Theoretical and Applied Climatology

MULTIFRACTALITY AND AUTOREGRESSIVE PROCESSES OF DRY SPELL LENGTHS IN EUROPE: AN APPROACH TO THEIR COMPLEXITY AND PREDICTABILITY

--Manuscript Draft--

Manuscript Number:	TAAC-D-14-00581R1
Full Title:	MULTIFRACTALITY AND AUTOREGRESSIVE PROCESSES OF DRY SPELL LENGTHS IN EUROPE: AN APPROACH TO THEIR COMPLEXITY AND PREDICTABILITY
Article Type:	Original Paper
Corresponding Author:	Xavier Lana, Ph.D. UNITED STATES
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	
Corresponding Author's Secondary Institution:	
First Author:	Xavier Lana, Ph.D.
First Author Secondary Information:	
Order of Authors:	Xavier Lana, Ph.D.
	August Burgueño, PhD
	Carina Serra, PhD
	Dolors Martínez, PhD
Order of Authors Secondary Information:	
Funding Information:	
Abstract:	Dry spell lengths, DSL, defined as the number of consecutive days with daily rain amounts below a given threshold, may provide relevant information about drought regimes. Taking advantage of a daily pluviometric database covering a great extension of Europe, a detailed analysis of the multifractality of the dry spell regimes is achieved. At the same time, an autoregressive process is applied with the aim of predicting DSL. A set of parameters, namely Hurst exponent, H, estimated from multifractal spectrum, , critical Hölder exponent, , for which reaches its maximum value, spectral width, W, and spectral asymmetry, B, permits a first clustering of European rain gauges in terms of the complexity of their DSL series. This set of parameters also allows distinguishing between time series describing fine- or smooth-structure of the DSL regime by using the Complexity Index, CI. Results of previous monofractal analyses also permits establishing comparisons between smooth-structures, relatively low correlation dimensions, notable predictive instability and anti-persistence of DSL for European areas, sometimes submitted to long droughts. Relationships are also found between the CI and the mean absolute deviation, MAD, and the optimum autoregressive order, OAO, of an ARIMA(p,d,0) autoregressive process applied to the DSL series. The detailed analysis of the discrepancies between empiric and predicted DSL underlines the uncertainty over predictability of long DSL, particularly for the Mediterranean region.
Response to Reviewers:	Manuscript TAAC-D-14-00581: Multifractality and autoregressive processes of dry spell lengths in Europe: An approach to their complexity and predictability (Authors: X. Lana, A. Burgueño, C. Serra and M.D. Martínez) Response to Reviewer 1: a)In agreement with recommendations of the Reviewer and the Editor, the Section 3.1 has been notably shortened, being only introduced the basic concepts concerning the

MF-DFA algorithm and the appropriate references (see also lines 162-165). The structure of Section 3.2 (Singularity spectrum) has not been changed, as many definitions of parameters and concepts used after are summarised in this Section. Additionally, in agreement with Reviewer 2, a detailed discussion about the range of qth-order has been added to this Section.

b)Effectively, the interpolation method has been the "inverse distance". We think that in our case the main problem of obtaining accurate plots of the spatial distribution of fractal parameters is the limited rain gauge density in some areas, as mentioned now in page 5 (lines 118-124).

c)In Section 5, lines 575-580, several alternatives to the AR(p) process are cited as possible improvements on DSL prediction for very long DSL. Certainly, an ARIMA(p,d,q) modelling and others methods based on the Poisson distribution and Monte Carlo algorithms would improve the mentioned very long DSL prediction, but we think they are beyond the scope of the present paper. It is also worth of mention that the relatively simple AR(p) process has led to good results when predicting monthly Western Mediterranean Oscillation index, as the authors of this manuscript have found (manuscript nowadays submitted to the International Journal of Climatology). With respect to these questions, we have to mention that in line 224-225 a mistake concerning the definition of AR(p) has been amended. AR(p) has to be properly defined as an ARIMA(p,d,q) with d=1 and q=0.

Response to Reviewer 2:
a)With the aim of a more complete description of the DSL series, new Figure 1c includes a histogram of the number of DSL, NDSL, for the 267 DSL series. Additionally, two very different examples of DSL series (Vaexjoe, Sweden, and Almeria, Spain) are shown in a new Figure 1d. The corresponding comments are developed in Section 2, lines 136-143.

b)A discussion about the appropriate range of the qth-order is developed in Section 3.2, lines 203-217. The paper suggested by the reviewer (*), the assumption that multifractal spectra should be fitted to a quadratic function taking as argument the Hölder exponent, and previous experience of the authors about this question have been used as reference points of this discussion.
(*) Ivanov P. Ch., Nunes Amaral, L.A., Goldenberg A.L., Haulin S., Rosenblum M.G., Stanley, H.E., Struzik, Z.B. (2001). From 1/f noise to multifractal cascades in hearthbeat dynamics. Chaos, 11, 641-652.
c)The process to obtain a configuration of 14 clusters is described with more detail in Section 4.3, lines 340-356. The explanation is partially based on the concept of similarity index, Lij, which is defined and quantified by the new Equation 9.

d)With respect to Figures 8, due to a technical problem with the writing software, the two first figures were plotted without appearing "C" codes designing the cluster number. The problem has been now solved.

MULTIFRACTALITY AND AUTOREGRESSIVE PROCESSES OF DRY SPELL LENGTHS IN EUROPE: AN APPROACH TO THEIR COMPLEXITY AND PREDICTABILITY

Lana X.[(1)], Burgueño A.[(2)], Serra C.[(1)], Martínez, M.D.[(3)]

[(1)] Departament de Física i Enginyeria Nuclear. ETSEIB, Universitat Politècnica de Catalunya. Av. Diagonal 647, 08028 Barcelona.

[(2)] Departament d'Astronomia i Meteorologia. Facultat de Física, Universitat de Barcelona. C. Martí Franquès 1, 08028 Barcelona.

[(3)] Departament de Física Aplicada. ETSAB, Universitat Politècnica de Catalunya. Av. Diagonal 649, 08028 Barcelona.

Corresponding author:

X. Lana, francisco.javier.lana@upc.edu

Abstract

Dry spell lengths, DSL, defined as the number of consecutive days with daily rain amounts below a given threshold, may provide relevant information about drought regimes. Taking advantage of a daily pluviometric database covering a great extension of Europe, a detailed analysis of the multifractality of the dry spell regimes is achieved. At the same time, an autoregressive process is applied with the aim of predicting DSL. A set of parameters, namely Hurst exponent, *H*, estimated from multifractal spectrum, $f(\alpha)$, critical Hölder exponent, α_0, for which $f(\alpha)$ reaches its maximum value, spectral width, *W*, and spectral asymmetry, *B*, permits a first clustering of European rain gauges in terms of the complexity of their DSL series. This set of parameters also allows distinguishing between time series describing fine- or smooth-structure of the DSL regime by using the Complexity Index, CI. Results of previous monofractal analyses also permits establishing comparisons between smooth-structures, relatively low correlation dimensions, notable predictive instability and anti-persistence of DSL for European areas, sometimes submitted to long droughts. Relationships are also found between the CI and the mean absolute deviation, MAD, and the optimum autoregressive order, OAO, of an ARIMA(p,d,0) autoregressive process applied to the DSL series. The detailed analysis of the discrepancies between empiric and predicted DSL underlines the uncertainty over predictability of long DSL, particularly for the Mediterranean region.

Keywords: DSL series, drought regime, multifractal detrended fluctuation analysis, Hölder and Hurst exponents, ARIMA process, European pluviometric network.

1. Introduction

Dry spell lengths, DSL, defined as the number of consecutive days with daily rain amounts lowering a certain threshold, represent a valuable magnitude to analyse several aspects of drought episodes. Besides studies based on “the standardised precipitation index”, SPI (McKee et al, 1993, 1995; Hayes et al., 1999; Lana et al, 2001), or “dry days since last rainy day”, DDSLR (Aviad et al. 2009; Reiser and Kutiel 2010; Lana et al., 2012), DSL have also been used along last years to characterise several patterns of the drought regime in Europe (Brunetti et al., 2002; Anagnostopoulou et al., 2003; Kostopoulou and Jones, 2005; Schmidli and Frei, 2005; Cindrić et al., 2010; Carvalho et al., 2013). The analysis of DSL series would permit to improve the knowledge about physical mechanisms and time trends governing drought regimes, particularly concerning southern Europe, as severe drought episodes frequently occur in the Mediterranean region (Lana et al., 2006, 2008b; Livada and Assimakopoulos, 2007; Diodato and Bellocchi, 2008; Nastos and Zerefos, 2009; García-Ruiz et al., 2011; Heinrich and Gobiet, 2011; Zolina et al., 2013). Specifically, the statistical distribution of DSL in Europe, with special emphasis on expected DSL for several return periods, and the assignment of statistical parent distributions to spatial clusters of rain gauges have been recently analysed (Serra et al, 2013, 2014). European partial duration series of DSL have also been recently studied (Serra et al., 2015).

A different approach may be the application of fractal theory. First, the rescaled range analysis (Turcotte, 1997) leads to obtain the Hurst exponent, which permits characterising the persistence, anti-persistence or randomness of DSL. Second, the self-affine character of DSL (Hausdorff exponent) and the possibility of modelling DSL series by fractional Gaussian noise series can be tested (Mandelbrot and Wallis, 1969; Malamud and Turcotte, 1999). And third, in agreement with the reconstruction theorem (Diks, 1999), it is possible to quantify several aspects of the physical mechanism governing DSL, like the complexity and chaotic behaviour (correlation and strange attractor dimensions), loss of memory (Kolmogorov entropy) and predictive instability (Lyapunov exponents and Kaplan-Yorke dimension). A complete analysis of the complexity and predictive instability of European DSL series, based on the rescaled range analysis and the reconstruction theorem, can be found in Lana et al. (2010).

The multifractal analysis of the DSL series (Kantelhardt et al., 2002) is now proposed through four parameters: the Hurst exponent, *H*, estimated from the multifractal spectral curve,

$f(\alpha)$, the range, *W*, of the Hölder exponent, α, the asymmetry, *B*, of $f(\alpha)$, and the critical Hölder exponent, α_0, for which the maximum of $f(\alpha)$ is reached. These parameters permit to characterise the complexity of every DSL series by using the Complexity Index, CI. After applying the Principal Component Analysis (Jolliffe, 1986; Preisendorfer, 1988) and clustering algorithms (Kalkstein et al., 1987; Davis and Kalkstein, 1990), DSL series may be spatially grouped in terms of their degrees of complexity, taking as variables for the classification those four multifractal parameters. In this way, regions with simple (smooth-structure) or complex (fine-structure) drought predictability can be distinguished.

The description and quantification of the predictability of DSL are complemented by the ARIMA(p,d,0) autoregressive process (Box and Jenkins, 1976), previously applied in Climatology, for instance, to the North Atlantic Oscillation, NAO, series (Stephenson et al., 2000; Mills, 2004). The mean absolute deviation, MAD, permits to quantify the goodness of fit between empiric DSL and those generated by the ARIMA(p,d,0) process, while the optimum autoregressive order, OAO, points out the order with minimum MAD. Additionally, valuable information can be obtained by analysing the discrepancies (residuals) between empiric and predicted DSL for every analysed rain gauge.

The main objectives of this paper are, first of all, to extend the multifractal analysis of DSL series to Europe. Specifically, areas of simple or complex predictability are well bounded and comparisons are made with previous results derived from the application of the reconstruction theorem and the rescaled range analysis (Lana et al., 2010). And second, to establish relationships between CI derived from multifractal theory and MAD and OAO values based on ARIMA(p,d,0) autoregressive algorithms, and to highlight advantages and shortcomings of predictability based on this autoregressive process.

The contents of the paper are organised as follows. The database is introduced in Section 2; the methodology to obtain the multifractal spectrum, the Complexity Index, CI, the spatial distribution of clusters and a description of the ARIMA(p,d,0) process are detailed in Section 3; results concerning multifractal spectrum parameters, complexity measures, clustering and autoregressive predictions are developed in Section 4. Finally, Conclusions are summarised in Section 5.

2. Database

Daily precipitation data for the years 1951-2000 have been compiled from 267 rain gauges in Europe and neighbouring countries. Most of these series (236) come from the European Climate Assessment and Dataset (ECA&D), http://eca.knmi.nl/. All these series are public and non-blended and their quality has been analysed by ECA&D (Klein Tank et al., 2002; Wijngaard et al. 2003; Klok and Klein Tank, 2009). The rest of series come from the *Agencia Estatal de Meteorología* (Spanish Ministry of Environment) and standard homogeneity and quality controls have been previously applied (Lana et al., 2008a).

Figure 1a depicts the spatial distribution of the available stations. Dense rain gauge coverage is observed in Western Europe, except for Italy, Great Britain and the north of the Scandinavian Peninsula, where it is not so dense. Results for Turkey and Israel are also included in the maps, but they only provide a broad approach for these regions. Interpolation of scarce spatial data could generate computational artefacts leading to some unrealistic spatial patterns, even taking into account additional data as height above sea level or orographic effects. All rain gauges have a minimum continuous recording period of 40 years, and the series are complete (50 years) for 102 out of 267 stations. Figure 1b shows, year by year, the number of stations with complete recordings. Most of data series are continuous for the period 1955-1990, the number of available records diminishing at the beginning (1951-1955), and especially at the end (1990-2000) of the recording period. If a DSL is likely to be incomplete due to lack of record continuity, it is rejected. This does not constitute a relevant shortcoming on account of the large enough statistical population of DSL.

Thresholds commonly used to define a DSL are 0.1, 1.0, 5.0 and 10.0 mm/day (Kutiel and Maheras, 1992; Martín Vide and Gómez, 1999; Anagnostopoulou et al. 2003; Serra et al. 2006, 2013, 2014; Lana et al 2008b; Cindrić et al., 2010). The current study is constrained to the threshold of 0.1 mm/day (the assumed resolution of pluviometers), with daily excess (shortage) defining a wet (dry) day, thus being obtained a long enough data basis. As a global description of the 267 DSL series, their numbers of dry spells, N_{DSL}, range from 1061 up to 3203, with an average of 2621 and a standard deviation of 477. The histogram of N_{DSL} (Figure 1c) clearly depicts a skewed distribution toward high N_{DSL}. Two examples of DSL series from northern and southern Europe are shown in Figure 1d. As expected, the southern Europe example is characterised by a relatively low N_{DSL} and a non negligible number of long DSLs.

The opposite example could be that corresponding to higher latitudes, with larger N_{DSL} and shorter DSLs.

The relevance of the DSL concept and its relationship with droughts is manifested, for example, by searching for the longest DSL, Lmax, obtained for every rain gauge along the recording period. In agreement with Serra et al. (2014), values of Lmax exceeding three months are detected in the southern Mediterranean coast, for latitudes south of 40ºN. The high spatial gradient of Lmax obtained in the Mediterranean region contrasts with the notable homogeneity and low Lmax values at northern latitudes.

3. Methodology

3.1 Multifractality.

Multifractals are complex self-similar objects that consist on differently weighted fractals with different non-integer dimensions. Thus, the fundamental characteristic of multifractality is that scaling properties may vary in different regions of the system (Dutta, 2010; Ghosh et al., 2012). If scaling properties are kept, whatever the region, the signal is known as monofractal. The multifractal detrended fluctuation analysis (MF-DFA) has been introduced as a reliable characterization of multifractal non-stationary and stationary time series (Kantelhardt et al., 2002). It is based on the identification of the scaling of the q^{th}-order moments depending on the signal length. The MF-DFA surpasses in quality and simplicity to previous algorithms such as the multifractal box counting (MF-BOX) (Feder, 1988) or the wavelet transform modulus maxima (WTMM) (Muzy et al., 1994) algorithms. A detailed description of the steps leading to the right application of the MF-DFA algorithm and the concepts of q^{th}-order fluctuation function, $F_q(s)$, also named standard partition function, can be found in Burgueño et al. (2013).

A variety of continuous multifractal time signals have been analyzed by the MF-DFA, including climatological series as hourly or daily wind speeds (Kavasseri and Nagarajan, 2005; Feng et al, 2009), daily temperatures (Lin and Fu, 2008; Yuan et al., 2013; Burgueño et al., 2013) and global monthly temperature anomalies (Mali, 2014), as well as not so typical magnitudes as lightning initiation process (Gou et al., 2010).

3.2 The singularity spectrum

According to Kantelhardt et al. (2002), the q^{th}-order fluctuation function, $F_q(s)$, follows a power law with exponent *h(q)*, which is known as the generalized Hurst exponent. The singularity spectrum, $f(\alpha)$, can be related to *h(q)* via a Legendre transform:

$$\alpha = h(q) + q\frac{d[h(q)]}{dq} \quad \overset{Legendre}{\leftrightarrow} \quad f(\alpha) = q\cdot[\alpha - h(q)] + 1 \qquad (1)$$

where α is the singularity strength or Hölder exponent, while $f(\alpha)$ denotes the dimension of the subset of the series. The multifractal scaling exponent is

$$\tau(q) = qh(q) - 1 \qquad (2)$$

α being expressed as

$$\alpha = \frac{d\tau(q)}{dq} \tag{3}$$

The characteristics of the singularity spectrum $f(\alpha)$ provide a new way of comparing signals, because it describes the dimensions of subsets of the series characterized by the same singularity strength α. Designing α_0 as the singularity strength with maximum spectrum, a small value of α_0 means that the underlying process loses fine-structure, that is, it becomes more regular in appearance; conversely, a large value of α_0 ensures more complexity. In this sense, the Hurst exponent depicts a linear relationship with α_0, as proved by the results. The shape of $f(\alpha)$ may be fitted to a quadratic function around the position α_0,

$$f(\alpha) = A(\alpha - \alpha_0)^2 + B(\alpha - \alpha_0) + C \tag{4}$$

where *C* is an additive constant equal to 1. Coefficient *B* indicates the asymmetry of the spectrum, being zero for a symmetric spectrum. A right-skewed spectrum, $B > 0$, indicates relatively strongly weighted high fractal exponents (with “fine-structure"), while left-skewed shapes, $B < 0$, point to lower ones (more regular or smooth-structure). Spectral width, *W*, is defined as

$$W = \alpha_2 - \alpha_1 \tag{5}$$

with $f(\alpha_1) = f(\alpha_2) = 0$, being α_2 larger than α_1, and the wider the spectral content, the stronger is the multifractality. In other words, the wider the range of possible fractal exponents, the “richer" is the process in structure. A signal with a high value of α_0, a wide range *W* of fractal exponents, and a right-skewed shape, $B > 0$, is more “complex" than one with the opposite characteristics (Shimizu et al., 2002). For monofractal series, the width of the spectrum would be zero and $h(q)$ would be independent of q. Hence, from Equation (1), there will be a unique value for both α and $f(\alpha)$, α being the Hurst exponent, *H*, and $f(\alpha)$ being equal to 1.

With respect to the appropriate range of the q^{th}-order, two questions have to be considered. First, an accurate revision of the goodness of fit of the q^{th}-order fluctuation functions, $F_q(s)$, to a power law is recommended. And second, a relevant property of the Legendre transform has to be considered. From a pure analytical point of view, the theoretical maximum (α_2) and minimum (α_1) Hölder exponents for which the multifractal spectrum is zero correspond to $q \to -\infty$ and $q \to +\infty$ respectively. In consequence, the expected range of the q^{th}-orders should be $(-\infty, +\infty)$. Computational instabilities for very high (positive and negative) q^{th}-orders may lead to departures from a power law for the q^{th}-order fluctuation function. Additionally, Burgueño et al. (2013) detected significant departures from the expected quadratic function (Equation 4) modeling the multifractal spectrum when high q^{th}-orders were applied. Considerations made by Ivanov et al. (2001), who assume that the appropriate range of q^{th}-orders depend on the series length, would be in agreement with the detection of q^{th}-orders for which the mentioned power law is not well accomplished. Then, the verification of spectral contents satisfying Equation 4 is suggested as an additional test to decide the optimum q^{th}-order range.

3.3 Autoregressive process

The autoregressive integrated moving average ARIMA(p,d,0) model (Box and Jenkins, 1976) assumes that

$$\Delta^d x(i) = \theta + \mu\, x(i-1) + \sum_{k=1}^{p} \delta_k \Delta^d x(i-k) + a_i \quad (i = p+2, \dots, N) \tag{6a}$$

Where $\{x\}$ is a set of N empirical data, Δx is the set of first differences $\Delta x(i) = [x(i+1) - x(i)]$, with $\Delta^d x(i-k) = [x(i-k+1) - x(i-k)]^d$, $\{\theta, \mu, \delta_1, \dots, \delta_p\}$ are the parameters of the autoregressive process, $\{a\}$ is a noise series and d is a real number. Alternatively, the ARIMA(p,d,0), with d = 1.0, can be written as

$$x(i) = \theta + \sum_{k=1}^{p} \delta_k x(i-k) + a_i \quad , i = p+1, \dots, N \tag{6b}$$

where time series $\{x\}$ are directly used instead of first differences. With the aim of avoiding singularities in the linear system of equations used to estimate $\{\theta, \mu, \delta_1, \dots, \delta_p\}$, parameter μ is implicitly included in parameter δ_1. The corresponding Equation (6b) is usually designed as autoregression, AR(p). The resulting system of linear equations, disregarding the stochastic component $\{a\}$, can be represented in terms of matrix formulation by

$$Z = AW \tag{7a}$$

with *Z* the $\{x(p+1), x(p+2), \dots, x(n)\}$ vector, n the number of empiric elements belonging to series $\{x\}$, the (n-p-1, p+1) matrix *A* with elements multiplying parameters $\{\theta, \delta_1, \dots, \delta_p\}$ and a p+1 dimension vector *W* containing the parameters to be solved from the linear system of

Equation (7a). The components of vector W can be estimated by multiplying Equation (7a) by the transposed A matrix, A^T

$$A^T Z = A^T A W \tag{7b}$$

and remembering that the symmetric matrix $A^T A$ can be decomposed in two triangular matrices. Then, it is straightforward to obtain the values of parameters $\{\theta, \delta_1, \ldots, \delta_p\}$ taking advantage of the Crout's algorithm (Press et al., 1992).

A convincing solution of Equation (6b) demands some criterion to decide the optimum autoregression order, OAO. The decision can be taken by searching for the OAO leading to a minimum of a convenient goodness of fit index. The selected index is the mean absolute deviation, MAD (Stephenson et al., 2000),

$$MAD = \frac{1.483}{(N-p-1)} \sum_{k=1}^{N-p-1} \left| x(k) - x^*(k) \right| \tag{8}$$

being $x(k)$ empiric DSL and $x^*(k)$ those predicted by parameters derived from the ARIMA(p,1,0) process.

4. Results

4.1 A representative example of multifractal spectrum

As an example, DSL multifractal spectrum parameters for the Fabra Observatory (NE Iberian Peninsula) are shown in Figures 2a-2d. It should be noted that their general features are common to the rest of DSL series obtained from the European pluviometric network records. Nevertheless, it is important to take into account that the different parameter values for every DSL series characterize their specific complexity. Figure 2a depicts three selected q^{th}-order fluctuation functions in a log-log scale. In spite of fluctuations, their evolution with length *s* is well fitted by a power-law. Figure 2b describes the evolution of the generalized Hurst exponents, *h(q)*, with *q*. Remembering that, for stationary series, *h(q=2)* is the Hurst exponent *H*, this DSL series is characterized by a strong randomness (*H* very close to 0.5). As a general rule, also observed for daily extreme temperature series (Burgueño et al., 2013), a second order polynomial on *q* fits well *h(q)*. The multiscaling fractal exponent, $\tau(q)$, and Hölder exponent, $\alpha(q)$, are shown in Figures 2c and 2d, being worth mentioning signs of linear behavior of τ and the clear linear evolution of $\alpha(q)$, with *q* within the ±15 range. The corresponding multifractal spectrum, $f(\alpha)$, is shown in Figure 3. It should be underlined that the empiric normalized value of 1.0 is reached for $\alpha = \alpha_0$, but the fit of empirical $f(\alpha)$ to a second order polynomial is not perfect. To summarize, DSL series corresponding to the Fabra Observatory are characterized by the following patterns:

1) Strong randomness, with weak signs of persistence, as Hurst exponent exceeds slightly 0.5.
2) A moderate Hölder exponent width, *W*, close to 0.38, with $\alpha_1 = 0.36$ and $\alpha_2 = 0.74$, thus suggesting a moderate multifractality in comparison with other DSL series.
3) A centered multifractal spectrum around $\alpha_0 \approx 0.55$ and a notable asymmetry (right-skewed shape) with *B* = +1.5. Both parameters indicate a better description of the "fine-structure" than the "smooth-structure" of the DSL series.

4.2 Spatial distribution of multifractal parameters

With the aim of obtaining a better picture of the multifractal variety of the different DSL series, Figure 4 shows the spatial distribution of the four parameters, *H*, α_0, *W* and *B*, describing multifractal spectra. Minimum α_1 and maximum α_2 Hölder exponents have not been considered as fundamental parameters, given that they could be substituted by the critical Hölder exponent α_0 and the spectral width *W*, as verified after through the Principal Component Analysis.

Figures 4a and 4b show quite similar *H* and α_0 patterns across Europe. It is worth mentioning some signs of a decreasing tendency from high to low latitudes. The strong negative gradient with latitude detected in Turkey (probably due to computational artifacts, attributable to low rain gauge density as mentioned before) and low values obtained in the SW of the Iberian Peninsula are indicative of a strong anti-persistence. Spatial patterns for the asymmetry shown in Figure 4c are difficult to interpret, as areas of "fine-structure" (*B* > 0) and "smooth-structure" (*B*< 0) of the DSL regimes are not generally distributed according to latitudinal or longitudinal factors or vicinity to Atlantic or Mediterranean coasts. The spectral width, *W* (Figure 4d), varies within the (0.2, 0.6) interval for most of Europe. The behavior of the DSL would be qualified as close to monofractal only for a very small area of the Iberian Peninsula, with *W* lowering 0.2. In short, it is difficult to obtain simplified spatial patterns describing the degree of multifractality and "fine/smooth" structure of the DSL regime. Alternatively, a clustering process (Davies and Kalkstein, 1990) to group DSL series with similar patterns is applied in Section 4.3.

Before the clustering process, some relationships between *H*, α_0, α_1, α_2 and *W* can be analyzed. Figure 5a depicts a clear linear relationship between the Hurst exponent and the critical Hölder exponent, such a pattern also found for daily extreme temperature regimes (Burgueño et al., 2013), which are characterized by strong persistence. Consequently, it could be proposed that increasing Hurst exponents would reinforce the multifractal character of the climatic series analyzed. With respect to minimum and maximum Hölder exponents, whereas α_2 tends to increase whit α_0 (Figure 5b), α_1 does not depict so clear evidences of a decreasing tendency. This fact would indicate that an increase of "fine-structure" of the DSL series would be accompanied by an increase of *W*, being reinforced then the multifractal character and the complex structure of the series. This hypothesis would be confirmed by some signs of an increasing tendency of *W* with α_0.

4.3 Principal Component Analysis

Previous to the clustering process, a revision of ranges corresponding to the Hurst and critical Hölder exponents, *H* and α_0, spectral asymmetry, *B*, and spectral width, *W*, is done, these parameters being obtained from the MF-DFA successful application to 258 out of 267 DSL series. The nine rain gauges for which it has not been possible to obtain the multifractal spectrum are detailed in Figure 1a. Taking into account they are widespread throughout Europe, the different pluviometric regimes would not be the reason for a successful or unsuccessful quantification of multifractal parameters. Most of H and α_0 values are within a relatively narrow interval from 0.4 to 0.6. Many of the asymmetries B are within the ±1.5 interval and a high number of multifractal spectra have a width W within the (0.25, 0.55) range. This distribution of the four multifractal parameters permits a better explanation and discussion of rain gauge spatial clusters obtained after the application of the Principal Component Analysis, PCA (Jolliffe, 1986; Preisendorfer, 1988), and the clustering process known as average linkage algorithm (Kalkstein et al., 1987; Davis and Kalkstein, 1990). Table 1 summarizes the results of the PCA. After the application of the PCA algorithm to the covariance matrix of original data, the four original variables, H, α_0, B, W, are substituted for three principal components, PC1, PC2 and PC3, which explain 99.8% of data variance. A review of the factor loadings, FL, quantifying the relationship among original variables and PCs, clearly manifests that H and α_0 are strongly correlated with PC1, W with PC2, and B with PC3. Consequently, the values of the four original variables are substituted for the factor score values, FSC, associated with the three chosen PCs.

Figure 6 describes the spatial distribution of the three FSCs throughout Europe. Only FSC1, essentially representing H and α_0, shows a similar pattern to those of Figures 4a and 4b. Spatial distributions of FSC2 and FSC3 are much more spotted, in such a way that establishing any spatial patterns becomes difficult. Consequently, it is very likely that the clustering process will lead to define homogeneous groups of rain gauges without excessive spatial coherence. This clustering process is based, as mentioned before, on the average linkage algorithm and, more specifically, on the similarity index, L_{ij}, concept. This index is defined as

$$L_{ij} = D_{ij}^{\ 2} + \nu_i + \nu_j \qquad (9)$$

with D_{ij} the Euclidean distance between centroids of clusters i and j, and ν_i and ν_j the corresponding within-group variances. L_{ij} systematically increases whatever the merged pair (i, j) of clusters. Then, in each step of the AL process, only the merging of the pair of clusters associated with the minimum increase of L_{ij} is considered. When an accepted merging of clusters is associated with a sharp increase on L_{ij}, the previous cluster configuration is chosen as the optimum and the clustering process finishes.

Figure 7 depicts the evolution of L_{ij} of the average linkage algorithm with the decreasing number of clusters. It should be remembered that every time the algorithm merges two clusters of very different characteristics, a step is clearly observed in the L_{ij} curve. Looking at Figure 7, two configurations of 9 or 14 clusters could be accepted. The configuration of 14 clusters is finally chosen. The 9 clusters configuration is rejected by having a large cluster and a high number of very small, or even singular, clusters. Additionally, the configuration of 9 clusters would neglect valuable spatial variety from the viewpoint of DSL multifractality.

The spatial distribution of 14 clusters is shown in Figure 8. The number of rain gauges belonging to every cluster, their centroid coordinates (average and standard deviation of *H*, α_0, *B* and *W*) and range of every multifractal parameter are given in Table 2. Rain gauges belonging to clusters C1, C2, C3 and C7 are located at latitudes approximately north of 40ºN. Clusters C4, C5 and C13 are mainly associated with areas of the Iberian Peninsula and the rest of clusters, all of them with a small number of elements, are spread throughout Europe, including Mediterranean countries. In agreement with the distribution of clusters, different European zones share more than one DSL regime. Then, to establish simple relationships between geographical factors (orography, latitudes, longitudes and vicinity to the Atlantic Ocean and the Mediterranean Sea, for instance) and DSL regimes becomes quite difficult. Nevertheless, in accordance with the well differentiated centroids summarized in Table 2, the classification of rain gauges within specific clusters becomes coherent.

4.4 Complexity Index

Another complementary point of view of the DSL regime in Europe consists in the application of the complexity index, CI (Burgueño et al., 2013), which is defined through the addition of the standardized α_0, $z(\alpha_0)$, *B*, *z(B)*, and *W*, *z(W)*, parameters, in agreement with the concept of complexity developed by Shimizu et al. (2002)

$$Z = z(\alpha_0) + z(B) + z(W) \tag{10}$$

For a better interpretation of this addition, *Z* is also normalized, *CI* being defined as

$$CI(Z) = \frac{Z - <Z>}{S(Z)} \tag{11}$$

<Z> being equal to zero and *S(Z)* the standard deviation of *Z*. Definition given by Equation (11) permits a straightforward quantification in terms of “high complexity” ($Z \geq 0$) and “low complexity” ($Z < 0$). Table 3 summarizes the number of DSL series belonging to different *Z* intervals given in standard deviation units. It is outstanding the concentration of *CI(Z)* within the ±1.0 interval (in standard deviation units), close to 70%, and the very low number of series with very smooth- (9 out of 258) and very fine-structure (2 out of 258). As a general feature, DSL series associated with fine-structure ($Z \geq 0$) are slightly predominant (close to 56%) in comparison with those related to smooth-structure ($Z < 0$), which represent close to 44% of series. The spatial distribution of *CI(Z)* is depicted in Figure 9. Similarly to the spatial distribution of clusters, a clear spatial coherence is not found, being detected small areas sharing DSL series with smooth and fine-structures, especially in Central and Western Europe. Nevertheless, south-western Iberian Peninsula, southern Italy and the Adriatic coast are characterized by smooth-structure of the DSL regime. North-eastern regions of Europe have a predominant (almost unique) fine-structure of the DSL regime.

4.5. Comparisons with monofractal results.

Comparisons with monofractal results (Lana et al., 2010) are aimed to assess if Hurst exponents derived from rescaled range, R/S, analysis and from MF-DFA are coincident and, in addition, to establish some relationship between DSL series characterized by smooth/fine-structure and monofractal properties derived from the reconstruction theorem, namely, correlation dimension, Kolmogorov entropy and Lyapunov exponents.

Hurst exponent values obtained from the R/S analysis (within the 0.2–0.7 interval) are generally slightly shifted from those derived from the MF-DFA (within the 0.1–0.6 range). A good number of Hurst exponents derived from R/S analysis are related to randomness with slight signs of persistence ($H \approx 0.6$) of DSL series, whereas a significant number of Hurst exponents estimated by MF-DFA lying within the (0.4, 0.6) interval points also to predominant randomness, but with weak signs of both persistence or anti-persistence. Anti-persistence is detected, both for R/S and MF-DFA, at south and south-western Iberian Peninsula and, additionally, at southern Italy and Sicily, in agreement with MF-DFA. The highest Hurst exponents are predominantly found at high latitudes (North-Eastern Europe), close to Atlantic Ocean and Baltic Sea. In short, the results concerning Hurst exponents derived from R/S and MF-DFA, and its spatial distributions, could be assumed as roughly coincident.

With respect to possible relationships among monofractal variables and fine/smooth-structures defined by CI, the incidence of the Kolmogorov entropy should be discarded as their values are scattered throughout Europe without clear spatial patterns. Consequently, the loss of memory of the physical system should be a common feature, but not a discriminating factor to distinguish between smooth and fine-structures. On the contrary, spatial distributions of low values of the correlation dimension (giving an estimation of the minimum number of nonlinear equations required to describe the physical process) and high values of the Lyapunov exponents (quantifying the predictive instability of the physical system) could be associated with smooth-structures. From the point of view of CI, south and south-west of the Iberian Peninsula, south of Italy and Sicily are characterised by smooth-structure. Almost the same areas are linked to the highest values of the first Lyapunov exponent, which strongly governs the predictive instability, and the lowest correlation dimensions. Consequently, it could be proposed the hypothesis that smooth-structures of the DSL regime would be described by a relatively simple system of nonlinear equations. Nevertheless, as a counterpart, the predictive instability would be notable. Given that many areas for the rest of Europe share smooth- and fine-structures, the conjecture of a certain relationship between fine-structure and high correlation dimensions and low Lyapunov exponents is difficult to be ascertained.

It is also worth of mention that, in Mediterranean countries, with quite usual drought episodes (some of them unusually long)**,** smooth-structures are compatible with low correlation dimensions and high Lyapunov exponents, as well as with Hurst exponents indicating anti-persistent behaviours.

4.6 Autoregression results

Figure 10 illustrates the dependence of MAD and OAO on latitude. A roughly linear decrease of MAD, from 20 days up to approximately 4 days, with increasing latitudes ranging from 30°N to 45°N, is clearly observed. For higher latitudes, this almost linear decreasing tendency is smoother, MAD reaching values close to 2 days for the highest latitudes. Even though OAO does not show a so clearly linear evolution, an increasing tendency is found up to latitudes close to 45°N. For higher latitudes, values of OAO remain within the 180-200 range, this last autoregressive order being the largest lag analyzed. These changes on tendencies close to 45ºN were also detected by Serra et al. (2015), who analyzed, among other questions, the evolution with latitude of the L-moment coefficient of variation of the longest DSL. For a better understanding of OAO, It has to be underlined that for the same OAO, the required recording period depends on the DSL, especially when comparing dry and wet pluviometric regimes or very different latitudes. As an example, an OAO value equal to 150 (150 consecutive dry spells) represents a wide range of recording periods for the different rain gauges of the European database. These periods vary from 10 months (Bjoernoeya, Norway, 74.59ºN, 19.02ºE) with a wet pluviometric regime and a short average DSL of 2.1 days, up to 5 years (Alicante, Spain, 38.01ºN, 0.71ºW) with a dry pluviometric regime and a longer average DSL of 12.5 days.

With respect to the distribution of MAD and OAO values, it is worth mentioning that MAD values are mainly concentrated within the 2-5 days interval (197 out of 267 series), without signs of a Gaussian distribution, and most OAO values equal to or exceeding 175 (233 out of 267 series). Figure 11 shows the histogram of the residuals of all DSL series, defined as the difference between empiric values and those reproduced with the OAO. These residuals are mainly concentrated within the (−4, 2) days range, representing close to 70% of DSL in Europe, and without signs of a Gaussian distribution.

Figure 12 describes the spatial distribution of MAD and OAO for the 267 DSL series. Both parameters have been classified in five intervals corresponding to 0-20th, 20-40th, 40-60th, 60-80th and 80-100th percentiles of their empirical cumulative distributions. The first class of MAD percentiles (0-20th) covers latitudes exceeding 50ºN. Most of series with MAD belonging to the second interval are within the (45-55ºN, 5-15ºE) latitude and longitude ranges. The third interval mostly corresponds to a long fringe (45-60ºN, 0-45ºE) with a low rain gauge spatial density for longitudes exceeding 15ºE. The fourth interval is predominantly associated with latitudes 40-60ºN, whatever the longitude. Finally, the fifth interval is mainly linked to the Iberian Peninsula and some Mediterranean coastal rain gauges, with latitudes south of 45ºN. The five classes of OAO percentiles are however, scattered throughout Europe, without any clear spatial pattern.

Two examples of MAD values and DSL residuals for very different latitudes and pluviometric regimes are shown in Figure 13a (Varexjoe, Sweden) and Figure 13b (Zaragoza, Spain).

Whereas MAD is close to 3 days for Varexjoe, it attains 7 days for Zaragoza. Additionally, the standard deviations of the residuals are close to 3 and 7 days respectively. After a detailed review of DSL residuals, it is observed that the percentage of overestimated DSL (negative residuals) for both series is quite similar (close to 66%). Nevertheless, a notable difference appears when the most negative residuals are compared. Whereas the largest DSL overestimation is 5 days for Varexjoe, it is 12 days for Zaragoza. Another similar feature is observed when comparing positive residuals (underestimated DSL) exceeding one standard deviation. In both cases the percentage of underestimations is relatively small (12%). Nevertheless, notable differences are detected once again when comparing the maximum underestimation. Whereas it could be assumed as notably high (30 days) for Varexjoe, it should be considered absolutely inappropriate for Zaragoza (70 days).

Table 4 summarizes the main features of the residuals for the 267 DSL series. It should be accepted that predicted DSL are not biased given that the average of the residuals is very close to zero, whatever the DSL series. Their standard deviation varies within a wide range from 1.7 to 29.5 days. The whole set of European DSL residuals would be then represented by a moderate average standard deviation of 4.6 days. In spite of this small length, at least one of the DSL series would be predicted by the AR(p) process without an excessive error, as manifested by the mentioned maximum standard deviation of 29.5 days**.** Skewness and kurtosis for every one of the 267 sets of residuals indicate that a Gaussian distribution should be discarded. These empirical skewness and kurtosis notably depart from values expected for a Gaussian distribution, which should be equal to 0.0 and 3.0 respectively.

Percentages $R(>2\sigma)$ of positive residuals (underestimated DSL) exceeding two standard deviations are not very relevant. Minimum and maximum percentages of 2.9% and 5.3% respectively, as well as an average of 4.7% for all DSL series, summarize the behavior of the largest positive residuals. An excessive overestimation of DSL is quantified by the percentage $R(<-\sigma)$ of negative residuals lowering at least minus one standard deviation. The overestimation of DSL is not a relevant problem, given that for the whole database analyzed, only 4.6% of DSL have been overestimated and the maximum detected ratio for a single DSL series attains 8.3%.

The hypothesis that the positive residuals (underestimated DSL) are higher than the negative residuals (overestimated DSL) is not only suggested by the results given in Table 4, but also searching for the maximum positive residual for every one of the 267 DSL series. These extreme residuals range from 14.7 to 286.8 days. The first length could be considered high, but in some way acceptable. The second length is unacceptable. Conversely, minima of negative residuals of the 267 series range from −30.0 to −2.7 days.

For a better comprehension of advantages and shortcomings of the AR(p) process applied to the DSL prediction, additional reasons for these large values of residuals have to be found, as well as for the high standard deviation detected for at least one series of residuals and to null ratios of residuals lowering −σ or within the range (σ, 2σ). Four DSL series with these characteristics are found. They come from pluviometric records in South-eastern Mediterranean, with longitudes within the (25.2ºE − 35.5ºE) range and low latitudes within a narrow fringe (31.8º – 35.3ºN), where dryness is well-known. All of them have a very high maximum underestimation of real DSL (from −152.9 to −286.8 days) and maximum outstanding overestimation of real DSL (from 25.3 to 30.0 days), accompanied by high standard deviation of residuals, varying from 18.7 to 29.5 days. It is also worth mentioning that the percentage of overestimated DSL lowering the residual −σ is null for two series, and almost null for the other two.

A possible relationship between CI and MAD is shown in Figures 14a. Most of MAD values range from 2 to 5 days and are associated with CI within the (−2.0, +2.5) interval. Figure 14b shows signs of a possible relationship between CI and OAO too. Whereas for values of CI within the ±2.5 interval, a concentration of high values of OAO (exceeding 145) is observed, a positive trend in OAO values lowering 145 is detected for the (−4.0, +1.5) CI range.

5. Discussion and conclusions

The present MF-DFA widens the monofractal analysis of the European DSL regime (Lana et al., 2010). The degree of complexity of DSL series derived from daily pluviometric records has been quantified through multifractal parameters *H*, α_0, *B and W*, and their synthesis, the complexity index, CI.

With respect to possible relationships between monofractal and multifractal parameters, it is worth mentioning that DSL regimes characterized by physical mechanisms with a low number of nonlinear equations (correlation dimension) and notable predictive instability (high positive Lyapunov exponents) are generally associated with CI negative values corresponding to smooth structures. On the contrary, fine structures (positive CI) are related to mechanisms requiring a high number of nonlinear equations. As a counterpart, the predictive instability would not be as high as for smooth structures. It is also worth mentioning that the loss of memory of the system (Kolmogorov entropy) is a common feature to all DSL regimes, without clear relationships with multifractal parameters.

From the point of view of multifractal parameters across Europe, signs of a North-east to South-west decreasing trend is only observed for the Hurst exponent, *H*, and the critical Hölder exponent, α_0. For the other two multifractal parameters, spectral width, *W*, and spectral asymmetry, *B*, clear spatial patterns are not obtained. Neither the quantification of DSL complexity, in terms of CI, nor a clustering process, based on a previous PCA, have permitted a spatially coherent distribution of fine- and smooth-structures. This fact could be considered a shortcoming of the multifractal analysis. Alternatively, strong dependence of the DSL regime on topographic parameters (height above sea level and orographic slope for instance) and on other local and atmospheric dynamic variables (dominant wind direction, vicinity to orographic barriers such as the Alps and the Pyrenees and vicinity to Atlantic or Mediterranean seas) could be an explanation to this lack of coherent spatial clustering.

With respect to the ARIMA process, several questions concerning the residuals of predicted DSL should be mentioned. First, they are not distributed according to a Gaussian model. Second, predicted DSL are not biased given that residual averages are for every one of the DSL series almost null. Third, whereas the overestimation of DSL is not a very relevant problem, their underestimation may become very important, especially for pluviometric series characterized by long DSL. Fourth, similar to spatial patterns of multifractal parameters, the spatial distribution of MAD is also complex and the OAO is not coherently distributed. And fifth, all DSL residuals characterized by high standard deviations and very high underestimations belong to pluviometric records associated with very dry regimes at low latitudes in the Mediterranean area. In other words, DSL for dry regimes (usually low latitudes) are more difficult to be predicted than for wet regimes (generally high latitudes). In fact, DSL belonging to the aforementioned very dry regimes should be qualified as very hardly predictable in agreement with results obtained from the ARIMA process. An acceptable prediction for these very dry regimes might be achieved with the ARIMA(p,d,q) model, applied before for instance to monthly rainfall (Wang et al., 2014). Also, methods based on conditional Poisson distribution and Monte Carlo algorithms (Jung et al., 2006) might be considered. In short, the prediction of these extreme DSL would be beyond the main objectives of this paper.

It is worth noticing the high OAO values for most of DSL series. These high values could be in agreement with the Kolmogorov entropy, which quantifies the loss of memory of the physical mechanisms along the time process. Prediction of DSL series affected by a high loss of memory would require then a notable number of previous steps (DSLs) to obtain acceptable predictions.

Similarly to comparisons made between monofractality and multifractality, coincidences and disagreements between multifractal and ARIMA results are worthy to be mentioned. A clear relationship between CI and MAD has not been found. Nevertheless, low MAD, from 2 to 5 days, are usually associated with fine and smooth structures with CI ranging from −2 to +2. The highest MAD are detected for CI within the range (−4, −3), corresponding to smooth structures from the multifractal point of view. With respect to a possible relationship between CI and OAO, a great number of DSL series are characterized by OAO exceeding 150 and belonging to fine or smooth structures. A few cases of OAO lowering 150 with CI varying from −4 (very smooth structures) to +1 (moderate fine structures) are characterized by a clear linear increasing trend.

Finally, latitudinal changes on MAD and OAO are outstanding, roughly at 45^{o}N, thus reinforcing the hypothesis of a certain dependence on latitude. As mentioned in Section 4.6, a similar behavior has been observed in recent statistical analyses of long DSL.

	EV(%)	FL_H	FL_{α_0}	FL_W	FL_B
PC1	49.1	0.992	0.970	0.180	-0.084
PC2	25.5	0.101	0.205	0.983	0.030
PC3	25.2	-0.035	-0.112	0.032	0.996

Table 1. Explained variance, EV, in percentage, of the original data for every one of the three first rotated Principal Components, PC. FL_H, FL_{α_0}, FL_W and FL_B represent the rotated factor loadings for every PC.

Cluster	$\langle H \rangle$	σ_H	R_H	$\langle \alpha_0 \rangle$	σ_{α_0}	R_{α_0}	$\langle W \rangle$	σ_W	R_W	$\langle B \rangle$	σ_B	R_B
C1(53)	0.504	0.055	0.262	0.519	0.056	0.282	0.323	0.055	0.223	−1.038	0.602	2.594
C2(47)	0.490	0.043	0.171	0.506	0.044	0.175	0.398	0.035	0.120	1.296	0.348	1.275
C3(85)	0.495	0.053	0.229	0.528	0.053	0.275	0.432	0.046	0.200	−0.145	0.575	3.126
C4(12)	0.326	0.054	0.163	0.346	0.053	0.160	0.445	0.025	0.097	1.068	0.287	0.915
C5(08)	0.183	0.058	0.157	0.190	0.058	0.164	0.258	0.068	0.186	0.742	0.495	1.523
C6(03)	0.125	----	----	0.122	----	----	0.157	----	----	−0.817	----	----
C7(25)	0.501	0.054	0.232	0.507	0.048	0.225	0.247	0.043	0.148	0.935	0.708	2.824
C8(01)	0.541	----	----	0.651	----	----	0.618	----	----	−1.132	----	----
C9(01)	0.508	----	----	0.407	----	----	0.591	----	----	3.364	----	----
C10(01)	0.469	----	----	0.612	----	----	0.599	----	----	−2.901	----	----
C11(06)	0.334	0.014	0.039	0.365	0.019	0.055	0.402	0.060	0.170	−0.612	0.392	1.133
C12(10)	0.474	0.044	0.123	0.492	0.045	0.131	0.533	0.034	0.126	1.034	0.248	0.709
C13(05)	0.251	0.015	0.036	0.258	0.016	0.037	0.223	0.051	0.113	1.782	0.560	1.305
C14(01)	0.177	----	----	0.197	----	----	0.367	----	----	−0.639	----	----

Table 2. Summary of average {$\langle H \rangle$, $\langle \alpha_0 \rangle$, $\langle W \rangle$, $\langle B \rangle$}, standard deviation {σ_H, σ_{α_0}, σ_W, σ_B} and range {R_H, R_{α_0}, R_W, R_B} of the four multifractal parameters representing every spatial cluster. The number of rain gauges belonging to every cluster is given in the first column within parentheses. Standard deviations for clusters with less than four rain gauges are not included.

CI	$Z < -2$	$-2 \le Z < -1$	$-1 \le Z < 0$	$0 \le Z < 1$	$1 \le Z < 2$	$Z \ge 2$
N_{DSL}	9	31	74	115	27	2

Table 3. Number of DSL series, N_{DSL}, with Complexity Index, CI, within several standard deviation intervals.

	σ(days)	Sk	K	R(σ, 2σ) (%)	R(>2σ) (%)	R(<-σ) (%)
Minimum	1.71	1.85	4.43	0.0	2.90	0.0
Maximum	29.48	5.19	31.20	9.21	5.27	8.28
Average	4.64	2.55	10.91	7.03	4.69	4.55

Table 4. Statistical summary (minimum, maximum and average) of the 267 samples of standard deviations, σ, skewness, Sk, kurtosis, K, and percentages of residuals within the (σ, 2σ), R(σ, 2σ), exceeding 2σ, R(>2σ), and lowering −σ, R(<−σ), ranges.

References

Anagnostopoulou Chr., Maheras P., Karacostas T., Vafiadis M. (2003). Spatial and temporal analysis of dry spells in Greece. Theor. Appl. Climatol., 74, 77–91.

Aviad Y, Kutiel H, Lavee H. (2009) Variation of the *dry day since last rain*, DDSLR, as a measure of dryness along a Mediterranean – Arid transect. J. Arids Environ, 73 , 658-665.

Box G.E.P., Jenkins G.M. (1976). Time Series Analysis: Forecasting and Control. Holden-Day, California, 575 pp.

Brunetti M., Maugeri M., Nanni T., Navarra A. (2002). Droughts and extreme events in regional daily Italian precipitation series. Int. J. Climatol., 22 (5), 543–558.

Burgueño A., Lana X., Serra C., Martínez M.D. (2013). Daily extreme temperature multifractals in Catalonia (NE Spain). Physics Letters A, 378, 874-885.

Carvalho J.R.P., Assad E.D., Evangelista S.R.M., Pinto H.S. (2013). Estimation of dry spells in three Brazilian regions – analysis of extremes. Atmos. Res., 132–133, 12–21.

Cindrić K., Pasarić Z., Gajić-Čapka M. (2010). Spatial and temporal analysis of dry spells in Croatia. Theor. Appl. Climatol., 102, 171-184.

Davis R.E., Kalkstein L.S. (1990). Development of an automatic spatial synoptic climatological classification. Int. J. Climatol., 10, 769-794.

Diks C. (1999). Non linear time series analysis. Methods and applications. World Scientific, New York, 207 pp.

Diodato N., Bellocchi G. (2008). Drought stress patterns in Italy using agroclimatic indicators. Clim. Res., 36 (1), 53–63.

Dutta S. (2010). EEG pattern of normal and epileptic rats: monofractal or multifractal? Fractals, 18, 425-431.

Feder J. (1988). Fractals. Plenum Press, New York – London.

Feng T., Fu Z., Deng X., Mao J. (2009). A brief description to different multi-fractal behaviors of daily wind speed records over China. Phys. Lett. A, 373, 4134–4141.

Feng H., Xu Y. (2012): Multifractal detrended fluctuaction analysis of WLAN. Wireless Pers Commun, 66, 385–395.

Ge E., Leung Y. (2012): Detection of crossover time scales in multifractal detrended fluctuation analysis. J Geogr Syst, DOI 10.1007/s10109-012-0169-9.

García-Ruiz J.M., López-Moreno J.I., Vicente-Serrano S.M., Lasanta-Martínez T., Beguería S. (2011) Mediterranean water resources in a global change scenario. Earth Sci. Rev., 105, 121–139.

Ghosh D., Dep A., Dutta S., Sengupta R., Samanta S. (2012). Multifractality of radon concentration fluctuation in earthquake related signal. Fractals, 20, 33-39.

Gou X., Chen M., Du Y., Dong W. (2010): Fractal dynamics analysis of the VHF radiation pulses during initial breakdown process of lightning. Geophysical Research Letters, 37, L11808, doi:10.1029/2010GL043178.

Govindan R.B., Vyushin D., Bunde A., Brenner S., Havlin S., Schellnhuber (2002): Global climate models violate scaling of the observed atmospheric variability. Physical Rev. Lett., 89, 028501.

Hayes MJ, Svoboda MD, Wilhite DA, Vanyarkho OV. (1999). Monitoring the 1996 drought using the standardised precipitation index. *Bulletin of the American Meteorological Society*, 80(3) , 429–438.

Heinrich G., Gobiet A. (2011). The future of dry and wet spells in Europe: A comprehensive study based on the ENSEMBLES regional climate models. Int. J. Climatol., 32, 1951-1970.

Ivanov P. Ch., Nunes Amaral L.A., Goldenberg A.L., Haulin S., Rosenblum M.G., Stanley H.E., Struzik Z.B. (2001). From 1/f noise to multifractal cascades in hearthbeat dynamics. Chaos, 11, 641-652.

Jolliffe I.T. (1986). Principal Components Analysis. Springer Series in Statistics. Springer verlag, New York, 271.

Jung R.C., Kukuk M., Liesenfeld R. (2006). Time series of count data: modelling, estimation and diagnostics. Computational Statistics and Data Analysis, 51, 2350-2364.

Kalkstein L.S., Tan, G., Skindlov, J.A. (1987). An evaluation of three clustering procedures for use in synoptic climatological classification. J. Clim. Appl. Meteorol., 26, 717-730.

Kantelhardt J.W., Zschiegner S.A., Koscielny-Bunde E., Havlin S., Bunde A., Stanley H.E. (2002): Multifractal detrended fluctuaction analysis of nonstationary time series. Physica A, 316, 87-114.

Kavasseri R.G., Nagarajan R. (2005): A multifractal description of wind speed records. Chaos, Solitons and Fractals, 24, 165–173.

Király A., Jánosi I.M. (2005): Detrended fluctuation analysis of daily temperature records: Geographic dependence over Australia. Meteorol. Atmos. Phys., 88, 119–128.

Klein Tank AMG, Wijngaard JB, Können GP, Böhm R, Demarée G,Gocheva A, Mileta M, Pashiardis S,Hejkrlik L, Kern-Hansen C, Heino R, Bessemoulin P, Müller-Westermeier G, Tzanakou M, Szalai S, Pálsdóttir T, Fitzgerald D, Rubin S, Capaldo M, Maugeri M, Leitass A, Bukantis A, Aberfeld R, Van Engelen AFV, Forland E, Mietus M, Coelho F, Mares C, Razuvaev V, Nieplova E, Cegnar T, Antonio López J, Dahlström B, Moberg A, Kirchhofer W, Ceylan A, Pachaliuk O, Alexander LV, Petrovic P (2002). Daily dataset of 20th century surface air temperature and precipitation series for the European Climate Assessment. Int. J. Climatol., 22, 1441–1453.

Klok E.J., Klein Tank A.M.G. (2009). Updated and extended European dataset of daily climate observations. Int. J. Climatol., 29, 1182-1191.

Koscielny-Bunde E., Roman H.E., Bunde A., Havlin S., Schellnhuber (1998): Long-range power-law correlations in local daily temperature fluctuactions. Philosophical Magazine B, 77, 1331-1340.

Koscielny-Bunde E., Kantelhardt J.W, Braund P., Bundea A., Havlin S.: (2006) Long-term persistence and multifractality of river runoff records: Detrended fluctuation studies. Journal of Hydrology, 322,120–137.

Kostopoulou E., Jones P.D. (2005). Assessment of climate extremes in the Eastern Mediterranean. Meteorol. Atmos. Phys., 89, 69–85.

Kutiel H, Maheras, P (1992). Variations interannuelles des séquences sèches et des situations synoptiques en Méditerranée. Publications de l'AIC, 5, 15-27.

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., Burgueño A., Martínez M.D., Serra, C. (2006). Statistical distributions and sampling strategies for the analysis of extreme dry spells in Catalonia (NE Spain). J. Hydrol., 324, 94–114.

Lana X., Martínez M.D., Burgueño A., Serra C., Martín-Vide J., Gomez L., (2008a). Spatial and temporal patterns of dry spell lengths in the Iberian Peninsula for the second half of the twentieth century. Theor. Appl. Climatol., 91, 99-116.

Lana X., Martínez M.D., Burgueño A., Serra C. (2008b). Return period maps of dry spells for Catalonia (North-eastern Spain) based on the Weibull distribution. Hydrol. Science Journal, 53(1), 48-64.

Lana X., Martínez M.D., Serra C., Burgueño A. (2010). Complex behaviour and predictability of the European dry spell regimes. Nonlinear processes in Geophysics, 17, 499-512.

Lana X., Burgueño A., Martínez M.D., Serra C. (2012). Some characteristics of a daily rainfall regime based on the DDSLR index. Theor. Apl. Climatol, 109, 153-174.

Lin G., Fu Z. (2008): A universal model to characterize different multi-fractal behaviors of daily temperature records over China. Physica A, 387, 573–579.

Livada I., Assimakopoulos V.D. (2007). Spatial and temporal analysis of drought in Greece using the Standard Precipitation Index (SPI). Theor. Appl. Climatol., 89, 143–153.

Malamud B.D., Turcotte D.L. (1999). Self-affine time series: I. Generation and analyses. Advances in Geophysics, 40. Edited by R. Domowska and V. Saltzman, Academic Press, 90 pp.

Mali P. (2014). Multifractal characterization of global temperature anomalies. Theor. Appl. Climatol., DOI 10.1007/s00704-014-1268-y

Mandelbrot B.B., Wallis J.R. (1969). Computer experiments with fractional Gaussian noises, Parts I, II , III. Water Resour. Res., 5, 228-267.

Martín-Vide J., Gómez L. (1999). Regionalization of peninsular Spain based on the length of dry spells. Int. J. Climatol., 19, 537–555.

McKee TB, Doesken NJ, Kleist J. (1993). The relationship of drought frequency and duration to time scales. Preprints, Eighth Conference on Applied Climatology. Anaheim, CA. American Meteorological Society, 179–184.

McKee TB, Doesken NJ, Kleist J. (1995). Drought monitoring with multiple time scales. Preprints, 9th Conference on Applied Climatology. Dallas, TX. American Meteorological Society, 233–236.

Mills T.C. (2004). Is the North Atlantic oscillation a random walk? : A comment with further results. Int. J. Climatol., 24, 377-383.

Movahed M. S., Hermanis E. (2008): Fractal analysis of river flow fluctuations. Physica A, 387, 915–932.

Muzy J.F., Bacry E., Arneodo A. (1994): The multifractal formalism revisited with wavelets. Internat. J. Bifur. Chaos, 4, 245–302.

Nastos P.T., Zerefos C.S. (2009). Spatial and temporal variability of consecutive dry and wet days in Greece. Atmos. Res., 94, 616–628.

Preisendorfer R.W. (1988). Principal Component Analysis in Meteorology and Oceanography. Development in in Atmospheric Sciences, 17. Elsevier, Amsterdam, 435 pp.

Press W.H., Teukolsky S.A., Vetterling W.T., Flannery B.P.(1992). Numerical Recipes in C. (Chapter 2). Cambridge University Press, Cambridge, 994 pp.

Reiser H, Kutiel H (2010) Rainfall uncertainty in the Mediterranean: dryness distribution. Theor. Appl. Climatol., 100, 123-135.

Serra C., Burgueño A., Martínez M. D., Lana X. (2006). Trends in dry spells across Catalonia (NE Spain) during the second half of the 20th century. Theor. Appl. Climatol., 85, 165–183.

Serra C., Martínez D., Lana X., Burgueño A. (2013). European dry spell length distributions, years 1951-2000. Theor. Appl. Climatol., 114, 531-551.

Serra C., Martínez M.D., Lana X., Burgueño A. (2014). European dry spell regimes (1951-2000): Clustering process and time trends. Atmos. Res., 144, 151-174.

Serra C., Lana X., Burgueño A., Martínez M.D. (2015). Partial duration series distribution of the European dry spell lengths for the second half of the twentieth century. Theor. Appl. Climatol., DOI 10.1007/s00704-014-1337-2.

Shimizu Y., Thurner S., Ehrenberger K. (2002): Multifractal spectra as a measure of complexity in human posture. Fractals, 10, 103-116.

Schmidli J., Frei C. (2005). Trends of heavy precipitation and wet and dry spells in Switzerland during the 20th century. Int. J. Climatol., 25, 753–771.

Schumann A.Y., Kantelhardt J.W. (2011): Multifractal moving average analysis and test of multifractal model with tuned correlations. Physica A, 390, 2637–2654.

Stephenson D.B., Pavan V., Bojariu R. (2010). Is the North Atlantic Oscillation a random walk? Int. J. Climatol., 20, 1-18.

Turcotte D.L. (1997). Fractal and chaos in Geology and Geophysics (2nd edition). Cambridge University Press, Cambridge, 398 pp.

Wang H.R., Wang C., Lin X., Kang J. (2014): An improved ARIMA model for precipitation simulations. Nonlin. Processes Geophys., 21, 1159–1168.

Wijngaard J.B., Klein Tank A.M.G., Konnen G.P. (2003). Homogeneity of 20th century European daily temperature and precipitation series. Int. J. Climatol., 23, 679-692.

Yuan N., Fu Z., Mao J. (2013). Different multi-fractal behaviors of diurnal temperature range over the north and the south of China. Theor. Appl. Climatol., 112, 673-682

Zolina O., Simmer C., Konstantin B., Sergey K. G., Peter K. (2013). Changes in the duration of European wet and dry spells during the last 60 years. J. Climate, 26, 2022-2047.

List of Figures

Figure 1. a) Spatial distribution of the 267 European rain gauges. Solid circles design rain gauges for which the multifractal spectrum has not been computed. b) Number of available gauges per year along the recording period. c) Histogram of N_{DSL}. d) DSL series for Vaexjoe (Sweden) and Almería (Spain).

Figure 2. a) Fluctuation function $F_q(s)$ for some q^{th}-orders, b) *h(q)* curve fitted to a polynomial function of q^{th}- orders (dashed lines correspond to the Hurst exponent, h(*q=2*)), c) multifractal scaling exponents, $\tau(q)$, and d) Hölder exponents, $\alpha(q)$, for DSL series corresponding to the Fabra Observatory (NW Mediterranean coast).

Figure 3. Multifractal spectrum (solid circles) and second order polynomial fit (solid line) for DSL series corresponding to the Fabra Observatory (NW Mediterranean coast).

Figure 4. Spatial distribution of a) Hurst exponent, *H*, b) critical Hölder exponent, α_0, c) asymmetry, *B*, and d) spectral width, *W*.

Figure 5. Relationships between a) *H* and α_0 and b) α_2 and α_0.

Figure 6. Spatial distribution of FSCs obtained after the PCA applied to parameters H, α_0, *B* and *W*.

Figure 7. Evolution of the similarity index along the clustering process.

Figure 8. Spatial distribution of the selected 14 clusters.

Figure 9. Spatial distribution of (a) negative CI and b) positive CI.

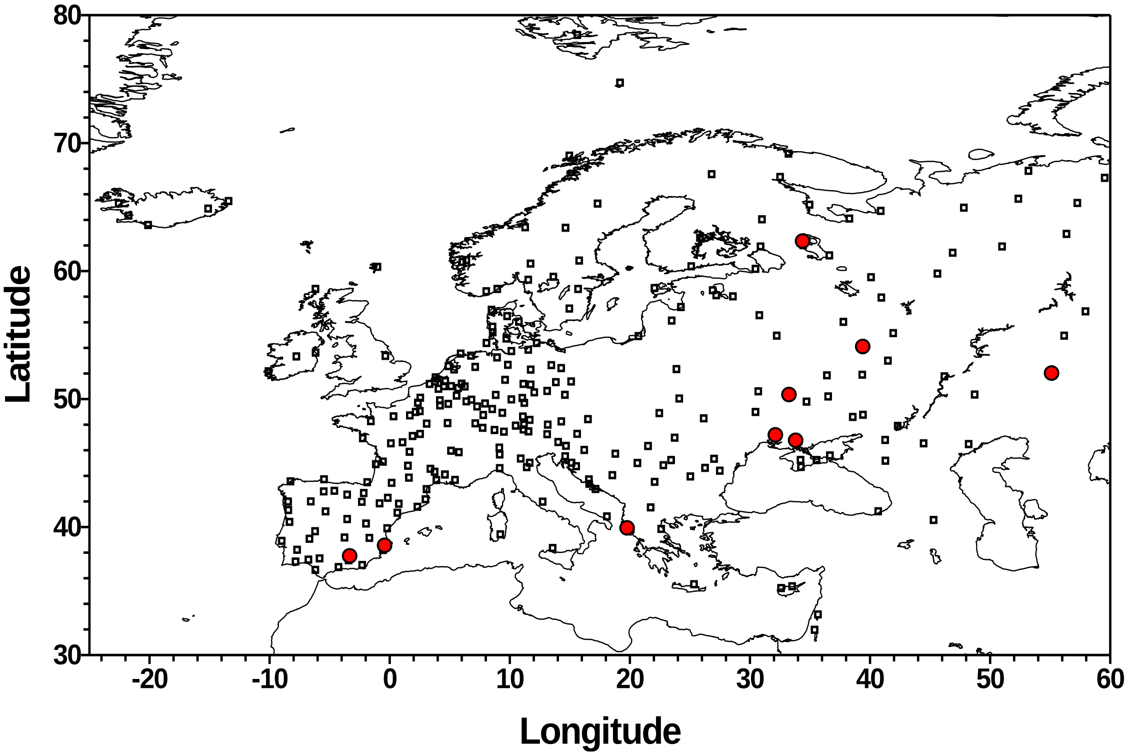
Figure 10. Dependence of (a) MAD and (b) OAO on latitude.

Figure 11. Histogram of DSL residuals for the 267 DSL series.

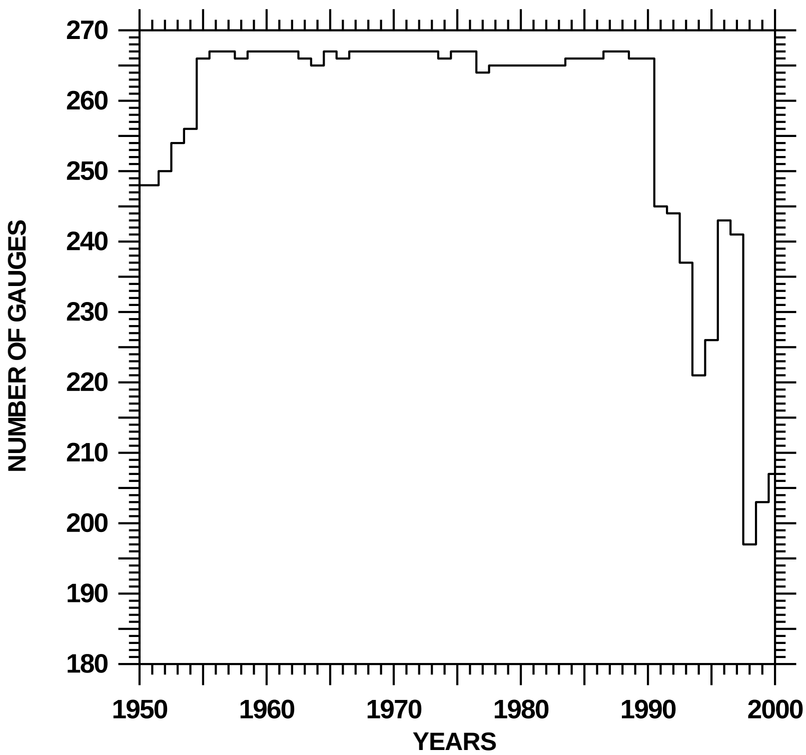
Figure 12. Spatial distribution of (a) MAD and (b) OAO, which have been classified as belonging to 0-20th, 20-40th, 40-60th, 60-80th and 80-100th percentile intervals. The corresponding thresholds of MAD and OAO are included, in increasing order, at the foot of every map.

Figure 13. Two examples of DSL residuals for DSL series belonging to (a) Vaexjoe (Sweden) and (b) Zaragoza (Spain). The evolution of the corresponding MAD with OAO is also included.

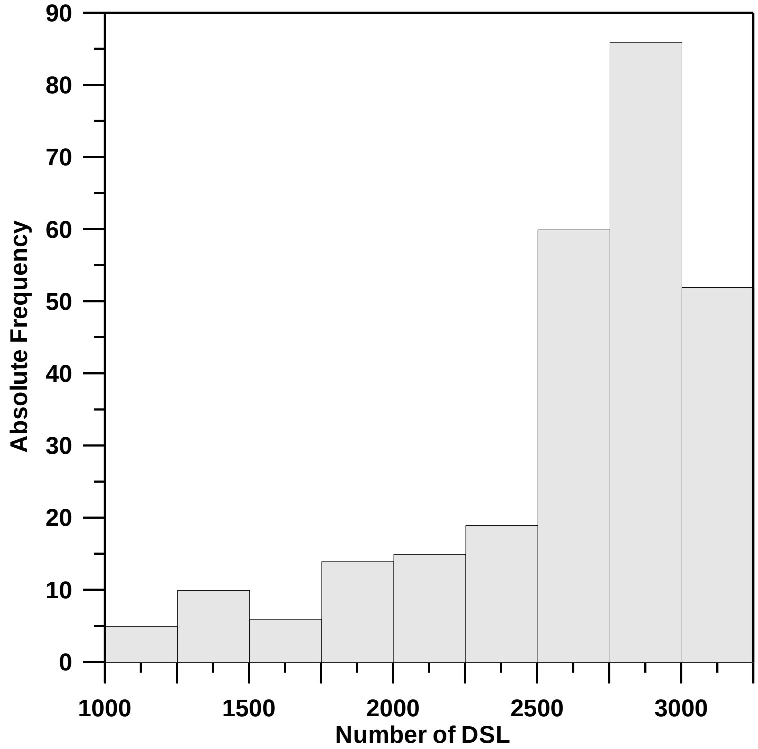
Figure 14. Dependence of (a) MAD and (b) OAO on CI.


Latitude
80
70
60
50
40
30
-20
-10
0
10
20
30
40
50
60
Longitude

(b)


NUMBER OF GAUGES
270
260
250
240
230
220
210
200
190
180
1950
1960
1970
1980
1990
2000
YEARS

c)


Absolute Frequency
Number of DSL
0
10
20
30
40
50
60
70
80
90
1000
1500
2000
2500
3000

d)

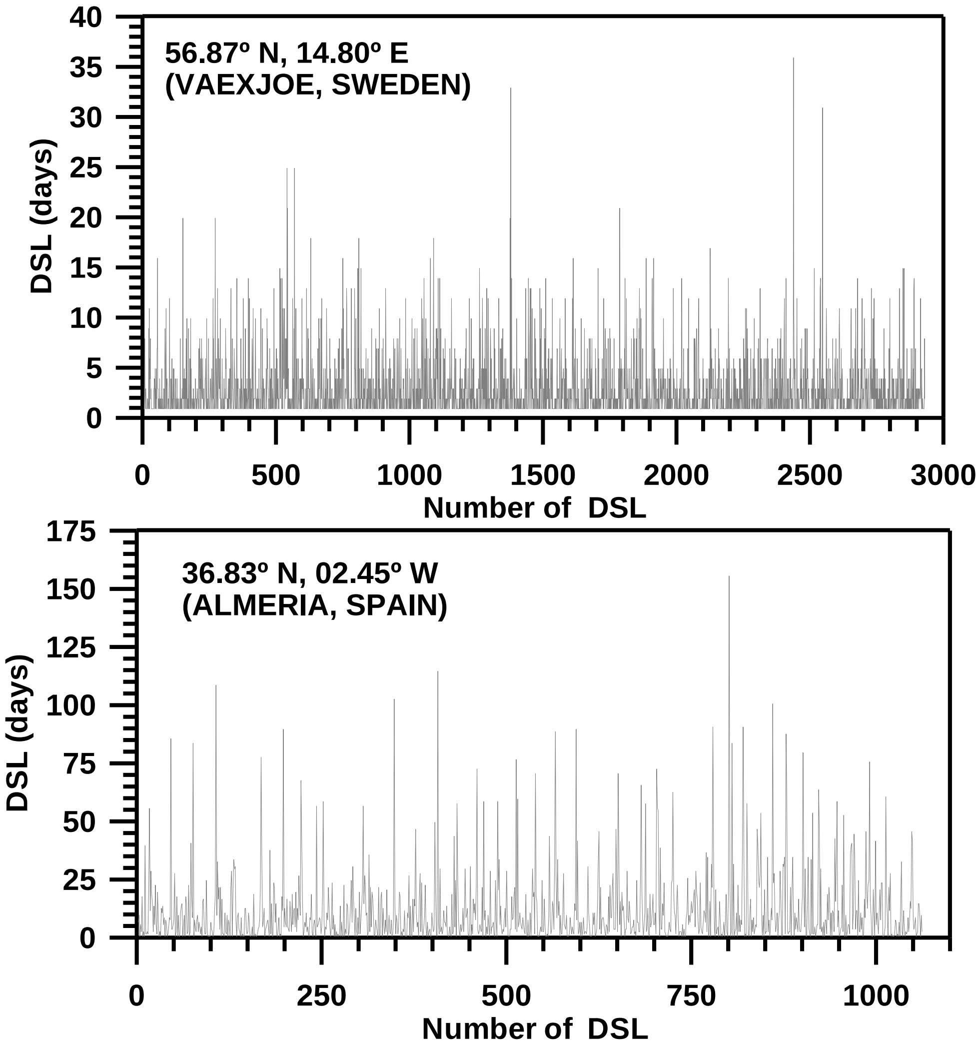
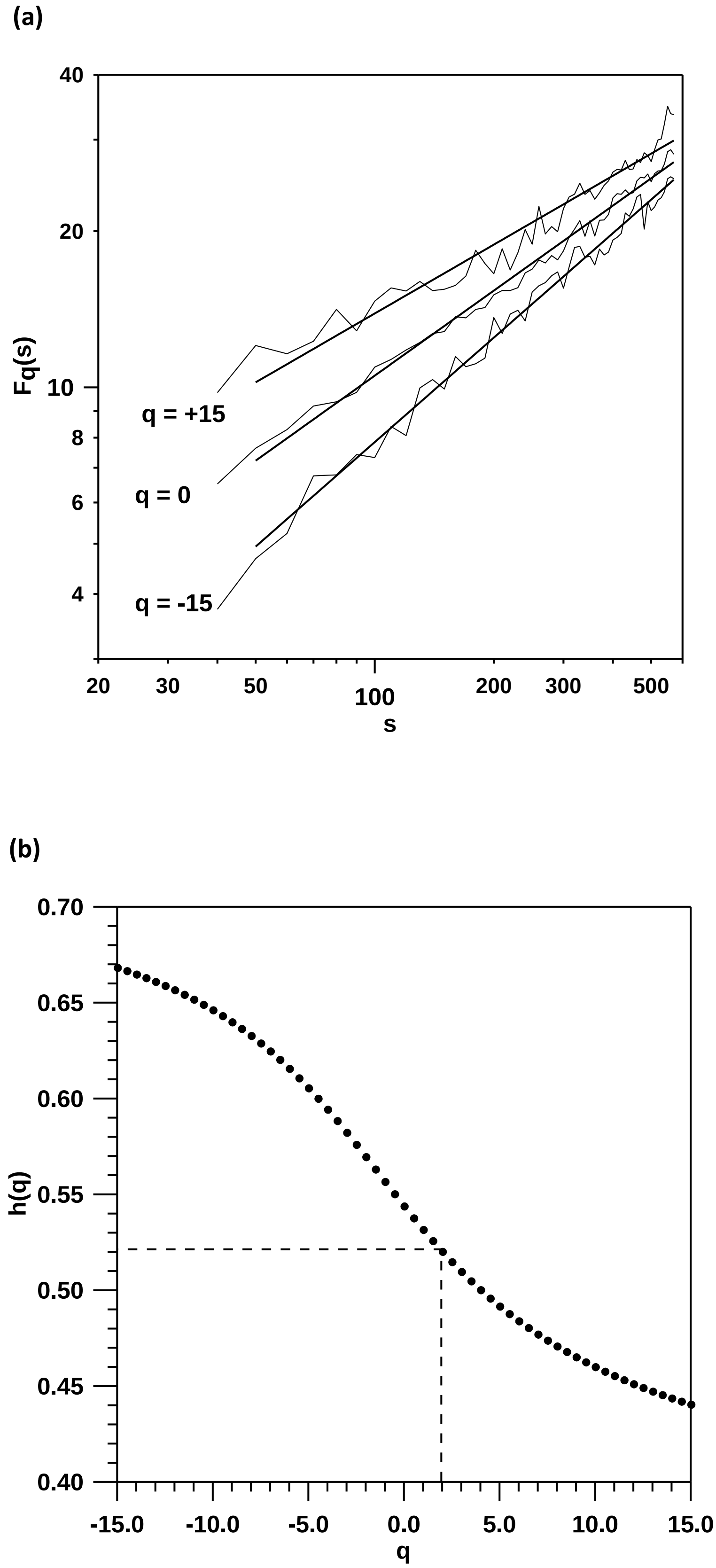
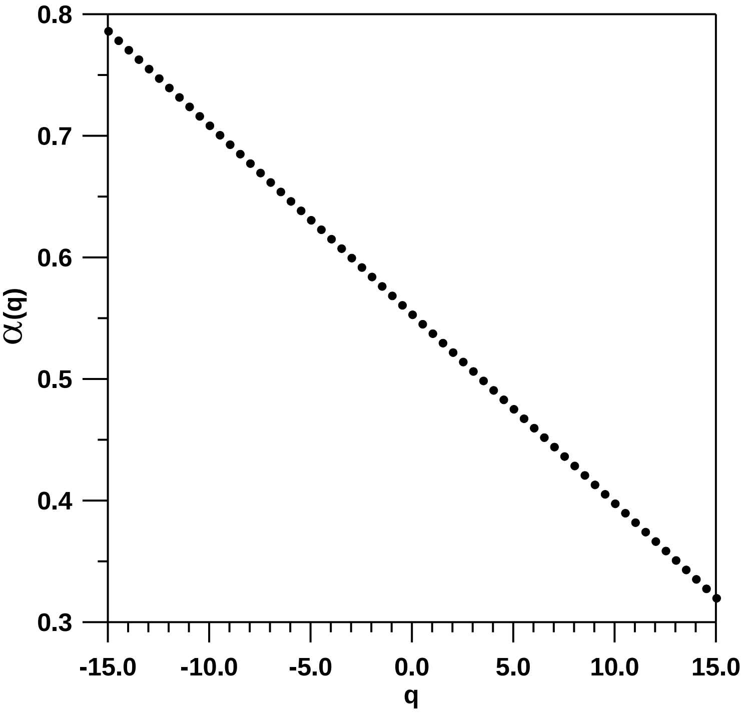

56.87º N, 14.80º E
(VAEXJOE, SWEDEN)
DSL (days)
Number of DSL
0
5
10
15
20
25
30
35
40
0
500
1000
1500
2000
2500
3000
36.83º N, 02.45º W
(ALMERIA, SPAIN)
DSL (days)
Number of DSL
0
25
50
75
100
125
150
175
0
250
500
750
1000

Figure 1


(a)
Fq(s)
40
20
10
8
6
4
q = +15
q = 0
q = -15
20
30
50
100
200
300
500
s
(b)
h(q)
0.70
0.65
0.60
0.55
0.50
0.45
0.40
-15.0
-10.0
-5.0
0.0
5.0
10.0
15.0
q

(c)


0.8
0.7
0.6
0.5
0.4
0.3
α(q)
-15.0
-10.0
-5.0
0.0
5.0
10.0
15.0
q

(d)

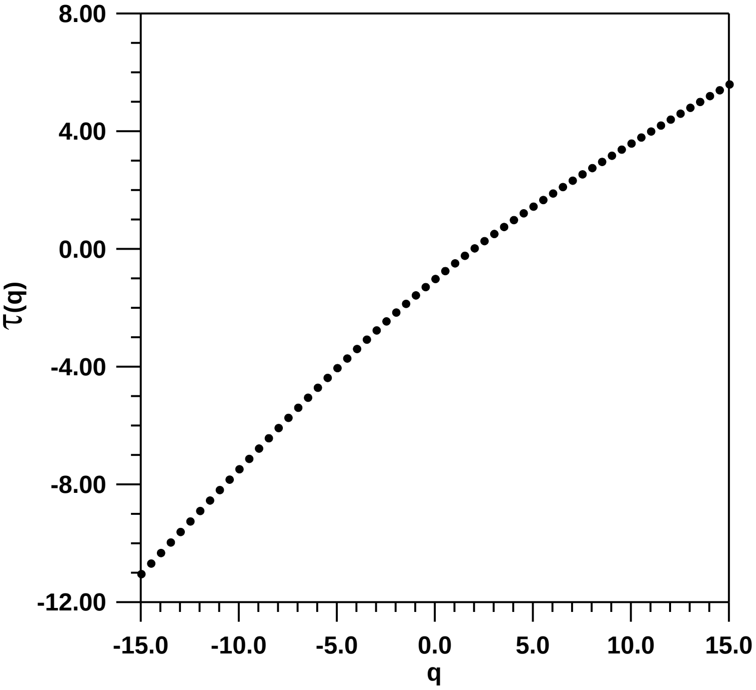

8.00
4.00
0.00
-4.00
-8.00
-12.00
τ(q)
-15.0
-10.0
-5.0
0.0
5.0
10.0
15.0
q

Figure 2

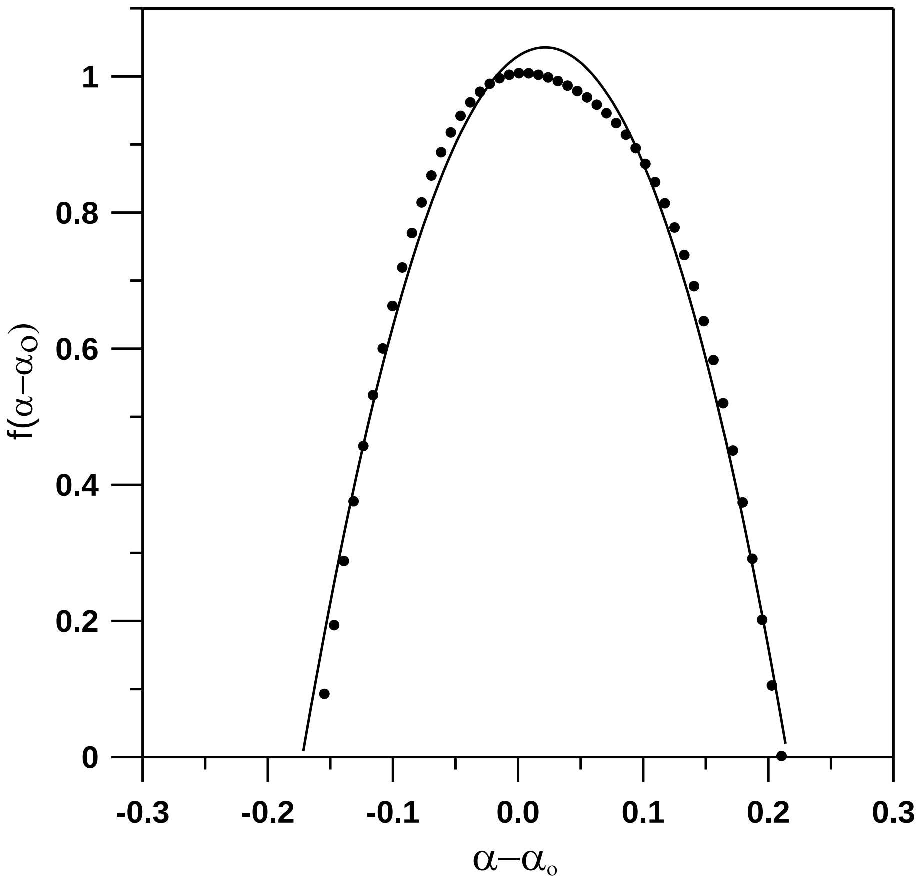
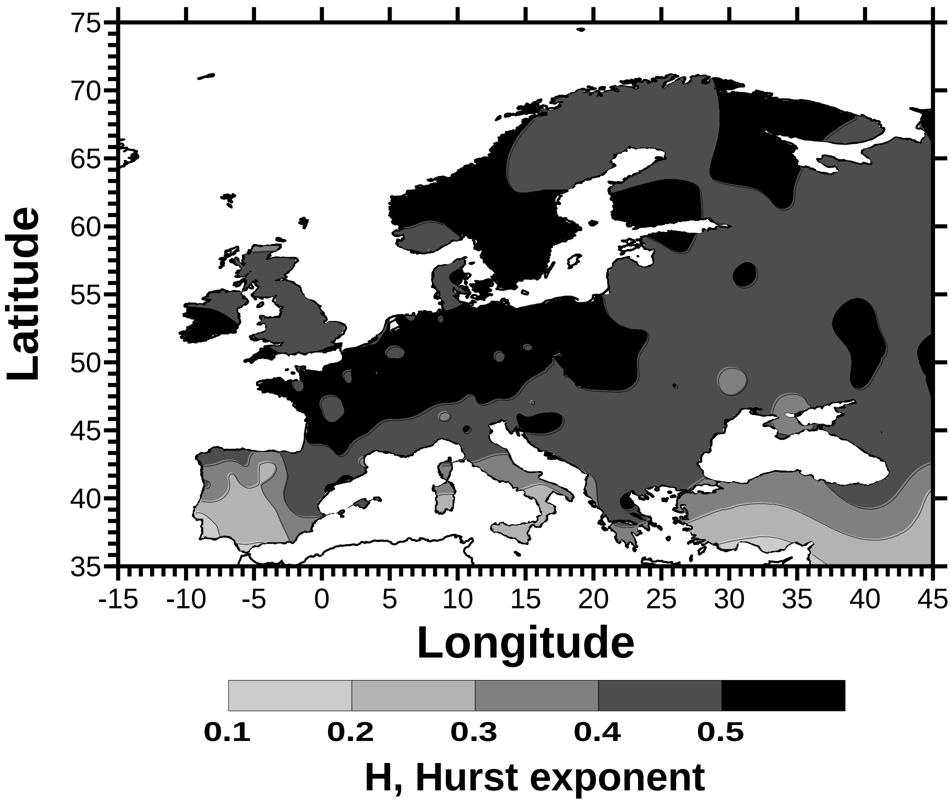


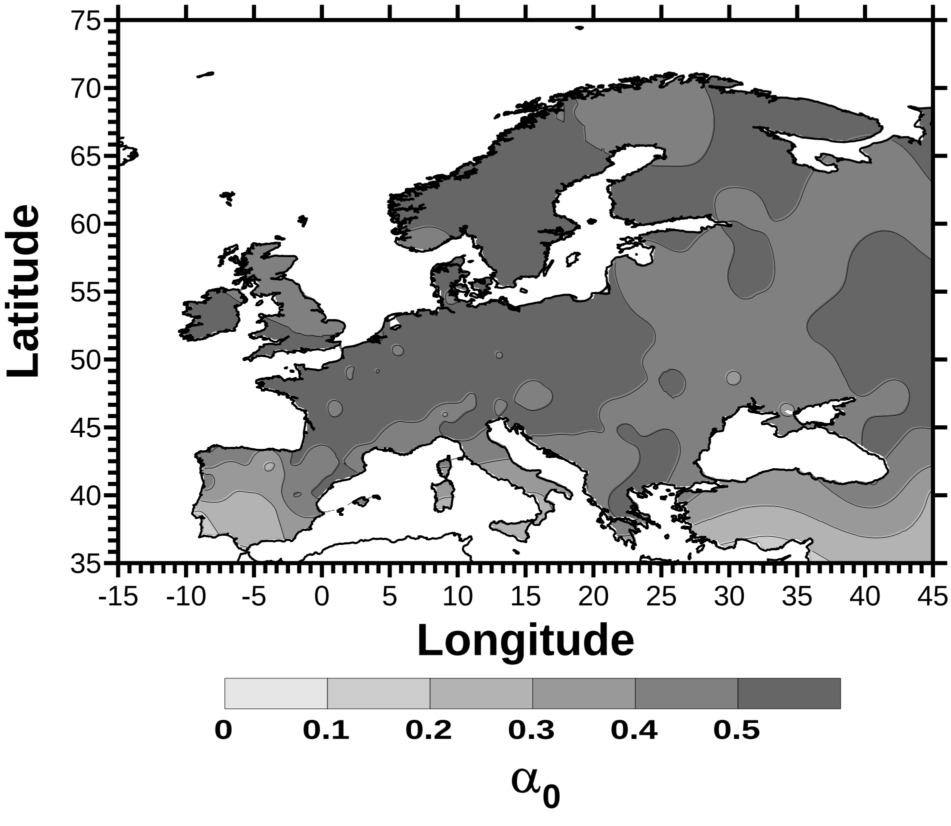
Figure 3

(a)



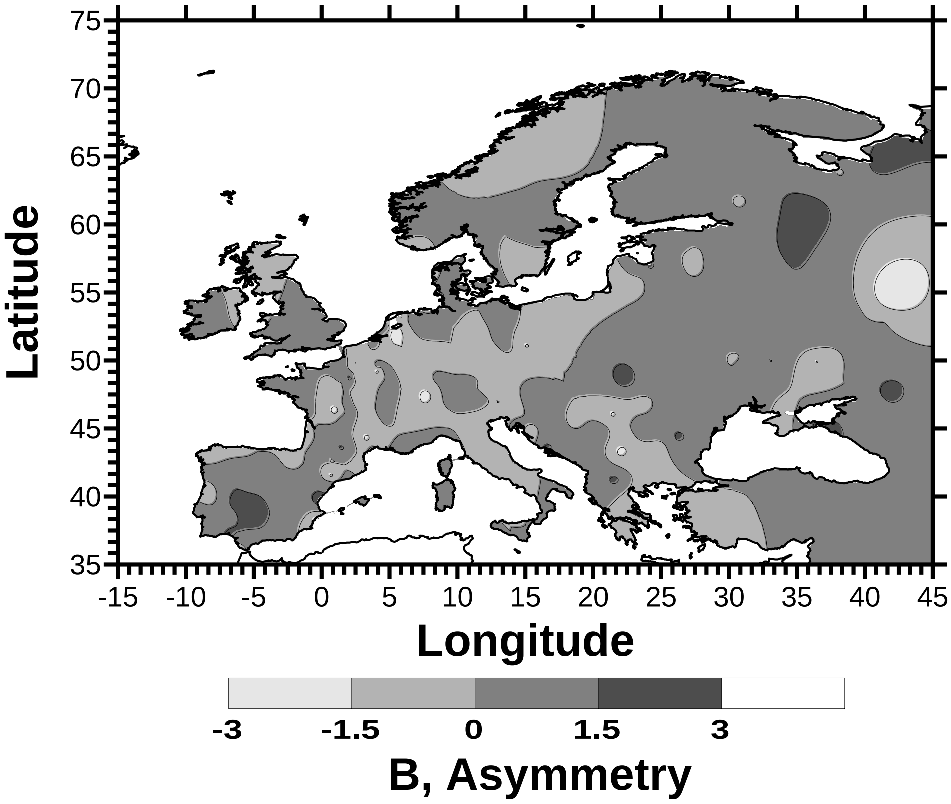
Latitude
75
70
65
60
55
50
45
40
35
-15 -10 -5 0 5 10 15 20 25 30 35 40 45
Longitude
0.1 0.2 0.3 0.4 0.5
H, Hurst exponent

(b)



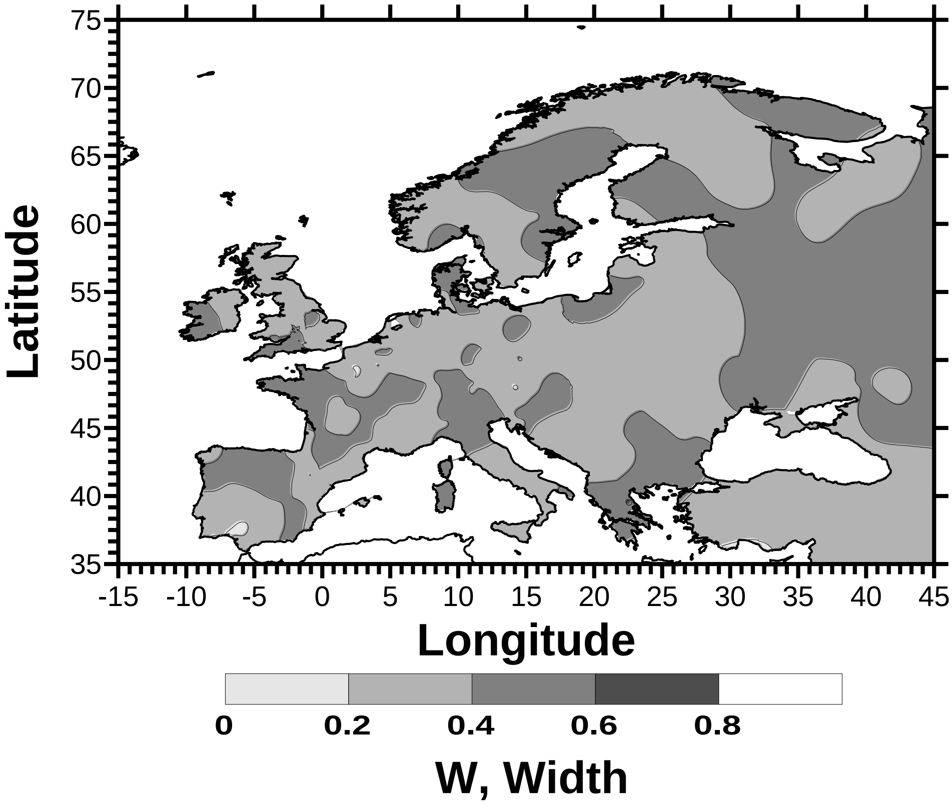
Latitude
75
70
65
60
55
50
45
40
35
-15 -10 -5 0 5 10 15 20 25 30 35 40 45
Longitude
0 0.1 0.2 0.3 0.4 0.5
α_0

(c)



Latitude
75
70
65
60
55
50
45
40
35
-15 -10 -5 0 5 10 15 20 25 30 35 40 45
Longitude
-3 -1.5 0 1.5 3
B, Asymmetry

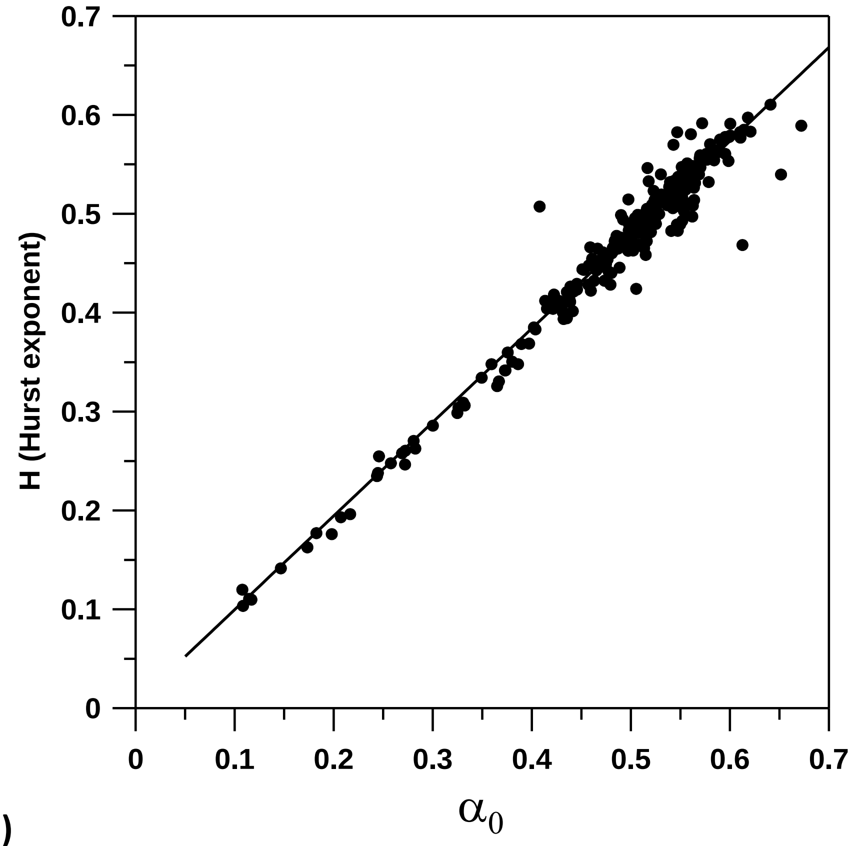
(d)



Latitude
75
70
65
60
55
50
45
40
35
-15 -10 -5 0 5 10 15 20 25 30 35 40 45
Longitude
0 0.2 0.4 0.6 0.8
W, Width

Figure 4

(a)


H (Hurst exponent)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
α_0

(b)

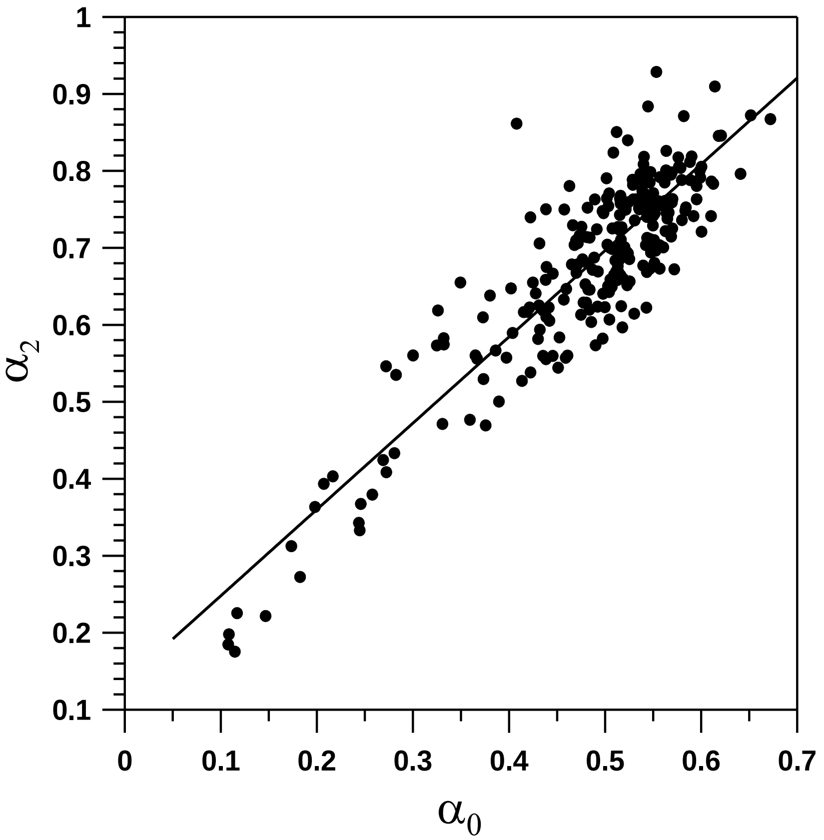

α_2
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
α_0

Figure 5

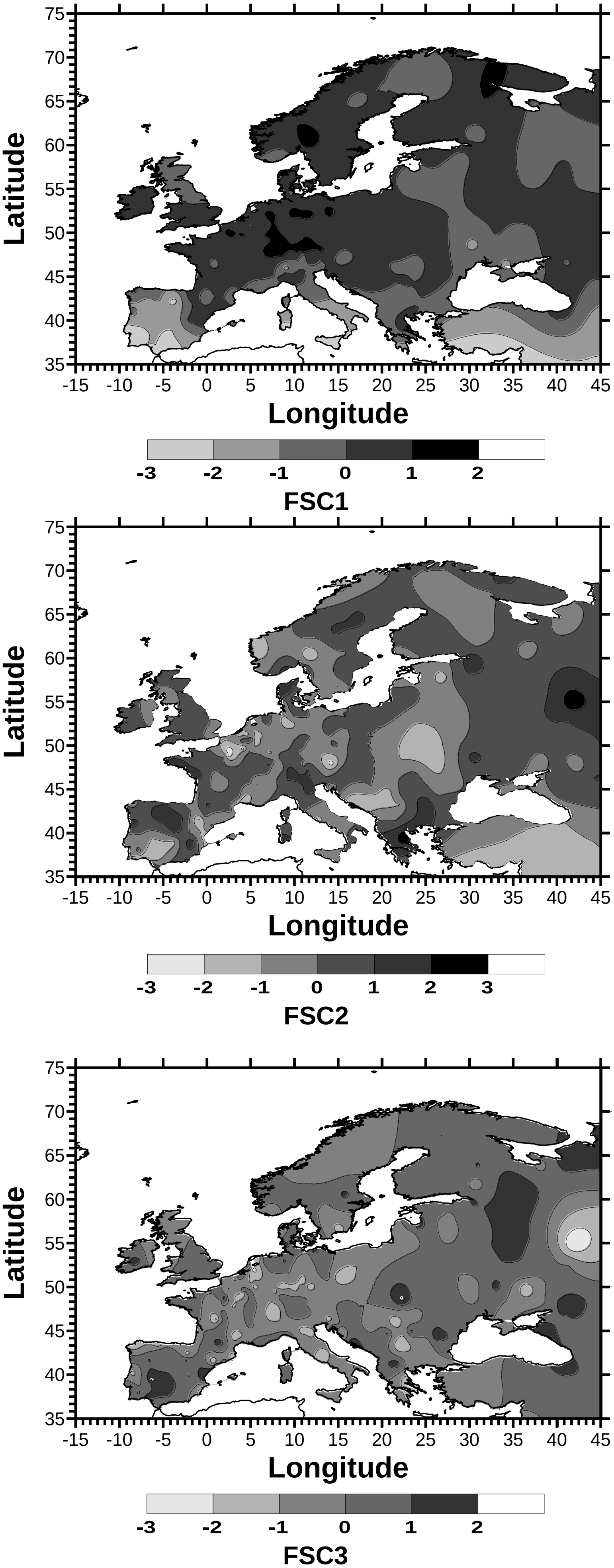

Latitude
Longitude
75
70
65
60
55
50
45
40
35
-15 -10 -5 0 5 10 15 20 25 30 35 40 45
-3 -2 -1 0 1 2
FSC1
-3 -2 -1 0 1 2 3
FSC2
-3 -2 -1 0 1 2
FSC3

Figure 6

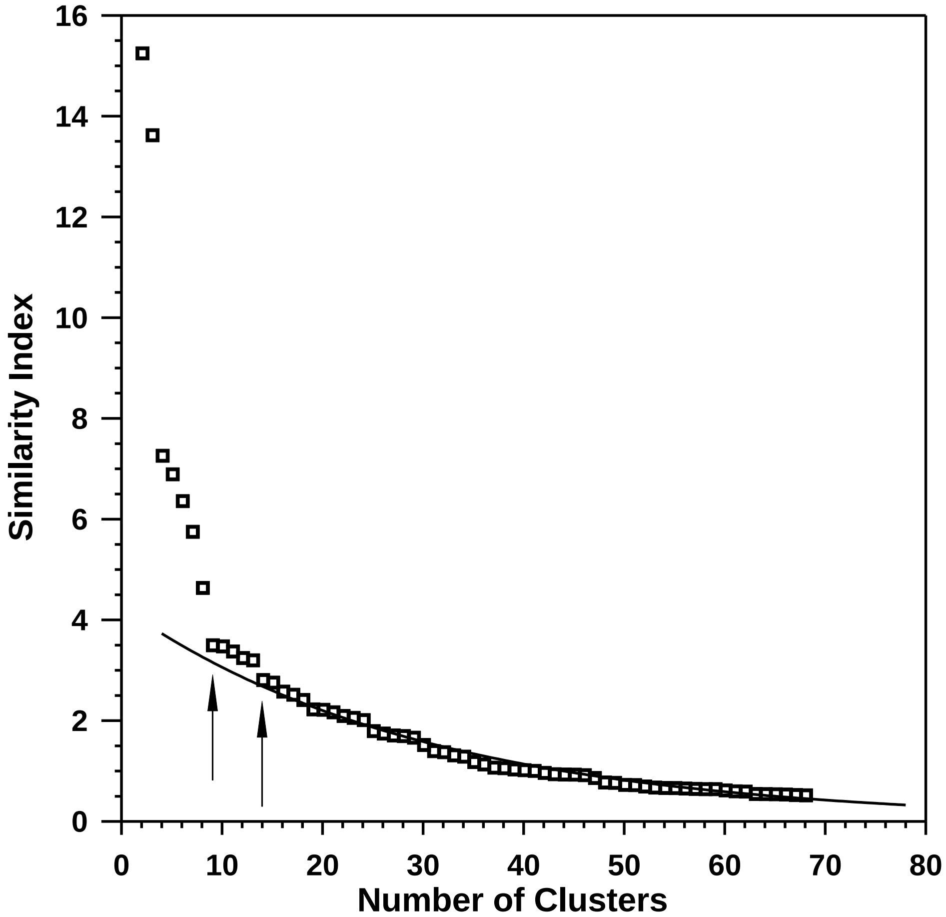

16
14
12
10
8
6
4
2
0
Similarity Index
0
10
20
30
40
50
60
70
80
Number of Clusters

Figure 7

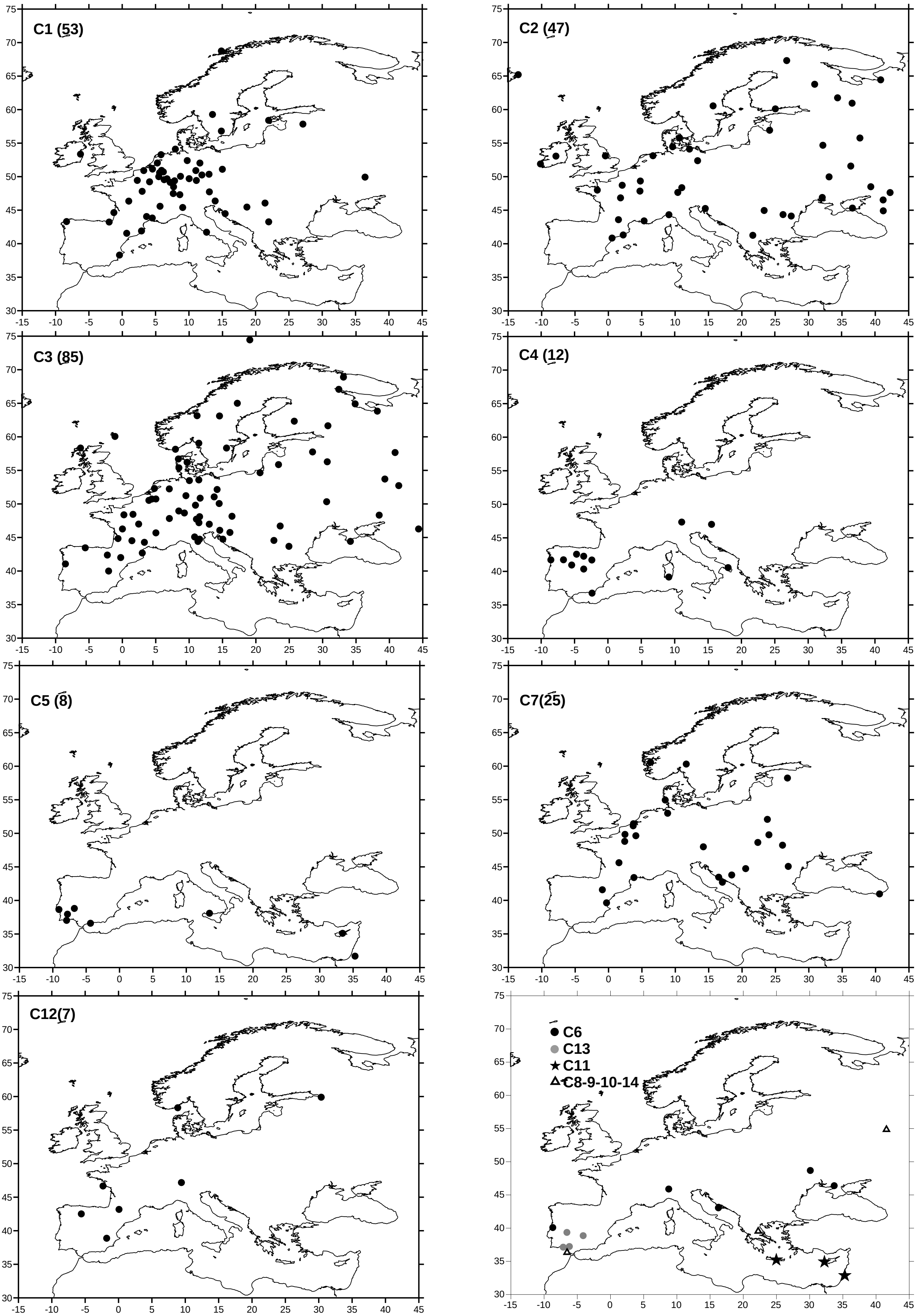
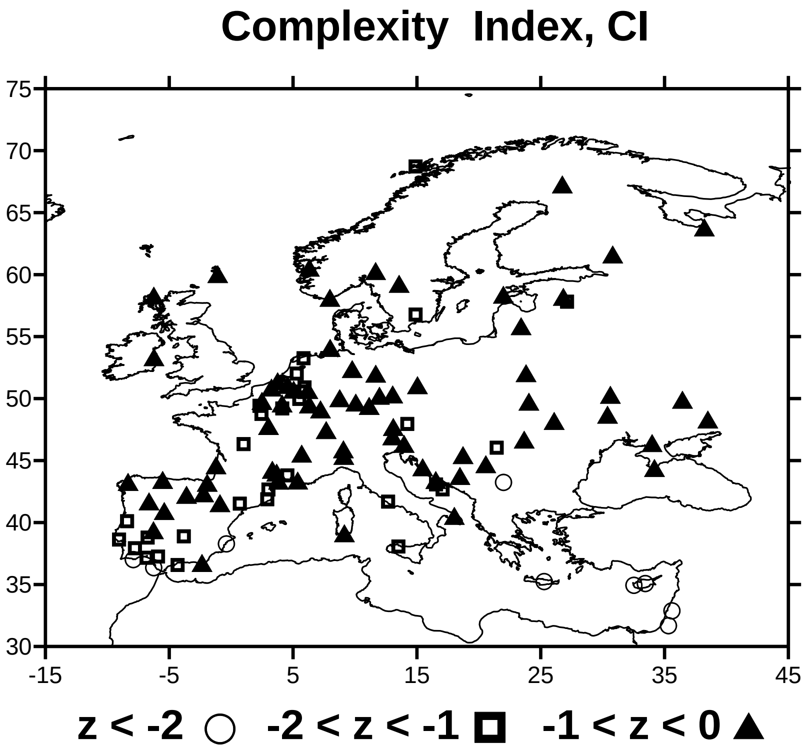


Figure 8

a)


Complexity Index, CI
75
70
65
60
55
50
45
40
35
30
-15
-5
5
15
25
35
45
z < -2
-2 < z < -1
-1 < z < 0

b)

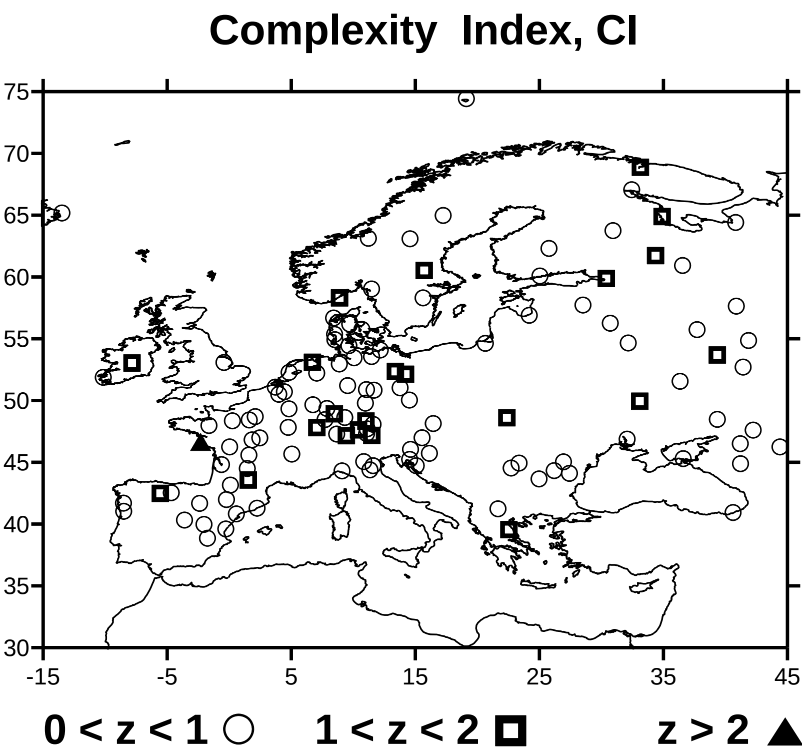
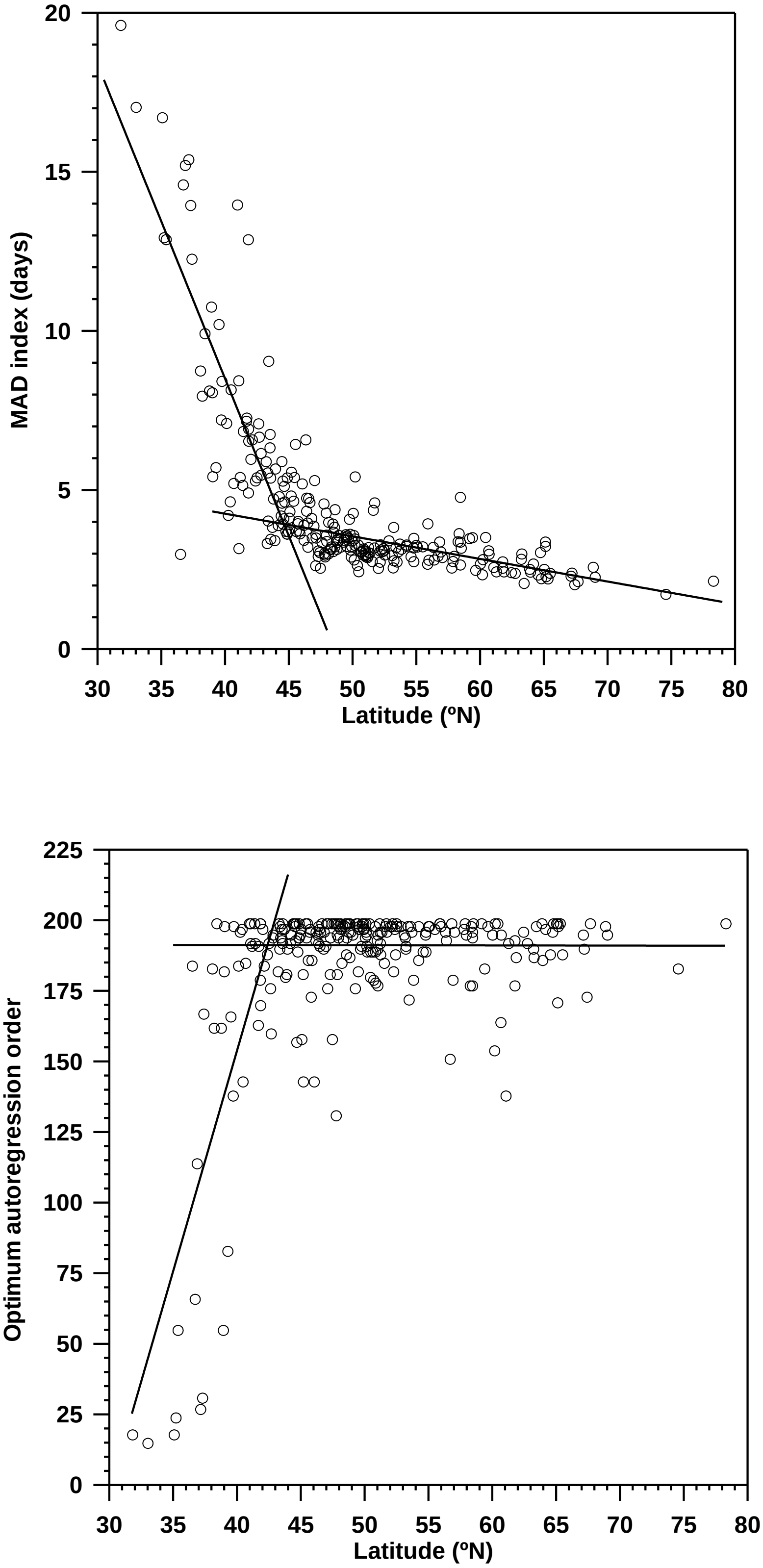

Complexity Index, CI
75
70
65
60
55
50
45
40
35
30
-15
-5
5
15
25
35
45
0 < z < 1
1 < z < 2
z > 2

Figure 9

(a)



MAD index (days)
20
15
10
5
0
30
35
40
45
50
55
60
65
70
75
80
Latitude (ºN)
Optimum autoregression order
225
200
175
150
125
100
75
50
25
0
30
35
40
45
50
55
60
65
70
75
80
Latitude (ºN)

(b)

Figure 10

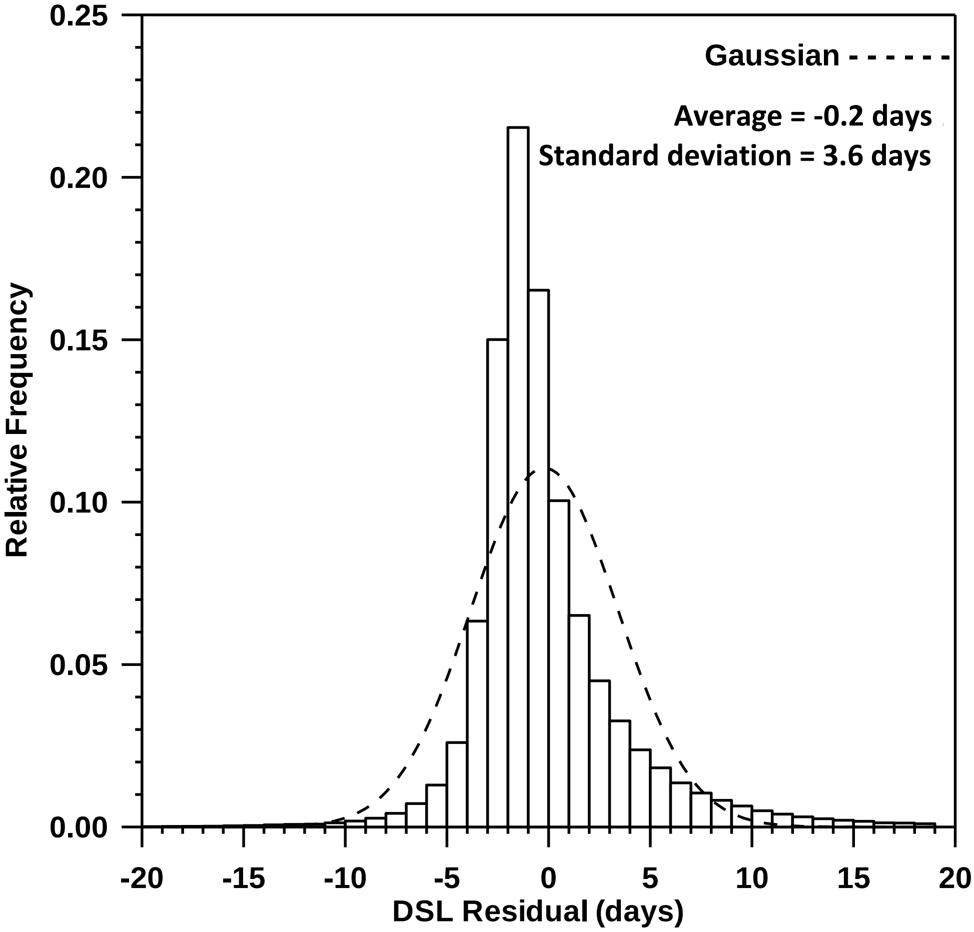

Gaussian
Average = -0.2 days
Standard deviation = 3.6 days
Relative Frequency
0.25
0.20
0.15
0.10
0.05
0.00
-20
-15
-10
-5
0
5
10
15
20
DSL Residual (days)

Figure 11

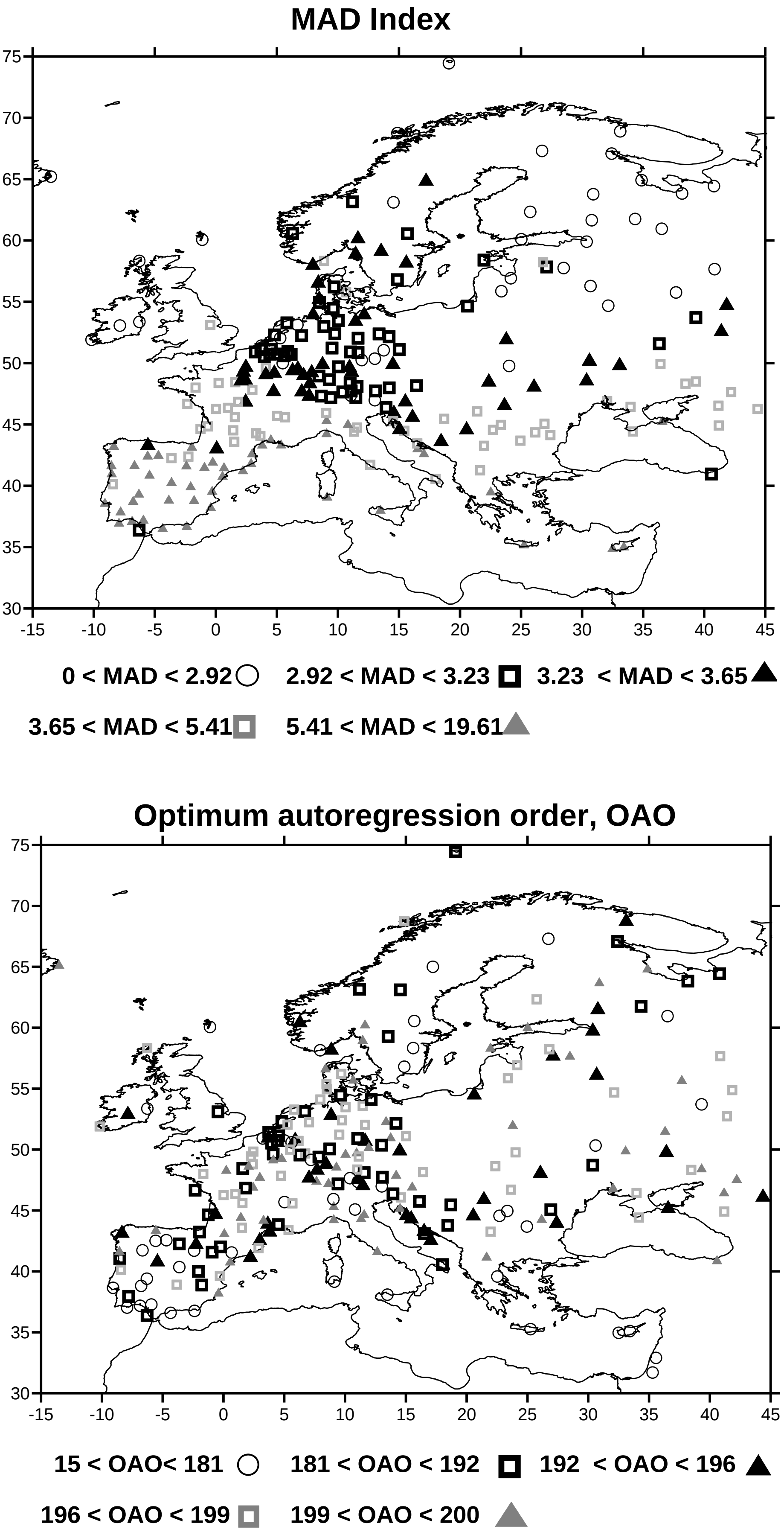
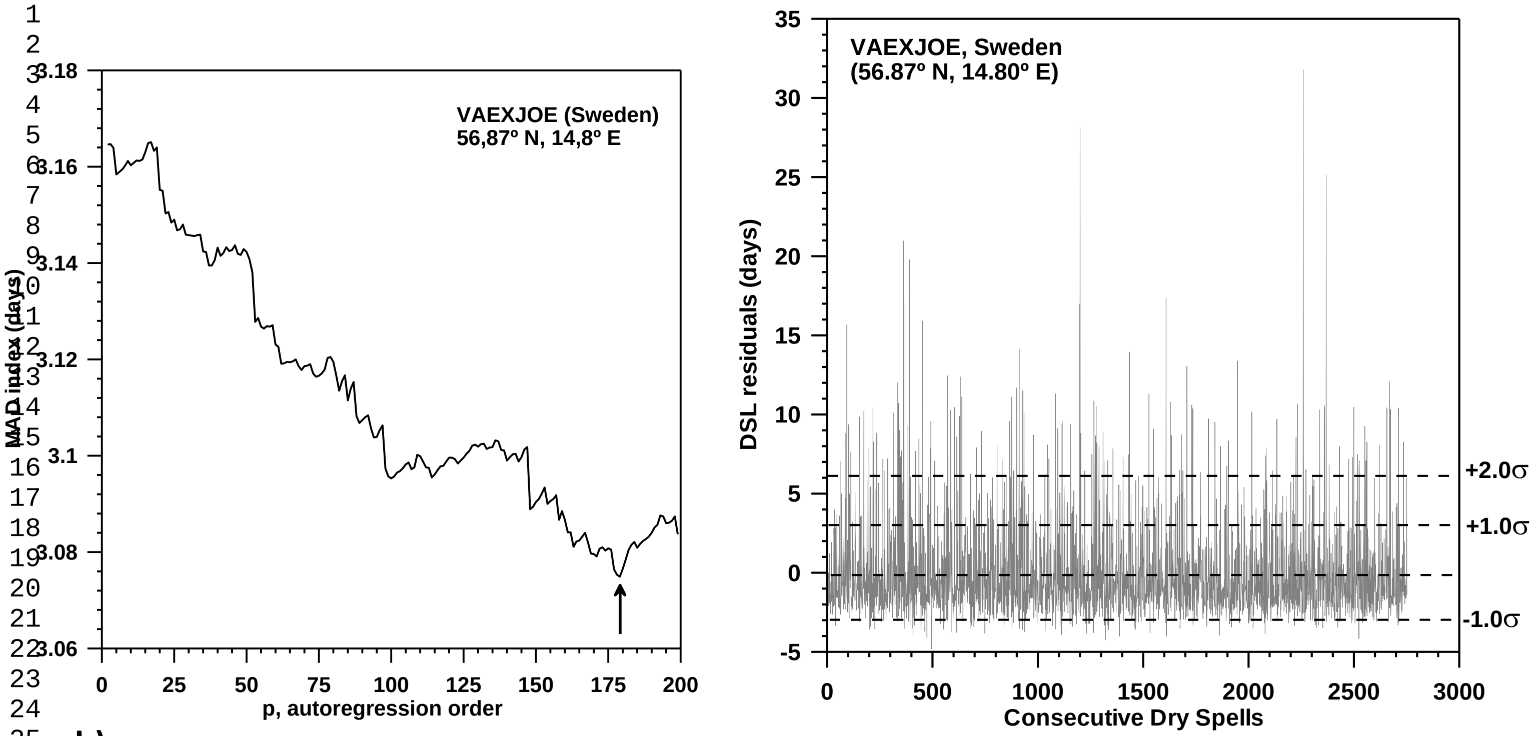

MAD Index
0 < MAD < 2.92
2.92 < MAD < 3.23
3.23 < MAD < 3.65
3.65 < MAD < 5.41
5.41 < MAD < 19.61
Optimum autoregression order, OAO
15 < OAO< 181
181 < OAO < 192
192 < OAO < 196
196 < OAO < 199
199 < OAO < 200

Figure 12

a)


VAEXJOE (Sweden)
56,87º N, 14,8º E
MAD index (days)
p, autoregression order
VAEXJOE, Sweden
(56.87º N, 14.80º E)
DSL residuals (days)
Consecutive Dry Spells
+2.0σ
+1.0σ
-1.0σ

b)

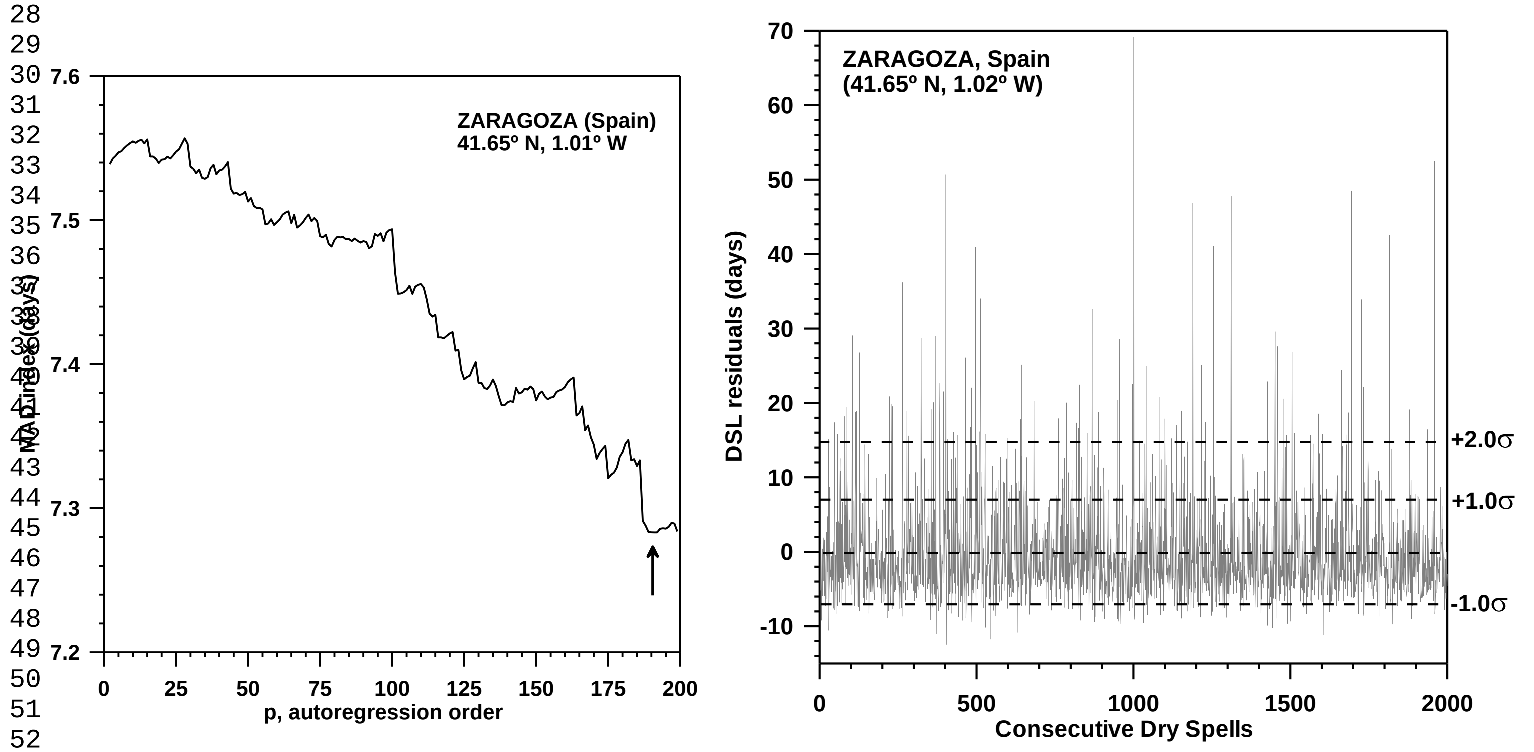
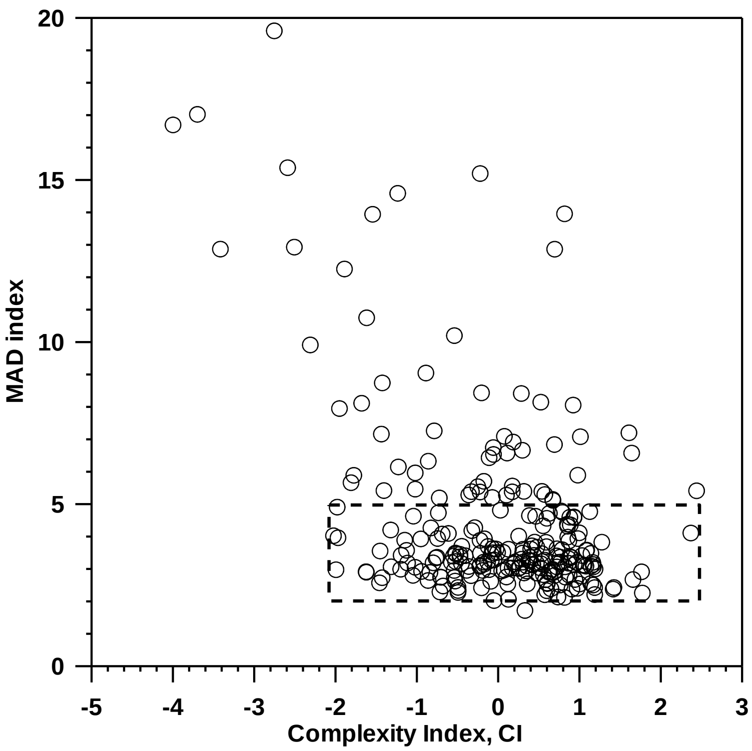

ZARAGOZA (Spain)
41.65º N, 1.01º W
MAD index (days)
p, autoregression order
ZARAGOZA, Spain
(41.65º N, 1.02º W)
DSL residuals (days)
Consecutive Dry Spells
+2.0σ
+1.0σ
-1.0σ

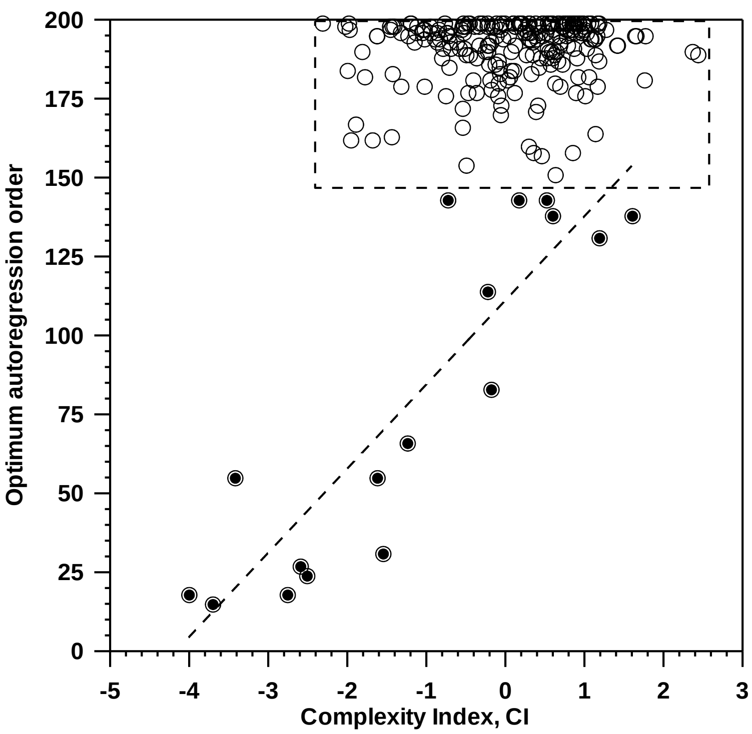
Figure 13

(a)



20
15
10
5
0
MAD index
-5
-4
-3
-2
-1
0
1
2
3
Complexity Index, CI

(b)



200
175
150
125
100
75
50
25
0
Optimum autoregression order
-5
-4
-3
-2
-1
0
1
2
3
Complexity Index, CI

Figure 14

Manuscript TAAC-D-14-00581: Multifractality and autoregressive processes of dry spell lengths in Europe: An approach to their complexity and predictability (Authors: X. Lana, A. Burgueño, C. Serra and M.D. Martínez)

Response to Reviewer 1:

a) In agreement with recommendations of the Reviewer and the Editor, the **Section 3.1** has been notably shortened, being only introduced the basic concepts concerning the MF-DFA algorithm and the appropriate references (see also **lines 162-165**). The structure of **Section 3.2** (Singularity spectrum) has not been changed, as many definitions of parameters and concepts used after are summarised in this Section. Additionally, in agreement with Reviewer 2, a detailed discussion about the range of q^{th}-order has been added to this Section.

b) Effectively, the interpolation method has been the "inverse distance". We think that in our case the main problem of obtaining accurate plots of the spatial distribution of fractal parameters is the limited rain gauge density in some areas, as mentioned now in **page 5 (lines 118-124)**.

c) In **Section 5, lines 575-580,** several alternatives to the AR(p) process are cited as possible improvements on DSL prediction for very long DSL. Certainly, an ARIMA(p,d,q) modelling and others methods based on the Poisson distribution and Monte Carlo algorithms would improve the mentioned very long DSL prediction, but we think they are beyond the scope of the present paper. It is also worth of mention that the relatively simple AR(p) process has led to good results when predicting monthly Western Mediterranean Oscillation index, as the authors of this manuscript have found (manuscript nowadays submitted to the International Journal of Climatology). With respect to these questions, we have to mention that in **line 224-225** a mistake concerning the definition of AR(p) has been amended. AR(p) has to be properly defined as an ARIMA(p,d,q) with d=1 and q=0.

Response to Reviewer 2:

a) With the aim of a more complete description of the DSL series, **new Figure 1c** includes a histogram of the number of DSL, N_{DSL}, for the 267 DSL series. Additionally, two very different examples of DSL series (Vaexjoe, Sweden, and Almeria, Spain) are shown in a **new Figure 1d**. The corresponding comments are developed in **Section 2, lines 136-143**.

b) A discussion about the appropriate range of the q^{th}-order is developed in **Section 3.2, lines 203-217**. The paper suggested by the reviewer (*), the assumption that multifractal spectra should be fitted to a quadratic function taking as argument the Hölder exponent, and previous experience of the authors about this question have been used as reference points of this discussion.

(*) Ivanov P. Ch., Nunes Amaral, L.A., Goldenberg A.L., Haulin S., Rosenblum M.G., Stanley, H.E., Struzik, Z.B. (2001). From 1/f noise to multifractal cascades in hearthbeat dynamics. Chaos, 11, 641-652.

c) The process to obtain a configuration of 14 clusters is described with more detail in **Section 4.3, lines 340-356**. The explanation is partially based on the concept of similarity index, L_{ij}, which is defined and quantified by the **new Equation 9**.

d) With respect to **Figures 8**, due to a technical problem with the writing software, the two first figures were plotted without appearing "C" codes designing the cluster number. The problem has been now solved.