Extended Abstract

Introduction

The Portugal mainland currently has a wide range of water gages, where the precipitation data can be collected, these are often used as the foundation of various hydrological studies. The collected data are organized and available online, for research purposes, at the Sistema Nacional de Informação de Recursos Hídricos's website, SNIRH, under Instituto da Água's responsibility, INAG. In its majority, the registries refer to daily precipitation values, even though since the beginning of the century hourly registries have been used more often, which can often become an important tool, within a project context, once shorter rainfall durations are considered.

One of the possible applications of those records is to estimate the intensity-duration-frequency curves or IDF that, in a watershed, relate the intensity of precipitation, with the associated duration of a certain return period. These curves have a special relevance to determine the maximum rainfall that are widely used on design of hydraulic works and engineering projects.

So, taking advantage of shorter rainfall events records, it is intended in this investigation to establish IDF curves, for durations shorter than a day with a generic but sufficiently rigorous methodology, based in the theory of Partial Duration Series (PDS) in continental Portugal. It is also intended to test their suitability by comparison with other already established and distinct methodological curves, thus proving that the proposal can be presented as a good alternative to the existing ones, more so when shorter recordkeeping periods are available.

With this purpose three water gages in the continental Portuguese territory were analyzed (Abrantes, Monchique and S.Julião do Tojal) and were considered a total of 289 296 records of hourly precipitation, comprised in periods of 10 to 12 hydrological years.

Methodology/Results

The IDF curves utilized in project, have as base a methodology allusive to maximum annual series, where in every hydrological year of study the maximum daily precipitations are taken into account. This method requires from the beginning the analysis of long periods as to have enough samples with a considerable dimension suitable to its statistic treatment.

The technique the present investigation proposes is based on partial duration series that by definition considers extremes all the records above a defined threshold, thus achieving samples with reasonable dimensions without the necessity of long periods of sampling. Figure 1 exemplifies the use of this method:

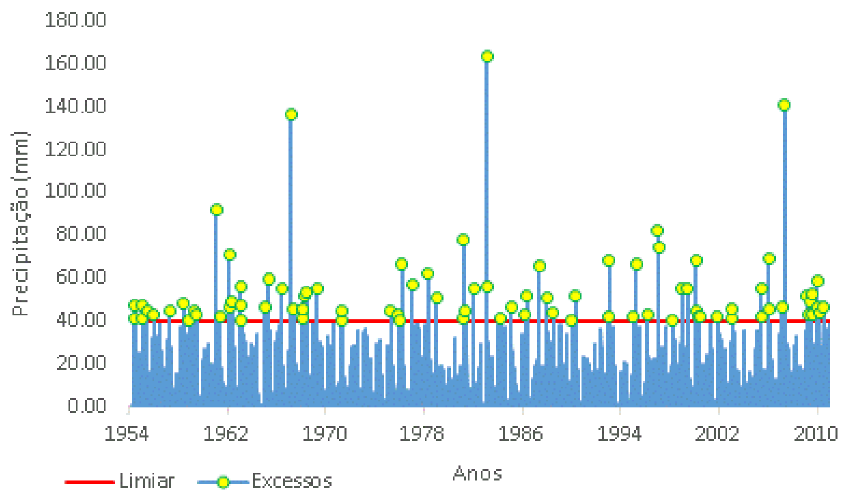

Figure 1- SDP sample for the water gage of S.Julião Tojal with threshold set u = 40mm and hydrologic years between 1954/55 and 2010/2011.

The consideration of this method encompasses still some subjectivity regarding its application, mainly when it comes to choosing the thresholds and the criteria of independent statistic to utilize, being however quite promising because it allows from the start the consideration of rainfall events that are really extreme regarding a reference value, and the rejection of others that could compromise the inference for reasons such as its low absolute value or statistical dependence.

The study began with the collection of hourly rainfall data to further analyze. The PDS theory assumes that the recorded values are continuous from an hydrologic point of view and statistically independent to be submitted to a Poisson process, meaning that the records on course to be analyzed must not contain gaps to ensure that only full hydrological years are considered.

The completion of gaps thus constitutes the first step to ensure the correct use of the available information. The collected data, show that three filling criteria can be adopted: consideration of null precipitation in rainfall free periods, use of a linear variation criteria among contiguous records in the absent periods and the utilization of a linear regression analysis with records on a geographically close water gage.

Secondly, an automated calculus program was used, whose functionality was to obtain the SDP and some statistical variables relevant to the analysis as well as the coefficients of autocorrelation, the mean number of events by hydrological year, the mean excedent and the respective limits of the confidence interval (CI) of 95%. The program has as a functioning principle the progressive increase of the limit values, u, with specification for the each limit in the respective sample

Durations of 1, 2, 4, 6, 8, 10, 12, 18 and 24h were adopted and for independence criteria was considered that the rainfall events must be spaced out by a rainfall period of at least 24h Furthermore for each water gage the program was executed, and obtained for each limit the respective results, and then run again with the introduction of the threshold leading to the desired sample according to selection criteria reported in the bibliography.

The selection of the limit constitutes an important step in the whole investigation. Three criteria were taken into account for its choice:

- Annual average number of events, λ, greater than three as mentioned by Lang, Ouarda, & Bobée, 1999.
- Approximately linear behavior (increasing, decreasing or constant) depending on the average exceedance of the threshold u according to Davison & Smith, 1990.
- Analysis of autocorrelation coefficients with increments of 1 and 2 inside the limits of the confidence interval, to assure the statistical independence of the sample as shown by Miquel, Chambolle, & Bernier, 1984.

The aforementioned criteria were applied by observation of the graph that relates the excedent means and their limits in the confidence interval of 95% with its respective limit as demonstrated in Figure 2. Said analysis is done for all considered durations and for each water gage in study.

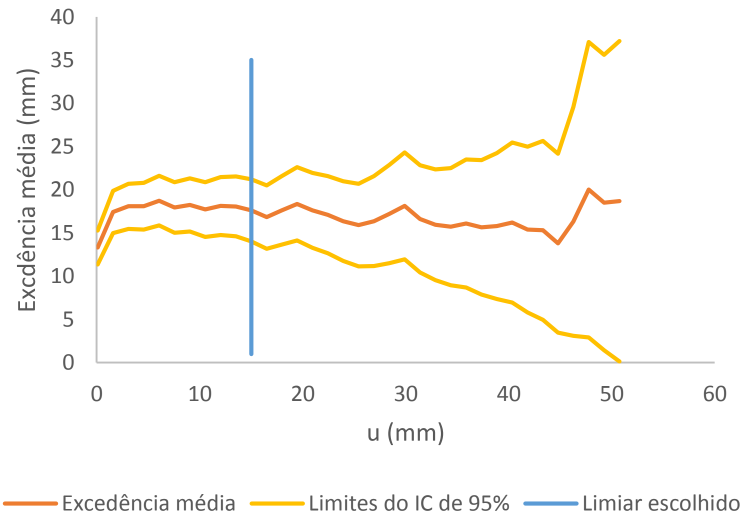

Figure 2 - Exceedance average, respective confidence intervals of 95% and average number of exceedances per year at S.Julião Tojal for the duration of 1 hour in the hydrological years between 2001/02 and 2011/12.

Once the selected threshold, each sample duration are obtained, thereafter being necessary to prove statistical independence of the data included. For this purpose, as mentioned above, it is through the increase of autocorrelation coefficients 1 and 2 which should be within the limits of the 95 % confidence intervals that are determined according to:

$$r_k = \frac{-1 \pm \mu_{1-\alpha/2}\sqrt{N-k-1}}{N-k} \tag{1}$$

Where $\mu_{1-\alpha/2}$represents the inverse value of the normal standard probability density function for α level of significance, *N* is the sample dimension, and *k* the increasing the autocorrelation coefficient .

The independence test is done by considering several dimensions of spaced samples of ten events with the respective limits of confidence interval of 95% and then, verified the inclusion or not of the autocorrelation values of coefficients for the range of sample sizes where inserted.

Once verified the independence of the records contained in the resulting samples, proceeds to his statistical analysis in order to obtain quantile probability. In this sense it refers to the generalized Pareto distribution (GPD) and the exponential distribution is the most appropriate to extreme model above a certain threshold *u*. In this research was used a methodology where exceedances are modeled by an exponential distribution leading to a Gumbel distribution in the field of annual maximum, with probability distribution function given by:

$$F(x) = \exp[-\lambda \exp\left(-\frac{x-u}{\sigma}\right)], \sigma > 0 \qquad (2)$$

Where *u* refers to a positional parameter which in this case assumes the threshold value set, *σ* is a parameter so that assumes the average value of the exceedance and *λ* is the average number of events per year.

In the case of PDS, being λ an estimator of the rate of occurrences of the Poisson distribution, equal to the average number of events per year above a threshold set *u*, it can be shown that the distribution function of maximum annual F(x), the PDS distribution function, H(x), and annual intensity λ occurrences of events defined above threshold u are related as follows:

$$F(x) = e^{-\lambda[1-H(x)]} \qquad (3)$$

Then, in order to quantile estimation probability annual basis with PDS samples, was followed the following methodology:

- Were defined as return periods T, 5, 10, 100, 500 and 1000 years.
- With F(x) = 1-1/T the frequency of annual events was obtained for each return period.
- With the average number of events per year, λ, for the threshold *u*, the given duration and the value of F(x) for a given return period T, gave H(x) by solving the equation (3).
- With the value of H(x) for each threshold and return period, it is calculated the probability quantile of the PDS for the return period on the maximum annual according to the equation (37).

Schematically, the methodology is summarized as follows:

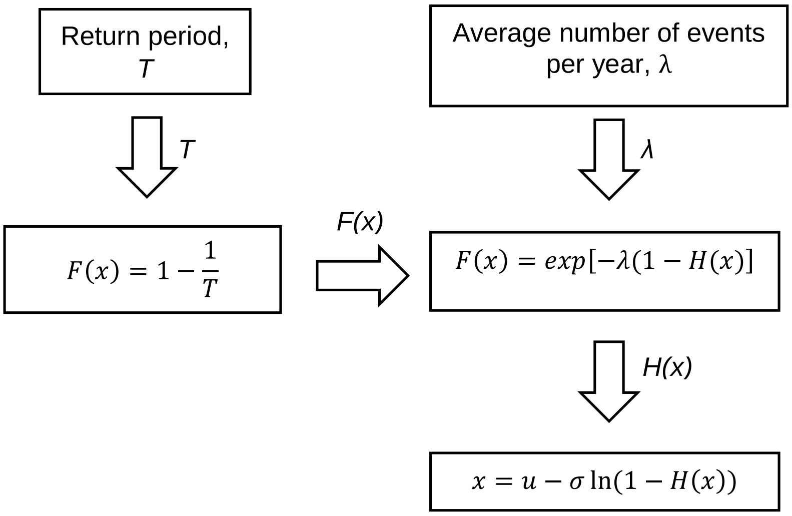

Figure 3 - Simplified scheme of the methodology adopted for the quantile estimation of probability of the PDS, with return periods for the annual maximum.

Finally proceed to the calculation of the parameters required for establishing IDF curves of type $i = at^n$ where we have the precipitations of values and respective durations or intensities, we apply one logarithmic operator to linearize and thus allow the estimation of the coefficients by analysis of simple linear regression:

$$i = at^n \Leftrightarrow \ln i = \ln a + \mathrm{n} \ln t \qquad (4)$$

In each water gage, based on the quantile probability with given return period and different durations resulting from the statistical analysis of the PDS, we obtained the regression that estimates the values of *a* and *n* parameters for *i* in mm / h, and *t* in minutes. The results of the above process are presented in Table 1.

With this we proceeded to the delimitation of the respective IDF curves and a comparative analysis is performed simultaneously with two background studies on the same water gages or some of them.

Comparative studies analysis are Pereira, Rodrigues, & Costa, 2001 which considers curves of type $i = at^b$ with a MAS based analysis and by applying breakdown coefficients for rainfall events below the day, and Correia (2008) that uses PDS of hourly rainfall estimates and $i = at^b c$ type curves based on the method proposed by NERC (1975) .

Table 1 - Parameters of the IDF curves for intensity of precipitation expressed in mm / h, and duration of rainfall in minutes, obtained with the methodology, based on SDP for the different return periods in water gages chosen for this study.

		Parâmetros das IDF	
Posto	T (anos)	a	n
Abrantes	5	256.684	-0.648
	10	295.382	-0.647
	100	416.645	-0.644
	500	499.948	-0.643
	1000	535.767	-0.643
Monchique	5	285.426	-0.560
	10	325.717	-0.557
	100	452.147	-0.553
	500	539.074	-0.551
	1000	576.461	-0.550
S.Julião do Tojal	5	2712.453	-0.969
	10	3674.546	-0.994
	100	6987.043	-1.042
	500	9414.755	-1.060
	1000	10481.812	-1.066

Figure 4 presents an example of the results from this research for S.Julião Tojal as comparative studies of reference.

Conclusions and recommendations

Observing the comparative analysis that was performed there are some differences between the present study and studies history reference, being the same described below:

- The smallest differences between rainfall intensities estimated by the method proposed in this dissertation and the Correia (2008) occur in Abrantes water gage for any return periods and durations considered. The water gages of Monchique and S.Julião do Tojal for rainfall durations of more than 12 hours is observed in comparison to the same study an approximation of the results, and such is also observed for all return periods in both.

- The smallest differences between rainfall intensities estimated by the proposed method in this thesis and Pereira Rodrigues, and Costa (2001) are observed in Monchique water gage for higher durations 12h for any return period considered. In S.Julião do Tojal was observed considerable similarities to return periods up to 100 years and up to 6 hours duration whereas for periods of 500 and 1000 years is true for durations longer than 12 hours. In bot, the differences increase almost proportionally to increase return period.

In general it appears that the method provides a reasonable agreement with the background being the major discrepancies found in comparison with the method of Pereira et al., 2001. This may be due to factors such as the use of separate periods for analysis, differences in estimation procedures in the case of this dissertation is underlying the use of PDS and some simplifications considered in establishing the coefficients a and b of the IDF curves in the study .

The differences between the obtained estimates and the Correia, 2008 may be justified for reasons such as the number of collected records, the thresholds selection in this study that is something subjective for most studied durations and own methodology separate estimation.

Then it is considered that the study carried out in this context, has an important contribution in the domain of extreme precipitation phenomena analysis, as that allows simple and accurate calculation of maximum precipitation at lower durations per day, for any return period and watershed of the Portuguese mainland.

Established IDF curves resulting from the research carried out, are presented as an alternative to existing and more used in project context, as they are involved in a methodology based on PDS, which can return more reliable results according to the bibliography. Comparison with previous studies based on MAS and PDS concluded that the proposed method of analysis should be considered, although it still lacks some modifications.

This subject is provided with some working complexity and therefore should be updated over time. In this sense it is important to recommend to the future some of the possible investigations and improvements to be made, following which was developed in this work:

- More water gages and an increase of their temporal amplitude.
- Investigate whether a study based on PDS be a minimum amount of entries in both dimensions of the samples formed as the number of years of records, from which it returns shows that the technique more reliable results.
- Improve the criteria used in this study for identifying independent rainfall events and make the PDS.
- Integrate in the computer program a simple way to handle the records and select the thresholds expeditiously.

Finally, it is important to refer to the importance of monitoring hydrologic variables. The PDS allows to use more efficiently the information on small samples and the results are all the more valid the more long, and therefore greater, are registers, and this fundamental sense to continue to monitoring for the future this methodology can be as a better alternative analysis.

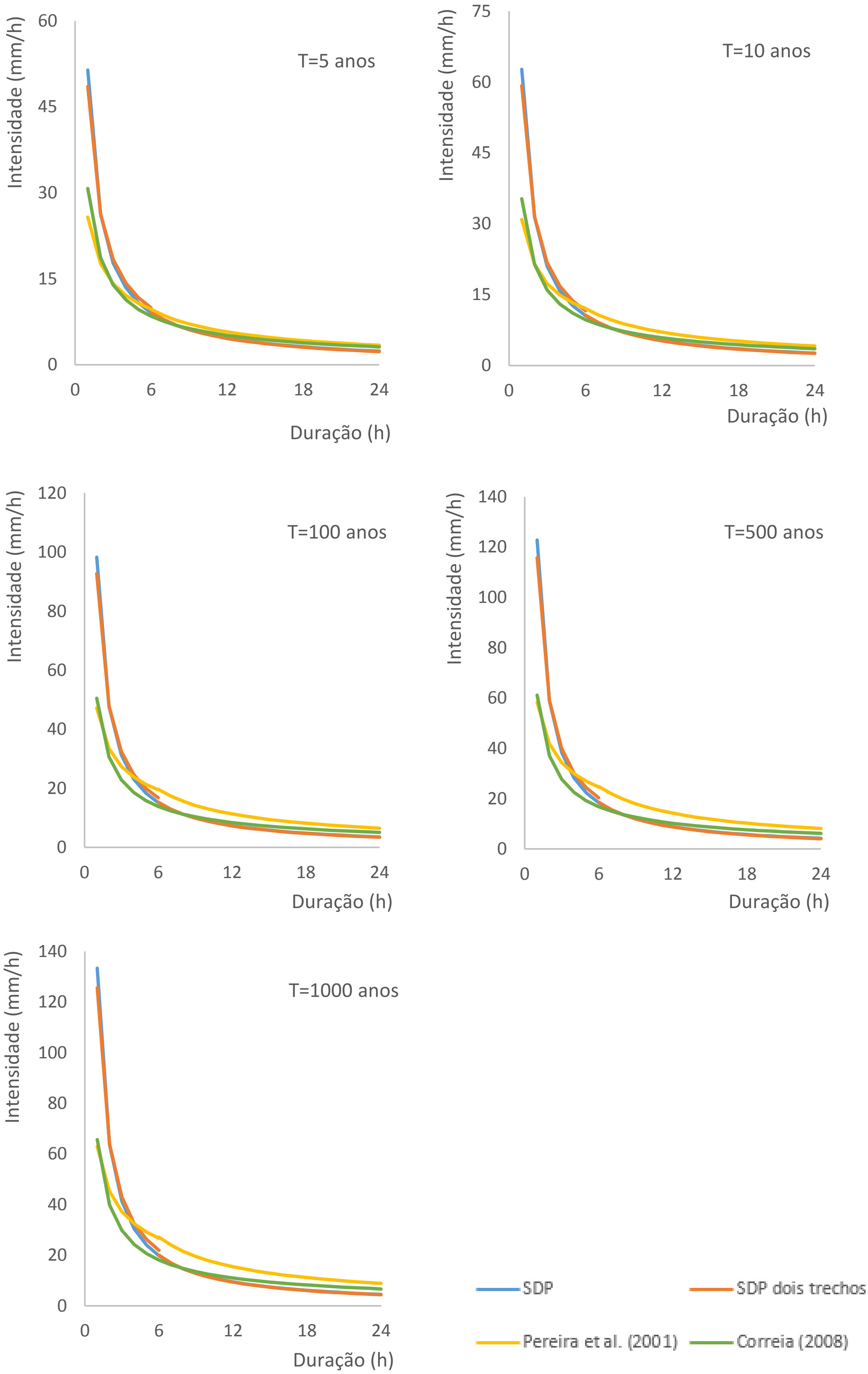

Figure 4- IDF curves for S.Julião Tojal for the different return periods and curves of Pereira et al. (2001) and Correia (2008).

Bibliography

Correia, M. G. P. (2008). *Análise da precipitação máxima com média duração em Portugal Continental*. Instituto Superior Técnico.

Davison, A. C., & Smith, R. L. (1990). Models for exceedances over high thresholds. J R Stat Soc Ser B. *Journal of the Royal Statistical Society. Series B: Methodological*, *52*(3), 393–442.

Lang, M., Ouarda, T. B. M. J., & Bobée, B. (1999). Towards operational guidelines for over-threshold modeling. *Journal of Hydrology*, *225*(3-4), 103–117.

Miquel, J., Chambolle, T., & Bernier, J. (1984). Guide pratique d'estimation des probabilités de crues, 160.

Pereira, C. C. L. B., Rodrigues, R., & Costa, J. (2001). Análise de fenómenos extremos. Precipitações intensas em Portugal Continental. *Lisboa, DSRH*, 64.