Entropy **2015**, *17*, 4155-4172; doi:10.3390/e17064155

OPEN ACCESS
entropy
ISSN 1099-4300
www.mdpi.com/journal/entropy

Article

2D Anisotropic Wavelet Entropy with an Application to Earthquakes in Chile

Orietta Nicolis [1] and Jorge Mateu [2,*]

[1] Institute of Statistics, University of Valparaiso, Av.Gran Bretaña 1111 , Valparaiso, Chile; E-Mail: orietta.nicolis@uv.cl

[2] Institute of Mathematics, University Jaume I of Castellón, Campus Riu Sec, Castellon 12071, Spain

* Author to whom correspondence should be addressed; E-Mail: mateu@uji.es.

Academic Editor: Raúl Alcaraz Martínez

Received: 12 May 2015 / Accepted: 4 June 2015 / Published: 16 June 2015

Abstract: We propose a wavelet-based approach to measure the Shannon entropy in the context of spatial point patterns. The method uses the fully anisotropic Morlet wavelet to estimate the energy distribution at different directions and scales. The spatial heterogeneity and complexity of spatial point patterns is then analyzed using the multiscale anisotropic wavelet entropy. The efficacy of the approach is shown through a simulation study. Finally, an application to the catalog of earthquake events in Chile is considered.

Keywords: anisotropy; earthquake events; entropy; spatial point patterns; wavelets

1. Introduction

The concept of entropy was first introduced by [1] in thermodynamics as a measure of the amount of energy in a system, and [2] was the first who gave a probabilistic interpretation. Shannon [3] applied the entropy concept to the information theory. According to the information theory, entropy is a measure of the uncertainty and unpredictability associated with a random variable, and the Shannon entropy quantifies the expected value of information generated from a random variable.

The definition of entropy has been widely used in many applications, such as neural systems [4], image segmentation through thresholding [5,6] and climatology and hydrology [7–10]. Nicholson *et al.* [11] introduced a spatial entropy to quantify the simplification of earthquake

distributions due to relocation procedures. A temporal definition of entropy was used by [12–16] to study the seismicity in different parts of the world.

For a process characterized by a certain number N of states or classes of events, the Shannon entropy [3] is defined as:

$$S = H_1 = -\sum_{i=1}^{N} p_i \log(p_i),$$

where p_i is the probability of occurrence of the events in each ith class. The choice of the base of the logarithm is arbitrary: for practical convenience, we use base two throughout this paper. For $p_i = 0$, $p_i \log p_i = 0$. The Shannon entropy is maximal when all of the outcomes are equally likely, that is $S = \log(N)$. The Shannon entropy S can also be defined in a continuous framework as:

$$S = -\int_0^{\infty} P(x) \log(P(x)) dx,$$

where $P(x)$ is the continuous probability density function of the variable x.

A generalization of the Shannon entropy was defined by [17] as:

$$H_\alpha = \frac{1}{1-\alpha} \log \left(\sum_{i=1}^{N} p_i^\alpha \right),$$

where α is a parameter ($\alpha = 1$ gives the Shannon entropy). Papoulis [18] extended the definition of entropy to a point set as:

$$S = -\int_V P(\mathbf{x}) log(P(\mathbf{x})) d\mathbf{x}, \quad (1)$$

where P is the spatial probability density, and $\mathbf{x}$ is a point in the spatial domain V. One method for estimating the density P is to divide the spatial domain into boxes and then count the number of points in each box (the box-counting method). Moreover, the optimal size of the boxes is an open problem: too big boxes produce a low resolution in measuring the variability of density, and too small boxes cause a large increase in density error. To avoid this problem, different approaches have been proposed [11,19].

However, a probability density is not the only type of distribution that can give information, and the definition of entropy can be extended to other types of distributions, such as the energy distribution based on wavelet coefficients [20]. A definition of wavelet entropy based on the energy distribution of the wavelet coefficients has been proposed by [21–26]. All of these authors propose or use a definition of the wavelet entropy in time, but an extension to the two-dimensional (2D) case has not been proposed so far.

The goal of this paper is to extend the wavelet entropy to the 2D case and to introduce new definitions of wavelet entropy using the Morlet directional wavelet in order to detect the spatial heterogeneity and complexity of spatial point patterns through scales and directions. The efficacy of the method is shown through a simulation study, and the directional wavelet entropy is then applied to the earthquake catalog of Chile for describing the spatial complexity of seismic events.

Wavelet analysis has succeeded in a variety of applications and held promise in the area of spatial pattern analysis (e.g., [27–29]). The main advantages are its ability to preserve and display locational information, while allowing for pattern decomposition, and it does not require the stationarity of the

data. Despite its advantages, wavelet analysis has only been involved in several works for the detection of patterns (e.g., [30–34]).

Spatial point process models are useful tools to model irregularly-scattered point patterns that are frequently encountered in many studies of natural phenomena. A spatial point pattern is a set of points $\{\mathbf{x}_i \in A : i = 1, \ldots, n\}$ in some planar region A. Very often, A is a sampling window within a much larger region, and it is reasonable to regard the point pattern as a partial realization of a stochastic planar point process, the events consisting of all points of the process that lie within A. Let N be this stochastic planar point process defined on $\mathbb{R}^2$, but observed on a finite observation window W. For an arbitrary set $A \in \mathbb{R}^2$, let $|A|$ and $N(A)$ denote the area of A and the number of events from N that are in A, respectively.

The study of spatial point patterns has a long history in ecology and forestry [35–38]. Spatial point patterns have also found application in fields as diverse as cosmology [39], geography [40], seismology [41] and epidemiology [42]. Recent textbooks related to the topic of analysis and modeling of point processes include [43–47]. A point process is stationary and isotropic if its statistical properties do not change under translation and rotation, respectively. Informally, stationarity implies that one can estimate properties of the process from a single realization on A, by exploiting the fact that these properties are the same in different, but geometrically similar, subregions of A; isotropy means that there are no directional effects.

The assumption of isotropy is often made in practice due to a simpler interpretation, ease of analysis and also to increase the power of statistical analyses. However, isotropy is many times hard to find in real applications. Many point processes are indeed anisotropic. There are many varied forms of anisotropy. Orientation analysis is the quantification of the degree of anisotropy in the case of non-isotropic point patterns and the detection of inner orientations in the case of isotropy [48–51]. Anisotropy can be present when the spatial point patterns contain points placed roughly on line segments [52]. A typical example of oriented point patterns is given by the distribution of earthquake epicenters after a "mainshock" event. "Aftershock" events are normally clustered along the linear directions or segments given by the active faults [53].

The arguments shown and the literature involved in the analysis and detection of spatial anisotropies sets a motivating research line in terms of the detection of linearities in spatial point patterns and, even more generally, in terms of testing for spatial anisotropy. Here, we understand spatial anisotropy as the presence of the main directions in the point pattern [54]. Ohser *et al.* [48], Brillinger [55] and Mateu *et al.* [34] proposed methods to assess isotropy (and to consequently detect anisotropy) for spatial point processes. The study of the spatial entropy of seismic events allows one to characterize the complexity of the earthquakes and, consequently, to assess their predictability.

The article is organized as follows. Section 2 provides some basic concepts of wavelets, and Section 3 gives a definition of wavelet entropy. Directional wavelets and the anisotropic wavelet entropy are introduced in Section 4. A simulation study is reported in Section 5. Finally, an application of the anisotropic wavelet entropy to the earthquake catalog of Chile is described in Section 6. The paper ends with some conclusions in Section 7.

2. Basics on Wavelets

One-dimensional wavelets are functions with zero mean and moderate decay, such that they are non-zero only over a small region. They can be defined as translations and re-scales of a single squared-integrable function $\psi(x) \in \mathbb{L}_2(\mathbb{R})$, called the wavelet function or the mother wavelet, as:

$$\psi_{a,b}(x) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{x-b}{a}\right). \tag{2}$$

where $a \in \mathbb{R}\backslash\{0\}$ and $b \in \mathbb{R}$ are the scale and shift parameters, respectively. Normalization by $\frac{1}{\sqrt{|a|}}$ ensures that the energy of the corresponding wavelet is independent of a and b, *i.e.*,

$$\int_{-\infty}^{\infty} |\psi_{a,b}(x)|^2 dx = \int_{-\infty}^{\infty} |\psi(x)|^2 dx.$$

For any function $f(x) \in \mathbb{L}_2(\mathbb{R})$, the continuous wavelet transform is given by:

$$W_f(a,b) = \langle f, \psi_{a,b}\rangle = \int_{\mathbb{R}} f(x)\overline{\psi_{a,b}(x)}dx, \tag{3}$$

where the overline denotes complex conjugate.

The two-dimensional extension of Equation (3) is straightforward: by denoting with $\mathbf{x} = (x, y)$ and $\mathbf{b} = (b_1, b_2)$ a spatial location and the translations, respectively, Equation (3) for $f(\mathbf{x}) \in \mathbb{L}_2(\mathbb{R}^2)$ can be written as:

$$W_f(a,\mathbf{b}) = \langle f, \psi_{a,\mathbf{b}}\rangle = \int_{\mathbb{R}^2} f(\mathbf{x})\overline{\psi_{a,\mathbf{b}}(\mathbf{x})}d\mathbf{x}, \tag{4}$$

where:

$$\psi_{a,\mathbf{b}}(\mathbf{x}) = \frac{1}{a}\psi\left(\frac{\mathbf{x}-\mathbf{b}}{a}\right).$$

is a 2D isotropic wavelet. An example of an isotropic wavelet is given by the Mexican hat whose wavelet mother is defined as $\psi_{MH}(\mathbf{x}) = (2 - |\mathbf{x}|^2 \exp\left(-\frac{1}{2}|\mathbf{x}|^2\right)$. Isotropic wavelets are normally used when no oriented feature is present or relevant in the signal or image.

3. Wavelet Entropy

The continuous wavelet entropy based on the energy distribution is defined by [22] as:

$$S_w(t) = \int_0^{\infty} -P_W(a,t)\log(P_W(a,t))da \tag{5}$$

where:

$$P_W(a,t) = \frac{|W_f(a,t)|^2}{\int |W_f(a,t)|^2 da}$$

is the wavelet energy probability distribution for each scale a at time t ($b = t$ in the wavelet definition of Equation (2)). Similarly to the classical definition, the wavelet entropy is maximum when the signal is a "white noise" process, and it is minimum when it is an ordered mono-frequency signal. In the latter case, approximatively 100% of the energy concentrates around one unique level (or scale), and the wavelet

entropy will be close to zero. Contrarily, when considering a "white noise", all levels will carry a certain amount of energy, and the wavelet entropy will be maximal.

Through this multiscale approach, one can detect the relevant scale levels that represent the highest complexity of the system.

Discretization choices of the parameters a and b can be used for defining a multiresolution wavelet entropy by allowing for a scale representation of the disorder of a system [56]. For example, the discretization $a = 2^j$ and $b = k2^{-j}$ produces an orthogonal basis given by $\{\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k), j, k \in \mathbb{Z}\}$ [57], and any function $f \in \mathbb{L}_2(\mathbb{R})$ can be represented as $f(x) = \sum_{j,k} d_{j,k}\psi_{j,k}(x)$, where $d_{j,k}$ are the multiresolution coefficients at the time k and scale j [58]. More general discretizations are given by $a = a_0^{-j}$, $b = kb_0a_0^{-j}$, $j, k \in \mathbb{Z}$, $a_0 > 1$, $b_0 > 0$.

In the discrete case, the energy of the signal can be approximated by $E_W(j,k) = |d_{j,k}|^2$. As previously seen, summing this energy for any discrete time k leads to an approximation of the energy content at scale j,

$$E_W(j) = \frac{1}{N}\sum_k E_W(j,k).$$

Following the Shannon entropy framework, the probability density distribution of energy across the scales is given by:

$$p_W(j) = \frac{E_W(j)}{\sum_j E_W(j)},$$

with $\sum_j p_W(j) = 1$ (because of the orthogonal representation).

Finally, the multiresolution wavelet entropy (MWE) at the scale j is defined by analogy with the continuous wavelet entropy:

$$MWE(j) = -p_W(j)\log(p_W(j)), \tag{6}$$

(see [20,23,24]).

The definition of wavelet entropy can be easily extended to 2D isotropic processes as follows:

$$S_W(\mathbf{b}) = \int_0^\infty -P_W(a,\mathbf{b})\log(P_W(a,\mathbf{b}))da \tag{7}$$

where:

$$P_W(a,\mathbf{b}) = \frac{|W_f(a,\mathbf{b})|^2}{\int |W_f(a,\mathbf{b})|^2 da}$$

corresponds to the wavelet energy probability distribution for each scale a at point $\mathbf{b}$.

Denoting by $E_W(a,\mathbf{b})$ the energy of the wavelet coefficients at point $\mathbf{b}$ and for scale a,

$$E_W(a,\mathbf{b}) = |W_f(a,\mathbf{b})|^2,$$

the energy content at scale a is given by:

$$E_W(a) = \int |W_f(a,\mathbf{b})|^2 d\mathbf{b}.$$

In order to follow the Shannon entropy framework, a probability density function must be defined as the ratio between the energy of each level and the total energy:

$$P_W(a) = \frac{E_W(a)}{\int E_W(a)da} = \frac{\int |W_f(a,\mathbf{b})|^2 d\mathbf{b}}{\int\int |W_f(a,\mathbf{b})|^2 d\mathbf{b}da}.$$

This corresponds exactly to the probability density distribution of energy across the scales where the following relation holds:

$$P_W(a) = 1.$$

Finally, the multiscale wavelet entropy (MWE) at the scale a is defined as:

$$MWE(a) = -P_W(a)\log(P_W(a)). \tag{8}$$

Integrating Equation (8) over all scales a, we obtain a measure of the global measure of wavelet entropy (GWE),

$$GWE = \int -P_W(a)\log(p_W(a))da. \tag{9}$$

In the discrete case, the 2D wavelet coefficients are denoted by $d_{j,\mathbf{k}}$, where $\mathbf{k}$ is a bidimensional vector, $\mathbf{k} \in \mathbb{R}^2$. Using the multiresolution analysis, the energy $E_W(j,\mathbf{k}) = |d_{j,\mathbf{k}}|^2$ can be used to define the 2D discrete multiscale wavelet entropy (DMWE),

$$DMWE(j) = -p_W(j)\log(p_W(j)), \tag{10}$$

where:

$$p_W(j) = \frac{E_W(j)}{\sum_j E_W(j)} = \frac{\sum_{\mathbf{k}=(k_1,k_2)} |d_{j,\mathbf{k}}|^2}{\sum_{\mathbf{k}=(k_1,k_2)} \sum_j |d_{j,\mathbf{k}}|^2}.$$

The discrete global measure of wavelet entropy (DGWE) is then obtained by summing the DMWE over all scales,

$$DGWE = \sum_j -p_W(j)\log(p_W(j)). \tag{11}$$

4. Directional Wavelet Entropy

When the aim of the study is to detect oriented features of a signal or an image, a directional wavelet has to be used. A directional wavelet is not rotation invariant, and its transform gives information about the best angular selectivity.

For $\mathbf{x} \in \mathbb{R}^2$ and any function $f(\mathbf{x}) \in \mathbb{L}_2(\mathbb{R}^2)$, the continuous directional wavelet transform for a scale a and an orientation θ is given by:

$$W_f(a,\mathbf{b},\theta) = \langle f, \psi_{a,\mathbf{b},\theta}\rangle = \int_{\mathbb{R}^2} f(\mathbf{x})\overline{\psi_{a,\mathbf{b}}(\mathbf{x},\theta)}d\mathbf{x}. \tag{12}$$

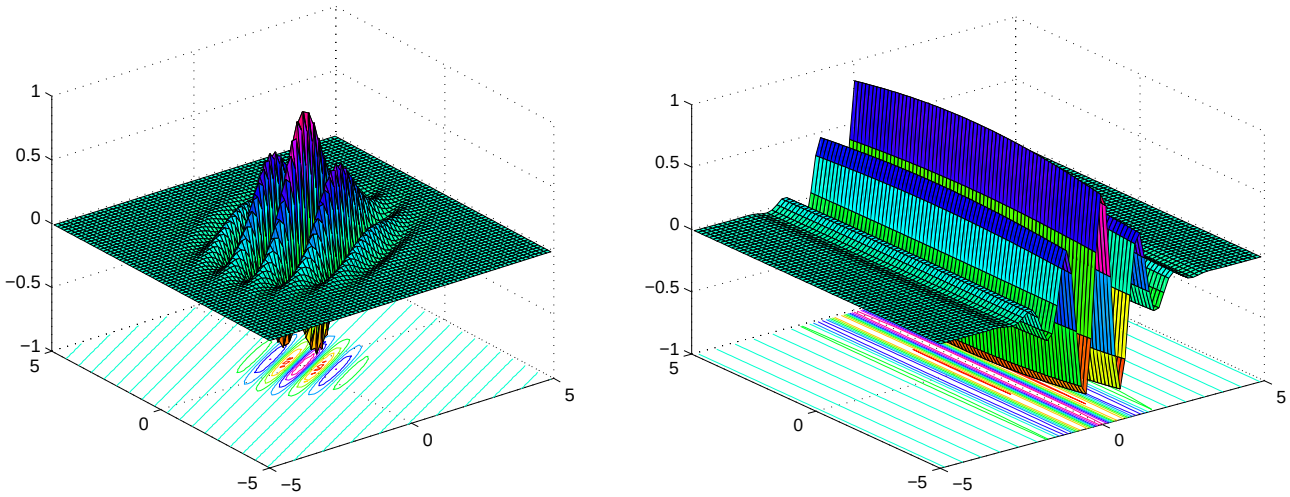
In the literature, a variety of directional wavelets $\psi_{a,\mathbf{b}}(\mathbf{x},\theta)$ have been proposed. In particular, [59] introduced a flexible function called the fully-anisotropic directional Morlet wavelet, given by:

$$\psi(\mathbf{x},\theta) = e^{i\mathbf{k_0}\mathbf{C}\mathbf{x}} e^{-1/2\mathbf{C}\mathbf{x}\mathbf{A}^T\mathbf{A}\mathbf{C}\mathbf{x}} \tag{13}$$

where $\mathbf{k_0} = (0, k_0)$ is a wave vector with $k_0 \geq 5.5$, $\mathbf{A} = \text{diag}(D, 1)$ denotes a diagonal matrix and D is the anisotropy ratio defined as the ratio of the length of the elliptical envelope in the y-direction to the length of the elliptical envelope in the x-direction. The matrix $\mathbf{C}$ is a linear transformation given by:

$$\mathbf{C} = \begin{pmatrix} \cos\theta & \sin\theta \\ \text{-}\sin\theta & \cos\theta \end{pmatrix}.$$

This transformation rotates the entire wavelet through an angle θ defined as positive in the counterclockwise direction. Two examples of this fully-anisotropic wavelet for directions $\theta = 30, 90$ are shown in Figure 1.



(a) (b)

Figure 1. Fully-anisotropic directional Morlet wavelet with parameters: (**a**) $D = 0.8$, $k_0 = 5.5$, $\theta = 30$; (**b**) $D = 0.1$, $k_0 = 5.5$, $\theta = 90$.

In order to identify the behavior of the process in different directions, [60] introduced two new functions, $\eta(a, \theta)$ and $\zeta(a, \theta)$, given by:

$$\eta(a, \theta) = \int |W_f(a, \mathbf{b}, \theta)|^2 d\mathbf{b} \tag{14}$$

and:

$$\zeta(a, \theta) = \frac{\eta(a, \theta)}{\int \eta(a, \theta) d\theta}. \tag{15}$$

The component $|W_f(a, \mathbf{b}, \theta)|^2$ in Equation (14) gives the distribution of the energy of a function at location $\mathbf{b}$, scale a and direction θ. Hence, $\eta(a, \theta)$ characterizes the distribution of the energy at different scales and directions, whereas $\zeta(a, \theta)$ provides the relative distribution of the energy in different directions at a particular scale. Mateu *et al.* [34] used fully-anisotropic wavelets for detecting linear patterns in spatial point processes.

Using the anisotropic wavelet of Equation (13), and denoting by:

$$E_W(a, \mathbf{b}, \theta) = |W_f(a, \mathbf{b}, \theta)|^2,$$

the energy of the wavelet coefficients at point **b**, scale a and with angle θ, one can define a probability density function for each direction θ and scale a as follows:

$$P_W(a,\theta) = \frac{E_W(a,\theta)}{\int_a \int_\theta E_W(a,\theta) da d\theta}, \tag{16}$$

where:

$$E_W(a,\theta) = \int_{\mathbb{R}^2} |W_f(a,\mathbf{b},\theta)|^2 d\mathbf{b}$$

and $\theta \in (0, 180)$. In this case, the multiscale directional wavelet entropy (MDWE) can be defined as:

$$MDWE(a,\theta) = -P_W(a,\theta)\log(P_W(a,\theta)). \tag{17}$$

Thus, by integrating the energy $E_W(a,\theta)$ over θ in Equation (16), that is,

$$P_W(a) = \frac{\int_\theta E_W(a,\theta) d\theta}{\int_a \int_\theta E_W(a,\theta) da d\theta}, \tag{18}$$

we obtain the isotropic multiscale wavelet entropy.

The spatial heterogeneity of the wavelet entropy can be achieved by decomposing the domain of 2D process into several boxes and calculating the entropy in each box.

In the analysis of discrete datasets, the above equations can be implemented by discretizing the parameters a and **b**. For example, [60] choose $a = 2^{m/4}$, $m = 0, 1, 2, \ldots$ and $\theta = k \cdot \Delta\theta$, where $\Delta\theta = \frac{\pi}{180}$.

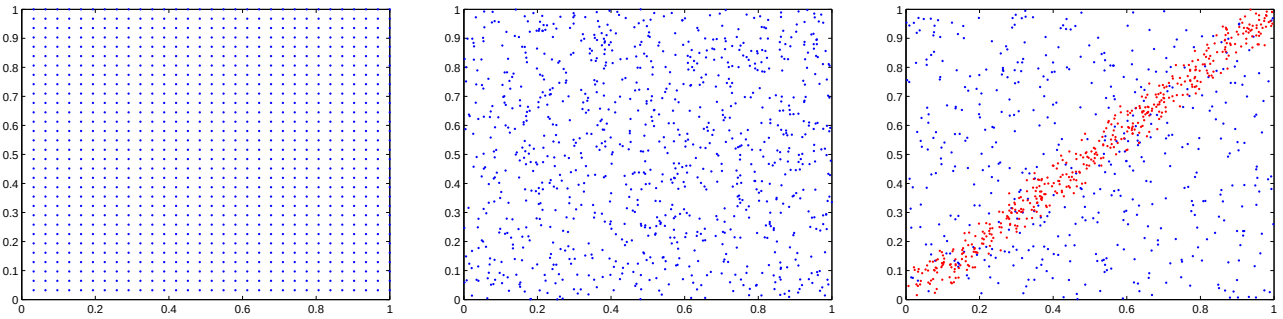
When the process is a spatial point pattern, the wavelet transform of Equation (12) is given by $W_f(a,\mathbf{b},\theta) = \int_{\mathbf{x}} \psi_{a,\mathbf{b}}(\mathbf{x},\theta) dN(\mathbf{x})$ and $dN(\mathbf{x}) = 1$ if a point falls in $d\mathbf{x}$ and zero otherwise (see [61,62]).

5. Simulation Study

For illustrative purposes, we considered three different scenarios of point patterns. In the first scenario, we considered a set of 961 spatial points distributed on a regular grid 31×31 (Figure 2a). In the second scenario, we simulated $N = 1000$ realizations from a uniform distribution (Figure 2b), and in the third scenario, we simulated $N_1 = 500$ points from a uniform distribution and $N_2 = 500$ points along a linear pattern with a slope of 45 degrees (Figure 2c). Further examples could be considered with a slope ranging between zero and 180 degrees as in [34].

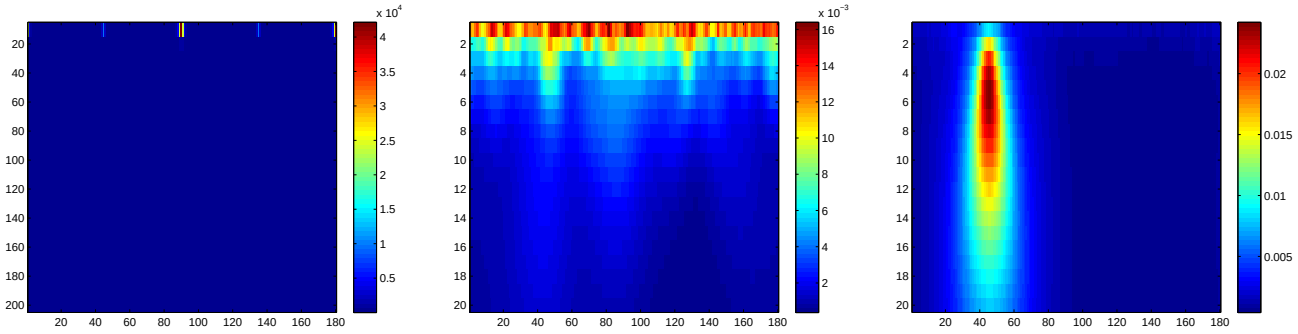
Figure 3 represents the directional multiscale wavelet entropy for the three datasets in Figure 2. In the first case, the entropy is approximatively equal to zero for each direction and for each scale. The higher values in the first level of resolution corresponding to the angles 0, 90 and 135 degrees are due to the border effects. In the second case (Figure 3b), the entropy seems higher in the first levels of resolution, but it is approximatively equal for each direction. In the third case, the entropy is higher in the area where the two random processes are superimposed, that is along the linear direction of 45 degrees.

If we evaluate the entropy for each scale of resolution (Figure 4) using Equation (18), we observe that the entropy for the regular grid of points shows a fast decrease in the first levels of resolutions, while in the case of a random pattern, the entropy measure has a slower decay. A notably different behavior is shown when the process is generated from two different random processes, as in Figure 2. In this case, the entropy reaches its maximum value at a certain level of resolution.



(a) (b) (c)

Figure 2. (**a**) Spatial points on a regular grid of size 31×31 ; (**b**) simulated spatial points ($n = 1000$) from a uniform distribution $U[0, 1]$; (**c**) simulated spatial points using a uniform distribution $U[0, 1]$ with $n = 500$ and an overlapped simulated linear pattern generated by a uniform distribution $U[0, 1]$ with $n = 500$ along the line with a slope of 45 degrees.



(a) (b) (c)

Figure 3. (**a**) Directional multiscale wavelet entropy for the: regular grid; (**b**) set of random points; (**c**) set of random points with a linear pattern.

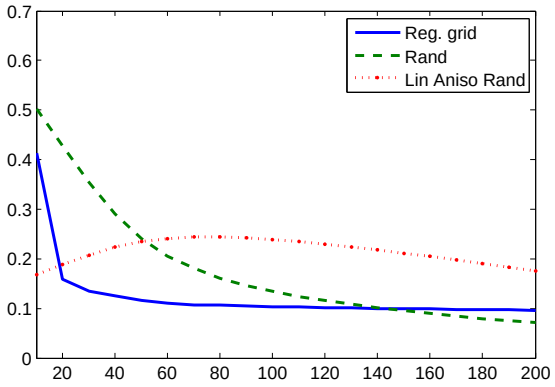
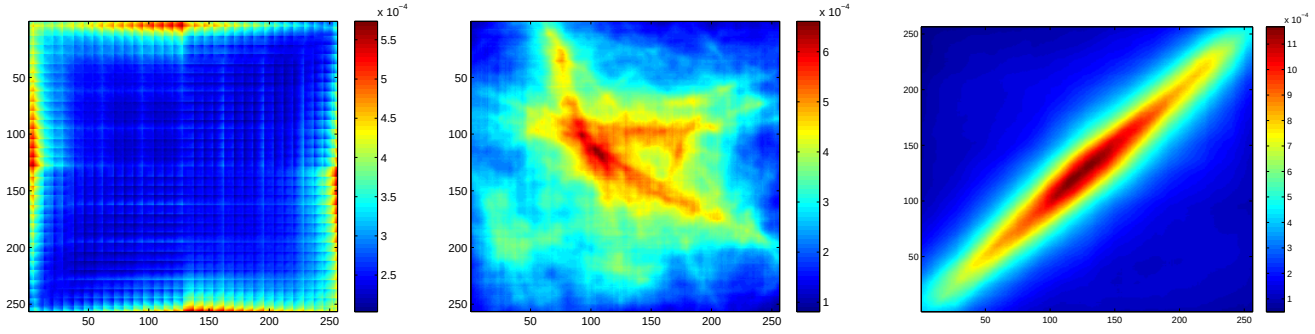


Figure 4. Multiscale wavelet entropy for the: regular grid (blue line); set of random points (green dashed line); set of random points with a linear pattern overlapped (red points).

Figure 5 represents the spatial entropy for all directions and all scales for the three scenarios: low entropy for a regular grid (Figure 5a), high entropy for the random process (Figure 5b) and directional entropy for the superpositions of both patterns (Figure 5c).



(a) (b) (c)

Figure 5. (**a**) Spatial wavelet entropy for the: regular grid; (**b**) set of random points; (**c**) set of random points with a linear pattern overlapped.

Table 1 shows the global measure of wavelet entropy for the three scenarios. As expected, the first process given by points on a regular grid has the lowest entropy, while the third process resulting from the overlapping of two uniform distributions shows the highest entropy.

Table 1. Global entropy for the three processes: **(a)** spatial points on a regular grid of size 31×31; **(b)** simulated spatial points ($n = 1000$) from a uniform distribution $U[0, 1]$; **(c)** simulated spatial points using a uniform distribution $U[0, 1]$ with $n = 500$ together with an overlapped linear pattern generated by a uniform distribution $U[0, 1]$ with $n = 500$ along the line $y = x$.

	Wavelet Entropy
Regular	2.47
Random	3.60
Random + linear	4.30

6. Application to Earthquake Data

We consider the Chilean earthquake catalog from January 2007 to August 2014, including 13,883 events with a magnitude larger than 3.0 and a depth less than 60 km. In order to show the spatial entropy, we focus on three main seismic areas represented in Figure 6. The first area (Figure 7a) is delimited by a rectangle with longitudes $(-72, -79)$ and latitudes $(-22.62, -18.62)$ and includes an earthquake event with magnitude 8.2 on 1 April 2014. The second zone (Figure 7b), including the town of Valparaiso, is delimited by coordinates with latitudes $(-33.79, -29.79)$ and longitudes $(-72, -69)$. Finally, the third area (Figure 7c), delimited by coordinates $(-74.5, -71.5)$ and $(-38.13, -34.13)$, includes a big earthquake (with magnitude 8.8) on 27 January 2010 in the town of Concepción. Although all areas extend along the Nazca plate, there are many other smaller faults along which earthquakes tend to cluster. The directional wavelet entropy can identify linear spatial directions with different entropy.

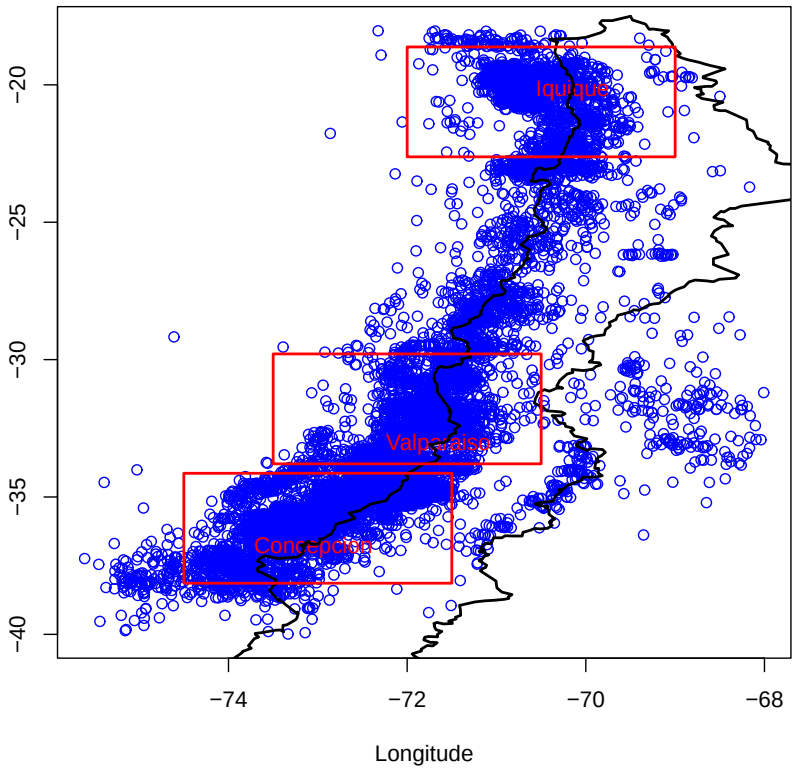
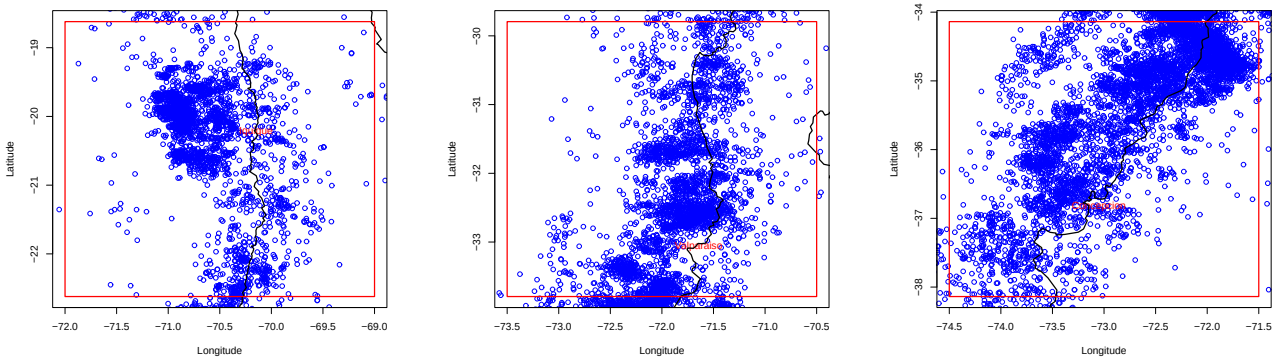


Figure 6. Earthquake catalog of Chile form January 2007 to August 2014 with a depth less than 60 km and a minimum magnitude of 3.0; red rectangles indicate three important seismic areas.

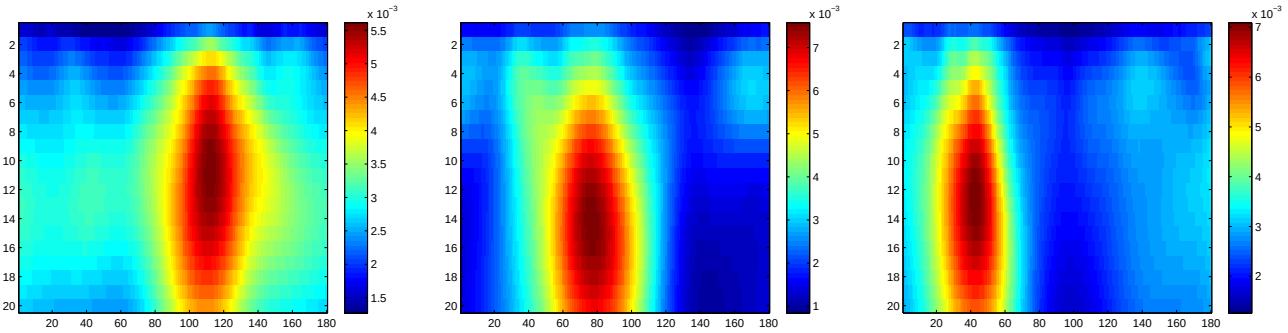


(a) (b) (c)

Figure 7. Earthquakes epicenters form January 2007 to August 2014 with a depth less than 60 km and a minimum magnitude of 3.0 for three areas of Chile: (**a**) the first area located in the North of Chile, including the town of Iquique, is delimited by latitudes $(-22.62, -18.62)$ and longitudes $(-72, -69)$; (**b**) the second area including the town of Valparaiso, is delimited by latitudes $(-33.79, -29.79)$ and longitudes $(-72, -69)$; (**c**) the third area including the town of Concepción, is delimited by latitudes $(-38.14, -34.14)$ and longitudes $(-74.5, -71.5)$.

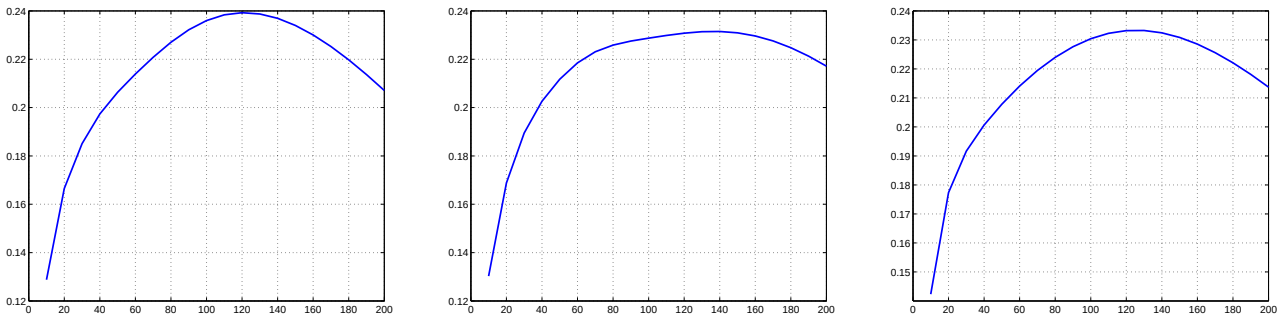
Figure 8 shows the wavelet entropy for each scale a and direction θ assessed on each dataset represented in Figure 7a–c, respectively. All plots show the maximum entropy along different dominant

linear directions: 110 degrees for the first plot, 80 degrees for the second plot and 45 degrees for the third one. The entropy analysis through scales (Figures 9 and 10) shows the presence of different degrees of complexity at different scales. While Figures 9a and 9c indicate the maximum entropy in correspondence to a certain scale, in Figure 9b, the maximum entropy extends for a range of scales. The different behavior of Figure 9b is probably due to the source of its seismic events: unlike the events of Figures 7a and 7c, which are clustered along the main fault that caused their big earthquakes, Figure 7b shows different anisotropic clusters, probably due to the ruptures of independent faults.



(a) (b) (c)

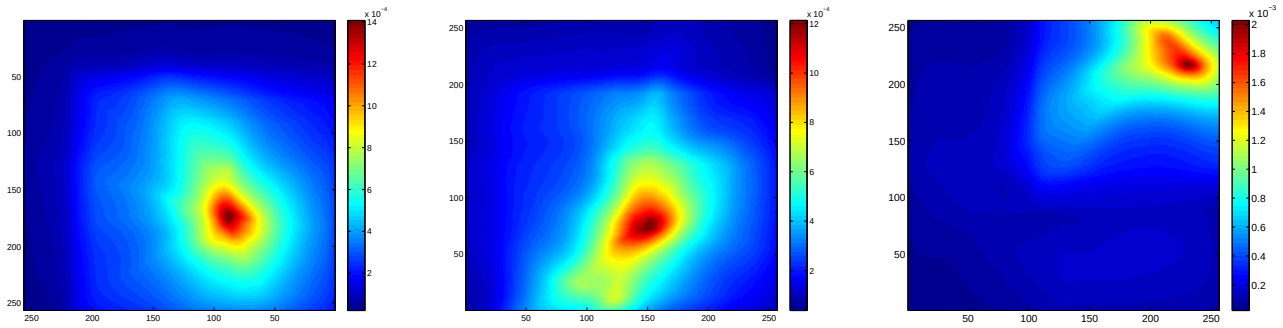
Figure 8. Directional multiscale wavelet entropy for the epicenters for each area of Figure 7.



(a) (b) (c)

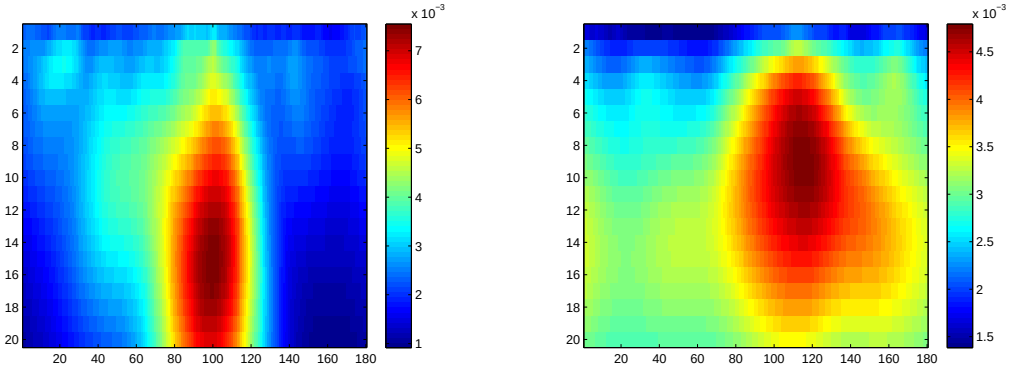
Figure 9. Wavelet entropy for each scale a and for all directions θ for three areas: (**a**) the area delimited by latitudes $(-22.62, -18.62)$ and longitudes $(-72, -69)$; (**b**) the area delimited by latitudes $(-33.79, -29.79)$ and longitudes $(-72, -69)$; (**c**) the area including the town of Concepción delimited by latitudes $(-38.14, -34.14)$ and longitudes $(-74.5, -71.5)$.

If we consider the global (isotropic) spatial entropy as described by Equation (11), it is not possible to distinguish the main directions where the entropy tends to be the same. However, we can observe that the higher values of entropy are not concentrated close to the epicenters of the big events, where most likely the predictability is larger. This is confirmed by the analysis in Figure 11, which shows the directional multiscale entropy before (a) and after (b) the Iquique earthquake with a magnitude 8.2 occurring in the North of Chile on 1 April 2014. Although the entropy seems to be concentrated along a particular direction before the big earthquake, the uncertainty seems higher than the period after the M8.2 earthquake. Similar results are shown in Table 2 representing the global entropy for each area.



(a) (b) (c)

Figure 10. Spatial wavelet entropy for three areas: (**a**) the area delimited by latitudes $(-22.62, -18.62)$ and longitudes $(-72, -69)$; (**b**) the area delimited by latitudes $(-33.79, -29.79)$ and longitudes $(-72, -69)$; (**c**) the area, including the town of Concepción, delimited by latitudes $(-38.14, -34.14)$ and longitudes $(-74.5, -71.5)$.



(a) (b)

Figure 11. Wavelet entropy for each scale a and for each angle θ for the earthquake events before (**a**) and after (**b**) the Iquique earthquake occurring on 1 April 2014.

Table 2. Global entropy for the earthquake catalog of Chile: "Window 1", "Window 2" and "Window 3" represent the wavelet entropy for the events in the rectangular windows represented in Figures 7a, 7b and 7c, respectively. The last two rows show the entropy values for the period before the Iquique earthquake and after the Iquique earthquake.

	Wavelet Entropy
Window 1	4.30
Window 2	4.55
Window 3	4.30
Window 1 (before)	4.31
Window 1 (after)	4.29

7. Conclusions

In this work, we have extended the concept of wavelet entropy to the two-dimensional case in order to detect the different complexity of a spatial point pattern. A new definition of wavelet entropy has been proposed for anisotropic processes using the directional Morlet wavelets. A simulation study has been used to show that the wavelet entropy is minimum when the process is a regular process, and it is maximum when the process is given by an overlapping of random point processes. Finally, an application to the earthquake catalog of Chile has been considered to detect the different spatial complexity. In particular, the results have shown that entropy is lower at spatial zones where a big earthquake happened than at areas characterized by different seismic activity, such as the area around the town of Valparaiso. Furthermore, the wavelet entropy is higher before a big earthquake than after. This means that aftershock events, which normally cluster along an active fault, have a lower degree of uncertainty than foreshock events, which are often characterized by a random distribution. The proposed methodology provides an important contribution to the study of earthquakes in Chile, given that it allows assessing the predictability of earthquakes in a given area.

Acknowledgments

The present work has been partially supported by Fondecyt Grant 1131147 from the Chilean government and by Grants MTM2013-43917-P and P1-1B2012-52.

Author Contributions

All authors have contributed to the study and preparation of the article. All authors have read and approved the final manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Clausius, R. On the Motive Power of Heat, and on the Laws which may be Deduced from it for the Theory of Heat. *Ann. Phys.* **1850**, *20*, 2010.
2. Boltzmann, L. Einige Allgemeine Satze Iiber Warmegleichgewicht unter Gas-molekulen. *Sitzungsber. Akad. Wiss. Wien.* **1871**, *63*, 679–711.
3. Shannon, C. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423.
4. Sengupta, B.; Stemmler, M.B.; Friston, K.J. Information and Efficiency in the Nervous System-A Synthesis. *PLoS Comput. Biol.* **2013**, *9*, doi:10.1371journal.pcbi.1003157.
5. Wong, A.K.C.; Sahoo, P.K. A Gray Level Threshold Selection Method Based on Maximum Entropy Principle. *IEEE Trans. Syst. Man. Cybern.* **1989**, *19*, 866–871.
6. Zhang, Y.; Wu, L. Optimal Multi-Level Thresholding Based on Maximum Tsallis Entropy via an Artificial Bee Colony Approach. *Entropy* **2011**, *13*, 841–859.

7. Oladipo, E.O. Spectral Analysis of Climatological Time Series: On the Performance of Periodogram, Non-integer and Maximum Entropy Methods. *Theor. Appl. Climatol.* **1988**, *39*, 40–53.
8. Koutsoyiannis, D. Uncertainty, Entropy, Scaling and Hydrological Stochastics, 2, Time Dependence of Hydrological Processes and time Scaling. *Hydrol. Sci. J.* **2005**, *50*, 405–426.
9. Li, S.; Zhou, Q.; Wu, S.; Dai, E. Measurement of Climate Complexity Using Sample Entropy. *Int. J. Climatol.* **2006**, *26*, 2131–2139.
10. AghaKouchak, A. Entropy-Copula in Hydrology and Climatology. *J. Hydrometeorol.* **2014**, *15*, 2176–2189.
11. Nicholson, T.; Sambridge, M.; Gudmundsson, O. On Entropy and Clustering in Earthquake Hypocentre Distributions. *Geophys. J. Int.* **2000**, *142*, 37–51.
12. Telesca, L.; Lapenna, V.; Lovallo, M. Information Entropy Analysis of Seismicity of Umbria-Marche Region (Central Italy). *Nat. Hazards Earth Syst. Sci.* **2004**, *4*, 691–695.
13. Telesca, L.; Lovallo, M.; Mohamed, A.E.E.A.; ElGabry, M.; El-hady, S.; Abou Elenean, K.M.; ElBary, R.E.F. Informational Analysis of Seismic Sequences by Applying the Fisher Information Measure and the Shannon Entropy: An Application to the 2004–2010 Seismicity of Aswan Area (Egypt). *Physica A* **2012**, *391*, 2889–2897.
14. Telesca, L.; Lovallo, M.; Babayev, G.; Kadirov, F. Spectral and Informational Analysis of Seismicity: An Application to the 1996–2012 Seismicity of the Northern Caucasus-Azerbaijan Part of the Greater Caucasus-Kopet Dag Region. *Physica A* **2013**, *392*, 6064–6078.
15. Telesca, L.; Lovallo, M.; Chamoli, A.; Dimri, V.P.; Srivastava, K. Fisher-Shannon Analysis of Seismograms of Tsunamigenic and Non-tsunamigenic Earthquakes. *Physica A* **2013**, *392*, 3424–3429.
16. Telesca, L.; Lovallo, M.; Martí Molist, J.; Lopez Moreno, C.; Abella Mendelez, R. Using the Fisher-Shannon Method to Characterize Continuous Seismic Signal during Volcanic Eruptions: Application to 2011–2012 El Hierro (Canary Islands) Eruption. *Terra Nova* **2014**, *26*, 425–429.
17. Rényi, A. On Measures of Entropy and Information. In Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability, Berkeley, CA, USA, 20 June–30 July 1960; Jerzy Neyman, Ed.; Berkeley: New York, NY, USA, 1961; pp. 547–561.
18. Papoulis, A. *Probability, Random Variables and Stochastic Processes*; McGraw-Hill: New York, NY, USA, 1984.
19. Telesca, L.; Cuomo, V.; Lapenna, V.; Macchiato, M. Statistical Analysis of Fractal Properties of Point Processes Modelling Seismic Sequences. *Phys. Earth Planet. Int.* **2001**, *12*, 65–83.
20. Labat D. Recent Advances in Wavelet Analyses: Part 1. A review of Concepts. *J. Hydrol.* **2005**, *314*, 275–288.
21. Rosso, O.; Blanco, S.; Yordanowa, J.; Kolev, V.; Figliola, A.; Schürmann, M.; Başar, E. Wavelet Entropy: A New Tool for Analysis of Short Duration Brain Electrical Signals. *J. Neurosci. Methods* **2001**, *105*, 65–75
22. Sello, S. Wavelet Entropy and the Multi-peaked Structure of Solar Cycle Maximum. *New Astron.* **2003**, *8*, 105–117.

23. Perez, D.G.; Zunino, L.; Garavaglia, M.; Rosso, O.A. Wavelet Entropy and Fractional Brownian Motion Time Series. *Physica A* **2006**, *365*, 282–288.
24. Zunino, L.; Perez, D.G.; Garavaglia, M.; Rosso, O.A. Wavelet Entropy of Stochastic Processes. *Physica A* **2007**, *379*, 503–512.
25. Lyubushin, A. Prognostic Properties of Low-Frequency Seismic Noise. *Nat. Sci.* **2012**, *4*, 659–666.
26. Lyubushin, A. How Soon would the next Mega-Earthquake Occur in Japan? *Nat. Sci.* **2013**, *5*, 1–7.
27. Donoho, D.L. Nonlinear Wavelet Methods for Recovery of Signals, Densities, and Spectra from Indirect and Noisy Data. In *Proceedings of Symposia in Applied Mathematics*; American Mathematical Society: Washington, WA, USA, 1993; pp.173–205.
28. Gao, W.; Li, B.L. Wavelet Analysis of Coherent Structures at the Atmosphere-Forest Interface. *J. Appl. Meteorol.* **1993**, *32*, 1717–1725.
29. Grenfell, N.; Bjørnstad, B.T.O.; Kappey, J. Travelling Waves and Spatial Hierarchies in Measles Epidemics. *Nature* **2001**, *414*, 716–723.
30. Saunders, S.C.; Chen, J.; Crow, T.R.; Brofoske, K.D. Hierarchical Relationships between Landscape Structure and Temperature in a Managed Forest Landscape. *Landsc. Ecol.* **1998**, *13*, 381–395.
31. Brosofske, K.D.; Chen, J.; Crow, T.R.; Saunders, S.C. Vegetation Responses to Landscape Structure at Multiple Scales across a Northern Wisconsin, USA, Pine Barrens Landscape. *Plant Ecol.* **1999**, *143*, 203–318.
32. Harper, K.A.; Macdonald, S.E. Structure and Composition of Riparian Boreal Forest: New Methods for Analyzing Edge Influence. *Ecology* **2001**, *82*, 649–659.
33. Perry, J.N.; Liebhold, A.M.; Rosenberg, M.S.; Dungan, J.; Miriti, M.; Jakomulska, A.; Citron-Pousty, S. Illustrations and Guidelines for Selecting Statistical Methods for Quantifying Spatial Pattern in Ecological Data. *Ecography* **2002**, *25*, 578–600.
34. Mateu, J.; Nicolis, O. Multiresolution Analysis of Linearly-Oriented Spatial Point Patterns. *J. Stat. Comput. Simul.* **2015**, *8*, 621–637.
35. Goodall, D.W. Some Considerations in the Use of Point Quadrats for the Analysis of Vegetation. *Aust. J. Sci. Res. B* **1952**, *5*, 1–41.
36. Matérn, B. Spatial Variation: Stochastic Models and their Application to some Problems in Forest Surveys and other Sampling Investigations. *Esselte* **1960**, *49*, 5.
37. Matérn, B. *Spatial Variation*; 2nd ed.; Springer: Berlin, Germany, 1986.
38. Ripley, B.D. *Spatial Statistics*; Wiley: New York, NY, USA, 1981.
39. Neyman, J.; Scott, E.L. Statisical Approach to Problems of Cosmology. *J. R. Stat. Soc. B* **1958**, *20*, 1–43.
40. Cliff, A.D.; Ord, J.K. *Spatial Processes: Models and Applications*; Pion: London, UK, 1981.
41. Ogata, Y. Space-Time Point Process Models for Earthquake Occurrences. *Ann. Inst. Stat. Math.* **1998**, *50*, 379–402.
42. Diggle, P.J.; Richardson, S. Epidemiological Studies of Industrial Pollutants: An Introduction. *Int. Stat. Rev.* **1993**, *61*, 203–206.

43. Stoyan, D.; Kendall, W.; Mecke, J. *Stochastic Geometry and its Applications*; 2nd ed.; Wiley: Chichester, UK, 1995.
44. Diggle, P.J. *Statistical Analysis of Spatial Point Patterns*; Hodder Education: London, UK, 2003.
45. Baddeley, A.; Gregori, P.; Mateu, J.; Stoica, R.; Stoyan, D., Eds. *Case Studies in Spatial Point Process Modeling (Lecture Notes in Statistics)*; Springer: Berlin, Germany, 2006.
46. Illian, J.; Penttinen, A.; Stoyan, H.; Stoyan, D. *Statistical Analysis and Modeling of Spatial Point Patterns*; Wiley: Chichester, UK, 2008.
47. Gelfand, A.E.; Diggle, P.J.; Fuentes, M.; Guttorp, P., Eds; *Handbook of Spatial Statistics*. CRC Press: Boca Raton, FL, USA, 2010.
48. Ohser, J.; Stoyan, D. On the Second-Order and Orientation Analysis of Planar Stationary Point Processes. *Biometr. J.* **1981**, *23*, 523–533.
49. Stoyan, D.; Benes, V. Anisotropy Analysis for Particle Systems. *J. Microsc.* **1991**, *164*, 159–168.
50. Mateu, J. Second-Order Characteristics of Spatial Marked Processes with Applications. *Nonlinear Anal. R. World Appl.* **2000**, *1*, 145–162.
51. Redenbach, C.; Särkkä, A.; Freitag, J.; Schladitz, K. Anisotropy Analysis of Pressed Point Processes. *Adv. Stat. Anal.* **2009**, *93*, 237–261.
52. Møller, J.; Rasmussen, J.G. A Sequential Point Process Model and Bayesian Inference for Spatial Point Patterns with Linear Structures. *Scand. J. Stat.* **2012**, *39*, 618–634.
53. Telesca, L.; Lovallo, M.; Golay, J.; Kanevski, M. Comparing Seismicity Declustering Techniques by Means of the Joint Use of Allan Factor and Morisita Index. *Stoch. Environ. Res. Risk Assess.* **2015**, doi:10.1007/s00477-015-1030-8.
54. Schenk, H.J.; Mahall, B.E. Positive and Negative Interactions Contribute to a North-South-Patterned Association between two Desert Shrub Species. *Oecologia* **2002**, *132*, 402–410.
55. Rosenberg, M.S. Wavelet Analysis for Detecting Anisotropy in Point Patterns. *J. Veg. Sci.* **2004**, *15*, 277–284.
56. Blanco, S.; Figliola, A.; Quiroga, R.Q.; Rosso, O.A.; Serrano, E. Time-Frequency Analysis of Electroencephalogram Series. III. Wavelet Packets and Information Cost Function. *Phys. Rev. E* **1998**, *57*, 932–940.
57. Vidakovic, B. *Statistical Modeling by Wavelets*; Wiley: New York, NY, USA, 1999.
58. Mallat, S. *A Wavelet Tour of Signal Processing*; Academic Press: Waltham, MA, USA, 1999.
59. Neupauer, R.M.; Powell, K.L. A Fully-Anisotropic Morlet Wavelet to Identify Dominant Orientations in a Porous Medium. *Comput. Geosci.* **2005**, *31*, 465–471.
60. Kumar, P. A Wavelet Based Methodology for Scale-Space Anisotropic Analysis. *Geophys. Res. Lett.* **1995**, *22*, 2777–2780.
61. Brillinger, D. Some Wavelet Analyses of Point Process Data. In *Proceedings of the Thirty-First Asilomar Conference on Signals, Systems and Computers*; IEEE: Universidade de Sao Paulo, Brazil, 1–4 November 1998; pp. 1087–1091.

62. Nicolis, O. Spatio-Temporal Analysis of Earthquake Occurrences Using a Multiresolution Approach. In *Mathematics of Planet Earth Lecture Notes in Earth System Sciences*; Springer: Berlin, Heidelberg, Germany, 2014; pp. 179–183.