Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Hydrology 324 (2006) 94–114
Journal of Hydrology
www.elsevier.com/locate/jhydrol

Statistical distributions and sampling strategies for the analysis of extreme dry spells in Catalonia (NE Spain)

X. Lana [a,*], A. Burgueño [b], M.D. Martínez [c], C. Serra [a]

[a] Departament de Física i Enginyeria Nuclear (ETSEIB), Universitat Politècnica de Catalunya, Av. Diagonal 647, 08028 Barcelona, Spain
[b] Departament d'Astronomia i Meteorologia, Facultat de Física, Universitat de Barcelona, C. Martí i Franquès 1, 08028 Barcelona, Spain
[c] Departament de Física Aplicada, Universitat Politècnica de Catalunya. Av. Diagonal 649, 08028 Barcelona, Spain

Received 15 October 2004; revised 29 June 2005; accepted 10 September 2005

Abstract

Statistical distributions of annual extreme and long dry spells, using daily rain amount thresholds of 0.1, 1.0 and 5.0 mm/day, are analysed from a database of 39 rain gauges in Catalonia (NE Spain), with a daily recording period extending from 1950 to 2000 and with little missing data. The generalised extreme value (GEV) and generalised Pareto (GP) distributions are considered to model the series of annual extreme (AE) dry spells. The same distributions are also assumed for the partial duration (PD) series, which are derived from dry spells exceeding the 95% percentile. In both cases, the three parameters of the GEV and GP distributions are fitted by means of the L-moments method, which offers a robust estimation of them. Additionally, the fit between empirical data and theoretical distributions is evaluated in terms of the empirical and theoretical skewness and kurtosis. Even though AE spells are commonly well fitted by the GEV model, the GP model is a better option for some rain gauges. Similarly, in the case of the PD series, a few cases are better modelled by the GEV distribution. This study establishes the basis for climatic drought risk assessment in Catalonia, given that dry spell lengths associated with return periods of 2, 5, 10 and 25 years are accurately revised by comparing results deduced from the GEV and GP models, both for the AE spells and PD series.

Keywords: Extreme dry spells; Partial duration series; Generalised Pareto and extreme value distributions; L-moments; Drought risk; Catalonia (NE Spain)

1. Introduction

According to Kiktev et al. (2003), for latitudes above 30°N and the period 1950–1995, most of Europe and East Asia are characterised by a positive trend in the annual maximum of consecutive days with daily precipitation below 1 mm. This general behaviour becomes of larger relevancy for the Mediterranean area, where an increase during 1951–1995 of extreme daily rainfall has been observed in spite of a decrease in total amounts (Alpert et al., 2002). Thus, in the Mediterranean area, the prospect is for a larger frequency of drought periods, with associated impacts on agriculture, on the profit from water resources and also on socio-economic activities (AMS statement, 2004).

* Corresponding author. Tel.: +34 934010825.
E-mail address: francisco.javier.lana@upc.es (X. Lana).

0022-1694/$ - see front matter
doi:10.1016/j.jhydrol.2005.09.013

With the prospect of a near future with a modified distribution of long dry spells, a complete analysis of the distributions of annual extreme (AE) dry spells and partial duration (PD) series of dry spells using several selected thresholds of daily rainfall is attempted from daily observations in Catalonia during the second half of the 20 century. The analysis and validity of both strategies to obtain reliable long dry spell statistics are the main objective of this research. Previous studies on spatial patterns of dry spells in the same region have characterised AE spells by means of the Jenkinson formulation (Jenkinson, 1969, 1975), the probability of transition from wet to dry spells by means of Markov chains, as well as the Poisson distribution of repeated long dry spells (Lana and Burgueño, 1998a–c). At a different time scale, with monthly instead of daily data, Lana et al. (2001) also analysed patterns of monthly shortage and excess of rain amounts by applying the standardised precipitation index.

AE dry spells are modelled both by means of the generalised extreme value (GEV) and the generalised Pareto (GP) distributions. The selection of just one extreme value for every year could have several shortcomings. Firstly, long spells, close to the extreme value in a year but, not the most extreme, would not be considered. Nevertheless, its inclusion in the series of annual extreme values could improve the estimation of the distribution parameters and modify results concerning return periods. Secondly, it would not be very logical that in wet years, an extreme dry spell, remarkably shorter than that of another year, was considered according to the AE strategy. In short, the estimation of return periods of AE dry spells could be subject to non-negligible errors and biases due to the sampling strategy.

The alternative way is based on the consideration of spell lengths exceeding a threshold level. This PD series will also be analysed by GEV and GP models. Examples of the implications of the AE and PD sampling strategies in the modelling of extreme hydrological events can be found for instance in Madsen et al. (1997a,b). In our case, spells for different return periods, derived from AE spells, are compared with those obtained from PD series and the parameters of the GEV and GP distributions are always estimated by means of L-moments (Hosking and Wallis, 1997). Two recent applications of L-moments can illustrate some advantages of this method. One of them is devoted to the analysis of the regional frequency of annual extreme rainfall in the United Kingdom (Fowler and Kilsby, 2003). The other is centred on the accurate evaluation of extreme dry spell risk in Aragón, Spain (Vicente-Serrano and Beguería-Portugués, 2003). The use of L-moments has very interesting features. The reliability of the fitted parameters is improved with respect to other standard estimation methods, such as those based on probability paper representation or on moments of the samples (Benjamin and Cornell, 1970). L-skewness and L-kurtosis can also be easily compared with the corresponding empirical values, because these theoretical L-moments depend exclusively on just one parameter of the GEV and GP models. Consequently, the best distribution for each rain gauge will be that leading to the lowest discrepancy between empirical and theoretical L-skewness and L-kurtosis. The AE and PD strategies are then compared in terms of return period results. Similar spatial features and discrepancies are analysed when dry spell lengths for return periods of 2, 5, 10, 25 and 50 years are considered.

The contents of the paper are structured as follows. Section 2 deals with the geographical and orographic characteristics that condition the rainfall regime of Catalonia, together with a brief specification of the dataset and sampling strategies and a description of the main climatological features of the precipitation in this region. The statistical models applied to the analysis of very long dry spells are stated in Section 3. Section 4 analyses the AE dry spells and PD series, applying a criterion for selecting the optimum probability distribution, and the resulting return period maps. Section 5 discusses the results by comparing them with previous research for the same region, but employing different methodologies.

2. Study area, database and regional climate

2.1. Study area

The complex orography of Catalonia, which could be a relevant factor to explain the spatial distribution of the parameters of the GEV and GP models, is schematically depicted in Fig. 1. It is worth mentioning the Pyrenees and Pre-Pyrenees Ranges,

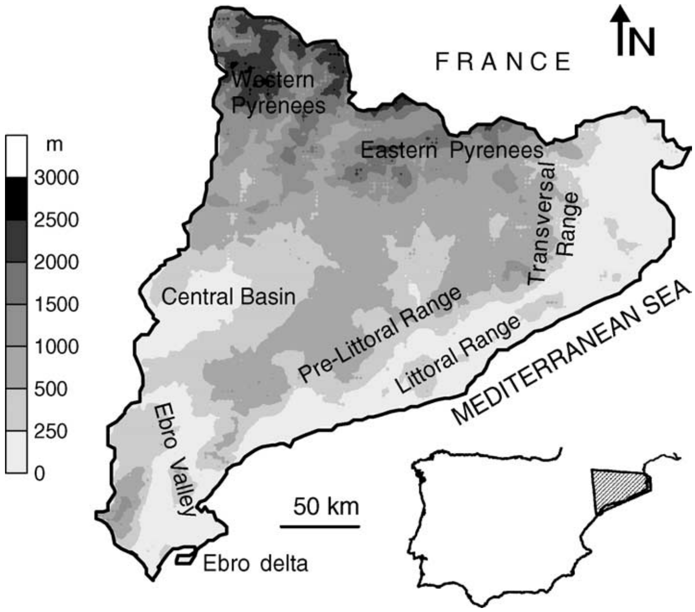

Fig. 1. Main orographic patterns of Catalonia and its location in the north-eastern Iberian Peninsula.

which mitigate the effects of northern advections for the rest of the country, the Littoral and Pre-Littoral Ranges, sheltering the Central Basin against eastern advections but, at the same time, enhancing their effects on the Littoral and Pre-Littoral domains, and the Central Basin itself, relatively far away from eastern advections and subject to the effects of frontal passages and western advections, mitigated by the remoteness to the Atlantic coast. Within the Pyrenees Range, distinctions have to be made between the Eastern Pyrenees regime, strongly influenced by the vicinity to the Mediterranean Sea, and the north face of the Western Pyrenees, with clear signs of Atlantic influence.

2.2. Database

The database is quite similar to that used in previous analyses of the daily pluviometric regime of Catalonia. Data have been obtained from 39 rain gauges selected, for reason of their recording continuity, from a previous set of 75 (Lana et al., 2004) belonging to the *Instituto Nacional de Meteorología* (Spanish Ministry of Environment), with a common recording period from 1950 to 2000 (Fig. 2). Missing data in some gauges can be up to 10 years, not usually distributed in very short periods, but in continuous periods close to 1 year. Consequently, a dry spell interrupted by a missing period is discarded for any kind of statistical analysis and the next dry spell begins after the last recorded rainy day, once the missing period is over. Alternatively, missing daily data could have been estimated by considering statistical methods and data corresponding to neighbouring gauges. Nevertheless, the strong spatial variability of the daily pluviometric regime of Catalonia (Lana et al., 2004) advises against this substitution of missing data. Even though the Western Pyrenees Range is poorly covered, the recording continuity of the remaining gauges is good. The results obtained from the GEV and GP models applied to the AE spells should not be affected by this lack of data. Due to the definition of the PD series, the influence of this lack of data is even smaller in this case.

2.3. Regional climate

The pluviometric regime of Catalonia is governed both by the synoptic circulation and by its complex orography (Serra et al., 1998; Sotillo et al., 2003; Lana et al., 2004). These facts, accompanied by the time

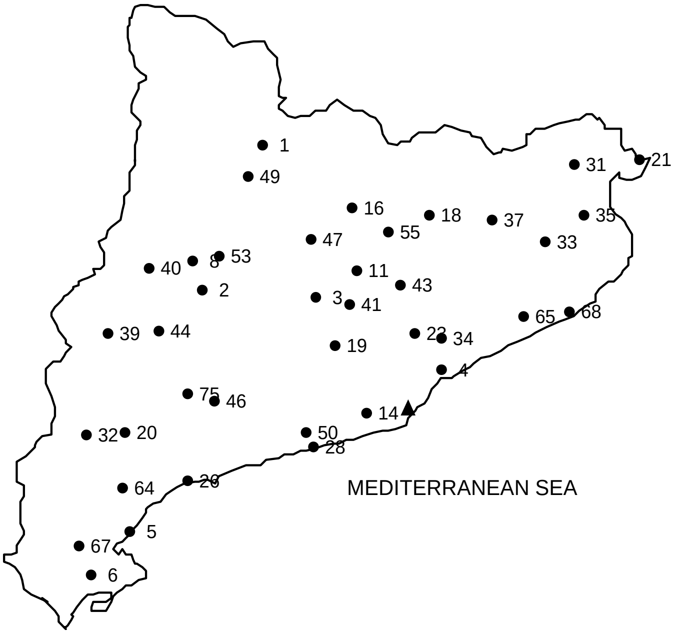

Fig. 2. Location of rain gauges used in the analysis of extreme dry spells. Fabra Observatory is indicated by a solid triangle.

irregularity of the Mediterranean climate (Barry and Chorley, 2003), generate a set of spatial and temporal patterns that makes an analysis of the drought phenomena relatively difficult, especially if the attention is focussed on the long dry spell lengths derived from daily data.

From the point of view of the synoptic circulation patterns that cause long dry spells, mention has to be made of the anticyclonic persistence in the mild and cold seasons, which usually affects the whole territory. The case for the hot season is quite different by reason of the convective phenomena, favoured by the orography and the income of humid air from the Mediterranean together with cold air masses at high levels in the atmosphere. Then, typical summer droughts can be interrupted in mountain ranges such as the Pyrenees and Pre-Pyrenees by some daily episodes of outstanding rainfall amount.

Other relevant synoptic circulations are the eastern and southern advections, which transport humid air masses from the Mediterranean and may contribute to rainfall amounts, sometimes with torrential episodes in spring and especially in autumn (Llasat and Puigcerver, 1994; Llasat et al., 1996; Romero et al., 1999). By the one hand, a lack of this circulation pattern in the mentioned seasons could be the cause of long and severe dry spells that would affect especially the Eastern Pyrenees and the Littoral and Pre-Littoral area. On the other hand, if these advections are persistent but of moderate effect, the western domains of Catalonia may also be subject to rainfall shortage, due to the sheltering effects of the Littoral and Pre-Littoral Ranges.

The effects on the rainfall amounts of front passages coming from the Atlantic and with dominant western component are usually small, due to the remoteness of the Atlantic coast and the Föhn effect generated by the set of orographic barriers to be crossed by these fronts on their way across the Iberian Peninsula. Consequently, a persistence of westerlies may cause a generalised and long dry spell in Catalonia.

The north-western fronts arriving from the Atlantic are relevant rainfall mechanisms for the north face of the Western Pyrenees, but they have only a minor contribution to the rain amounts for the rest of the territory, Consequently, a period of consecutive north-western front passages favours normal to high rainfall in the Western Pyrenees, but can contribute to a long dry spell for other parts of the country. Only nuclei of low-pressure systems crossing Catalonia, instead of fronts coming from the north-west, can produce relevant rain amounts in a large extent of the territory. Sometimes, these low-pressure systems are reactivated when they arrive to the Mediterranean and generate eastern advections. Consequently, a low-pressure system crossing Catalonia habitually ends a dry spell. Finally, northern outbreaks in the cold season do not contribute to significant rain episodes, except in the Pyrenees Range. These air masses are cold and dry due to their track along the European continent. However, as the Pyrenees is the main orographic barrier, very often the precipitation concentrates on the north face of these mountains due to the Föhn effect.

3. Statistical models

Although it is quite common to use the Gumbel (extreme value Type I) distribution for the analysis of extremes, a better option could sometimes be the Type II model, strictly restricted to positive values. A more advisable option is the GEV distribution, which includes Types I and II as special cases:

$$y = -\frac{1}{\kappa}\ln\{1-\kappa(x-\xi)/\alpha\}, \quad \kappa \neq 0 \tag{1a}$$

$$y = (x-\xi)/\alpha, \quad \kappa = 0 \tag{1b}$$

$$F(y) = \exp\{-\exp(-y)\} \tag{2}$$

where x represents the extremes to be analysed. Scale α, shape κ and location ξ are the three parameters of the distribution and $F(y)$ is the cumulative distribution function. The case of κ equal to 0 corresponds to the Gumbel distribution. Nevertheless, the substitution of the Gumbel model for a more general formulation does not remove a set of shortcomings, mentioned before in Section 1, and related to the sampling strategy.

A solution to these shortcomings could be the modification of the sampling strategy. Instead of generating series of extreme values by choosing the largest value for each sampling period, a PD series can be obtained by taking into account values exceeding a high enough percentile of the empirical data, which will be designed as a truncation percentile. In this way, instead of an extreme for every year, a series of high values are generated in order to achieve a more accurate statistical description. As a counterpart, the concept of return period has to be slightly modified.

In the case of the PD series, the alternative to the GEV distribution could be the GP distribution:

$$y = -\frac{1}{\kappa}\ln\{1-\kappa(x-\xi)/\alpha\}, \quad \kappa \neq 0 \tag{3a}$$

$$y = (x-\xi)/\alpha, \quad \kappa = 0 \tag{3b}$$

$$F(y) = 1-\exp\{-y\} \tag{4}$$

Now x represents the values exceeding a selected percentile and the three parameters have the same meaning as those of the GEV model.

A relevant question is the best strategy, either AE or PD series, to analyse the return periods. This decision is closely related to the estimation of the distribution parameters in terms of the L-moments method. According to Hosking and Wallis (1997); Hosking et al. (1985), the parameters of the GEV distribution can be determined as:

$$\kappa \approx 7.8590c + 2.9554c^2 \tag{5}$$

with

$$c = \frac{2}{3+\tau_3} - \ln(2)/\ln(3) \tag{6}$$

and

$$\alpha = \lambda_2\kappa/\{(1-2^{-\kappa})\Gamma(1+\kappa)\} \tag{7}$$

$$\xi = \lambda_1 - \alpha\{1-\Gamma(1+\kappa)\}/\kappa \tag{8}$$

τ_3 is the L-skewness, λ_1 the L-location, λ_2 the L-scale and $\Gamma(\kappa)$ the gamma function. In the case of the GP distribution, the three parameters are given by:

$$\kappa = (1-3\tau_3)/(1+\tau_3) \tag{9}$$

$$\alpha = (1+\kappa)(2+\kappa)\lambda_2 \tag{10}$$

$$\xi = \lambda_1 - (2+\kappa)\lambda_2 \tag{11}$$

An interesting feature of the L-moments formulation is the dependence of the L-skewness, τ_3, and L-kurtosis, τ_4, only on parameter κ (Hosking and Wallis, 1997). In this way, a measure of the fit between empirical data and a theoretical distribution function is given by the Euclidean distance:

$$D = [\{\tau_3(\kappa) - \tau_3\}^2 + \{\tau_4(\kappa) - \tau_4\}^2]^{1/2} \quad (12)$$

with $\tau_3(\kappa)$ and $\tau_4(\kappa)$ the L-skewness and L-kurtosis for the GEV or GP model, κ the shape parameter deduced from equation (5) or equation (9) and τ_3 and τ_4 the skewness and kurtosis derived from the empirical data. Obviously, the lowest value of D will indicate the best distribution function.

After the determination of the best distribution function and the corresponding parameters, the computation of the value x_T linked to a return period T_r is straightforward by remembering that:

$$T_r = 1/\{1 - F(x_T)\} \quad (13)$$

In the case of the GEV distribution, this will be quantified by:

$$x_T = \frac{\alpha}{\kappa}\{1 - g(T_r)\} + \xi \quad (14)$$

with

$$g(T_r) = \{-\ln[1 - 1/T_r]\}^{\kappa} \quad (15)$$

and in the case of GP by:

$$x_T = \frac{\alpha}{\kappa}\{1 - T_r^{-\kappa}\} + \xi \quad (16)$$

Eqs. (14) and (16) are accurate for the AE strategy, where one extreme value is considered for every year. For the PD sampling strategy T_r has to be substituted by βT_r, with β being defined as the average number of values per year. This is, β is equal to 1 for the AE strategy.

4. Results

4.1. Rain amount thresholds and truncation spell length for the PD series

The dry spells have been computed with reference to three daily rainfall thresholds: 0.1, 1.0 and 5.0 mm/day (Hounam et al., 1975). The length of a dry spell is identified as the number of consecutive days with daily amounts below the assumed threshold. The selection of these common daily rainfall thresholds for all rain gauges is due to the fact that they are very close to the 0, 25 and 50% percentiles of daily rain amounts for the Fabra Observatory (Fig. 2), whose rainfall data have been deeply analysed (Lana et al., 2003). Additionally, while the two first thresholds are usually employed to analyse wet-dry periods, the last threshold is related to rainy episodes which could balance rain deficits generated by drought episodes.

The selection of the best truncation spell length to generate an optimum PD series is not a well solved problem. The truncation value should be high enough to ensure the independence of observations, but not too high, in order to avoid the loss of information and the increase of uncertainty. The simplest method, such as Beguería (2005) discusses for daily rainfall series, is to choose the truncation spell length according to a fixed frequency, for example, a length repeated 1.2–5 times a year. An alternative method to choose the optimum truncation level makes use of the mean excess plot, which consists of representing the mean excess over a threshold against the value of the threshold. The mean excesses should have a linear evolution with the truncation threshold up to the optimum truncation level. According to this strategy, Fig. 3 shows an example of the mean excess plot of gauge number 1 for every daily rainfall amount threshold. A noticeable feature is the almost perfect linear behaviour up to the truncation threshold, which remains above the 90th percentile. Due to the discrete character of dry spell lengths, it is very common that they repeat many times, especially the shortest ones. Thus, to avoid distortion of the linear character of the mean excess plots, just the spell with the highest cumulative probability is considered for every different dry spell length. This smoothing should be the cause of the outstanding linear behaviour of the plots. The application of this technique to the 39 rain gauges, accepting a square regression coefficient of at least 0.9975 as criterion for the linear fit, leads to average truncation levels of 97th, 96th and 92th percentiles for the respective daily rain amount thresholds (0.1, 1.0 and 5.0 mm/day). The standard deviations are 2, 3, and 4%, respectively. In our analyses, a common 95th percentile has been finally chosen to define the truncation spell length given that it is quite close to that suggested by the mean excess plot.

Fig. 4 summarises the spatial distribution of the truncation spell lengths. Whereas the range of values

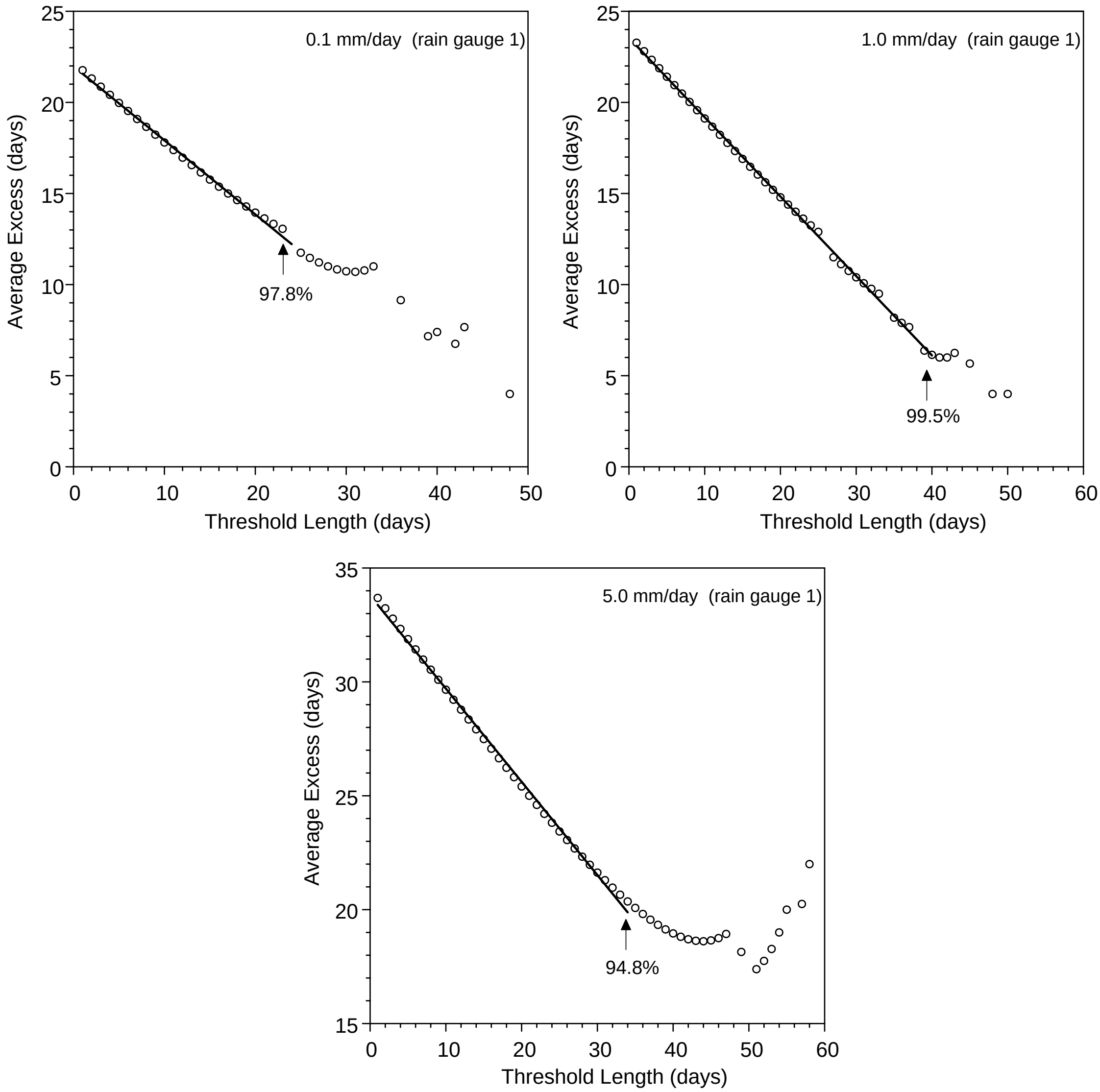

Fig. 3. Mean excess plots for rain gauge 1 and three threshold levels of daily rainfall.

(10–50 days) and the spatial distribution is quite similar for 0.1 and 1.0 mm/day, the case of 5.0 mm/day is characterised by a slightly wider range, from 20 to 70 days. A common feature of the three plots is a north–south gradient and signs of a drift of the longest spells from the southern end of Catalonia towards the Central Basin. Additionally, the shortest dry spells for the 95th percentile are obtained in all cases in the Pyrenees and Pre-Pyrenees Ranges, places where the most copious annual rain amounts are usually recorded.

4.2. Annual extreme dry spells and partial duration series

Fig. 5 compares the empirical L-skewness and L-kurtosis (triangles) of the AE series for each rain gauge with the theoretical GEV (solid line) and GP (dashed line) models for dry spell lengths derived from threshold levels of 0.1, 1.0 and 5.0 mm/day. The open squares represent the Gumbel distribution, which is a particular case of the GEV model.

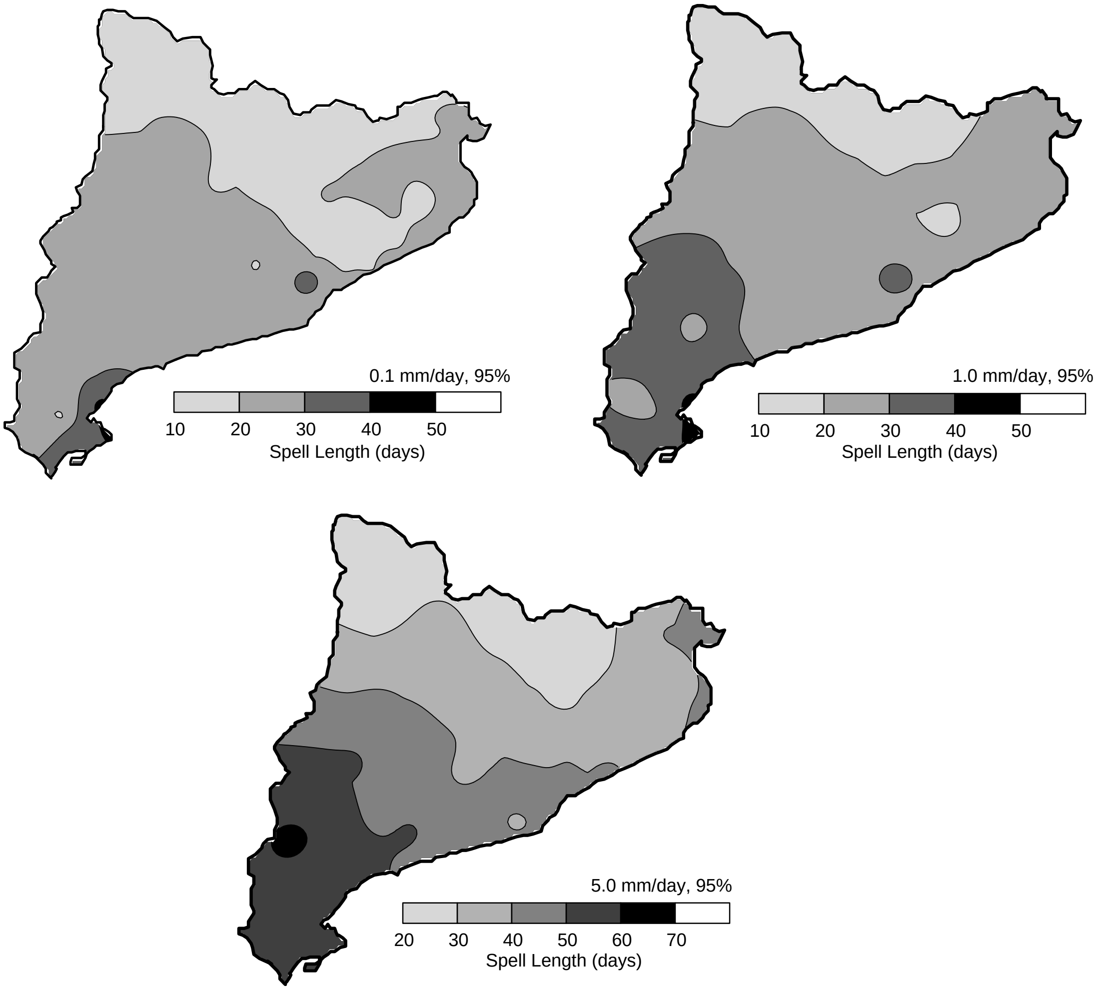

Fig. 4. 95th percentile dry spell lengths for threshold levels of 0.1, 1.0 and 5.0 mm/day.

The first remarkable feature is that, although the GEV model offers a better fit for most gauges, a non-negligible number of gauges satisfies better the GP model. Secondly, the vicinity of the empirical L-moments to the Gumbel model is questionable in most of cases. Thirdly, the optimal (minimum) distance either to the GEV or to the GP model is characterised, according to Eq. (12), by an average value of 0.031 and a standard deviation of 0.024 for the 0.1 mm/day level. The corresponding values for the other two levels are quite similar, with averages ranging from 0.034 to 0.035 and standard deviations from 0.015 to 0.025. The optimum statistic model is then selected in each case according to the minimum L-skewness-kurtosis distance D given by Eq. (12).

Fig. 6 depicts the same comparison as for Fig. 5, but now for the PD series, generated by taking into account spell lengths exceeding a threshold level corresponding to the 95th percentile of the empirical spells. A clear shift towards the GP distribution is observed in all the threshold levels. Nevertheless, a GEV distribution would give a better fit than a GP model for a non-negligible number of gauges. The Gumbel model should be discarded for most of the rain gauges. The average distance, D, range from 0.031 to 0.050, the standard deviations from 0.024 to

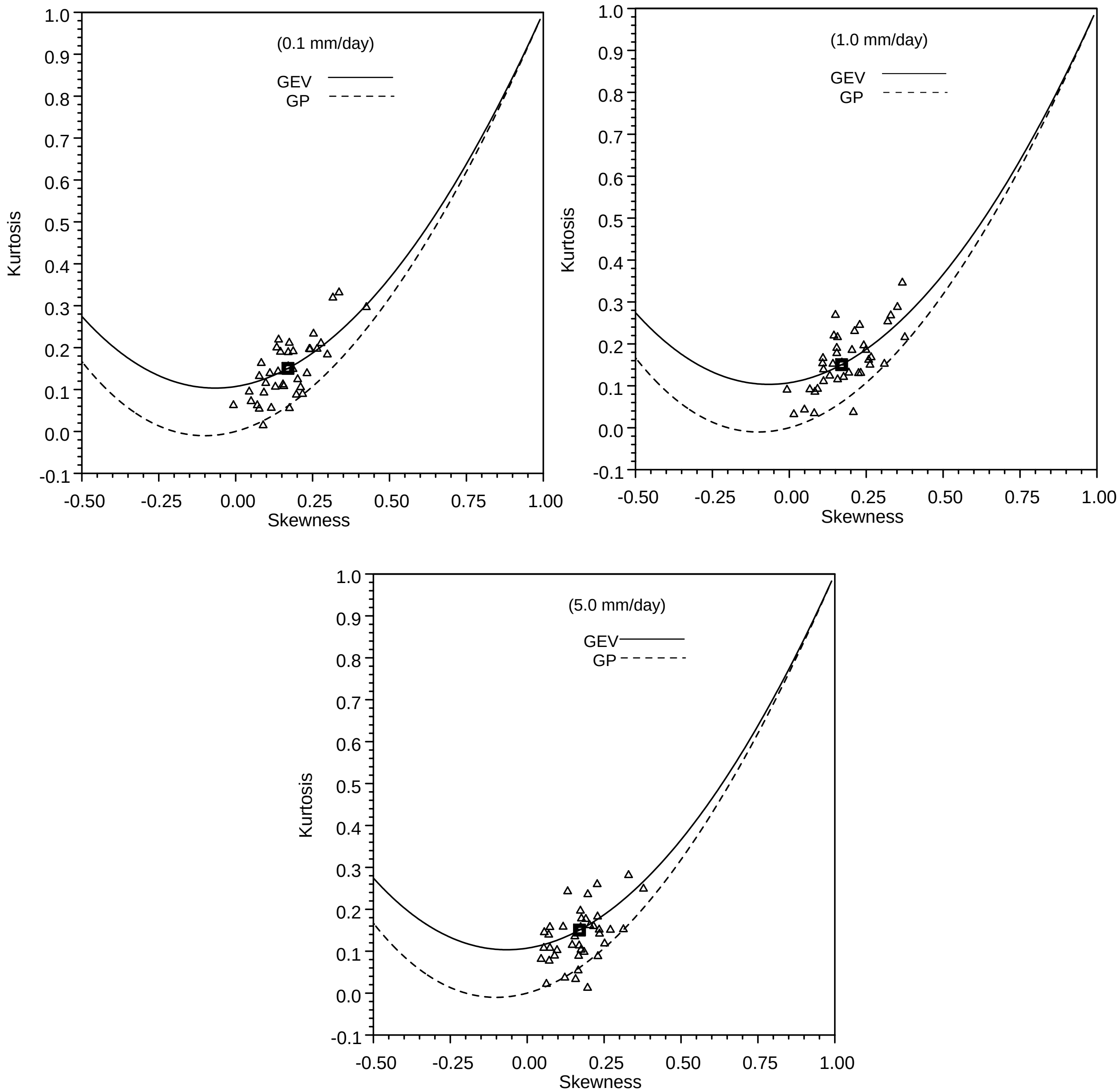

Fig. 5. Empirical (open triangles) and theoretical L-skewness and L-kurtosis for threshold rain amounts of 0.1, 1.0 and 5.0 mm/day according to the AE option. The solid line represents the GEV distribution, dashed line the GP distribution and the open square the Gumbel model.

0.032 and, once more, the optimal statistical model is chosen in each case according to the minimum L-skewness-kurtosis distance *D*. It has to be mentioned that average distances are slightly higher for PD than for AE series.

Fig. 7 shows rain gauges that follow the GEV (codified by 1) and the GP (codified by 2) model for the AE series and the three threshold levels of daily rainfall. The number of gauges better represented by the GP model increases from 8

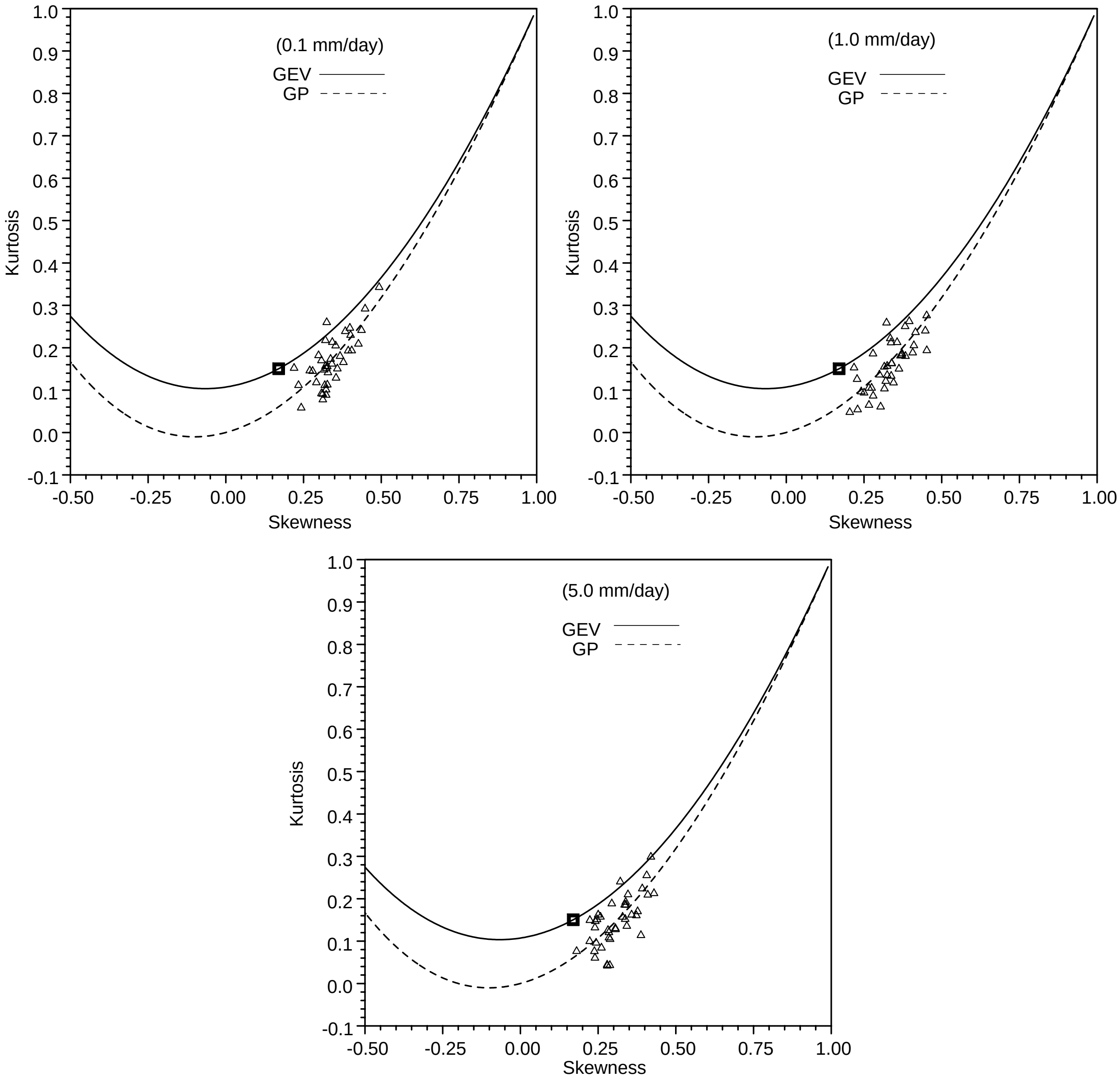

Fig. 6. Empirical (open triangles) and theoretical L-skewness and L-kurtosis for threshold rain amounts of 0.1, 1.0 and 5.0 mm/day according to the PD option. The solid line represents the GEV distribution, dashed line the GP distribution and the open square the Gumbel model.

(0.1 mm/day) to 12 (5.0 mm/day). In this way, there appears to be a tendency to favour the GP model to the detriment of the GEV model with an increasing threshold level of daily rainfall. Nevertheless, the number of gauges following the GEV model is always the highest. Additionally, there is not a spatial pattern to explain the location of gauges favouring one option or another. The spatial distribution corresponding to the PD series is not included, but a close analysis of Fig. 6 suggests that most of gauges satisfy the GP model and only a few cases follow the GEV model.

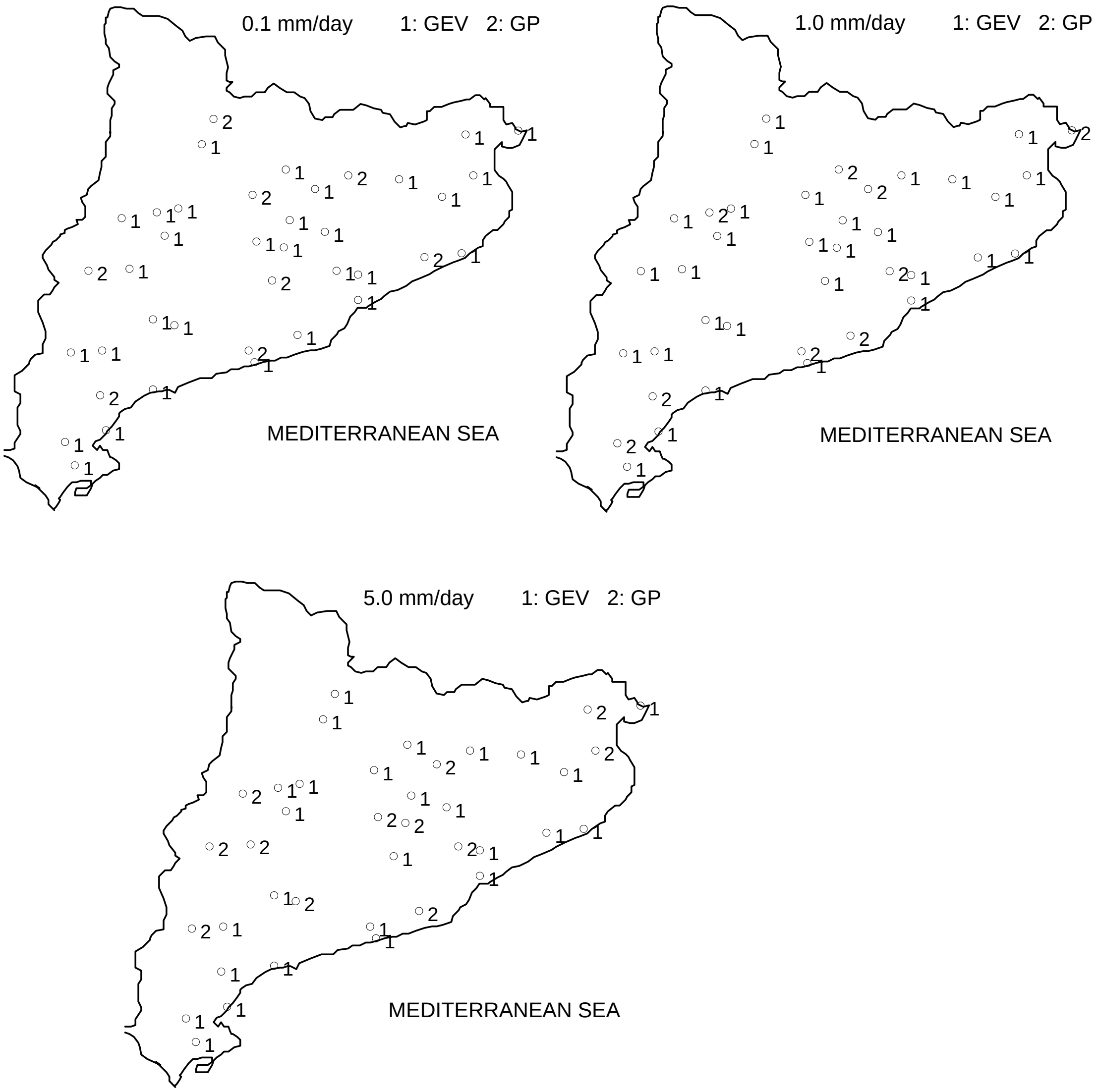

Fig. 7. Rain gauges for which either the GEV (1) or the GP (2) models better fit the cumulative distribution of extreme dry spells generated according to the AE strategy.

4.3. Some examples of GEV and GP distributions

Fig. 8 depicts the fit obtained with the GEV and GP models for the AE series, for a few rain gauges, together with the 95% confidence bands (curved dashed lines) of the Kolmogorov-Smirnov test (Benjamin and Cornell, 1970). Fig. 9 shows the same examples for the PD series and the GP model. Empirical cumulative probabilities out of these confidence bands indicate that empirical distributions would not be well fitted by GEV or GP distributions. Rain gauges 14 and 50 are located in the Central Mediterranean coast and rain gauge 21 in the northern end of the Mediterranean coast (Fig. 2). For rain gauge 14, the empirical probabilities for the AE series are reasonably well explained both by the GEV and the GP models, and are also quite well fitted by the Gumbel distribution (Fig. 8). The best fit, according to distance D, is given by the GEV model ($D=0.040$). Nevertheless, the difference with respect to the GP model ($D=0.060$) is almost negligible. The Gumbel

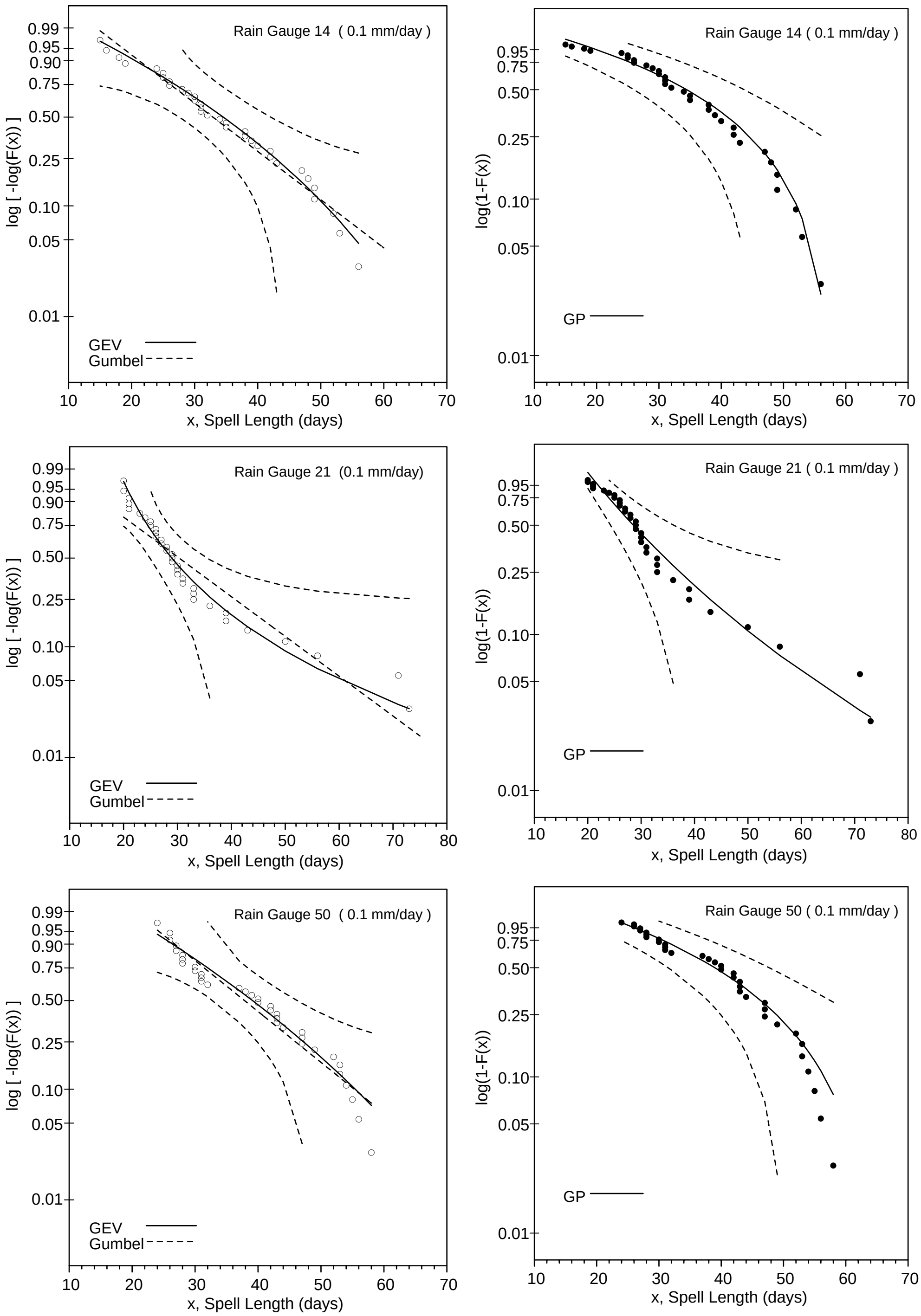

Fig. 8. Some examples of fitting between empirical and theoretical cumulative distributions of extreme dry spells generated by the AE strategy: open and solid circles represent empirical data, solid lines correspond to both the GEV and GP distributions, straight dashed lines correspond to the Gumbel model and curved dashed lines represent the Kolmogorov-Smirnov 95% confidence bands.

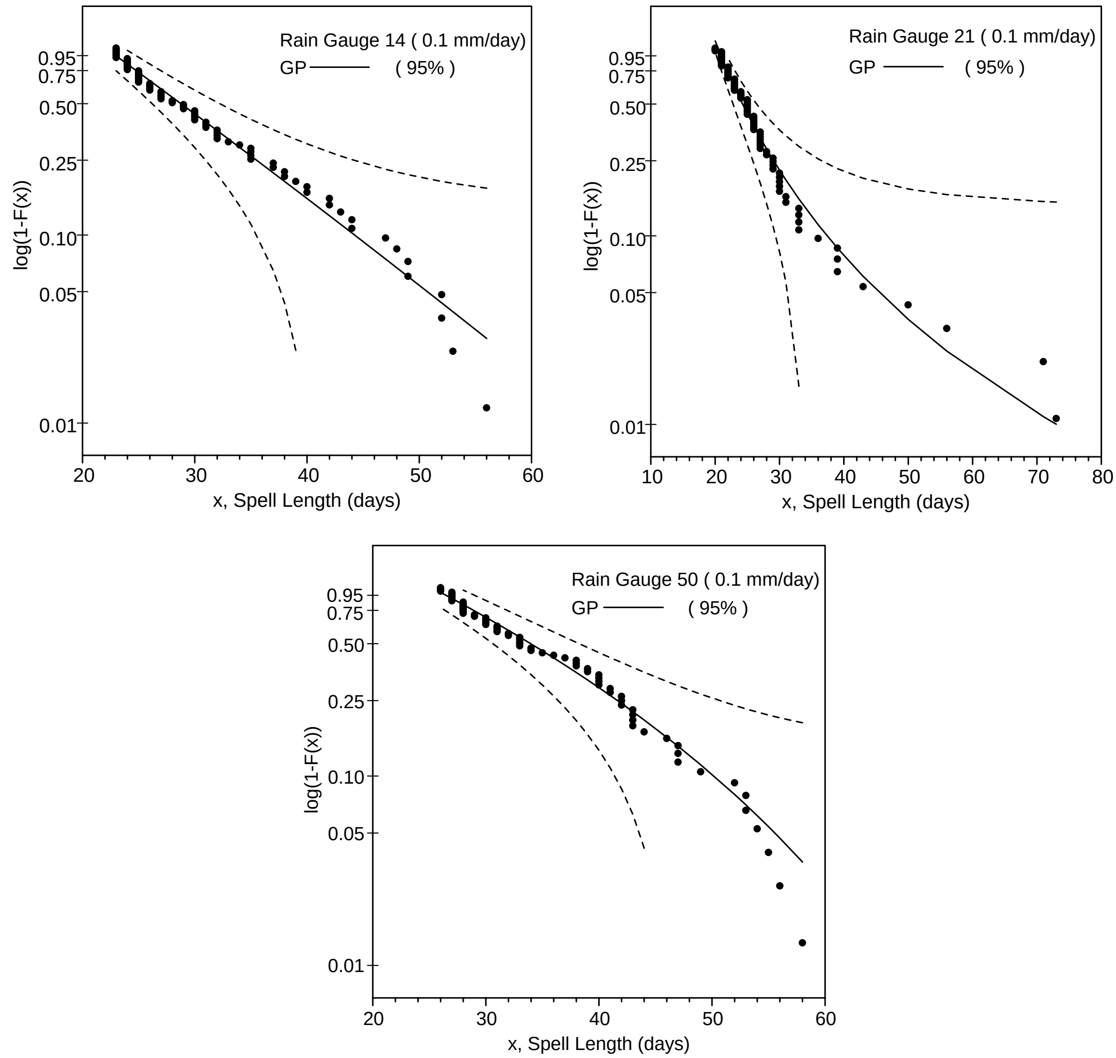

Fig. 9. Some examples of fits between empirical and theoretical cumulative distributions of extreme dry spells generated by the PD strategy: solid circles represent empirical data, solid lines correspond to the GP model and curved dashed line represent the Kolmogorov-Smirnov 95% confidence bands.

distribution is the worst option, with $D=0.142$. For rain gauge 21, there is a clear difference between distances for the GEV ($D=0.003$) and the Gumbel ($D=0.295$) models, while the GP model ($D=0.054$) also fits very well the empirical data. Finally, for rain gauge 50, both the GEV and GP models have some difficulty fitting the longest spells. The GP model is the best ($D=0.009$), in comparison with GEV ($D=0.107$) and Gumbel ($D=0.156$) distributions. Fig. 9 illustrates the fits of the PD series offered by the GP model for the same rain gauges shown in Fig. 8. The three cases depict a good fit, with D ranging from 0.034 to 0.066, which are distances quite similar to those obtained for AE spells and for the GP and GEV models, in spite of some difficulties to fit the longest spells.

According to the Kolmogorov-Smirnov test, the 95% confidence bands of the theoretical cumulative distribution depicted in Figs. 8 and 9 are given by the statistic $1.36/n^{1/2}$, where n is the number of empirical data. Given that n is not excessively high for the selected samples, even for PD series, the discriminating power of the test is low, as can be observed in the case of rain gauge 21 for the AE spells. This test would accept both the GEV model and its simplification (Gumbel), but the distances computed in terms of the shape parameter κ permit to discriminate between these two possibilities. This fact reinforces the utility of the L-moments, especially when the number of data is not very high. By one hand, the distance D in terms of L-skewness and L-kurtosis is more discriminating. On the other hand, if the available data is not very numerous, the parameter estimation from statistical moments of the empirical data or from probability paper representation is less robust than those derived from L-moments.

4.4. Return periods

Fig. 10a–c illustrate the spatial distribution of the spell lengths for the different threshold levels (0.1, 1.0 and 5.0 mm/day) for the PD strategy and return periods of 2, 5, 10 and 25 years. All these maps have

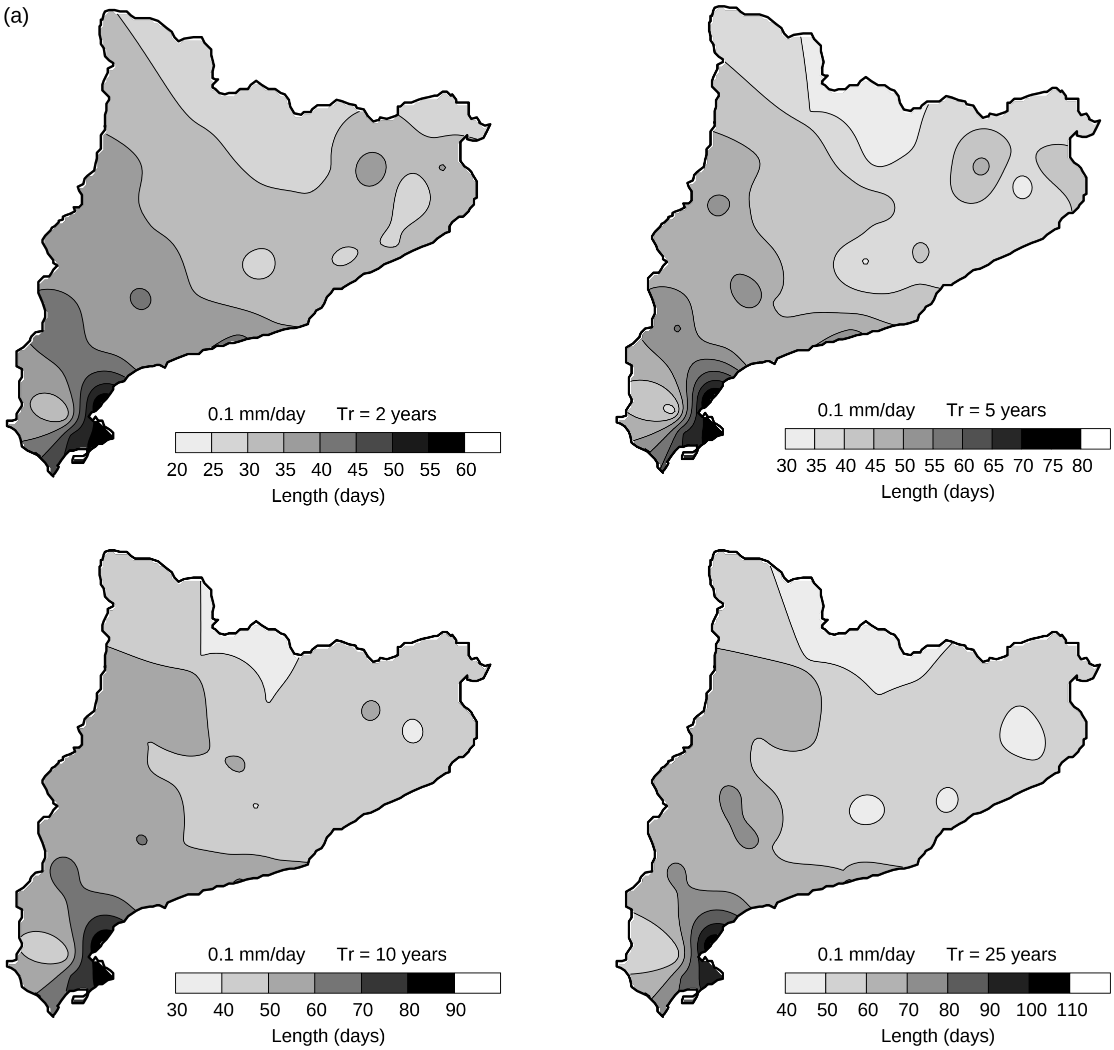

Fig. 10. Dry spell lengths for return periods of 2, 5, 10 and 25 years and threshold levels of (a) 0.1, (b) 1.0 and (c) 5.0 mm/day, when the PD sampling strategy is considered.

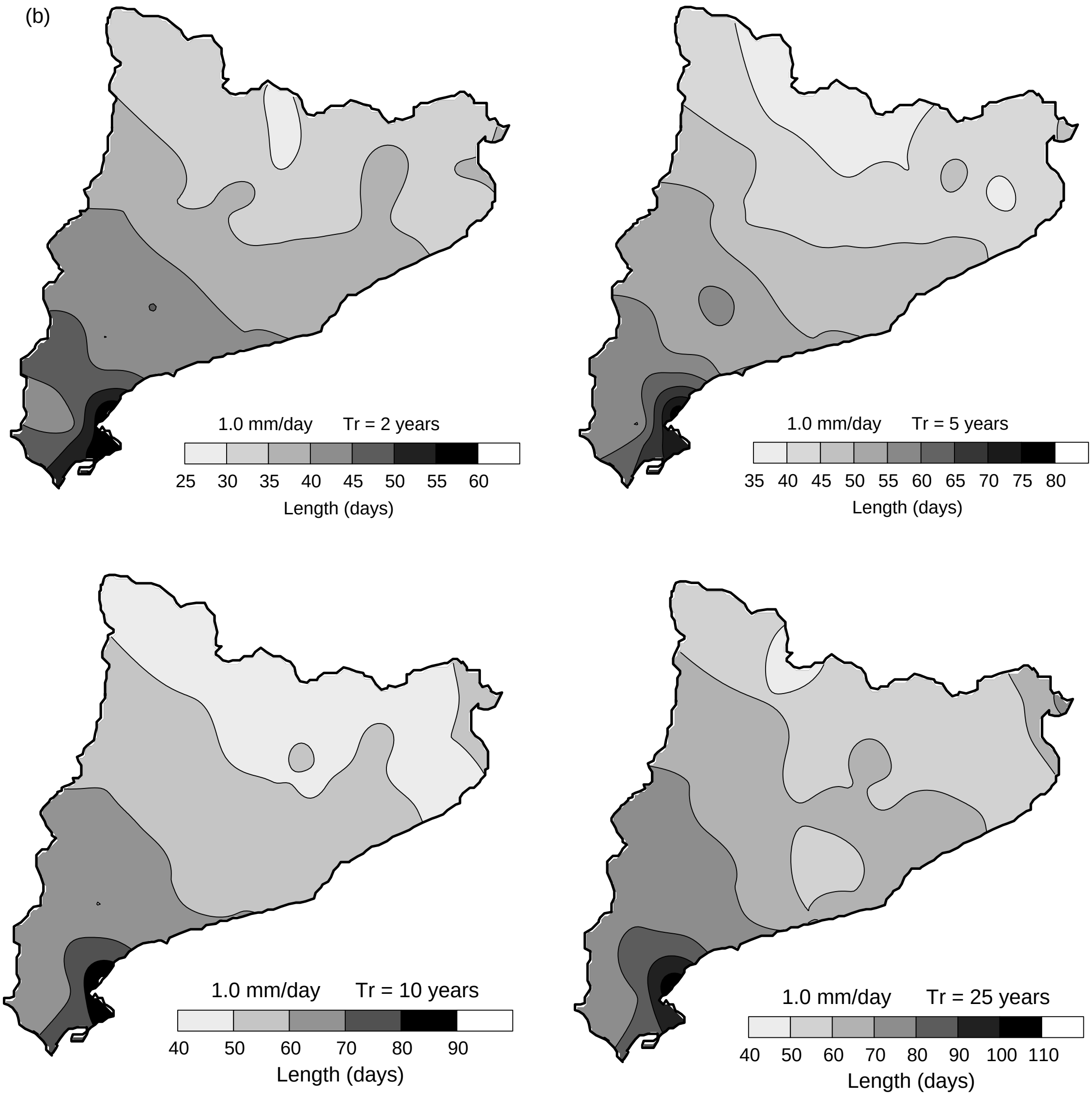

Fig. 10 (*continued*)

been built up by choosing the best option from the GEV and GP models, according to the minimum L-skewness-kurtosis distance *D*, for each rain gauge. A common pattern is the presence in all cases of a clear north to south gradient, with a nucleus of maximum length in the southern corner of Catalonia for 0.1 and 1.0 mm/day, which drifts towards the Central Basin for 5.0 mm/day. Another interesting and expected fact is that the general features of these return period maps are quite similar to the spatial distribution of average annual amounts in Catalonia (Clavero et al., 1996), in the sense that the humid areas of Catalonia depict the shortest spell lengths and the driest domains are associated with the longest spell lengths, whatever the daily rainfall level or return period considered.

The return period maps for 2, 5, 10 and 25 years and the AE strategy depicts quite similar patterns to

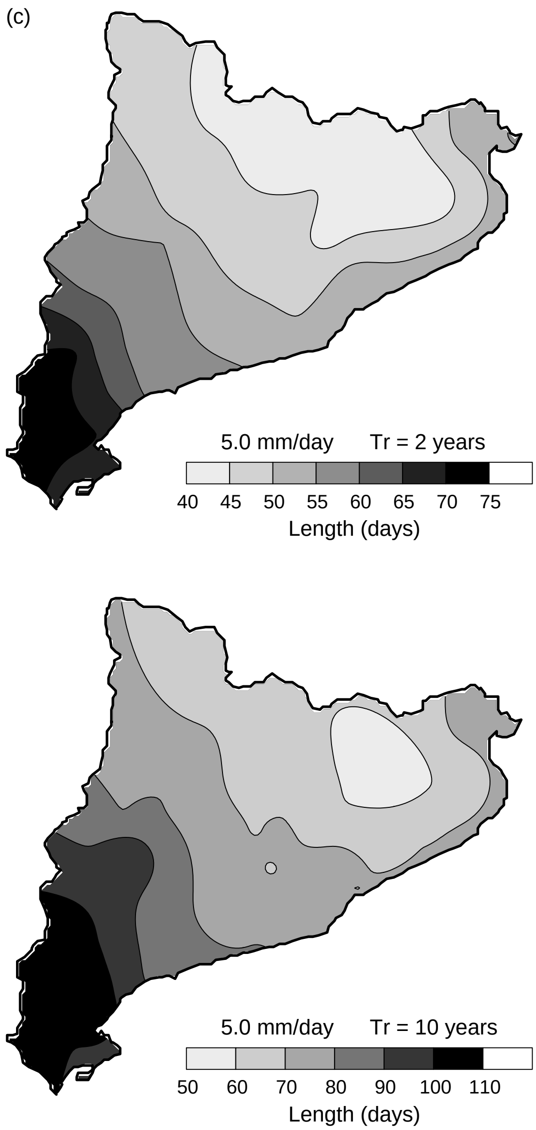

Fig. 10 (*continued*)

those described in the set of Figs. 10 and they are not shown. Nevertheless, it is interesting to discuss the relative discrepancies between the dry spell lengths generated according the AE and the PD strategies. In all cases, percentages are computed as differences between the length of dry spells generated by the AE and PD strategies divided by the AE spell.

The return period of 2 years shows discrepancies, most of them from 0 to −5%, for threshold levels of 0.1 and 1.0 mm/day. This pattern changes for 5.0 mm/day, for which the predominant discrepancies between the AE and PD options range from −5 to +5%. For the 5 year return period, the relative discrepancies in dry spells range from 0 to 5% for a great part of Catalonia, at least for threshold levels of 0.1 and 1.0 mm/day. For 5.0 mm/day, these relative differences increase and reach values within the 0–10% interval for a wide area. The results for the 10 year return period suggest a quite similar

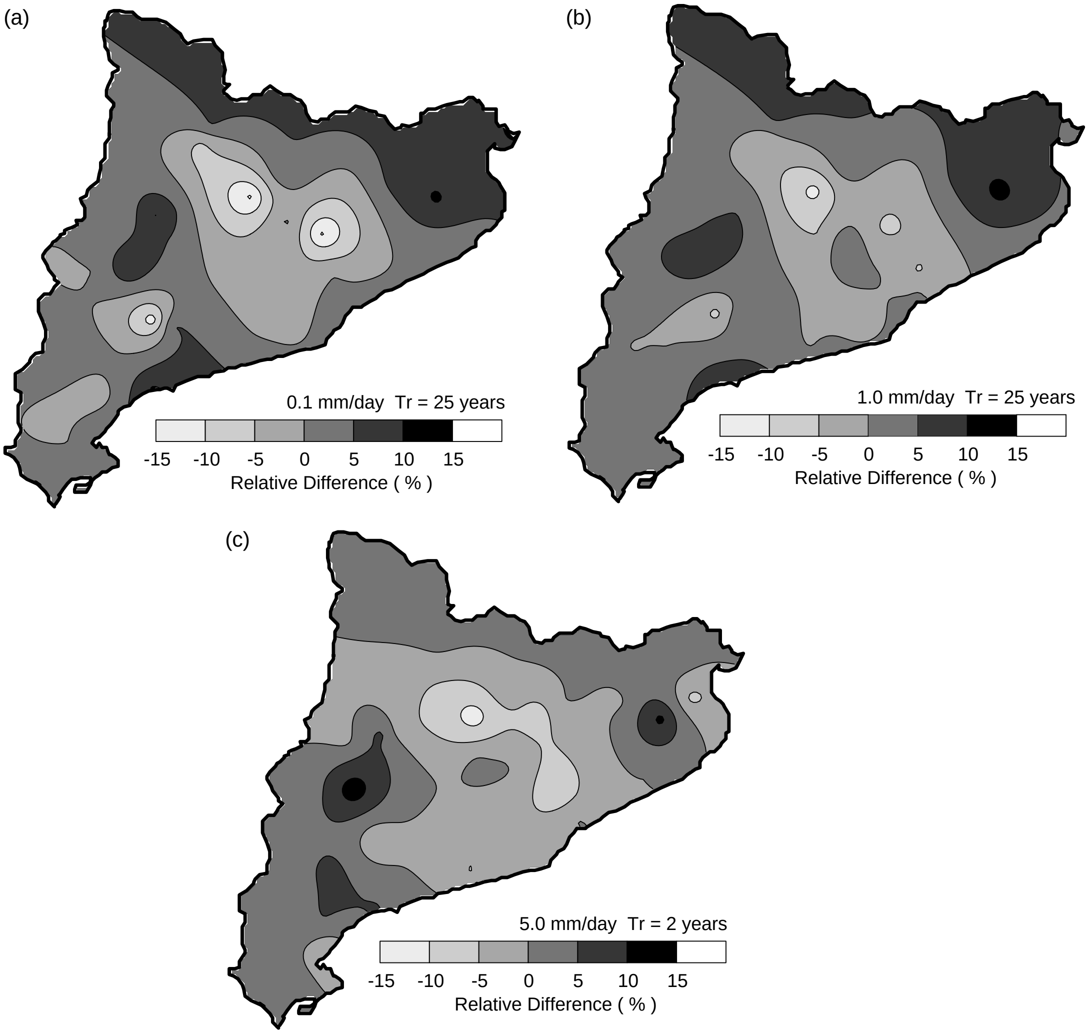

Fig. 11. Three examples of the widest range of discrepancies in spell lengths when the AE and the PD sample strategies are considered. (a) 0.1 mm/day and 25 years; (b) 1.0 mm/day and 25 years; (c) 5.0 mm/day and 2 years.

behaviour. Most of Catalonia depicts longer dry spells for the AE option, with predominant discrepancies ranging from 0 to 5% for the three threshold levels. Finally, the 25 year return period is characterised by maximum discrepancies, arriving up to ±15%, attained for 0.1 and 1.0 mm/day.

Figs. 11a–c illustrate some examples of the spatial distribution of these relative discrepancies. Every figure corresponds to a threshold level and the return period depicting the widest range of discrepancies. First, it is worth mentioning that in all cases discrepancies can reach values of approximately ±15%, which is a non-negligible discrepancy on dry spell lengths. Second, the widest range of discrepancies is not always detected for the highest return period of 25 years. An example is the case of 5.0 mm/day, for which the widest range is observed for 2 years.

5. Discussion of the results

A first question to discuss is the specific percentile considered to define the PD series. According to

Vicente-Serrano and Beguería-Portugués (2003), who treated a very similar problem, relatively small changes in the truncation level could generate relevant changes in the distribution parameters and return period results. A good strategy could be that proposed by the mentioned authors, who tested several percentiles ranging from 90th to 99.75th and assumed as the best parameters the average derived from the different percentiles. Given that the goal is the optimal sample of long dry spells, the 90th percentile could be a low percentile for some rain gauges and dry spells not strictly classified as extremes could be considered. Opposite to that, it is very likely that percentiles close to the 99th percentile only include maximum lengths in all cases. Nevertheless, two shortcomings could then appear. On the one hand, some extreme events could be excluded by the use of so high percentile. On the other hand, the number of samples exceeding this percentile is likely to be so small that advantages offered by L-moments could be balanced by numerical estimations derived from an insufficient number of data. Consequently, in the present research a common 95th percentile has been chosen for all rain gauges, with the aim of including only extreme dry lengths, but in a number high enough to derive reliable parameters and return periods. In effect, the number of dry spells generated by the PD strategy for every rain gauge is higher than that of extreme annual dry spells of the AE series. The mean excess plot technique (Beguería, 2005) has also been applied and the results reinforce the use of the 95th percentile to define PD series. An alternative procedure has been recently proposed (Beguería, 2005) based on average parameters derived from an increasing set of truncation percentiles, starting with the lowest value accomplishing the mean excess plot.

Another question is the advantage offered by the PD over the AE strategy. According to Moreno and Roldán (1999); Alila (2000), among others, the PD strategy usually leads to good results for the stochastic modelling of hydrological extremes. In fact, Cunnane (1973) and Madsen et al. (1997a,b), among others, have illustrated the shortcomings of using AE over the PD option. A complementary point of view is offered by an accurate search of the statistical distribution that better fits the empirical data. At least for AE spells, there is not an optimal single distribution. The GP model is the best option for some rain gauges, but, for others, the GEV and even the Gumbel model accomplishes better the requirement imposed by distance D computed in terms of L-skewness and L-kurtosis. Nevertheless, a clear preference of the GEV is observed in the case of the AE spells and of the GP for the PD series. The Gumbel distribution is discarded in many cases, especially for the PD option. Consequently, many analyses of AE spells based on the classical Gumbel approach should be revised. However, spell lengths are sometimes difficult to predict by the Gumbel model or the GEV and GP distributions. Even though it is only a hypothesis, these very long spells, difficult to forecast using statistical models, could be attributable to some specific synoptic situations, such as a persistent lack of eastern advections during an autumn season.

The spatial distribution of gauges using either the GEV or GP distributions, which can change with the threshold level, is not of importance to our main objective. It should be remembered that we are interested in the best sampling strategy, AE spells in front of PD series, by considering the best statistical distribution for every rain gauge and daily amount threshold.

Possible influences originated by the orography and the vicinity to the Mediterranean Sea are not so evidently manifested here as in other kind of daily rainfall analyses (Lana et al., 2004). The Pre-Pyrenees and Pyrenees Ranges, the most important orographic systems in Catalonia, are characterised by the shortest dry spells associated with the 95th percentile used in the PD strategy, whatever the daily rainfall threshold considered. The effects of the Littoral and Pre-Littoral Ranges have not been detected, whereas the effects of the Transversal Range are only detected for 5.0 mm/day. Additionally, east–west gradients, attributable to the distance to the Mediterranean Sea, are not observed.

With respect to other strategies applied to the dry spell series derived for Catalonia, even though the recording period and the rain gauge network are not exactly the same, comparisons can be made with results derived by Lana and Burgueño (1998a,c), given that the spatial coverage of the rain gauges is very similar in all cases. The AE option was used by Lana and Burgueño (1998a) to characterise annual extreme droughts for the 0.1 mm/day threshold level. Besides the Gumbel distribution, the Jenkinson formulation with three parameters was also considered (Jenkinson, 1969, 1975), which better fitted the AE spells for some rain gauges. Return period maps with optimum values derived from either the Gumbel or Jenkinson formulation depicted quite similar patterns to those deduced here for the threshold of 0.1 mm/day and return periods of 2, 5, 10 and 25 years. Both signs of a north–south gradient and a nucleus of long spells in the southern corner of Catalonia were observed. Nevertheless, results derived from the Gumbel or Jenkinson formulation have some discrepancies with respect to return period maps obtained here. On the one hand, maximum dry spell lengths in the southern corner of Catalonia were overestimated approximately by 20 days for the 25 year return period. These discrepancies decrease for shorter return periods, achieving as much as 5 days for the 2 year return period. On the other hand, the rest of dry spells deduced by Lana and Burgueño (1998a) are quite similar when they are compared with lengths deduced here. The western domains of the Central Basin became roughly defined by the isolines of 40, 55, 70 and 80 days, for 2, 5, 10 and 25 years, respectively, but this pattern has not been detected in return period maps derived from either the AE option or the PD strategy.

It is also interesting to compare return period maps obtained here with those derived by Lana and Burgueño (1998c), who assumed a Poisson distribution to quantify the probabilities of repeated long dry episodes in Catalonia. In this case, comparisons can be made for return periods of 5, 10 and 25 years for a threshold level of 0.1 mm/day. Once more a north–south gradient and a nucleus of long spells close to the Ebro Delta were detected. Additionally, both the length range and the spatial distribution of spells for every return period are quite similar to that obtained here. More concretely, the delimitation of the western end of the Central Basin, observed in Lana and Burgueño (1998a), is not detected in Lana and Burgueño (1998c), in agreement with the present results. In short, improvements achieved with the GEV and GP distributions for both the AE and PD strategies agree better with the Poisson approach than with the Gumbel and Jenkinson models. Possibly, the use of the Gumbel model in Lana and Burgueño (1998a) biased their return period maps.

6. Conclusions

Some statistical drought patterns of Catalonia during the second half of the 20th century have been analysed. Long dry spells with return periods of 2, 5, 10 and 25 years for daily rain thresholds of 0.1, 1.0 and 5.0 mm/day have been obtained from the AE and PD sampling strategies. Gumbel, GEV and GP distributions have been tested for both strategies. The Gumbel model has been discarded for most of the rain gauges, threshold levels and the two sampling strategies. The GEV and GP models are usually much more convenient, with a preference of the GEV distribution for the AE strategy and the GP model for the PD option.

In agreement with the main orographic features of Catalonia, a north to south gradient is detected for the return period maps of 2, 5, 10 and 25 years, both for the AE and PD strategies. The discrepancies in spell lengths for both options range from -5 to $+5\%$ in most cases, with large discrepancies from -15 to $+15\%$ in some places. With respect to the PD strategy, the selected 95th percentile threshold is in agreement with the average truncation percentile suggested by the mean excess plot for daily rainfall thresholds of 0.1, 1.0 and 5.0 mm/day.

From an applied point of view, the estimation of return periods of long dry spells could represent a valuable aid to prevent problems concerning water demands generated by agriculture or socio-economic activity when drought episodes associated with short return periods of 2 or 5 years are considered. Similarly, management of water resources (reservoirs and ground surface waters) should consider results concerning return periods of 5, 10 and 25 years.

From a methodological point of view several questions have to be underlined:

AE and PD sampling strategies have been compared. Given that the PD strategy is not submitted to the same shortcomings affecting the AE sampling method, the better option for an accurate estimation of spells for different return periods should be the PD strategy. Nevertheless, the selection of the percentile for which the PD series are defined is in some way debatable. The chosen truncation threshold has to define only extreme dry lengths, but in a high enough number. Thus, advantages offered by L-moments were not cancelled because of an insufficient number of data. Consequently, the chosen percentile is a trade-off between using a high enough number of samples and the results offered by the mean excess plot.

Instead of the estimation of Gumbel, GEV or GP distribution parameters by probability paper, maximum likelihood or statistical moments, the L-moments method is proposed here due to its robustness. Fits offered by the different distribution models are easily evaluated by remembering that the L-moments, skewness and kurtosis, solely depend on the shape parameter κ. This property permits us to define a distance, D, between models and empirical distributions, the minimum distance, D, thus providing a complementary criterion to the Kolmogorov-Smirnoff test, which is not very discriminatory when the number of observations is small.

Acknowledgements

We are indebted to the *Instituto Nacional de Meteorologia* (Spanish Ministry of Environment) for providing the pluviometric data series.

References

Alila, Y., 2000. Regional rainfall depth-duration-frequency equations for Canada. Water Resour. Res. 36, 1767–1778.

Alpert, P., Ben-Gai, T., Baharad, A., Benjamini, Y., Yekutieli, D., Colacino, M., Diodato, L., Ramis, C., Homar, V., Romero, R., Michaelides, S., Manes, A., 2002. The paradoxical increase of Mediterranean extreme daily rainfall in spite of decrease in total values. Geophys. Res. Lett. 29 (101029/2001GL013554).

AMS statement, 2004. AMS statement on meteorological drought. Bulletin Am. Meteorol. Soc. May, 771–773.

Barry, R.G., Chorley, R.J., 2003. Atmosphere, Weather and Climate, eighth ed. Routledge, London p. 421.

Beguería, S., 2005. Uncertainties in partial duration series modelling of extremes related to the choice of the threshold value. J. Hydrology 303, 215–230.

Benjamin, J.R., Cornell, C.A., 1970. Probability, Statistics and Decision for Civil Engineers. McGraw-Hill Inc., New York NY.

Clavero P., Martín Vide J., Raso J. 1996. Atles climàtic de catalunya: termopluviometria, Departament de Medi Ambient, Generalitat de Catalunya, 42 pp.

Cunnane, C., 1973. A particular comparison of annual maxima and partial duration series methods of flood frequency prediction. J. Hydrology 18, 257–271.

Fowler, H.J., Kilsby, C.G., 2003. A regional frequency analysis of United Kingdom extreme rainfall from 1961 to 2000. Int. J. Climatol. 23, 1313–1334.

Hosking, J.R.M., Wallis, J.R., 1997. Regional Frequency Analysis. An Approach Based on L-moments. Cambridge University Press, London p. 224.

Hosking, J.R.M., Wallis, J.R., Wood, E.F., 1985. Estimation of the generalised extreme-value distribution by the methods of probability weighted-moments. Technometrics 27, 251–261.

Hounam, C., Burgos, J., Kalik, M., Palmer, W., Rodda, J. (1975). Drought and Agriculture, World Meteorological Organization, Technical Note 138, Geneva, Switzerland, 128 pp.

Jenkinson, A.F. 1969. Statistics of extremes of Maximum floods, WMO Technical Note, 98. World Meteorological Organization, Geneva, pp. 183–228.

Jenkinson, A.F., 1975. Extreme value analysis in meteorology, Fourth Conference on Probability and Statistics in Atmospheric Sciences. American Meteorological Society, Tallahassee, FL.

Kiktev, D., Sexton, D.M.H., Alexander, L., Folland, C.K., 2003. Comparison of modelled and observed trends in indices of Daily Climate Extremes. J. Clim. 16, 3560–3571.

Lana, X., Burgueño, A., 1998a. Spatial and temporal characterization of annual extreme droughts in Catalunya (NE Spain). Int. J. Climatol. 18, 93–110.

Lana, X., Burgueño, A., 1998b. Daily dry–wet behavior in Catalonia (NE Spain) from the viewpoint of first and second order Markov chains. Int. J. Climatol 18 (7), 793–815.

Lana, X., Burgueño, A., 1998c. Probabilities of repeated long dry episodes based on the Poisson distribution. An example for Catalonia (NE Spain). Theor. Appl. Climtol. 60, 111–120.

Lana, X., Serra, C., Burgueño, A., 2001. Patterns of monthly rainfall shortage and excess in terms of the standardized precipitation index for Catalonia (NE Spain). Int. J. Climatol. 21, 1669–1691.

Lana X., Serra C., Burgueño A., 2003. Trends affecting pluviometric indices at the Fabra Observatory (Barcelona, NE Spain) from 1917 to 1999. Int. J. Climatol. 23, 315–332.

Lana, X., Martínez, M.D., Serra, C., Burgueño, A., 2004. Spatial and temporal variability of the daily rainfall regime for Catalonia (NE Spain), 1950–2000. Int. J. Climatol. 24, 613–641.

Llasat, M.C., Puigcerver, M., 1994. Meteorological factors associated with floods in the North-Eastern part of the Iberian Peninsula. Nat. Hazards 9, 81–93.

Llasat, M.C., Ramis, C., Barrantes, J., 1996. The meteorology of high-intensity rainfall events over the west mediterranean region. Remote Sensing Rev. 14, 51–90.

Madsen, H., Pearson, C.P., Robsjerg, D., 1997a. Comparison of annual maximum series and partial duration methods for modelling extreme hydrological events. 1. At-site modelling. Water Resour. Res. 33, 759–769.

Madsen, H., Pearson, C.P., Robsjerg, D., 1997b. Comparison of annual maximum series and partial duration methods for modelling extreme hydrological events. 2. Regional modelling. Water Resour. Res. 33, 771–790.

Moreno, F., Roldán, J., 1999. Regional daily precipitation stochastic model parameters. Application to the Guadalquivir valley in southern Spain. Phys. Chem. Earth, Part B: Hydrol. Oceans Atmos. 24, 35–47.

Romero, R., Ramis, C., Guijarro, J.A., 1999. Daily rainfall patterns in the Spanish Mediterranean area: an objective classification. Intern. J. Climatol. 19, 95–112.

Serra, C., Fernández-Mills, G., Periago, C., Lana, X., 1998. Surface synoptic circulation and daily precipitation in Catalonia. Theor. Appl. Climatol. 59, 24–49.

Sotillo, M.G., Ramis, C., Romero, R., Alonso, S., Homar, V., 2003. Role of orography in the spatial distribution of precipitation over the Spanish Mediterranean zone. Clim. Res. 23, 247–261.

Vicente-Serrano, S.M., Beguería-Portugués, S., 2003. Estimating extreme dry-spell risk in the middle Ebro valley (NE Spain): a comparative analysis of partial duration series with general Pareto distribution and annual maxima series with a Gumbel distribution. Int. J. Climatol. 23, 1103–1118.