1 VISIBILITY GRAPHS OF GROUND-LEVEL OZONE TIME SERIES: A MULTIFRACTAL ANALYSIS

2

3 AUTHORS: Carmona-Cabezas Rafael[1], Ariza-Villaverde Ana B.[1], Gutiérrez de Ravé

4 Eduardo[1], Jiménez-Hornero Francisco J.[1]

5 [1]University of Córdoba (Spain)

6 CORRESPONDING AUTHOR: Carmona-Cabezas Rafael

7

8

1   VISIBILITY GRAPHS OF GROUND-LEVEL OZONE TIME SERIES: A MULTIFRACTAL ANALYSIS

2   AUTHORS: Carmona-Cabezas Rafael, Ariza-Villaverde Ana B., Gutiérrez de Ravé Eduardo,

3   Jiménez-Hornero Francisco J.

4   ABSTRACT

5   A recent method based on the concurrence of complex networks and multifractal

6   analyses is applied for the first time to explore ground-level ozone behavior. Ozone time

7   series are converted into 2-D complex networks for their posterior analysis. The

8   searched purpose is to check the suitability of this transformation and to see whether

9   some features of these complex networks could constitute a preliminary analysis before

10   the more thorough multifractal formalism.

11   Results show effectively that the exposed transformation stores the original

12   information about the ozone dynamics and gives meaningful knowledge about the time

13   series. Based on these results, the multifractal analysis of the complex networks is

14   performed. Looking at the physical meaning of the multifractal properties (such as

15   fractal dimensions and singularity spectrum), a relationship between those and the

16   degree distribution of the complex networks is found.

17   In addition to all the promising results, this novel connection between time series

18   and complex networks can deal with both stationary and non-stationary time series,

19   overcoming one of the main limitations of multifractal analysis. Therefore, this

20   technique can be regarded as an alternative to give supplementary information within

21   the study of complex signals.

Abbreviations: VG (Visibility graph), SBA (Sandbox Algorithm)

22 <u>KEYWORDS</u>

23 - Ground-level ozone

24 - Visibility graphs

25 - Multifractal analysis

26 - Sandbox method

27

28 1. <u>INTRODUCTION</u>

29 Many studies have been performed about ground-level ozone over the last decades.

30 The importance of ozone characterization and analysis lies on the fact that it is one of

31 the main photochemical oxidants (due to its abundance). This irritant gas has serious

32 repercussions for human health and harvests when its concentration is high (Doherty et

33 al., 2009). Those exposed damages have an impact from the economical point of view,

34 and according to Miao et al. (2017), they lead every year to losses of several billions of

35 dollars.

36 Ozone is a secondary pollutant whose chemical formation and destruction

37 mechanisms are known to be photochemical and nonlinear processes (Graedel and

38 Crutzen, 1993; Trainer et al, 2000). These mechanisms highly depend on meteorological

39 variables such as temperature, wind direction and mainly solar radiation (Graedel and

40 Crutzen, 1993; Guicherit and Van Dop 1977), as it has been studied for the Mediteranean

41 area in several works (Güsten et al., 1994; Kouvarakis et al., 2000; Ribas y Peñuelas

42 2004). In addition to that, ozone concentration depends as well on chemical precursors,

43 such as nitrogen oxides and volatile organic compounds derived from the urban and

44   industrial activity (Sillman, 1999). All these factors make the analysis of the temporal

45   evolution of ozone a very complex task indeed.

46   Due to the facts exposed above, ozone studies based on traditional statistical

47   analysis may provide a limited description of more complex dynamics of time series

48   where the variability is high. The reason for this limitation is that these methods

49   approximate and smooth the signal by means of fitting to functions, with the derived

50   loss of information. Besides, they base their results on one (time) scale, while the

51   physical phenomenon can appear at several scales due to the number of variables in

52   play (Zeleke and Si 2006). On the contrary, multifractal methods can be used to save this

53   drawback, since they work directly with the raw data extracting the information from

54   their singularities. Furthermore, fractals (and multifractals) are characterized for being

55   self-similar when divided into smaller parts (i.e. they are scale-independent) or at least

56   their statistical properties (Mandelbrot 1982). That way, if a natural phenomenon can

57   be characterized by means of multifractal parameters, these will be able to describe it

58   for a range of scales.

59   In the presented work, a link between the multifractal analysis and complex

60   networks has been tested for the description of ground-level ozone dynamics. To that

61   purpose, ozone concentration time series have been transformed into 2-D visibility

62   graphs (VG) (whose topology inherits the features of the associated time series) and

63   then evaluated using two methods for multifractal analysis: the sandbox method in

64   order to compute the generalized fractal dimensions (Rényi spectrum) and the approach

65   introduced by Chhabra and Jensen (1989) for the calculation of the α-spectrum as an

66   independent value from the other.

67   With this study, the main purpose is to check the suitability of the multifractal

68   analysis performed over the VGs by connecting their degree distribution with Rényi and

69   α-spectra. It could be expected, since the resulting graph stores indeed much of the

70   original information and properties of the original time series (Lacasa et al., 2008).

71

72   2. <u>MATERIALS AND METHODS</u>

73   2.1. <u>Data</u>

74   The information that has been used in the analysis of this work corresponds to a 10-

75   min ozone concentration data collected for the months of January, April, July and

76   October in 2007. These time series of ozone concentration can be seen in Figure 1, for

77   the four months analyzed in this study.

78   The chosen region is the western part of Andalusia (Spain) since as exposed by

79   Domínguez-López et al. (2014), this area meets the weather conditions (high

80   temperatures and solar radiation), orographic (the valley of the Guadalquivir river) and

81   anthropic (four capitals and two important industrial centers such as the chemical pole

82   of Huelva and the Bay of Algeciras) to be potentially vulnerable to pollution by surface-

83   level ozone. The measurements were performed at the urban station located in Lepanto,

84   Córdoba (37.53° N, 4.47° W). The cited station belongs to the regional network in charge

85   of controlling the air quality in Andalusia, co-finanzed by the Consejería de

86   Medioambiente (Regional Environmental Department) and the European Union. This

87   station is located at 117 m of altitude and the average temperature for each moth is 8,

88   15, 28 and 18 °C (January, April, July and October, respectively). The average direct solar

89   radiation is 310.2, 577.8, 991.4 and 419.3 MJ/m^2 (again for the four months ordered).

90   The climate of the zone of study, according to the Köppen-Geiger classification, is

91   defined as Csa with warm average temperatures and hot and dry summer.

92   As it can be seen in Figure 1, the ozone concentrations are especially high in summer

93   (July) and low in Winter (January). As it has been previously discussed for this region by

94   Adame et al. (2007) and Jiménez-Hornero et al. (2010a), that is due the fact that the

95   conditions for the ozone creation are more suitable from the end of January. The

96   progressive raise of temperature and solar radiation that reaches its peak in July allows

97   those higher creation rates and thus its concentration. One of the reactions that governs

98   the ozone production can be found below (Graedel and Crutzen, 1993).

$$NO_2 + O_2 \leftrightarrow O_3 + NO \tag{1}$$

99   It is a reversible photochemical reaction which tends to the ozone production when

100   there is energy available in form of light (right sense of the arrow) and in the other way

101   when there is not. For that reason, the higher values of ozone are always found during

102   the day and vice versa, happening the same with summer and winter respectively, as

103   discussed before.

104

105   2.2. Visibility graph

106   One of the main fields of application of the multifractal analysis is referred to the

107   study of graphs and complex networks. A graph can be defined as a set of vertices, points

108   or nodes connected to each other by lines that are usually called *edges*. A tool to

109   transform time series into a graph was presented by Lacasa et al. (2008). This new

110   complex network receives the name of *visibility graph* and has been proven to inherit

111   many of the properties of the original series (being some multifractal properties among

112   them).

113   In order construct the visibility matrix which contains the information of all the

114   nodes of the new system, it is necessary to stablish a criterion to discern whether two

115   points would be connected or not. This criterion reads as follows: two arbitrary data

116   from the time series (t_a, y_a) and (t_b, y_b) have visibility (and would become two

117   connected nodes in the graph) if any other data point (t_c, y_c) between them $(t_a < t_c < t_b)$

118   fulfills the following condition:

$$y_c < y_a + (y_b - y_a)\frac{t_c - t_a}{t_b - t_a} \tag{2}$$

119   In Figure 2, an application of this method for a simple time series is given as

120   illustration. As it can be seen, the original time series has been transformed into a

121   complex network. The complexity of the original series is inherited by the new graph, as

122   it has been found by Lacasa et al. (2008, 2010), meaning that for instance a periodic time

123   series, would lead to a regular graph.

124   The result of applying this visibility method is a NxN adjacency binary matrix, being

125   N the number of points in the set. Each row of the matrix contains the information of a

126   different node, such as $a_{ij} = 1$ means that the node i and j have visibility; whereas

127    $a_{ij} = 0$ means that there is no edge between them. The resulting matrix has several

128   properties that can be used to simplify the algorithm and thus reduce the computational

129   required time.

- Hollow matrix: All the elements in the diagonal are zero ($a_{ii} = 0$), because there is no visibility of an element with itself, since there are no intermediate points to fulfills the criterion.

- Symmetric matrix: The elements satisfy $a_{ij} = a_{ji}$, due to the reciprocity of the visibility between two nodes.

- Nearest neighbors: Because each point always sees the closest previous and next node, the elements surrounding the diagonal are always 1 ($a_{ij} = 1$ for $j = i \pm 1$).

Taking all of that into account, the visibility matrix A has a general form as shown below:

$$A = \begin{pmatrix} 0 & 1 & \dots & a_{1,N} \\ 1 & 0 & 1 & \vdots \\ \vdots & 1 & \ddots & 1 \\ a_{N,1} & \dots & 1 & 0 \end{pmatrix}$$

The degree of a node (k_i) can be defined as the number nodes that have reciprocal visibility with the first one ($k_i = \sum_j a_{ij}$). In Figure 2, the degree of the first node is $k = 3$, for the second one $k = 2$, for the third one $k = 3$, and so on. From the degree of each one of the nodes present in the VG, it is possible to obtain the degree distribution of the sample ($P(k)$), which is nothing but the probability that every degree has.

Previous works have shown how the analysis of this degree distribution built from the VG can effectively describe the nature of the time series (Lacasa et al., 2008; Mali et al., 2018), distinguishing between periodic, random or fractal series for

151   instance. Therefore, by studying the degree distribution, one can get a first insight

152   on the behavior of the ozone time series before stepping into a more complex

153   multifractal analysis. As stated by Lacasa et al. (2009, 2010), time-series that have

154   VGs with degree distributions that follow power laws such as $P(k) \propto k^{-\gamma}$, can be

155   considered as scale free.

156

2.3. Multifractal measurements

158   While fractal analysis is based on the complexity of a fractal set, the multifractal

159   approach can describe the distribution of a given measure over a fractal object

160   (Mandelbrot, 1974; Halsey et al., 1986). It implies the possibility of having different

161   densities depending on the region of application.

162   There are two ways of representing multifractals: the generalized fractal dimensions

163   D_q and the singularity or multifractal spectrum ($f(\alpha)$). Both of them are discussed below

164   separately. Typically, the multifractal analysis has been widely performed by means of

165   the fixed-size algorithms (FSA), that rely on the subdivision of the system into smaller

166   parts with equal size and then that size is changed iteratively. The method used in this

167   work is the "sand-box algorithm" and will be discussed later on.

168   As stated before, the first of the measurements when it comes to multifractal

169   analysis are the generalized fractal or Rényi dimensions D_q, which describe the scaling

170   exponents of the qth moments of the system and can be defined (Feder, 1988) as:

$$D_q = \frac{1}{q-1} \lim_{\delta \to 0} \frac{\ln Z_q(\delta)}{\ln \delta} \quad \forall\, q \neq 1 \tag{3}$$

171

$$D_1 = \lim_{\delta \to 0} \frac{\sum_{i=1}^{N_c(\delta)} \mu_i(\delta) \ln \mu_i(\delta)}{\ln \delta} \qquad (4)$$

172

173   With q the moment order, δ the size of the used cells to cover the system, $Z_q(\delta)$ the

174   partition function, $N_c(\delta)$ the number of cells with length δ and $\mu_i(\delta)$ the probability

175   measurement of each cell. The expression for D_1 is obtained taking the limit of D_q when

176     $q \to 1$.

177   From these generalized fractal dimensions, it must be pointed out that $D_{q=0}$

178   corresponds to the fractal dimension of the given system or *box-counting dimension*,

179     $D_{q=1}$ is the so-called *information entropy* and $D_{q=2}$ the *correlation dimension.* The limits

180   of D_q when q goes to $-\infty$ and $+\infty$ describe the scaling properties of the regions where

181   the measure is more rarified and concentrated, respectively.

182   The other set of multifractal parameters is the so-called singularity spectrum ($f(\alpha)$),

183   as commented previously. A frequent method to determine is based on the use of a

184   Legendre transform from mass exponents $\tau(q)$. However, some authors as Chhabra and

185   Jensen (1989) and Veneziano et al (1995), state that it can lead to some errors due to

186   the inclusion of spurious points and error amplification from the derivative.

187   Furthermore, it does not yield independent measurements from the Rényi spectrum, as

188     $\tau(q) = (1 - q)D_q$. As an alternative to this and more focused on experimental data,

189   Chhabra and Jensen (1989) proposed a direct method to determine the α-spectrum,

190   overcoming the drawbacks referred before. This technique relies on the normalized

191   measure $\beta_i(q)$, μ_i in the original work, for computing the probabilities of the boxes of

192   radius r:

$$\beta_i(q,r) = [P_i(r)]^q / \sum_j [P_j(r)]^q \tag{5}$$

193   With $P_i(r)$ the different fractal measurements for each box of radius r (number of

194   nodes in this case). From this, $f(\alpha)$ and α can be retrieved using the next expressions:

$$f(q) = \lim_{r \to 0} \frac{\sum_i \beta_i(q,r) \log[\beta_i(q,r)]}{\log r} \tag{6}$$

195

$$\alpha(q) = \lim_{r \to 0} \frac{\sum_i \beta_i(q,r) \log[P_i(r)]}{\log r} \tag{7}$$

196   Being α known as the Lipschitz-Hölder exponent. In practice, those quantities are

197   computed from the slope of $\sum_i \beta_i(q,r) \log[\beta_i(q,r)]$ over $\log r$ for $f(q)$; and

198    $\sum_i \beta_i(q,r) \log[P_i(r)]$ over $\log r$ for $\alpha(q)$. This slope is determined by means of a linear

199   regression in the same range of radii where the other fractal measures are computed.

200

201   2.4. <u>Sandbox algorithm</u>

202   The sandbox algorithm (SBA) was firstly introduced by Tél et al. (1989) and

203   developed by Vicsek et al. (1990), as an improvement of the previously used fixed-size

204   box-counting methods for computing the generalized fractal dimensions. The main

205   advantage of this method with respect to other box-counting FSA is that it is capable of

206   properly determine the side corresponding to the negative values of q from both the

207   Rényi and singularity spectra.

208   The basic idea behind the SBA is that for each radius, a number of randomly placed

209   boxes are selected, and they are always centered in a non-zero point of the system (a

210   node). In that way the entire network is covered with those boxes by choosing a

211   sufficiently high number of them. For each box (B), the probability measurement used

212   is computed as shown in equation (8).

$$\mu(B) = \frac{M(B)}{M_0} \tag{8}$$

213   Once that quantity is computed for each sandbox for a given radius, the generalized

214   fractal dimensions can be obtained as explained previously in the multifractal

215   measurement part. Applying it to equation (3), the following formula is obtained for

216   D_q^{sb}:

$$D_q^{sb} = \frac{1}{q-1} \lim_{r \to 0} \frac{\ln\langle \mu(B)^{q-1} \rangle}{\ln r} \quad \forall\, q \neq 1 \tag{9}$$

217   And the expression when $q = 1$ for the SBA also can be adapted from equation (4):

$$D_1^{sb} = \lim_{r \to 0} \frac{\langle \ln \mu(B) \rangle}{\ln r} \tag{10}$$

218   With the aim of implementing the SBA on the VG, the steps to follow are:

219   1) The original time series is transformed into a VG, resulting on a matrix as

220   described in the previous section.

221   2) A range of different radii is set in order to cover the entire network. These

222   radii are chosen between 1 and N, being N the total number of nodes. For

223   this case, the typical values are $r \in [1, 100]$, because larger values have

224   proven to give the same information, since the curves start to saturate.

225   3) Then, N_c centers are selected for each radius. This number is inversely

226   proportional to the radius itself, because the greater the box is, the smaller

227   number of boxes will be needed to cover the network. The location of the

228   centers is randomly chosen within the nodes of the graph.

229   4) For each sandbox, the amount of nodes inside the box connected to the

230   center (M) are counted, giving N_c values of $M(B)$. From here, the quantity

231     $\mu(B)$ (for every box) is computed by means of dividing the previous $M(B)$ by

232   the total amount of nodes connected to the center M_0. Then $\mu(B)$ is used to

233   compute the partition function for all the q values chosen, being the average

234   in equations (9) and (10) over all the sandboxes created.

235   5) The steps 3) and 4) are then repeated for all the radii considered, obtaining

236   a value of the partition function for each one of them and q.

237

238   3. RESULTS AND DISCUSSION

239   3.1. Degree distribution

240   Firstly, a fast method to check the fractal nature of the ozone concentration time

241   series, before performing a deeper study, is described here. This method consists on

242   analyzing the degree distribution of the VGs.

243   The mentioned distribution is computed as the number of nodes that have a

244   given degree and divided by the total number of them in the VG. It is clear that the

245   greater the degree is, the less likely to be repeated within the network it will be;

246   because large degrees are exclusive of nodes with the highest concentration (which

247   we will refer as *hub*) due to their typically high visibility. As those hubs correspond

248   normally to the extreme values of ozone concentration of each day, their likeliness

249   will be lower, since the most repeated ones will be the ones close to the average.

250   That leads to an expected distribution with a negative trend, as can be seen in Figure

251   3.

252   Once the degree distributions of the different months are computed, a clear

253   fractal behavior is observed since all of them follow a power law $P(k) \propto k^{-\gamma}$ as

254   expected, with a linear part in the last part of the log-log plot. After computing the

255   slope for the values of $k \geq 30$ in all the cases, the γ parameter can be determined,

256   obtaining that the biggest and the lowest ones correspond to July and January

257   respectively. Although this is a fast and direct method to determine whether the

258   series is fractal or not, it does not give much detailed information about the

259   differences between each moth, since all of them share a very similar degree

260   distribution. Hence a deeper analysis devoted to the multifractal properties of the

261   series has been done with that aim, as will be presented in the following section, in

262   order to give some light to the usability of this degree distribution and some of its

263   parameters.

264

265   3.2. Multifractality

266   In this part, the authors present two independent methods to study the

267   multifractality of the signal, both previously described: the generalized fractal

268   (Rényi) dimensions and the α spectra.

269   For the analysis of the Rényi dimensions of the samples, the quantities

270    $\ln\langle\mu(B)^{q-1}\rangle \, / \, (q-1)$ for $q \neq 1$ and $\langle\ln\mu(B)\rangle$ for $q = 1$ are plotted against $\ln r$, as

271   seen in Figure 4. The interval used for the radii goes from $r = 1$ to $r = 100$, as it was

272   proven to be enough in order to reproduce the expected linear behavior (after

273 several tests). It must be pointed out that the interval taken for the linear regression

274 is always selected so that the Pearson correlation coefficient is $r \geq 0.99$. For April,

275 July and October, the chosen range has been from $\ln r = 3$ up to end of the curve;

276 whereas for January the range had to be taken such as $\ln r \in [2,4]$ because the

277 linear part is larger and to avoid an artifact that appears for higher radii and made

278 the results misleading (can be observed in Figure 7). In all those cases, the values

279 correspond with the higher radii used in the SBA, in accordance with previous studies

280 that used the same technique (Ariza-Villaverde et al., 2015; De Bartolo et al., 2004;

281 Jiménez-Hornero et al. 2013). After the linear regressions are performed for all

282 values of q of each month, the Rényi spectrum can be constructed following

283 equations (9) and (10). In Figure 5 the result of such procedure is shown in the form

284 of the generalized fractal dimensions. After observing the graph, it can be inferred

285 that the VG properly reproduces the multifractal behavior of the series, which was

286 previously demonstrated to be so (He, 2017; Jiménez-Hornero et al., 2010a, 2010b;

287 Pavón-Domínguez et al., 2013), since for all the studied months $D_0 > D_1 > D_2$ (see

288 Table 1). The difference between the maximum value of D_q and the minimum (ΔD_q)

289 is usually taken as a measure of the multifractality degree of the signal (ozone

290 concentration in this case) (Ariza-Villaverde, 2013; Telesca et al., 2004). Looking at

291 Table 1, this degree seems to be sensibly higher for July than the rest of the months,

292 followed by April and October (with a very similar value) and then the last one is

293 January.

294 Now, the fractal dimensions themselves from the Rényi spectrum (D_0, D_1 and

295 D_2), are being discussed. Firstly, the capacity dimension or "box-counting"

296 dimension D_0 is related with the number of boxes needed in order to cover the

297   fractal object. The more boxes are needed, the more extended is the fractal object

298   and then the grater would be D_0. Since in this case the fractals are different for each

299   iteration (each node sees a different visibility network), the final behavior would be

300   expected to be an average of the whole system. In the context of these complex

301   networks, a high extension of the object would mean a bigger degree (then related

302   to hubs). Following that logic, a higher value of D_0 is related with greater number of

303   hubs in the system, meaning that one could expect the average degree $\bar{k}$ of the

304   distribution increased. This is found in the results as can be observed in Table 1.

305   Values for the months of January and April are very close to each other, while the

306   one for July is sensibly the highest and October has an intermediate value. This result

307   is in accordance to the behavior of the ozone concentration, since for instance the

308   month of July is the one with the greatest temperature and UV radiation rates and

309   therefore elevated values are reached more often, producing a major number of

310   hubs in the VG.

311   The second parameter is the entropy dimension D_1, which is usually known as

312   the dimension that is related to the uniformity of the data and how different is the

313   probability of certain regions to be visited by a randomly chosen box with respect to

314   others. A way of measuring its uniformity is looking at the difference between D_0

315   and D_1, because when they are equal it means that the sample is uniform. The

316   greater $(D_0 - D_1)$, the less uniform it is. Also, another reason to take the difference

317   instead of the absolute value is because we are interested on comparing the curves

318   and their behavior, and since the D_0 of each one is different, it is necessary to

319   stablish a reference point. When it is translated to the context of VG, a greater

320   uniformity would mean less difference between the degrees of the sample, hence

321   the standard deviation of the degree would be decreased. In that case, what should

322   be observed in the original ozone data is that the differences encountered in the

323   data are less significative (what is indeed observed for the months with lower $(D_0 -$

324    $D_1)$. The most uniform would be January and again the other extreme is found for

325   July. The reason for this is that in July the differences encountered between day and

326   night concentrations of ozone are much greater than the other months, since during

327   the night they remain at their minimum values for all months. In contrast, the

328   maximum values are much higher in summer than the rest of the year because of

329   the higher UV exposures and the opposite case for winter.

330   The last parameter that has been computed from the Rényi spectrum is the

331   correlation dimension D_2. For the same objective reason as in $(D_0 - D_1)$, here $(D_0 -$

332    $D_2)$ values are discussed as well. Once again, January and July exhibit the extreme

333   cases whereas the other two months are located in between. In this case the authors

334   have not been able to directly relate this magnitude with any property of the degree

335   distribution of the VGs as for $(D_0 - D_1)$ and D_0.

336   As commented before, another important feature that can be studied is the

337   spectrum of α, following the equations (6) and (7). Applying a similar approach to

338   the one used to compute D_q, a linear regression is performed for each value of q to

339   compute the corresponding magnitude (α in Figure 6 and $f(\alpha)$ in Figure 7). For each

340   month, the range for the regression has been chosen to be the same that was used

341   for the Rényi dimensions, in order to conserve the scale used. In this case, the

342   obtained curves have as well a positive trend that increases with q for the case of

343   the curves of Figure 6; and first increases and then decreases for Figure 7 (as

344   expected, because $f(\alpha)$ should have a maximum and then drop). As commented

345   before, the artifact that made us change the chosen interval for the regression in

346   January can be observed in Figure 7 for high values of radius.

347   Once both α and $f(\alpha)$ are resolved for all the possible values of q, the spectrum

348   of every month can be plotted, being the ones shown in Figure 8. All the spectra as

349   can be seen have their maximum at $f(\alpha) = 1$ and start at $\alpha = 0$, while the overall

350   shape depends on each case. Several properties of the underlying signal can be

351   extracted from the spectra. The width of the curve W is shown in Table 1, as well as

352   the position of the curve maximum α_0. As in the case of generalized fractal

353   dimensions D_q, July is the month with the widest spectrum and highest α_0. January

354   as well exhibits the lowest values, whereas April and October have intermediate

355   ones. This width W is related with the multifractality degree of the signal (Ariza-

356   Villaverde, 2013; Telesca et al., 2004).

357   According to the shapes of the curves, the spectra are strongly non-symmetric,

358   being the right tail much more pronounced than the left one for all the different

359   months. The right side is usually associated to the homogeneity of low values in a

360   temporal distribution of data, whereas the left one is related with the same feature

361   of the high values instead. Therefore, the behavior of the four months is very similar

362   for high values (left tail) while it differs significantly for low ones (right tail). In Figure

363   8, all of them show an heterogenous behavior since the α distribution is not uniform

364   along the curve, being concentrated on the left and right extremes. Nevertheless, it

365   is possible to extract from that figure that this heterogeneity is more pronounced in

366   July. This fact suggests an influence of the higher UV radiation that creates a greater

367   difference between day and night ozone concentrations. This is in accordance with

368   what was shown in the previous analysis described above for $(D_0 - D_1)$.

369

370   4. <u>CONCLUSIONS</u>

371   The results of this work show that the multifractal analysis of VGs from ground-

372   level ozone concentration time series is a suitable tool to describe the seasonal

373   dynamics of this air pollutant. VGs have proven to have several advantages such as:

374   i) their topology inherits the features of the associated time series, which ends up

375   resulting on supplementary information through the degree distribution; ii) they can

376   be use for both stationary and non-stationary time series, removing the multifractal

377   analysis requirement of dealing only with stationary processes. Besides, VGs filter

378   out trends in the signal, making unnecessary to apply detrended methods. iii) Also,

379   they can be applied to multivariate time series, which can be very helpful in order to

380   find correlations between tropospheric ozone and its precursors for instance; iv) and

381   finally, this novel connection between time series and complex networks opens a

382   broad range of possibilities within the study of complex signals.

383   When it comes to the multifractal analysis performed in this work, both the SBA

384   and the Chhabra and Jensen method for the Rényi and singularity spectra

385   respectively, where chosen based on their advantages with respect to other. The

386   SBA overcomes the drawbacks of the box-counting algorithm for the computation

387   of the generalized fractal dimensions for negative probability moment orders q.

388   Furthermore, the Chhabra and Jensen method for the α-spectrum does not need a

389   Legendre transform to be applied (with the associated error to it) and gives and

390   independent measurement from the SBA as well, being more robust for comparison.

391   After applying the methods mentioned above, clear and coherent results that fits

392   the expected behavior of the ozone dynamics were found for the different months.

393   Furthermore, several properties can be directly inferred from the degree

394   distribution almost at first sight, meaning a powerful tool for predicting results

395   before any more complex data treatment is performed. The same relations between

396   multifractal parameters such as capacity dimension, W of the α-spectrum amongst

397   others are found in simple statistical parameters as the average or standard

398   deviation of the degree distribution of the VG.

399   To conclude, authors would like to point out that this promising technique could

400   be extended to other applications due to the many possibilities that complex

401   networks have. One possible target would be the relation between ozone and its

402   precursors, which could be looked by means of relating the VG of each one

403   separately or using multi-layer networks as proposed by Lacasa et al. (2015).

404

410

411   6. <u>REFERENCES</u>

412   Adame, J.A., Lozano, A., Bolívar, J.P., De la Morena, B.A., Contreras, J., Godoy, F., 2008.

413   Behavior, distribution and variability of surface ozone at an arid region in the

414   south of Iberian Peninsula (Seville, Spain). Chemosphere 70, 841–849.

415   https://doi.org/10.1016/j.chemosphere.2007.07.009

416

417   Ariza-Villaverde, A.B., Jiménez-Hornero, F.J., Gutiérrez de Ravé, E., 2015. Influence of

418   DEM resolution on drainage network extraction: A multifractal analysis.

419   Geomorphology 241, 243–254.

420   https://doi.org/10.1016/j.geomorph.2015.03.040

421

422   Ariza-Villaverde, A.B., Jiménez-Hornero, F.J., Gutiérrez de Ravé, E., 2013. Multifractal

423   analysis applied to the study of the accuracy of DEM-based stream derivation.

424   Geomorphology 197, 85–95.

425   https://doi.org/10.1016/j.geomorph.2013.04.040

426

427   Chhabra, A., Jensen, R.V., 1989. Direct determination of the $f(\alpha)$ singularity spectrum.

428   Physical Review Letters 62, 1327–1330.

429   https://doi.org/10.1103/PhysRevLett.62.1327

430

431   De Bartolo, S.G., Gaudio, R., Gabriele, S., 2004. Multifractal analysis of river networks:

432   Sandbox approach. Water Resources Research 40.

433   https://doi.org/10.1029/2003WR002760

434

435   Feder, J., 1988. Fractals, Physics of Solids and Liquids. Springer US.

436

437   Graedel, T.E., Crutzen, P.J., 1993. Atmospheric change: an earth system perspective.

438   Atmospheric change: an earth system perspective.

439

440   Guicherit, R., van Dop, H., 1977. Photochemical production of ozone in Western Europe

441   (1971–1975) and its relation to meteorology. Atmospheric Environment (1967)

442   11, 145–155.

443   https://doi.org/10.1016/0004-6981(77)90219-0

444

445   Güsten, H., Heinrich, G., Weppner, J., Abdel-Aal, M.M., Abdel-Hay, F.A., Ramadan, A.B.,

446   Tawfik, F.S., Ahmed, D.M., Hassan, G.K.Y., Cvitaš, T., Jeftić, J., Klasinc, L., 1994.

447   Ozone formation in the greater Cairo area. Science of The Total Environment

448   155, 285–295.

449   https://doi.org/10.1016/0048-9697(94)90507-X

450

451   Halsey, T.C., Jensen, M.H., Kadanoff, L.P., Procaccia, I., Shraiman, B.I., 1986. Fractal

452   measures and their singularities: The characterization of strange sets. Phys. Rev.

453   A 33, 1141–1151.

454   https://doi.org/10.1103/PhysRevA.33.1141

455

456   He, H., 2017. Multifractal analysis of interactive patterns between meteorological

457   factors and pollutants in urban and rural areas. Atmospheric Environment 149,

458   47–54. https://doi.org/10.1016/j.atmosenv.2016.11.004

459

460   Jiménez-Hornero, F.J., Ariza-Villaverde, A.B., De Ravé, E.G., 2013. Multifractal

461   description of simulated flow velocity in idealised porous media by using the

462   sandbox method. Fractals 21, 1350006.

463   https://doi.org/10.1142/S0218348X13500060

464

465   Jiménez-Hornero, F. J., Gutiérrez de Ravé, E., Ariza-Villarverde, A.B., Giráldez, J.V.,

466   2010a. Description of the seasonal pattern in ozone concentration time series by

467   using the strange attractor multifractal formalism. Environmental Monitoring

468   and Assessment 160, 229–236.

469   https://doi.org/10.1007/s10661-008-0690-y

470

471   Jiménez-Hornero, Francisco J., Jiménez-Hornero, J.E., Gutiérrez de Ravé, E., Pavón-

472   Domínguez, P., 2010b. Exploring the relationship between nitrogen dioxide and

473   ground-level ozone by applying the joint multifractal analysis. Environmental

474   Monitoring and Assessment 167, 675–684.

475   https://doi.org/10.1007/s10661-009-1083-6

476

477   Kouvarakis, G., Tsigaridis, K., Kanakidou, M., Mihalopoulos, N., 2000. Temporal

478   variations of surface regional background ozone over Crete Island in the

479   southeast Mediterranean. Journal of Geophysical Research: Atmospheres 105,

480   4399–4407. https://doi.org/10.1029/1999JD900984

481

482   Lacasa, L., Luque, B., Ballesteros, F., Luque, J., Nuño, J.C., 2008. From time series to

483   complex networks: The visibility graph. Proceedings of the National Academy of

484   Sciences 105, 4972–4975.

485   https://doi.org/10.1073/pnas.0709247105

486

487 Lacasa, L., Luque, B., Luque, J., Nuño, J.C., 2009. The visibility graph: A new method for

488   estimating the Hurst exponent of fractional Brownian motion. EPL (Europhysics

489   Letters) 86, 30001.

490   https://doi.org/10.1209/0295-5075/86/30001

491

492 Lacasa, L., Toral, R., 2010. Description of stochastic and chaotic series using visibility

493   graphs. Physical Review E 82.

494   https://doi.org/10.1103/PhysRevE.82.036120

495

496 Mali, P., Manna, S.K., Mukhopadhyay, A., Haldar, P.K., Singh, G., 2018. Multifractal

497   analysis of multiparticle emission data in the framework of visibility graph and

498   sandbox algorithm. Physica A: Statistical Mechanics and its Applications 493,

499   253–266.

500   https://doi.org/10.1016/j.physa.2017.10.015

501

502 Mandelbrot, B.B., 1982. The fractal geometry of nature. W.H. Freeman, San Francisco.

503

504 Pavon-Dominguez, P., Jiménez-Hornero, F.J., Gutierrez de Rave, E., 2013. Multifractal

505   analysis of ground–level ozone concentrations at urban, suburban and rural

506   background monitoring sites in Southwestern Iberian Peninsula. Atmospheric

507   Pollution Research 4, 229–237.

508   https://doi.org/10.5094/APR.2013.024

509

510   Ribas, À., Peñuelas, J., 2004. Temporal patterns of surface ozone levels in different

511   habitats of the North Western Mediterranean basin. Atmospheric Environment

512   38, 985–992.

513   https://doi.org/10.1016/j.atmosenv.2003.10.045

514

515   Sillman, S., 1999. The relation between ozone, NOx and hydrocarbons in urban and

516   polluted rural environments. Atmospheric Environment 33, 1821–1845.

517   https://doi.org/10.1016/S1352-2310(98)00345-8

518

519   Tél, T., Fülöp, Á., Vicsek, T., 1989. Determination of fractal dimensions for geometrical

520   multifractals. Physica A: Statistical Mechanics and its Applications 159, 155–166.

521   https://doi.org/10.1016/0378-4371(89)90563-3

522

523   Telesca, L., Colangelo, G., Lapenna, V., Macchiato, M., 2004. Fluctuation dynamics in

524   geoelectrical data: an investigation by using multifractal detrended fluctuation

525   analysis. Physics Letters A 332, 398–404.

526   https://doi.org/10.1016/j.physleta.2004.10.011

527

528   Trainer, M., Parrish, D.D., Goldan, P.D., Roberts, J., Fehsenfeld, F.C., 2000. Review of

529   observation-based analysis of the regional factors influencing ozone

530   concentrations. Atmospheric Environment 17.

531

532   Veneziano, D., Moglen, G.E., Bras, R.L., 1995. Multifractal analysis: Pitfalls of standard

533   procedures and alternatives. Phys. Rev. E 52, 1387–1398.

534   https://doi.org/10.1103/PhysRevE.52.1387

535

536 Vicsek, T., Family, F., Meakin, P., 1990. Multifractal Geometry of Diffusion-Limited

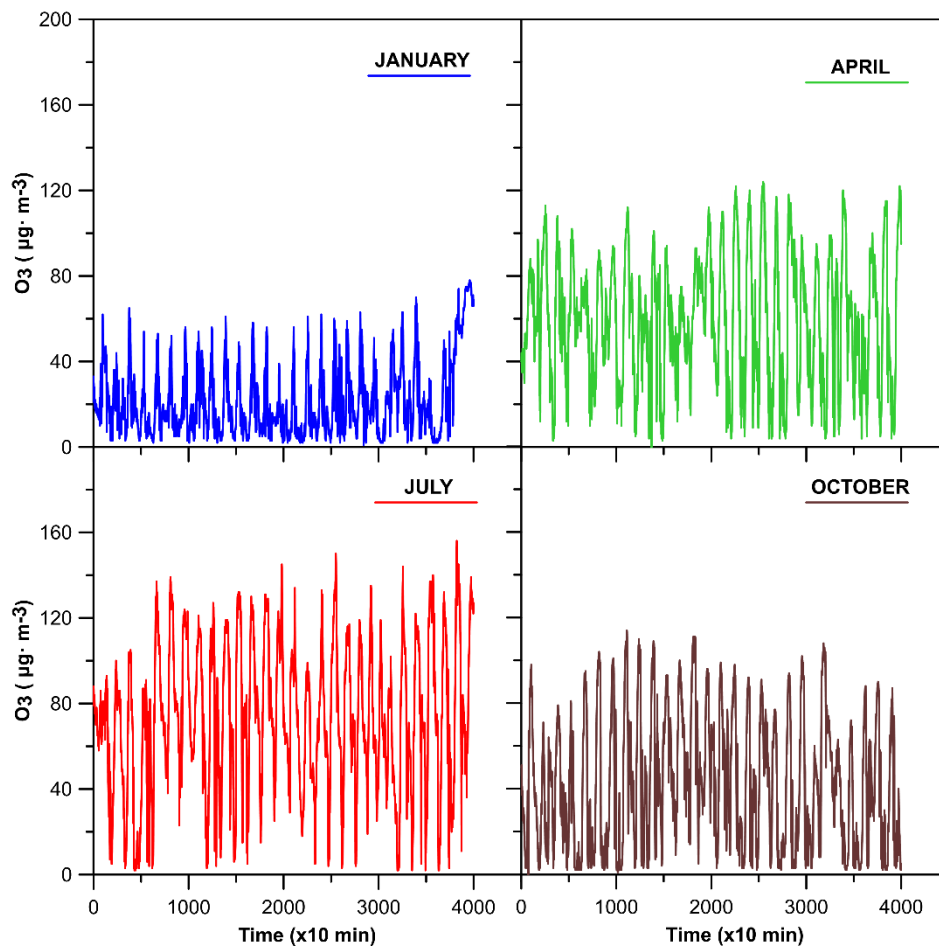537   Aggregates. Europhysics Letters (EPL) 12, 217–222.

538   https://doi.org/10.1209/0295-5075/12/3/005

539

540 Zeleke, T.B., Si, B.C., 2006. Characterizing scale-dependent spatial relationships between

541   soil properties using multifractal techniques. Geoderma 134, 440–452.

542   https://doi.org/10.1016/j.geoderma.2006.03.013

543

544

545   7. <u>FIGURES</u>

546   8.

547   *Figure 1: Ozone time series for the different analyzed months.*
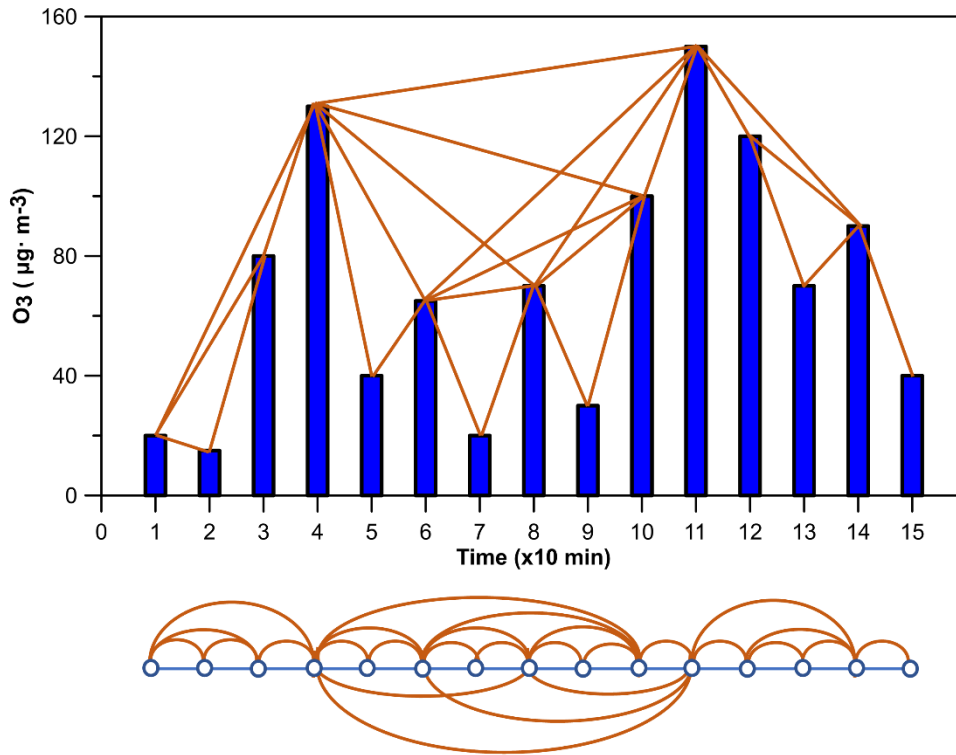
Figure 2: Sample time series transformed into a complex network through the visibility algorithm. Below, all the connections are shown in a more visual way.
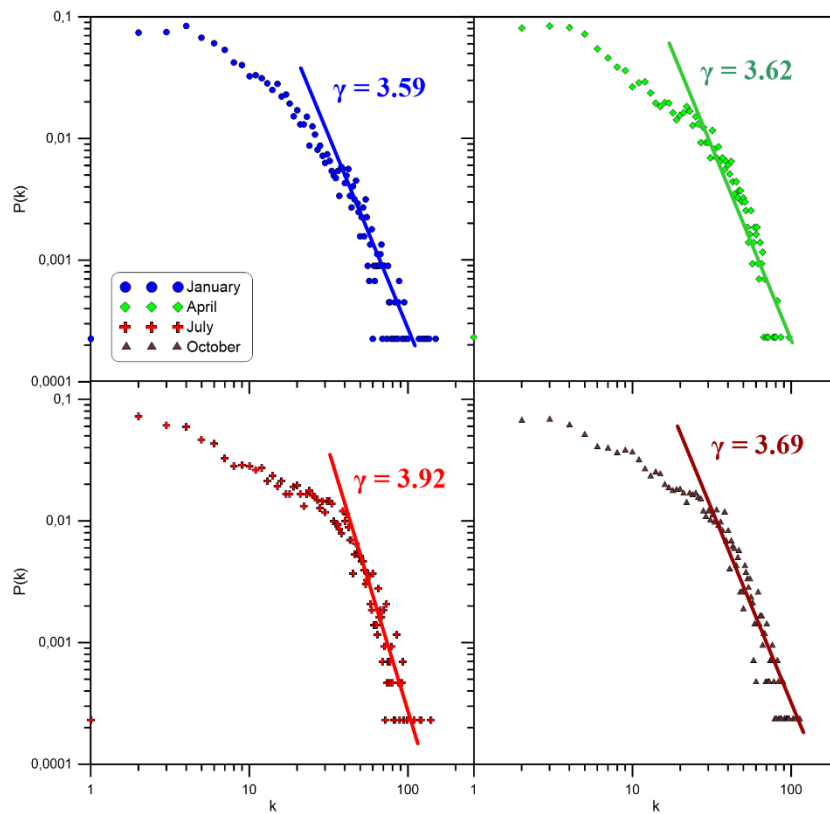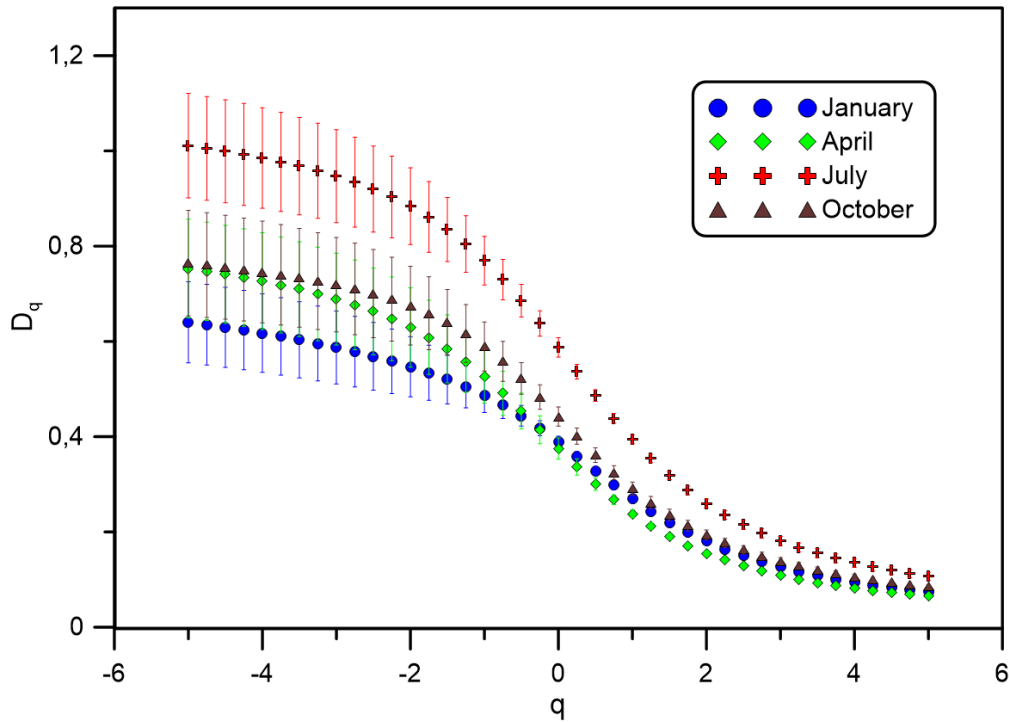
553   *Figure 3: Degree distribution of the visibility graph from each month in logarithmic*

554   *scale. As it can be observed, the tail of the distribution shows a scale-free behavior*

555   *because it can be fitted by a power law.*
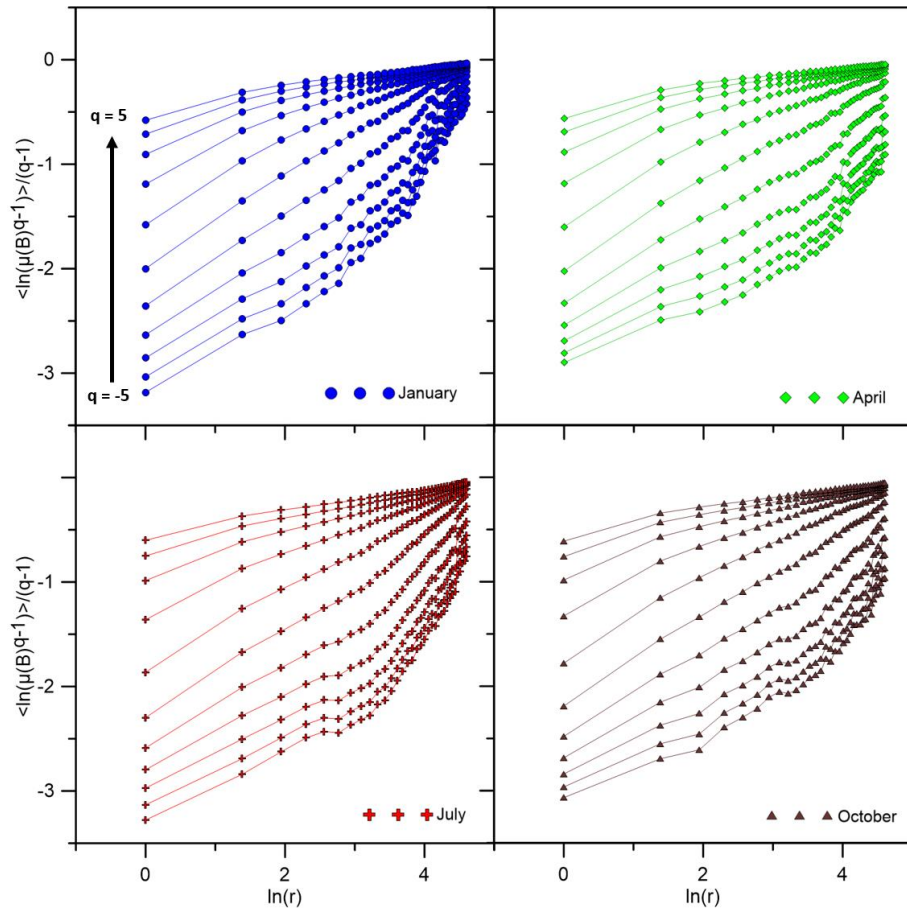
556

557   *Figure 4: Curves of the partition function against ln(r) obtained after applying the*
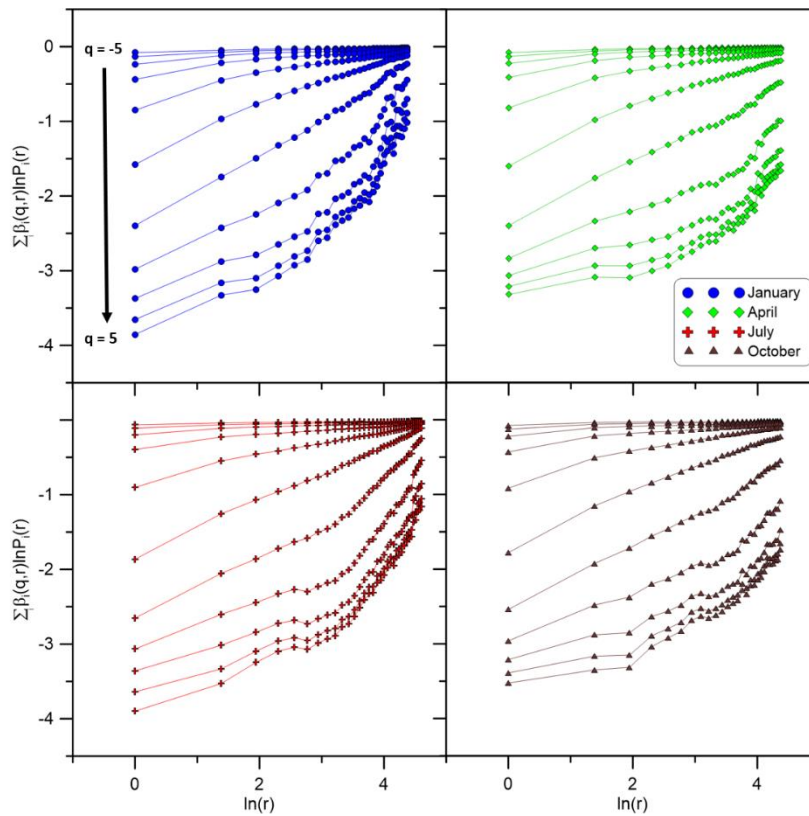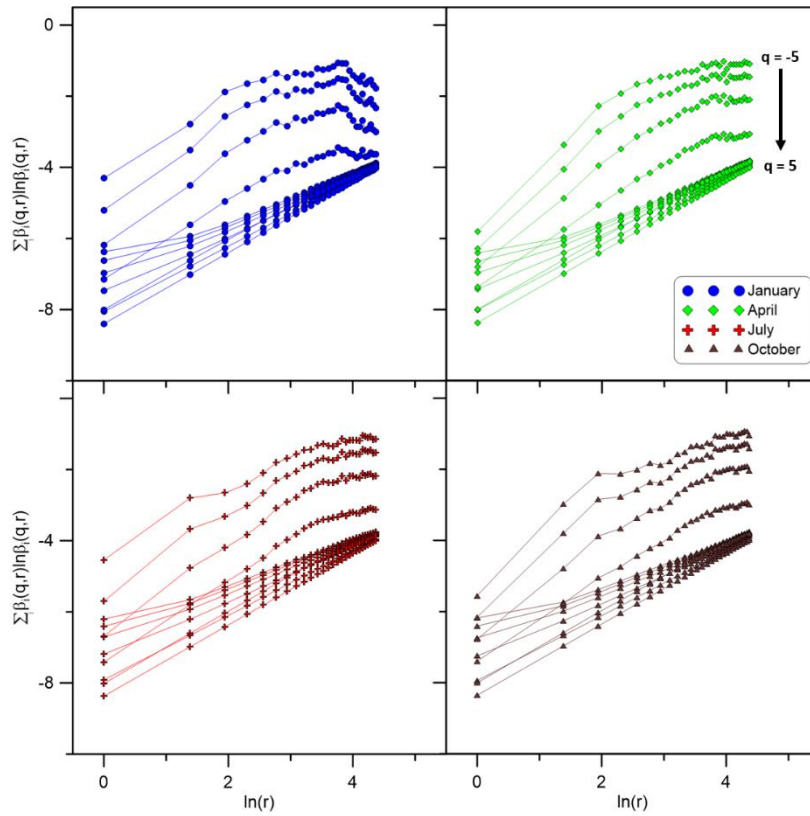558   *sandbox algorithm.*

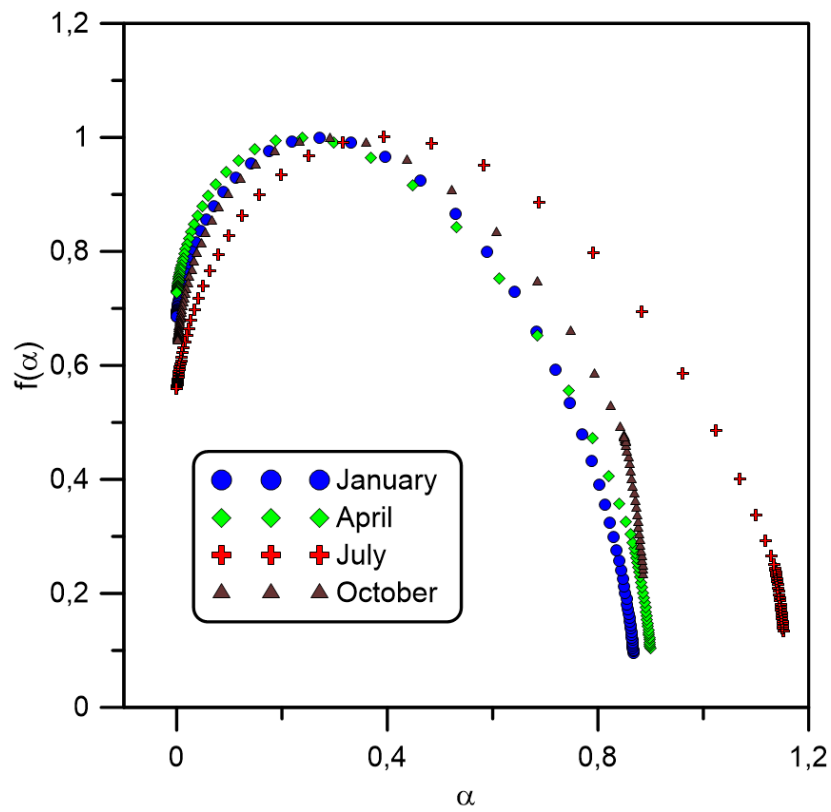559

Figure 5: Rényi dimensions for each month.

Figure 6: Curves from where α is computed by regression.

564

565   *Figure 7: Curves from where f(α) is computed using linear regressions.*

566

567   *Figure 8: α-spectra of all the months.*

568

569 9. <u>TABLES</u>

570 *Table 1: Relevant values for each month.*

MONTH	Average Direct Radiation (MJ/m^2)	γ	$\bar{k}$	σ_k	D_0	$D_0 - D_1$	$D_0 - D_2$	ΔD_q	α_0	W
January	310.2	3.59	14.78	15.00	0.389±0.012	0.12±0.03	0.21±0.02	0.56±0.10	0.291	0.779
April	577.8	3.62	15.12	13.98	0.37±0.03	0.13±0.04	0.22±0.04	0.69±0.12	0.246	0.934
July	991.4	3.92	19.94	17.69	0.59±0.02	0.20±0.04	0.33±0.03	0.90±0.12	0.443	1.152
October	419.3	3.69	17.90	15.91	0.44±0.02	0.15±0.04	0.25±0.04	0.68±0.12	0.333	0.841

571

1   **<u>HIGHLIGHTS</u>**

2   -   Ground-level ozone concentration time series have a multifractal behavior.

3   -   Visibility graphs can be used to analyze multifractality of ozone time series.

4   -   Many aspects of ozone dynamics can be observed through the degree

5   distribution.

6   -   This technique gives supplementary information within the study of complex

7   signals.

1   **<u>HIGHLIGHTS</u>**