TÉCNICO LISBOA

Intensive rainfall. Characterization based on partial duration series.

Extended Abstract

Filipe Cerejo Correia

Dezembro 2013

Intensive rainfall. Characterization based on partial duration series.

Filipe Cerejo Correia

Scope

When designing some hydraulic structures it is common in Portugal to use the Annual Maximum Series technique, AMS, to estimate the appropriate quantile of probability with a given return period. This technique is based on a sample built upon the variable`s highest value occurred in each year. However, when there are not enough years of records, the AMS technique becomes partly or totally useless.

The main purpose of this research is to compare the AMS technique with another one used with the same objective. This other technique is called Partial Duration Series, PDS, and it makes use of all the variable`s independent values which are higher than a certain threshold. Besides being a theoretically supported alternative when there is lack or insufficient data, it makes possible to analyse events that, being extreme and even higher than annual maxima of others years, would not be included in the AMS technique since they are not annual maxima. It also makes possible to disregard events that, despite not being extremes, would be included in the AMS analysis because they are annual maxima.

It turns out that there is, in the state of the art, no universal and unambiguous rule for the selection of the threshold which incorporates as much information as desirable without compromising the independence ‘requirements between occurrences over time. This just so introduce some subjectivity in the PDS approach.

Notwithstanding, the orientations for the application of PDS of LANG, et. Al., 1999, developed in the context of flow series, were followed. In general terms, the average annual exceedances number was controlled, the average exceedance was analysed and the dispersion index had to be validated to allow each constituted sample to a PDS based analysis. Further, the autocorrelation coefficients were calculated to assess the occurrences ´independence within the sample, MIQUEL, 1984.

The study applied both techniques to the daily rainfall series of 11 rain gages and several estimations were made based on both applications. Once the PDS constituted samples are expected to be larger than the AMS ones, it was studied if the PDS technique can be applied to a less amount of data with the same reliability.

In the scope of this research, the Gumbel’s distribution function, recognized has the best intense rainfall phenomena characterizing distribution in the Portuguese mainland (PEREIRA, 1995, pag.132), was applied to the AMS technique, whereas for the PDS technique, the Exponential distribution associated to the Poisson model was adopted to describe the exceedances.

Data Basis

Even if 24 rain gages throughout the continental Portuguese territory were analysed, this research only made use, by many reasons, of 11 rainfall samples in the rain gages of **Tables 1** and **2.** Concerning the disregarded rain gages, some lacked information and the others ones mostly do not meet the set criteria.

Table 1 – Location of the rain gages and periods with records of annual maximum daily rainfalls, P_{dma}.

Rain Gage		Location coordinates			Periods with records	Number of records, N
Name	Code	X (m)	Y (m)	Altitude (m)		
Castro D´Aire	08J/04G	216484.12	435731.34	584	1916/17-2000/01	85
Gestosa	02O/01UG	281524.00	546423.00	706	1932/33-2002/03	71
Penafiel	06H/01UG	186148.00	471437.00	175	1913/14-1996/97	84
Vinhais	02O/02UG	294648.00	540463.00	636	1913/14-2000/01	88
Pragança	18C/01G	119572.95	248285.96	183	1928/29-2001/02	74
Alvaiázere	15G/01UG	178815.91	317665.13	335	1932/33-2005/06	74
Gaivão	17J/01UG	216891.00	277275.00	273	1932/33-2004/05	73
Muge	19E/01UG	149977.00	237501.00	14	1932/33-2002/03	71
Pavia	20I/01G	210363.00	214322.00	189	1930/31-2002/03	73
Pernes	17F/01UG	154325.79	269299.22	81	1914/15-2005/06	92
Serpa	26L/01UG	246521.75	108565.59	209	1931/32-2005/06	75

Table 2 – Statistical characteristics of the annual maximum daily rainfall series, P_{dma}.

Rain Gage		Watershed	Pdma average (mm)	Standard deviation (mm)	Variation coefficient (-)	Skewness coefficient (-)
Name	Code					
Castro D´Aire	08J/04G	rio Douro	90.86	27.88	0.31	1.87
Gestosa	02O/01UG	rio Douro	50.88	13.07	0.26	1.54
Penafiel	06H/01UG	rio Douro	83.04	25.47	0.31	1.09
Vinhais	02O/02UG	rio Douro	60.51	18.63	0.31	0.7
Pragança	18C/01G	Ribeiras do Oeste	57.44	21.43	0.37	2.24
Alvaiázere	15G/01UG	rio Tejo	61.35	14.95	0.24	0.62
Gaivão	17J/01UG	rio Tejo	51.54	17.61	0.34	0.77
Muge	19E/01UG	rio Tejo	39.55	14.95	0.38	1.88
Pavia	20I/01G	rio Tejo	39.60	17.22	0.43	1.15
Pernes	17F/01UG	rio Tejo	52.53	16.24	0.31	1.04
Serpa	26L/01UG	rio Guadiana	44.72	19.07	0.43	1.29

The first table contains the rain gages' identification and location (in terms of the watersheds where they are inserted and its coordinates), as well as the periods with complete hydrological years (and respective lengths, N) and **Table 2** refers to the statistical characteristics of P_{dma} samples: average, standard deviation, variation coefficient (defined by the ratio between the standard deviation and the average) and skewness coefficient.

All records used in the research were obtained through the SNIRH – from the Portuguese, *Sistema Nacional de Informação de Recursos Hídricos* – which were a public data basis of the former national water institute, INAG – from the Portuguese, *Instituto da Água*.

Procedures

After collecting the daily rainfall samples of the 11 rain gages, and knowing that the selected gages contained 70 to 92 of complete hydrological years of record, these periods were divided into 4 periods between 20 and 23 years. This was the minimum sample ´dimension considered compatible with the AMS technique. Subsequently, it was intended to compare both techniques applied in a few data scenario.

Afterwards, the AMS technique was applied to the global and the partial samples and different quantiles of probability, namely the 0.50, 0.75, 0.90, 0.95 and 0.99 quantiles, which correspond to the return periods of 2, 4, 10, 20 and 100 years, respectively, were calculated. To do so, it was used the following expression (CHOW, 1964):

$$x_T = \bar{x} + K_{DIST}^{F}.\sigma^2$$

where x_T is the estimated value, $\bar{x}$ is the sample average, σ^2 is the sample ´standard deviation and K_{DIST}^{F} is Gumbel´s probability factor given by:

$$K_{GUMBEL}^{F} \cong -\frac{\sqrt{6}}{\pi}\left[\delta + \ln\left(ln\frac{T}{T-1}\right)\right]$$

where $\delta = 0.577216$ is Euler-Mascheroni´s constant and T is the chosen return period.

The AMS technique was followed by the PDS one applied to the same quantiles of probability. In what concerns the PDS technique, the selection of the threshold was made by varying its value and assessing each correspondent sample under the following conditions.

According to the PDS technique, the sample is made by extracting all occurrences above a certain threshold which consist in independent values. To guarantee this independence, not only each daily rainfall value must be separated from the next one by one day of zero rainfall, but also the sample´s 1st and 2nd autocorrelation coefficients had to be within the 95%´s confidence interval, MIQUEL, 1984. This means that every selected threshold respects the following expression:

$$\rho_{X,Y,k} = \frac{COV\ (X_i, Y_i)}{\sigma_{X_i} \cdot \sigma_{Y_i}} \in \left[1 - 1.96\frac{\sqrt{N-2}}{N-1}, 1 + 1.96\frac{\sqrt{N-2}}{N-1}\right]$$

where k is the increment value, X_i are the samples' values and $Y_i = X_{i+k}$.

Then, accordingly to LANG, et. al., 1999, the average annual exceedances number λ, must be greater than 3 for the PDS technique to present better results compared to the AMS. Plus, DAVIDSON, SMITH, 1990, recommend plotting the average exceedance (mean residual plot). According to this author, the selected threshold must be located in a nearly linear portion of the graph thus obtained (increasing, decreasing or constant).

Concerning all presented factors, a range of thresholds which are within the terms is obtained. Ultimately, the Dispersion Index, DI, is a measure used to analyse if the selected thresholds obey to the conditions of a certain standard statistical model, in this case, the Poisson model, CUNNANE, 1979. The DI expression is as follows:

$$ID = \frac{\sigma^2}{\mu}$$

where μ is the average of the sample.

If DI is significantly close to 1 (with a significance level of 5%), the sample can be modelled by a Poisson process and so the correspondent threshold is not rejected.

Having the threshold´s range selected and verified by the DI test, the annual maxima rainfall estimations were made by resorting the expression:

$$x = u + \sigma\{\ln(\lambda) - \ln(-\ln[F(x)])\}$$

where $F(x) = 1 - \frac{1}{T_p}$. And T_p is the return period in the PDS context calculated by:

$$T_p = \frac{1}{\lambda[1 - H_u(x)]}$$

where $H_u(x)$ is the probability function of the Partial Duration Series, in this work, the Exponential distribution.

Results

The results achieved through the application of both techniques to each period of record (and sub-period) collected from the rain gages are synthesized in **Table 3**. For each rain gage, this table indicates the limiting values of the threshold ‘range, u, and the average annual exceedances number, λ, which are according to the terms, as well as the PDS estimations corresponding to the median, maximum and minimum threshold and the AMS estimations, for a return period of 10 and 100 years.

Figures 1 to **3** exemplify the results summarized in Table 3 for one sub-period of record. For that purpose the first sub-period of Pernes´rain gage (rio Tejo watershed) was selected.

Table 3 – AMS and PDS applied to the rain gages, for each period and sub-period and for the 0.90 and 0.99 probability quantiles (respectively, the return period of 10 and 100 years). Limit values of the threshold and the average annual exceedances number values corresponding to the threshold limits. Median, minimum and maximum value of daily rainfall estimation among all the estimations made within the limits.

Rain Gage	Period/ Dimension	Limit Values		Estimations							
				T = 10 years				T = 100 years			
				PDS			AMS	PDS			AMS
		u	λ	Median	Minimum	Maximum		Median	Minimum	Maximum	
		(mm)	(mm)	(mm)	(mm)	(mm)	(mm)	(mm)	(mm)	(mm)	(mm)
Castro Daire	1/22	27.68 - 43.38	9.82 - 5.32	136.12	132.50	137.20	132.18	192.28	185.90	194.22	188.79
	2/22	21.59 - 34.24	10.55 - 6.32	133.86	132.27	135.41	125.38	191.21	188.40	193.83	170.71
	3/22	15.01 - 32.41	15.18 - 7.59	132.40	128.46	134.34	117.15	187.97	181.23	191.20	156.25
	4/22	21.08 - 40.83	11.91 - 5.00	125.89	122.79	128.44	130.93	177.99	172.42	182.41	191.24
	Global/85	40.56 - 50.28	5.36 - 3.69	127.86	125.32	129.56	127.23	179.66	174.89	182.76	178.31
Gestosa	1/20	20.12 - 30.50	8.90 - 3.75	72.65	70.51	74.84	72.60	100.12	96.43	103.83	102.30
	2/20	3.62 - 13.48	26.10 - 12.55	75.38	73.53	77.11	62.58	105.92	102.93	108.58	79.36
	3/20	7.26 - 18.71	18.80 - 7.95	74.56	73.24	75.87	64.28	104.90	102.73	106.97	82.55
	4/20	8.10 - 16.82	19.25 - 9.70	72.59	71.31	74.00	71.27	101.55	99.51	103.73	99.55
	Global/71	20.60 - 31.10	7.65 - 3.10	71.56	70.96	72.64	67.92	99.38	98.33	101.49	91.86
Penafiel	1/22	23.37 - 33.71	11.68 - 7.09	122.54	121.01	125.39	114.76	172.12	169.59	177.05	154.94
	2/22	18.95 - 31.51	12.59 - 7.05	116.61	114.56	119.53	119.37	164.55	161.00	169.40	170.97
	3/22	18.74 - 31.17	13.32 - 6.86	105.75	102.32	107.70	110.80	147.73	142.13	150.90	156.45
	4/22	18.37 - 35.33	13.59 - 5.82	116.04	114.23	118.68	117.96	163.12	160.14	167.69	164.04
	Globa/84	24.35 - 42.54	10.40 - 4.20	115.41	114.85	116.61	116.29	161.94	161.03	164.06	162.17
Vinhais	1/22	20.78 - 34.50	9.45 - 4.41	94.96	92.11	98.05	93.27	133.58	128.36	138.86	133.22
	2/22	31.79 - 35.61	4.36 - 3.55	91.74	91.15	93.28	90.81	129.24	128.12	132.16	127.45
	3/22	20.26 - 32.60	8.00 - 3.27	81.40	80.02	83.51	79.59	114.66	112.23	118.57	112.10
	4/22	17.17 - 20.42	10.91 - 9.09	82.32	80.88	83.38	71.13	115.18	112.75	116.91	93.19
	Global/88	40.16 - 44.61	2.39 - 1.92	-	-	-	84.48	-	-	-	118.38
Serpa	1/20	10.45 - 16.48	10.70 - 6.05	63.79	62.28	65.50	80.21	90.92	88.34	93.94	131.03
	2/20	4.29 - 15.83	21.15 - 6.70	57.39	56.16	58.59	55.66	80.89	78.81	82.84	78.33
	3/20	12.26 - 16.69	9.05 - 5.75	56.19	55.41	57.18	61.19	79.33	77.97	81.02	89.10
	4/20	8.23 - 16.35	12.45 - 6.70	72.32	70.26	74.64	90.83	103.62	100.14	107.63	137.66
	Global/75	12.06 - 17.19	9.43 - 5.83	62.78	61.70	63.84	75.07	89.29	87.46	91.17	116.10
Pragança	1/20	22.97 - 25.81	6.45 - 5.40	69.80	68.25	70.70	72.84	96.36	93.58	97.97	103.89
	2/20	7.87 - 14.60	22.70 - 14.60	90.92	89.65	92.03	85.74	127.28	125.26	129.01	118.38
	3/20	18.81 - 30.59	8.90 - 3.80	93.70	87.74	97.44	100.47	134.53	124.25	141.48	152.57
	4/20	16.73 - 27.31	9.40 - 4.55	76.15	74.46	78.40	70.51	106.86	103.79	110.67	96.54
	Global/74	30.09 - 34.23	3.95 - 3.11	85.44	84.65	86.17	84.96	121.14	119.60	122.54	123.95
Alvaiázere	1/20	12.76 - 23.53	16.00 - 8.50	93.63	90.21	94.36	82.24	131.37	125.77	132.53	107.29
	2/20	18.23 - 39.28	10.40 - 3.35	99.81	97.82	101.96	88.47	141.24	137.88	144.79	118.95
	3/20	9.36 - 19.45	18.45 - 10.75	85.80	82.14	90.46	77.26	119.96	113.99	127.36	104.24
	4/20	20.23 - 30.29	9.05 - 4.40	81.09	79.60	82.32	75.68	113.07	110.45	115.11	104.14
	Global/74	16.79 - 25.70	12.07 - 6.64	88.57	87.36	89.61	80.95	124.07	122.02	125.79	109.55
Gaivão	1/20	6.86 - 17.35	19.60 - 8.55	72.26	70.94	73.78	66.56	101.70	99.68	104.17	90.67
	2/20	9.51 - 19.78	15.60 - 7.80	83.90	82.67	84.58	88.31	118.95	116.94	120.07	124.92
	3/20	7.51 - 17.89	16.55 - 7.65	71.72	70.14	73.84	75.17	101.33	98.81	104.88	107.94
	4/20	12.64 - 15.28	10.85 - 9.00	67.94	67.66	69.14	65.03	95.81	95.33	97.78	94.85
	Global/73	9.91 - 21.69	14.66 - 5.95	73.50	72.80	74.26	74.10	103.77	102.51	105.01	106.41
Muge	1/20	11.51 - 21.88	10.85 - 4.20	49.86	45.05	55.03	50.45	68.43	59.82	77.17	71.76
	2/20	7.19 - 19.63	16.65 - 5.30	59.12	56.45	60.38	57.50	83.14	78.49	85.17	83.23
	3/20	20.89 - 24.17	4.15 - 3.15	55.63	54.88	56.84	63.51	77.56	76.14	79.84	95.38
	4/20	15.49 - 18.91	7.15 - 5.00	60.56	59.92	61.66	69.91	85.78	84.67	87.76	106.87
	Global/71	17.35 - 23.09	6.37 - 3.39	54.92	54.10	55.66	59.06	76.42	74.93	77.80	86.45
Pavia	1/20	7.41 - 15.20	13.10 - 5.55	50.93	49.91	52.23	50.58	72.01	70.35	74.22	74.42
	2/20	13.27 - 20.80	8.65 - 4.50	63.08	62.04	64.11	64.94	89.54	87.69	91.36	94.38
	3/20	9.40 - 21.36	12.85 - 4.90	67.45	66.18	68.74	75.32	95.85	93.52	98.05	112.62
	4/20	12.20 - 23.20	9.55 - 3.50	58.69	57.67	59.75	68.04	82.72	80.78	84.55	102.50
	Global/73	5.07 - 15.01	18.10 - 7.10	59.63	59.23	59.93	63.28	84.56	83.90	85.05	93.71
Pernes	1/23	7.18 - 19.78	18.26 - 7.65	80.73	78.98	82.09	78.05	114.17	111.31	116.33	109.83
	2/23	11.36 - 17.74	13.74 - 8.70	74.72	73.66	75.72	65.07	105.23	103.44	106.84	88.32
	3/23	5.21 - 12.51	19.13 - 11.48	76.91	75.69	77.70	74.81	109.30	107.31	110.55	103.72
	4/23	14.17 - 22.21	10.61 - 5.48	66.87	65.50	68.65	65.52	93.49	91.11	96.48	92.85
	Global/92	10.61 - 22.08	13.83 - 6.07	74.61	73.70	76.36	71.31	105.21	103.62	108.04	99.75

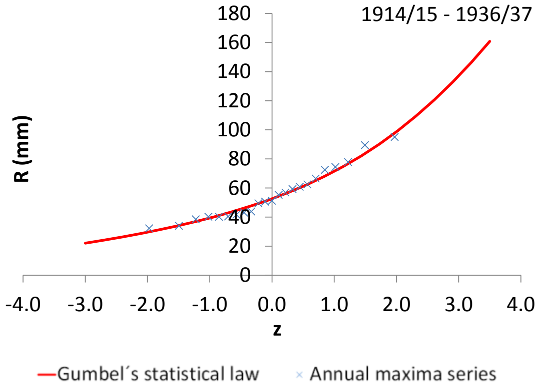

Figure 1 – Pernes rain gage. Sub-period from 1914/15 to 1936/37. AMS technique application.

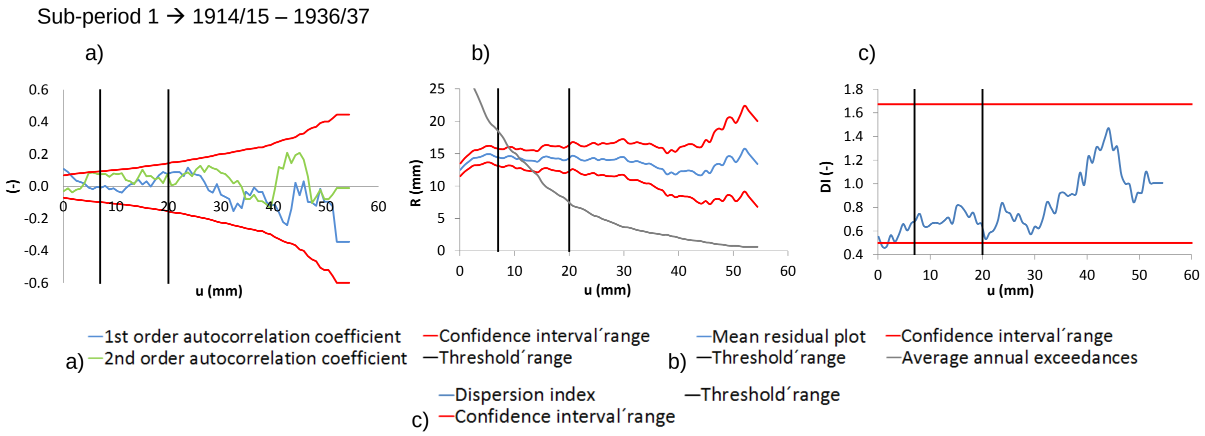

Figure 2 – Pernes rain gage. Sub-period from 1914/15 to 1936/37. Threshold ‘range criteria selection. As a threshold function: a) 1st and 2nd autocorrelation coefficients; b) mean residual plot and average annual exceedance number; c) dispersion index.

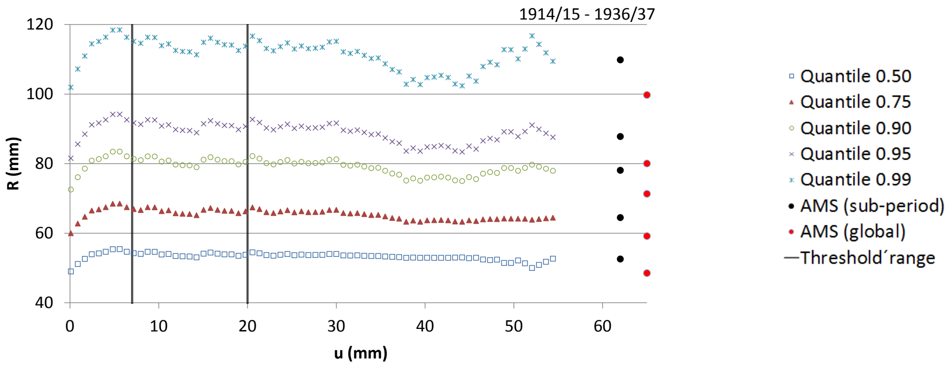

Figure 3 – Pernes rain gage. Sub-period from 1914/15 to 1936/37. Visual comparison between PDS estimations for varying thresholds and AMS estimations for the sub-period and the global period. Threshold ‘range selected.

Conclusions and recommendations

The research carried out showed that, based on the samples of extreme rainfall adopted as case studies it was not possible to clearly identify one of the estimations techniques as being the best one. However, due to the larger PDS `samples compared to the AMS ones when based on the same amount of data, one may confer a higher reliability to this technique compared to AMS, since its analysis is based on a larger amount of data.

Additionally, though the major size of the PDS constituted samples, the results of the application of both techniques are closer to each other when they are applied to the same period of record than when the AMS technique is applied to a larger period of record. **Figures 4** and **5** demonstrates this by dividing the median estimation based on PDS analysis, concerning each sub-period and the global period of record, with the median AMS estimation, for the global period of record, for Figure 4 and for each sub-period, Figure 5.

Finally, a closer look to Table 4 reveals that, in the case of limited amount of data available, the use of the PDS technique tends to result in higher estimations than the AMS technique.

Regarding future research on the area, some recommendations are made, namely for PDS and AMS comparison.

Firstly, it would be appropriate to compare both techniques, this time associated to different distribution functions, like the Normal distribution, Pearson and Pearson III to AMS technique and the generalized Pareto distribution to the PDS technique. One pretends not only to analyse the influence of the distribution `choice to the results but also to understand which distribution is more appropriate.

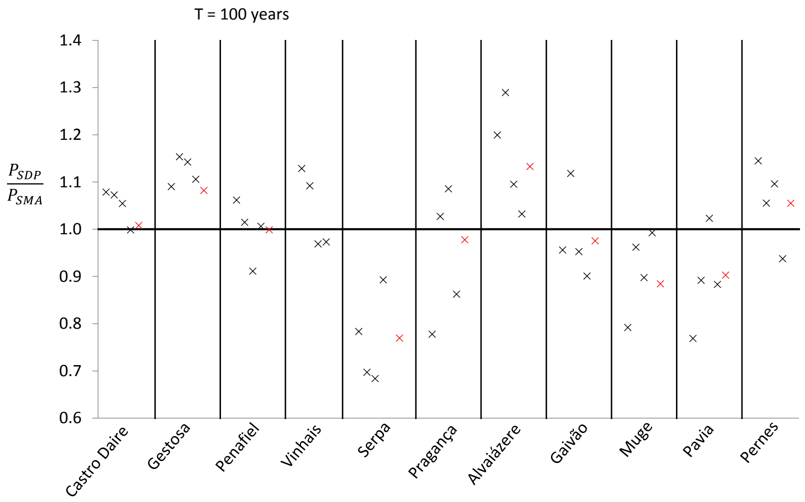

Figure 4 – Ratio between the median estimations of maximum annual rainfall produced by PDS and by the AMS for the return period of 100 years. In black: the ratios between the median estimations of PDS for the different sub-periods and median AMS estimations for the global period; in red: the ratios between the median estimations of PDS and AMS, both referring to the global period.

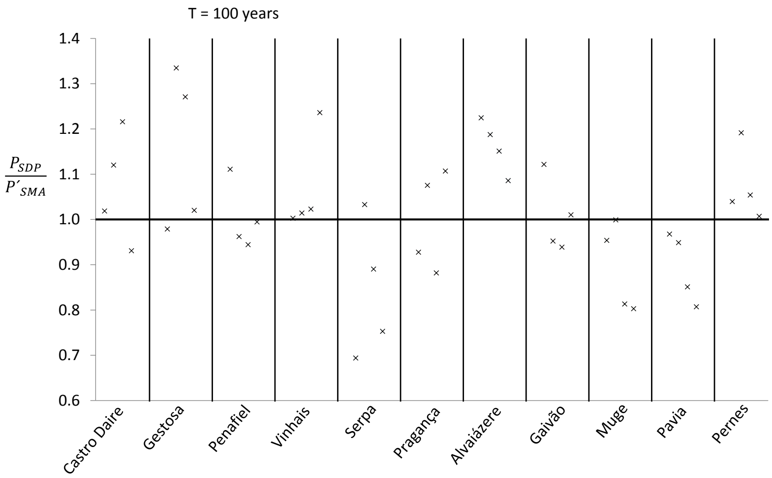

Figure 5 - Ratio between the median estimations of maximum annual rainfall produced by PDS and by the AMS for the corresponding sub-period, for the return period of 100 years.

Plus, it would be relevant to study the minimum amount of data, either the number of years of record or the samples size based on these, from which the PDS technique presents consistent .results

Table 4 – Comparison between the estimations provided by the PDS technique applied to the different sub-periods and the estimations of AMS to the global period. The greater estimation is highlighted.

Rain gage	Period	T = 10 years		T = 100 years	
		Median PDS	AMS	Median PDS	AMS
Castro Daire	1				
	2				
	3				
	4				
	Global				
Gestosa	1				
	2				
	3				
	4				
	Global				
Penafiel	1				
	2				
	3				
	4				
	Global				
Vinhais	1				
	2				
	3				
	4				
	Global	-	-	-	-
Serpa	1				
	2				
	3				
	4				
	Global				
Pragança	1				
	2				
	3				
	4				
	Global				

Rain gage	Period	T = 10 years		T = 100 years	
		Median PDS	AMS	Median PDS	AMS
Alvaiázere	1				
	2				
	3				
	4				
	Global				
Gaivão	1				
	2				
	3				
	4				
	Global				
Muge	1				
	2				
	3				
	4				
	Global				
Pavia	1				
	2				
	3				
	4				
	Global				
Pernes	1				
	2				
	3				
	4				
	Global				
Σ	Sub-periods	27	17	24	20
	Global	5	5	4	6

.

Bibliography

CHOW, V. T., 1964. Section 8-I, *Statistical and Probability Analysis of Hydrologic Data, Part I – Frequency Analysis.* In: Handbook of Applied Hydrology. McGraw-Hill, USA.

CUNNANE, C., 1979. *´A note on the Poisson assumption in partial duration series models´.* Water Resources Bulletin, 15 (2), pp. 489-494.

DAVIDSON, A. C., SMITH, R. L., 1990. *´Models for exceedances over high thresholds´.* Journal of the Royal Statistical Society B, 53 (3), pp. 393-442.

LANG, M., OUARDA, T. B. M. J., BOBBÉ, B., 1999. *´Towards operation guidelines for over-thresholds analysis´.* Journal of Hydrology, 225, pp. 103-177.

MIQUEL, J., 1984. *Guide pratique d´estimates des probabilités de crue.* Eyrolles, Paris.

PEREIRA, C. B., 1995, *Análise de precipitações intensas.* Dissertação para obtenção do grau de mestre em Hidráulica e Recursos Hídricos. Universidade Técnica de Lisboa, IST. Lisboa.