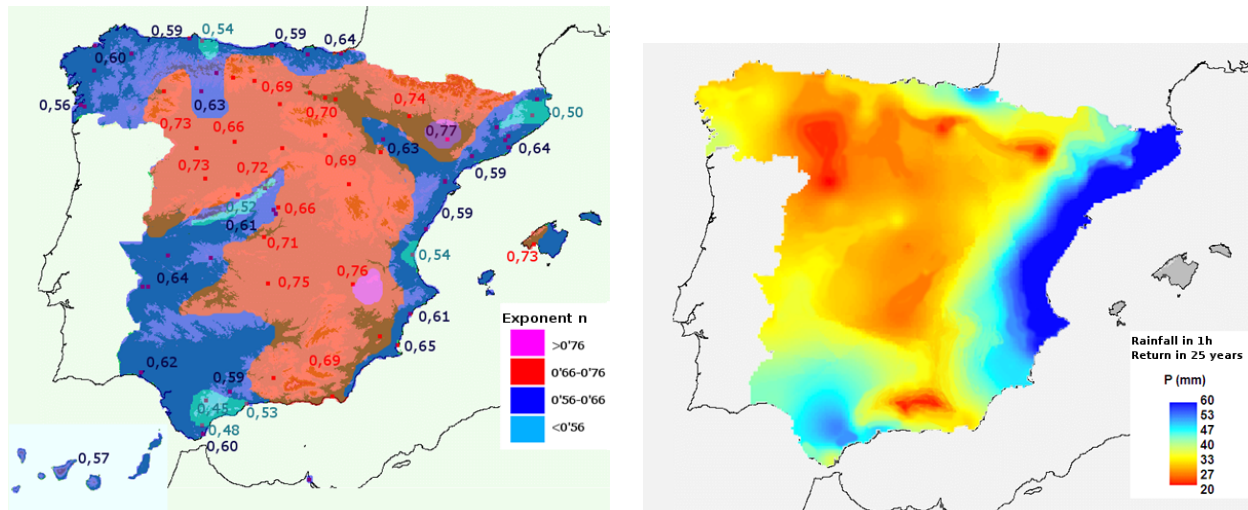
Finally, considering both Equation A5 and B3, we obtain a general expression for a maximum accumulation  $P$ , associated with a return period  $\rho$  and a duration  $t$ :

$$P(\rho, t) \approx P_o(\rho_o, t_o) \left( \frac{\rho}{\rho_o} \right)^m \left( \frac{t}{t_o} \right)^{1-n} \quad (\text{B4})$$

This equation requires only three parameters,  $m$ ,  $n$  and  $P_o(\rho_o, t_o)$ ; unlike the most commonly used distribution, the SQRT-ET<sub>max</sub> (Etoh et al., 1986), which requires the setting of two parameters for each time. In addition, parameter  $m$  is approximately constant for the whole of the Peninsula ( $0.24 \pm 0.06$ , Moncho et al., 2009), and the other parameters are known (Figure B1).

## References

- AEMet, 2010: Valores climáticos extremos, Agencia Estatal de Meteorología (WEB), 11/VI/2010.
- Anderson, T. W. and Darling, D. A., 1952: *Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes*, *Annals of Mathematical Statistics*, **23**, 193–212.
- Beguera, S., 2005: *Uncertainties in partial duration series modelling of extremes related to the choice of the threshold value*, *J Hydrol*, **303**, 215–230.
- Ben-Zvi, A., 2009: *Rainfall intensity-duration-frequency relationships derived from large partial duration series*, *J Hydrol*, **367**, 104–114, ISSN: 0022–1694.
- Chow, V. T., 1962: Hydrological determination of waterway areas for the design of drainage structures in small drainage basins, Engineering Experimental Station, Bulletin 462, University of Illinois, Urbana, USA.
- Corder, G. W. and Foreman, D. I., 2009: *Nonparametric Statistics for Non-Statisticians: A Step-by-Step Approach*, Wiley, ISBN: 978047045461.
- de Salas, L. and Fernández, J. A., 2006: *Nueva metodología para el análisis de la variable intensidad máxima anual de precipitación*, *Ecología*, **20**, 35–44.
- Etoh, T., Murota, A., and Nakanishi, M., 1986: SQRT-Exponential Type Distribution of Maximum, Hydrologic Frequency Modelling, V. P. Shing (Ed.), Reidel Pub. Com, Louisiana State University, USA, Proceedings of the International Symposium on Flood Frequency and Risk Analyses, 14-17 May 1986, 253-264.
- Ferrer, F. J., 1996: El modelo de función de distribución SQRT-ET<sub>max</sub> en el análisis regional de máximos hidrológicos: aplicación a lluvias diarias, Tesis Doctoral, Universidad Politécnica de Madrid.
- Font-Tullot, I., 2000: *Climatología de España y Portugal*, Universidad de Salamanca, 2a Edición, 422 pp.
- Kundzewicz, Z. W. and Robson, A., 2000: World Climate Programme - Water detecting trend and other changes in hydrological data, United Nations Educational Scientific and Cultural Org., Geneva, May, WCDMP, 45 WMO/TD - No. 1013 (PDF).
- Lana, X., Burgueño, A., Martínez, M. D., and Serra, C., 2009: *A review of statistical analyses on monthly and daily rainfall in Catalonia*, *Tethys*, **6**, 15–29, doi: 10.3369/tethys.2009.6.02.
- Llasat, M. C., Ramis, C., and Barrantes, J., 1996: *The meteorology of high-intensity rainfall events over the West Mediterranean region*, *Remote Sensing Reviews*, **14**, 51–90.
- Marco, J., 1999: Análisis Hidrológico y sus implicaciones en la planificación de una defensa integral frente a las inundaciones de la Ribera del Júcar, CONFEDERACIÓN HIDROGRÁFICA DEL JÚCAR: Estrategias de actuación en la zona inundable de la Ribera del Júcar, Ponencias Técnicas, Volumen I, Alzira, 29-30/III/1999.
- Martín-Vide, J., 2004: *Spatial distribution of a daily precipitation concentration index in peninsular Spain*, *Int J Climatol*, **24**, 959–971, doi: 10.1002/joc.1030.
- Ministerio de Fomento, 1999: Máximas lluvias diarias en la España Peninsular, Ministerio de Fomento, Madrid, ISBN: 978-84-498-0419-9 (PDF).
- Moncho, R., 2008: *Análisis de la intensidad de precipitación. Método de la intensidad contigua*, RAM3, **Enero**.
- Moncho, R., Belda, F., and Caselles, V., 2009: *Climatic study of the exponent n of the IDF curves of the Iberian Peninsula*, *Tethys*, **6**, 3–14, doi: 10.3369/tethys.2009.6.01.



**Figure B1.** (a) (left) Spatial distribution of the index  $n$  of Equation B4, where there is a distinction between climates with persistent maximum rainfall (blue) and anti-persistent (red). (b) (right) Maximum rainfall in one hour with a punctual return period of 25 years (Moncho et al., 2009).

Pereyra-Díaz, D., Pérez-Sesma, J. A. A., and Gómez-Romero, L., 2004: *Ecuaciones que Estiman las curvas Intensidad-Duración-Período de retorno de la lluvia*, GEOS, **24**, 46–56 (PDF).

Pérez-Cueva, A., 1983: *Precipitaciones extraordinarias en la España Peninsular*, Agricultura y sociedad, **28**, 189–203, ISSN: 0211–8394 (PDF).

Ruíz-Urrestarazu, E., 1983: Estado actual de las fuentes de los datos climáticos en Álava, Guipúzcoa y Vizcaya y su posible planificación, Eusko Ikaskuntza (Ed.), Encuentro de Geografía Euskal Herria-Catalunya, Donostia-San Sebastián, 395–419, ISSN: 0212–3223 (PDF).

Siegel, S., 1986: *Estadística no paramétrica aplicada a las Ciencias de la Conducta*, Editorial Trillas, México D.F.

Soro, G. E., Goula Bi, T. A., Kouassi, F. W., and Srohourou, B., 2010: *Update of intensity-duration-frequency curves for precipitation of short durations in tropical area of West Africa (Cote D'ivoire)*, J Appl Sci, **10**, 704–715, doi: 10.3923/jas.2010.704.715.

Témez, J. R., 1978: *Cálculo Hidrometeorológico de caudales máximos en pequeñas cuencas naturales*, Dirección General de Carreteras, Madrid, p. 111.

Ugarte, F. M. and González, J. A., 1984: *Precipitaciones de agosto de 1983 en la cuenca del río Oka (ría de Mundaka-Vizcaya): repercusiones geomorfológicos*, Cuadernos de investigación geográfica, **10**, 201–214, 7 Ref., ISSN: 0211–6820 (PDF).