Entropy **2013**, *15*, 4844-4888; doi:10.3390/e15114844

OPEN ACCESS

entropy

ISSN 1099-4300

www.mdpi.com/journal/entropy

Review

Statistical Mechanics and Information-Theoretic Perspectives on Complexity in the Earth System

Georgios Balasis [1,*], Reik V. Donner [2,3], Stelios M. Potirakis [4], Jakob Runge [2,5], Constantinos Papadimitriou [1,6], Ioannis A. Daglis [1,6,†], Konstantinos Eftaxias [7] and Jürgen Kurths [2,5,8]

[1] Institute for Astronomy, Astrophysics, Space Applications and Remote Sensing, National Observatory of Athens, I. Metaxa & Vas. Pavlou St., GR-15236 Penteli, Greece; E-Mails: constantinos@noa.gr (C.P.); daglis@noa.gr (I.A.D.)

[2] Research Domain IV—Transdisciplinary Concepts & Methods, Potsdam Institute for Climate Impact Research, 14473 Potsdam, Germany; E-Mails: redonner@pik-potsdam.de (R.V.D.); jakob.runge@pik-potsdam.de (J.R.); juergen.kurths@pik-potsdam.de (J.K.)

[3] Department of Biogeochemical Integration, Max Planck Institute for Biogeochemistry, 07745 Jena, Germany

[4] Department of Electronics Engineering, Technological Education Institute (TEI) of Piraeus, 250 Thivon and P. Ralli, GR-12244 Aigaleo, Athens, Greece; E-Mail: spoti@teipir.gr

[5] Department of Physics, Humboldt University, 12489 Berlin, Germany

[6] Section of Astrophysics, Astronomy and Mechanics, Department of Physics, University of Athens, Panepistimioupoli Zografou, 15784 Athens, Greece

[7] Section of Solid State Physics, Department of Physics, University of Athens, Panepistimioupoli Zografou, 15784 Athens, Greece; E-Mail: ceftax@phys.uoa.gr

[8] Institute of Complex Systems and Mathematical Biology, University of Aberdeen, Aberdeen AB24 3UE, UK

† The author's primary affiliation is 6.

* Author to whom correspondence should be addressed; E-Mail: gbalasis@noa.gr; Tel.: +30-2108109114; Fax: +30-2106138343.

Received: 16 September 2013; in revised form: 22 October 2013 / Accepted: 22 October 2013 / Published: 7 November 2013

Abstract: This review provides a summary of methods originated in (non-equilibrium) statistical mechanics and information theory, which have recently found successful applications to quantitatively studying complexity in various components of the complex

system Earth. Specifically, we discuss two classes of methods: (i) entropies of different kinds (e.g., on the one hand classical Shannon and Rényi entropies, as well as non-extensive Tsallis entropy based on symbolic dynamics techniques and, on the other hand, approximate entropy, sample entropy and fuzzy entropy); and (ii) measures of statistical interdependence and causality (e.g., mutual information and generalizations thereof, transfer entropy, momentary information transfer). We review a number of applications and case studies utilizing the above-mentioned methodological approaches for studying contemporary problems in some exemplary fields of the Earth sciences, highlighting the potentials of different techniques.

Keywords: entropy measures; symbolic dynamics; non-extensive statistical mechanics; causality; Earth sciences

1. Introduction

Complexity is, nowadays, a frequently used, yet (at least quantitatively speaking) poorly defined, concept [1]. It tries to embrace a great variety of scientific and technological approaches of all types of natural, artificial and social systems [1]. The concept of complexity in the Earth system pertains to processes that are not deterministic, but can be modeled with a statistical element, including fractal phenomena and deterministic chaos [2].

As in other areas of science, the last few decades of geoscientific research have been shaped by a rapid increase in the amount of existing data on relevant phenomena and processes obtained from extensive measurement campaigns (including both ground-based observations and remote sensing applications), as well as detailed simulation models making use of the meanwhile available powerful computation infrastructures. The resulting big datasets call for novel techniques, not just for data acquisition and storage, but, most prominently, also for their appropriate analysis, which is a necessary prerequisite for sophisticated process-based interpretations [3].

There is an ongoing transition from classical, yet powerful, statistical methods to such characterization of the subtle aspects of the complex dynamics displayed by the recorded spatio-temporal fluctuations of relevant variables in different components of the Earth system, such as atmosphere, biosphere, cryosphere, lithosphere, oceans, the near-Earth electromagnetic environment, *etc.* In fact, many novel concepts originated in dynamical systems or information theory have been developed, partly motivated by specific research questions from the geosciences, and have found a variety of different applications [4]. This continuously extending toolbox of nonlinear time series analysis highlights the paramount importance of dynamical complexity as a key to understanding the behavior of the complex system Earth and its components.

It is important to note that there is no way to find an "optimum" organization or complexity measure [5]. Moreover, a combination of complexity measures, which refer to different aspects of a system, such as structural *versus* dynamical properties, is the most promising way to deal with the complexity of a system. For instance, several well-known techniques have been used to extract

signatures of incipient changes in the system dynamics (*i.e.*, precursors of dynamical transitions) hidden in time series representing natural signals of the Earth system. Notable examples include applications to time series data originated in the near-Earth electromagnetic environment (magnetospheric variations associated with magnetic storms), lithosphere (preseismic electromagnetic emissions associated with earthquakes), atmosphere (meteorological or hydrological signals), as well as other areas.

As a consequence of the described scientific "trends", there is a large variety of methodological frameworks allowing us to study, characterize and, thus, understand complexity based on observational or model data. Among the main groups of conceptual approaches, methods originating in statistical mechanics and information theory belong to the most prominent and fruitful tools of modern complex systems-based nonlinear time series analysis [6,7]. Although statistical mechanics and information theory are based on somewhat different foundations, many concepts used in both areas are conceptually closely related. Most prominently, the notion of entropy naturally arises in both disciplines and provides a versatile basis for characterizing various complementary aspects associated with structural and dynamic complexity.

In this review, we discuss some recent developments, making use of statistical mechanics and information-theoretic notions of entropy for studying complexity in the Earth system. Specifically, we do not restrict ourselves to measuring the complexity of individual variables, but the wide-spread problem of characterizing dynamical interrelationships among a set of variables. Besides reviewing the algorithmic foundations of different approaches and highlighting their conceptual similarities and differences in Section 2, we provide an overview on recent successful applications in some exemplary fields of geoscience in Section 3. A discussion and summary of the potentials of statistical mechanics and information-theoretic approaches in the Earth sciences concludes this contribution.

2. Methods

2.1. Entropy Measures from Symbolic Sequences

2.1.1. Symbolic Dynamics

Symbolic time series analysis is a useful tool for modeling and characterization of nonlinear dynamical systems. It provides a rigorous way of looking at "real" dynamics with finite precision [8,9]. Briefly, it is a way of coarse-graining or simplifying the description.

The basic idea is quite simple. One divides the phase space into a finite number of classes and labels each class with a symbol (e.g., a letter from some alphabet). Instead of representing the trajectories by sequences of (typically real-valued) numbers (iterations of a discrete map or sampled points along the trajectories of a continuous flow), one watches the alteration of symbols. Of course, in doing so, one loses a considerable amount of detailed information, but some of the invariant, robust properties of the dynamics are typically kept, such as periodicity, symmetry or the chaotic nature of an orbit [8].

In the framework of symbolic dynamics, time series are transformed into a series of symbols by using an appropriate partition, which results in relatively few symbols. After symbolization, the next step is the construction of "symbol sequences" ("words" in the language of symbolic dynamics) from the symbol series by collecting groups of symbols together in temporal order.

To be more precise, the simplest possible coarse-graining of a time series is given by choosing a single threshold (usually the mean value or median of the data considered [10]), and assigning the symbols "1" and "0" to the signal, depending on whether it is above or below the threshold (binary partition) (The generalization to obtaining a classification using more than one threshold is trivial. Moreover, in addition to this kind of "static" encoding, there are other possibilities for obtaining a "dynamic" symbolization of observed time series data, such as using order patterns (*cf.* Section 2.1.4) or even mixed strategies [11]). Thus, we generate a symbolic time series from a two-letter ($\lambda = 2$) alphabet (0,1), e.g., $0110100110010110\ldots$. We can read this symbolic sequence in terms of distinct consecutive blocks of a length of n. For example, using $n = 2$, one obtains $01/10/10/01/10/01/01/10/\ldots$. This reading procedure is called "lumping".

The number of all possible kinds of words is $\lambda^n = 2^2 = 4$, namely 00, 01, 10, 11. The required probabilities for the estimation of an entropy, p_{00}, p_{01}, p_{10}, p_{11}, are the fractions of the different blocks (words), 00, 01, 10, 11, in the symbolic time series, namely, 0, 4/8, 4/8 and 0, correspondingly for the specific example. Based on these probabilities, we can estimate, for example, the probabilistic entropy measure, H_S, introduced by Shannon [12]:

$$H_S = -\sum_{j=1}^{\lambda^n} p_j \log_2 p_j \tag{1}$$

where p_j are the probabilities associated with the microscopic configurations. The binary logarithm is taken here for convenience in the case of a binary symbolization (giving rise to an interpretation of the Shannon entropy as the negative mean amount of information, measured in bits), but may be replaced by another basis (e.g., using $\log_\lambda$) without loss of generality.

In the case of a real-world time series, "experimental" conditions are often rather different (e.g., in terms of the system's energy content), which can lead to ambiguities when quantitatively comparing entropy estimates, e.g., obtained from different persons in clinical studies. In order to correct for this problem, the *renormalized entropy* [5,13] has been introduced as the difference between the Shannon entropies obtained for the time series under study and some reference data.

A generalization of the Shannon entropy is provided by the spectrum of Rényi entropies [14]:

$$H_q = \frac{1}{1-q} \log_2 \sum_{j=1}^{\lambda^n} p_j^q \tag{2}$$

with $q \in \mathbb{R}$, which give different weights to rare and frequent symbols. In this spirit, Rényi entropies allow for the studying of scaling properties associated with a non-uniform probability distribution of symbols or words as arising, e.g., in the case of multi-fractal processes. Notably, the Shannon entropy, H_S, is a special case of H_q for $q \to 1$. In a similar way, for $q \to 0$, H_q is the Hartley entropy (or Hartley information), one of the most classical information-theoretic measures of uncertainty [15]. Note that for the case of a uniform probability distribution of n-blocks, all Rényi entropies coincide with each other and, hence, the Shannon and Hartley entropies.

2.1.2. Block Entropies

In general, one can statistically analyze a symbolic time series of N symbols, $\{A_i\}, i = 1, 2, ..., N$, in terms of the frequency distribution of blocks of a length of n $(n < N)$. For an alphabet consisting

of λ letters, considering a word of a length of n, there are λ^n possible combinations of the symbols that may constitute a word, here, $\lambda^n = 2^n$. The probability of the occurrence, $p_j^{(n)}$, of the j-th combination of symbols ($j = 1, 2, ..., 2^n$) (*i.e.*, the j-th n-word) can be approximated by the number of occurrences of this n-word in the considered symbolic sequence divided by the total number of n-words.

Based on these probabilities, one can estimate the probabilistic entropy or other related characteristics. Various tools of information theory and entropy concepts can be used to identify statistical patterns in the symbolic sequences, onto which the dynamics of the original system under analysis has been projected. For detection of an anomaly, it suffices that a detectable change in the pattern represents a deviation of the system from nominal behavior [16]. Recently published work has reported novel methods for detection of anomalies in complex dynamical systems, which rely on symbolic time series analysis, e.g., [17,18]. Entropies depending on the word-frequency distribution in symbolic sequences are of special interest, extending Shannon's classical definition of the entropy and providing a link between dynamical systems and information theory. These entropies take a large/small value if there are many/few kinds of patterns, *i.e.*, they decrease while the organization of patterns is increasing. In this way, these entropies can measure the complexity of a signal.

Extending Shannon's classical definition of the entropy of a single state [12] to the entropy of a succession of states [19,20], one reaches the definition of the *n-block entropy*, $H(n)$, which, for a two-letter ($\lambda = 2$) alphabet and word length n, is given by:

$$H(n) = -\sum_{j=1}^{2^n} p_j^{(n)} \log_2 p_j^{(n)} \tag{3}$$

$H(n)$ is a measure of uncertainty and gives the average amount of information necessary to predict a subsequence of length n. Consequently, $H(n)/n$ can be interpreted as the mean uncertainty per symbol and should converge for $n \to \infty$ to some stationary value if the observed dynamics is deterministic. Moreover, from a practical perspective, one is often interested in quantifying the mean information gain when increasing the word length, measured by the *conditional (or differential) block entropies*:

$$h(n) = H(n+1) - H(n) \quad \text{for} \quad n \geq 1; \quad h(0) = H(1) \tag{4}$$

For stationary and ergodic processes, the limit of $h(n)$ for $n \to \infty$ provides an estimator of the Kolmogorov-Sinai entropy or source entropy h of the dynamical system under study [21,22] (In a similar spirit, $H(n)/n$ provides another estimate of the source entropy for $n \to \infty$.). For a broad class of systems, $h(n)$ converges exponentially fast, with the characteristic scaling parameter being related to the third-order Rényi entropy, H_3. In general, exploiting the corresponding convergence properties gives rise to a widely applicable probabilistic complexity measure, the *effective measure complexity* [23]:

$$EMC = \sum_{n=0}^{\infty} [h(n) - h] \tag{5}$$

A more extensive discussion and classification of complexity measures based on symbolic dynamics concepts has been given elsewhere [24].

2.1.3. Non-Extensive Tsallis Entropy

It has been established that physical systems characterized by long-range interactions or long-term memory or being of a multi-fractal nature are best described by a generalized statistical-mechanical formalism proposed by Tsallis [1,25,26]. More precisely, inspired by multi-fractal concepts, Tsallis introduced an entropic expression characterized by an index, q, which leads to non-extensive statistics [25,26]:

$$S_q = k\frac{1}{q-1}\left(1-\sum_{i=1}^{W} p_i^q\right) \tag{6}$$

where p_i are probabilities associated with the microscopic configurations (*i.e.*, the empirical frequency distribution of symbols), W is their total number, q is a real number and k is a constant, *i.e.*, the Boltzmann's constant from statistical thermodynamics. Notably, there is a striking conceptual similarity between Tsallis' entropy definition and the notion of Rényi entropies.

The entropic index, q, describes the deviation of Tsallis entropy from the standard Boltzmann-Gibbs entropy. Indeed, using $p_i^{(q-1)} = e^{(q-1)\ln(p_i)} \sim 1+(q-1)\ln(p_i)$ in the limit $q \to 1$, we recover the usual Boltzmann-Gibbs entropy:

$$S_1 = -k\sum_{i=1}^{W} p_i \ln(p_i) \tag{7}$$

as the thermodynamic analog of the information-theoretic Shannon entropy.

For $q \neq 1$, the entropic index, q, characterizes the degree of non-extensivity reflected in the following pseudo-additivity rule:

$$\frac{S_q(A+B)}{k} = \frac{S_q(A)}{k} + \frac{S_q(B)}{k} + (q-1)\frac{S_q(A)}{k}\frac{S_q(B)}{k} \tag{8}$$

where A and B are two subsystems. If these subsystems have special probability correlations, extensivity does not hold for $q = 1$ ($S_1 \neq S_1(A) + S_1(B)$), but may occur for S_q, with a particular value of the index, $q \neq 1$. Such systems are called *non-extensive* [1,25]. The cases $q > 1$ and $q < 1$ correspond to sub-additivity or super-additivity, respectively. As in the case of Rényi entropies, we may think of q as a bias parameter: $q < 1$ privileges rare events, while $q > 1$ highlights prominent events [27].

We note that the parameter, q, itself is not a measure of the complexity of the system, but measures the degree of the non-extensivity of the system. In turn, the time variations of the Tsallis entropy, S_q, for some q quantify the dynamic changes of the complexity of the system. Specifically, lower S_q values characterize the portions of the signal with lower complexity.

In terms of symbolic dynamics, the Tsallis entropy for a two-letter ($\lambda = 2$) alphabet and word length n reads [18,28,29]:

$$S_q(n) = k\frac{1}{q-1}\left(1-\sum_{j=1}^{2^n}\left[p_j^{(n)}\right]^q\right) \tag{9}$$

Broad symbol-sequence frequency distributions produce high entropy values, indicating a low degree of organization. Conversely, when certain sequences exhibit high frequencies, low values of $S_q(n)$ are produced, indicating a high degree of organization.

2.1.4. Order-Pattern Based Approaches

In the symbolic dynamics framework introduced above, the transformation from a continuous to a discrete-valued time series (necessary for estimating the Shannon entropy of a discrete distribution or related characteristics) has been realized by coarse-graining the range of the respective observable. This procedure has the advantage of being algorithmically simple. However, unless block lengths $n \geq 2$ are used, the resulting entropies follow directly from the specific discretization and, thus, exclusively reflect statistical, but not dynamical, characteristics. In turn, dynamics comes into play when considering blocks of symbols.

An alternative way of introducing a symbolization in an explicitly dynamic way is making use of order patterns and related concepts from ordinal time series analysis [30]. In the simplest case (*i.e.*, a pattern of order two), one considers a binary discretization of the underlying time series of a continuous random variable according to the sign of the difference between two subsequent values, e.g., using the symbols "0" if $s_{i+1} - s_i < 0$ and "1" if $s_{i+1} - s_i > 0$ (disregarding the "unlikely" case $s_{i+1} = s_i$). Notably, this kind of symbolization corresponds to the classical symbolic dynamics approach applied to the time series of increments (*i.e.*, a first-order difference filter applied to the original data) with the increment value of zero discriminating the two classes of the resulting binary encoding. Based on this discretization, one can again define all entropies described above.

Besides using binary encodings based on order patterns of order two, the above framework can be extended to patterns of a higher order. In this case, sequences of observations are transformed into ordinal sequences (e.g., the case $s_i < s_{i+1} < s_{i+2}$ would correspond to the sequence $(1, 2, 3)$). Obviously, for a pattern of order q, there are $q!$ possible permutations of $(1, \ldots, q)$ and, hence, at most $q!$ different patterns (Depending on the type of observed dynamics, there might be forbidden patterns. We do not further discuss the implications of this fact here.). Using the frequency distribution of the possible order patterns for calculating a Shannon-type entropy leads to the notion of *permutation entropy* [31–33].

Besides its direct application as a proxy for dynamical disorder based on time series data, Rosso *et al.* [34,35] suggested utilizing permutation entropy in combination with some associated measure of complexity for defining the so-called *complexity-entropy causality plane*. Here, the entropic quantity is the (Shannon-type) permutation entropy normalized by its maximum value, $\log_2 q!$:

$$H_S^*[P_\pi] = -\frac{1}{\log_2 q!} \sum_{i=1}^{q!} p(\pi_i) \log_2 p(\pi_i) \tag{10}$$

where π_i denote the different possible permutations of $(1, \ldots, q)$. As one possible complexity measure associated with the permutation entropy, the *Jensen-Shannon complexity*, $C_{JS}[P_\pi]$, is defined as [36]:

$$C_{JS}[P_\pi] = Q_J[P_\pi, P_e] H_S^*[P_\pi] \tag{11}$$

with the Jensen-Shannon divergence:

$$Q_J[P_\pi, P_e] = Q_0 \log_2 q! \left(H_S^*[(P_\pi + P_e)/2] - \frac{H_S^*[P_\pi]}{2} - \frac{1}{2} \right) \tag{12}$$

between the observed distribution, P_π, and the uniform distribution $P_e = (1/q!, \ldots, 1/q!)$, where Q_0 is a normalization constant, such that $0 \leq Q_J \leq 1$.

2.2. Entropy Measures from Continuous Data

2.2.1. Approximate Entropy

Approximate entropy ($ApEn$) has been introduced by Pincus [37] as a measure for characterizing the regularity in relatively short and potentially noisy data. Specifically, $ApEn$ quantifies the degree of irregularity or randomness within a time series (of length N) and has already been widely applied to (non-stationary) biological systems for dynamically monitoring the system's "health" status (see [38,39] and references therein). The conceptual idea is rooted in the work of Grassberger and Procaccia [40] and makes use of distances between sequences of successive observations. More specifically, $ApEn$ examines time series for detecting the presence of similar epochs; more similar and more frequent epochs lead to lower values of $ApEn$.

From a qualitative point of view, given N points, $ApEn$ is approximately equal to the negative logarithm of the conditional probability that two sequences that are similar for m points remain similar, that is, within a tolerance, r, at the next point. Smaller $ApEn$ values indicate a greater chance that a set of data will be followed by similar data (regularity). Thus, smaller values indicate greater regularity. Conversely, a larger value of $ApEn$ points to a lower chance of similar data being repeated (irregularity). Hence, larger values convey more disorder, randomness and system complexity. Consequently, a low/high value of $ApEn$ reflects a high/low degree of regularity. Notably, $ApEn$ detects changes in underlying episodic behavior not reflected in peak occurrences or amplitudes [41].

In the following, we give a brief description of the calculation of $ApEn$. A more comprehensive discussion can be found in [37–39].

For a time series, $\{s_k\}_k$ with $s_k = s(t_k)$, $t_k = kT$, $k = 1, 2, ..., N$, and T being the sampling period, we can define $N - m + 1$ vectors, each one consisting of m consecutive samples of this time series as:

$$\mathbf{X}_i^m = \{s_i, s_{i+1}, ..., s_{i+m-1}\}, i = 1, 2, ..., N - m + 1 \tag{13}$$

The main idea is to consider a window of length m running through the time series and forming the corresponding vectors, $\mathbf{X}_i^m$. The similarity between the formed vectors is used as a measure of the degree of organization of the time series. A quantitative measure of this similarity, $C_i^m(r)$, is given by the average number of vectors, $\mathbf{X}_j^m$, within a distance, r, from $\mathbf{X}_i^m$. Here, $\mathbf{X}_j^m$ is considered to be within a distance r from $\mathbf{X}_i^m$ if $d_{ij}^m \leq r$, where d_{ij}^m is the maximum absolute difference of the corresponding scalar components of $\mathbf{X}_i^m$ and $\mathbf{X}_j^m$ (*i.e.*, the two vectors have a distance smaller than r according to their supremum norm). By calculating $C_i^m(r)$ for each $i \leq N - m + 1$ and then taking the mean value of the corresponding natural logarithms:

$$\phi^m(r) = (N - m + 1)^{-1} \sum_{i=1}^{N-m+1} \ln C_i^m(r) \tag{14}$$

the $ApEn$ is defined as:

$$ApEn(m, r) = \lim_{N \to \infty} [\phi^m(r) - \phi^{m+1}(r)] \tag{15}$$

which, for a finite time series, can be estimated by the statistic:

$$ApEn(m, r, N) = \phi^m(r) - \phi^{m+1}(r) \tag{16}$$

By tuning r, one can reach a reasonable degree of similarity for "most" vectors, $\mathbf{X}_i^m$, and, hence, a reliable statistics, even for relatively short time series.

In summary, the presence of repetitive patterns of fluctuation in a time series renders it more predictable than a time series in which such patterns are absent. A time series containing many repetitive patterns has a relatively small $ApEn$; a less predictable (*i.e.*, more complex) process has a higher $ApEn$.

2.2.2. Sample Entropy

Sample entropy ($SampEn$) was proposed by Richman and Moorman [42] as an alternative that would provide an improvement of the intrinsic bias of $ApEn$ [43]. Specifically, when calculating the similarity within a time series using the vectors defined by Equation (13), $ApEn$ does not exclude self-matches, since the employed measure of similarity, $C_i^m(r)$, which is proportional to the number of vectors within a distance, r, from $\mathbf{X}_i^m$, does not exclude the $\mathbf{X}_i^m$ itself from this count. $SampEn$ is considered to be an evolution of $ApEn$ that displays relative consistency and less dependence on data length (see, e.g., [44] and the references therein).

First of all, we consider only the first $N-m$ vectors of a length of m of Equation (13), ensuring that, for $1 \leq i \leq N-m$, $\mathbf{X}_i^m$ and $\mathbf{X}_i^{m+1}$ are defined. Then, we define $B_i^m(r)$ as $(N-m-1)^{-1}$ times the number of vectors, $\mathbf{X}_j^m$, within a distance, r, from $\mathbf{X}_i^m$, where $j = 1, 2, ..., N-m$, but also $j \neq i$, excluding in this way the self-matches, and set:

$$B^m(r) = (N-m)^{-1} \sum_{i=1}^{N-m} B_i^m(r) \tag{17}$$

Correspondingly, we define $A_i^m(r)$ as $(N-m-1)^{-1}$ times the number of vectors, $\mathbf{X}_j^{m+1}$, within a distance, r, from $\mathbf{X}_i^{m+1}$, where $j = 1, 2, ..., N-m$ with $j \neq i$, and:

$$A^m(r) = (N-m)^{-1} \sum_{i=1}^{N-m} A_i^m(r) \tag{18}$$

The $SampEn(m, r)$ is then defined as:

$$SampEn(m, r) = \lim_{N \to \infty} \{-\ln[A^m(r)/B^m(r)]\} \tag{19}$$

which, for finite time series, can be estimated by the statistic:

$$SampEn(m, r, N) = -\ln\left[\frac{A^m(r)}{B^m(r)}\right] \tag{20}$$

2.2.3. Fuzzy Entropy

Fuzzy entropy ($FuzzyEn$) [43,44], like its ancestors, $ApEn$ and $SampleEn$ [44], is a "regularity statistic" that quantifies the (un)predictability of fluctuations in a time series. For the calculation of $FuzzyEn$, the similarity between vectors is defined based on fuzzy membership functions and the vectors' shapes. The gradual and continuous boundaries of the fuzzy membership functions lead to a series of advantages, like the continuity, as well as the validity of $FuzzyEn$ at small values, higher accuracy, stronger relative consistency and even less dependence on the data length. $FuzzyEn$ can

be considered as an upgraded alternative of $SampEn$ (and $ApEn$) for the evaluation of complexity, especially for short time series contaminated by noise.

Like $SampEn$, $FuzzyEn$ excludes self-matches. However, it follows a slightly different definition for the employed first $N-m$ vectors of a length of m, by removing a baseline, $\bar{s}_i$:

$$\bar{s}_i = m^{-1} \sum_{j=0}^{m-1} s_{i+j} \tag{21}$$

i.e., for the $FuzzyEn$ calculations, we use the first $N-m$ of the vectors:

$$\mathbf{X}_i^m = \{s_i, s_{i+1}, ..., s_{i+m-1}\} - \bar{s}_i, \;\; i = 1, 2, ..., N-m+1 \tag{22}$$

Then, the similarity degree, D_{ij}^m, between each pair of vectors, $\mathbf{X}_j^m$ and $\mathbf{X}_i^m$, being within a distance, r, from each other, is defined by a fuzzy membership function:

$$D_{ij}^m = \mu\left(d_{ij}^m, r\right) \tag{23}$$

where d_{ij}^m is, as in the case of $ApEn$ and $SampEn$, the supremum norm difference between $\mathbf{X}_i^m$ and $\mathbf{X}_j^m$. For each vector, $\mathbf{X}_i^m$, we calculate the average similarity degrees with respect to all other vectors, $\mathbf{X}_j^m$, $j = 1, 2, ..., N-m+1,$ and $j \neq i$ (*i.e.*, excluding itself):

$$\phi_i^m(r) = (N-m-1)^{-1} \sum_{i=1, j\neq i}^{N-m} D_{ij}^m \tag{24}$$

Then, we evaluate:

$$\varphi^m(r) = (N-m)^{-1} \sum_{i=1}^{N-m} \phi_i^m(r) \tag{25}$$

and:

$$\varphi^{m+1}(r) = (N-m)^{-1} \sum_{i=1}^{N-m} \phi_i^{m+1}(r) \tag{26}$$

The $FuzzyEn(m, r)$ is then defined as:

$$FuzzyEn(m, r) = \lim_{N\to\infty} \left[\ln \varphi^m(r) - \ln \varphi^{m+1}(r)\right] \tag{27}$$

which, for finite time series, can be estimated by the statistic:

$$FuzzyEn(m, r, N) = \ln \varphi^m(r) - \ln \varphi^{m+1}(r) \tag{28}$$

We note that by making use of vectors of different lengths, $ApEn$, $SampEn$ and $FuzzyEn$ are based on a similar rationale as the conditional block entropies (*i.e.*, changes in the information content as the sequence length increases).

2.3. Measures of Statistical Interdependence and Causality

The above-mentioned concepts make use of the general ideas of information-theoretic or thermodynamic entropy to gain information on the statistical complexity of a single time series. A natural extension of this idea is utilizing comparable approaches for studying complex signatures of interactions between two time series, or even among $K > 2$ records. In the following, we briefly review the corresponding basic concepts as used in the geoscientific context, so far.

2.3.1. Mutual Information

While Shannon entropy is a measure of the uncertainty about outcomes of one process, *mutual information* (MI) is a measure of the reduction thereof if another process is known (There are several other, conceptually-related symmetric measures for the strength of interdependencies between two variables, e.g., the maximal information coefficient [45]. For brevity, we do not further discuss these alternative measures here.). The Shannon-type MI can be expressed as:

$$I(X;Y) = H_S(Y) - H_S(Y|X) = H_S(X) - H_S(X|Y) \tag{29}$$

$$= H_S(X) + H_S(Y) - H_S(X,Y) \tag{30}$$

i.e., as the difference between the uncertainty in Y and the remaining uncertainty if X is already known (and *vice versa*). MI is symmetric in its arguments, non-negative and zero, if and only if X and Y are independent. The lagged cross-MI for two time series is given by:

$$I_{XY}(\tau) \equiv I(X_{t-\tau}; Y_t) \tag{31}$$

For $\tau > 0$, one measures the information in the past of X that is contained in Y, and *vice versa* for $\tau < 0$. In analogy, the auto-MI is defined as $I(Y_{t-\tau}; Y_t)$ for $\tau > 0$ (for $\tau = 0$, this gives the classical Shannon entropy, $H_S(Y)$).

Just like the entropy of a time series can be estimated based on symbols or patterns, MI can be estimated by simply plugging in the corresponding marginal and joint entropies into Equation (30). In contrast to pattern-based estimators, for binning estimators (*i.e.*, estimators using frequency distributions based on a bivariate symbolization with $\lambda \gg 2$ symbols for each time series, but typically considering individual symbols, *i.e.*, $n = 1$), an adaptive strategy is preferable in which the bins are chosen, such that the marginal distributions are uniform [46–48]. For continuous time series, a useful class of estimators is based on nearest-neighbor statistics (see Section 2.4.1.). For example, making use of the entropy estimators developed in [49], Kraskov *et al.* introduced a nearest-neighbor estimator of MI [47].

In the limiting case of a bivariate Gaussian distribution, the MI is simply given by:

$$I_{\text{Gauss}}(X;Y) = -\tfrac{1}{2} \ln \left(1 - \rho(X;Y)^2\right) \tag{32}$$

where $\rho(X;Y)$ is the linear Pearson correlation coefficient.

There are several straightforward generalizations of the mutual information concept. On the one hand, one can characterize nonlinear bivariate interrelationships between two variables by a spectrum

of generalized mutual information functions obtained by replacing the Shannon entropy, H_S, by Rényi entropies, H_q, of arbitrary order q [50] (Note that this replacement does not conserve the non-negativity property of the classical MI function [50].). On the other hand, the idea behind mutual information can be directly extended to sets of $K > 2$ variables, $X_1, \ldots, X_K$, leading to the so-called *(generalized) redundancies* [51–53]:

$$I_q(X_1; \ldots; X_K) = \sum_{k=1}^{K} H_q(X_k) - H_q(X_1, \ldots, X_K) \tag{33}$$

2.3.2. Conditional Mutual Information and Transfer Entropy

An important generalization of MI is the *conditional mutual information* (CMI):

$$I(X;Y|Z) = H_S(Y|Z) - H_S(Y|X,Z) = H_S(X|Z) - H_S(X|Y,Z) \tag{34}$$

$$= H_S(X,Z) + H_S(Y,Z) - H_S(Z) - H_S(X,Y,Z) \tag{35}$$

which measures the mutual information between X and Y that is *not* contained in a third variable, Z, which can also be multivariate. CMI shares the properties of MI and is zero, if and only if X and Y are mutually independent *conditionally on* Z. Moreover, it is possible to use any of the existing estimators of entropic characteristics (including binning, order patterns or k-nearest neighbors) for defining CMI and generalized versions thereof. Like in Equation (32), the CMI for multivariate Gaussian processes can be expressed in terms of the *partial correlation*, where the correlation, $\rho(X;Y)$, is replaced by $\rho(X;Y|Z)$.

Of particular interest is the use of information-theoretic measures to quantify causal relations between variables based on time series data. A first step is to assess the directionality of information transfer between two time series. While lagged MI can be used to quantify whether information in Y has already been present in the past of X, this information could have existed in the common past of both processes and, therefore, is not a sign of a transfer of unique information in X to Y. To overcome this limitation, *transfer entropy* (TE) [54] measures the information in the past of X that is shared in the present of Y and *not* already contained in the past of Y:

$$I^{\mathrm{TE}}_{X\to Y} \equiv I(X_t^-; Y_t \,|\, Y_t^-) \tag{36}$$

with the upper script "$^-$" denoting the (infinite) past vector, e.g., $X_t^- = (X_{t-1}, X_{t-2}, \ldots)$. It should be noted that the notion of causality employed here is in the predictive sense of *Granger causality*, which was introduced for linear processes in [55] and later generalized to nonlinear frameworks [56,57]. TE and Granger causality can actually be shown to be equivalent for Gaussian variables [58].

2.3.3. Graphical Models

The definition of TE leads to the problem that infinite-dimensional densities have to be estimated, which is commonly called the "curse of dimensionality". In the usual naive estimation of TE, the infinite vectors are simply truncated at some $\tau_{\max}$, which has to be chosen at least as large as the supposed maximal coupling delay between X and Y. This can lead to very large estimation dimensions, since the delay is not known *a priori*. As shown in [59,60], a high estimation dimension severely affects the

reliability of the inference of directionality. In [59], the problem of high dimensionality is overcome by utilizing the concept of *graphical models*.

While the TE discussed before is a bivariate measure of directionality, a more comprehensive causality assessment requires the inclusion of multiple variables to be able to exclude common drivers and indirect causalities. In the *graphical model* approach [61–63], the conditional independence properties of a multivariate process are visualized in a graph; in our case, a time series graph. As depicted in Figure 1a, each node in that graph represents a subprocess of a multivariate discrete-time process, $\mathbf{X}$, at a certain time, t. Nodes $X_{t-\tau}$ and Y_t are connected by a directed link "$X_{t-\tau} \to Y_t$", if and only if $\tau > 0$ and:

$$I(X_{t-\tau}; Y_t | \mathbf{X}_t^- \setminus \{X_{t-\tau}\}) > 0 \tag{37}$$

i.e., if they are not independent conditionally on the past of the whole process, which implies a lag-specific causality with respect to $\mathbf{X}$. If $Y \neq X$, we say that the link "$X_{t-\tau} \to Y_t$" represents a *coupling at lag* τ, while for $Y = X$, it represents an *auto-dependency at lag* τ. Nodes X_t and Y_t are connected by an undirected contemporaneous link "$X_t - Y_t$" (visualized by a line, *cf.* in Figure 1a) [63], if and only if:

$$I(X_t; Y_t \,|\, \mathbf{X}_{t+1}^- \backslash \{X_t, Y_t\}) > 0 \tag{38}$$

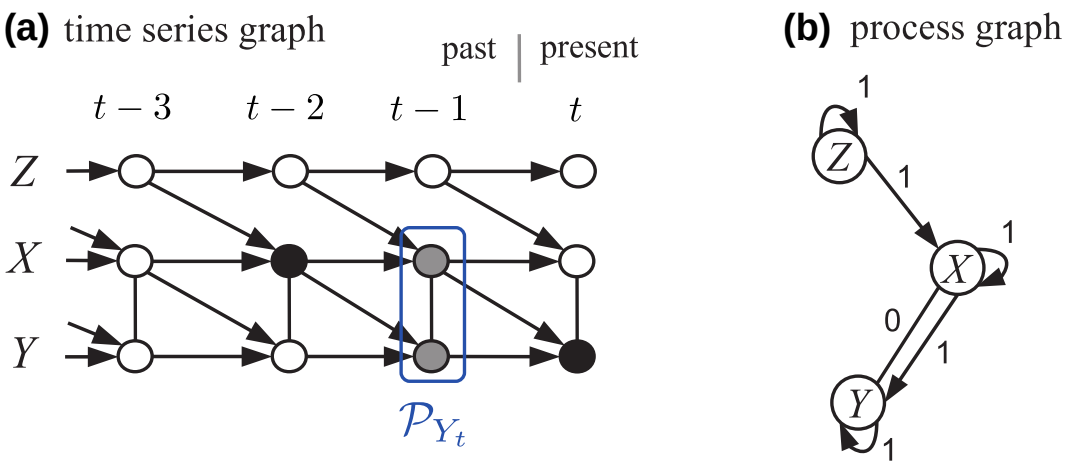
where also the contemporaneous present $\mathbf{X}_t \backslash \{X_t, Y_t\}$ is included in the condition. Note that stationarity implies that "$X_{t-\tau} \to Y_t$" whenever "$X_{t'-\tau} \to Y_{t'}$" for any t'. In the resulting graph, the parents and neighbors of a node Y_t are defined as:

$$\mathcal{P}_{Y_t} \equiv \{Z_{t-\tau} : \; Z \in \mathbf{X}, \; \tau > 0, \; Z_{t-\tau} \to Y_t\} \tag{39}$$

$$\mathcal{N}_{Y_t} \equiv \{X_t : X \in \mathbf{X}, X_t - Y_t\} \tag{40}$$

In [59], an algorithm for the estimation of time series graphs by iteratively inferring the parents is introduced. The supplementary material of [59] also provides a suitable shuffle test and a numerical study on the detection and false positive rates of the algorithm.

Figure 1. (Color online) Causal interactions in a multivariate process $\mathbf{X}_t = (X, Y, Z)_t$. (**a**) The time series graph (see definition in text). The set of *parents*, $\mathcal{P}_{Y_t}$ (blue box, gray nodes), separates Y_t from the past of the whole process, $\mathbf{X}_t^- \backslash \mathcal{P}_{Y_t}$, which is used in the algorithm to estimate the graph. (**b**) The process graph, which aggregates the information in the time series graph for better visualization (edge labels denote the associated lags).



Time series graphs encode the existence of a lag-specific causality, but they are not meant to assess the causal strength in a meaningful way. This topic is discussed in [64], where the concept of *momentary information transfer* (MIT) [65] is utilized for attributing a well-interpretable coupling strength solely to the inferred links of the time series graph. MIT between X at some lagged time, $t-\tau$, in the past and Y at time t is the CMI that measures the part of the source entropy of Y that is shared with the source entropy of X:

$$I^{\mathrm{MIT}}_{X\to Y}(\tau) \equiv I(X_{t-\tau}; Y_t | \mathcal{P}_{Y_t} \backslash \{X_{t-\tau}\}, \mathcal{P}_{X_{t-\tau}}) \tag{41}$$

The attribute *momentary* is used, because MIT measures the information of the "moment" $t-\tau$ in X that is transferred to Y_t. Similar to the definition of contemporaneous links in Equation (38), we can also define a contemporaneous MIT:

$$I^{\mathrm{MIT}}_{X-Y} \equiv I(X_t; Y_t | \mathcal{P}_{Y_t}, \mathcal{P}_{X_t}, \mathcal{N}_{X_t} \backslash \{Y_t\}, \mathcal{N}_{Y_t} \backslash \{X_t\}) \tag{42}$$

MIT is well interpretable, because it excludes past information, not only from Y, but also from X, *i.e.*, information transmitted by serial correlations. Often X and Y are entangled by multiple links, and only the exclusion of both pasts truly yields a coupling measure that depends solely on the mutual coupling strength. This is demonstrated analytically and numerically in [64].

2.4. Linkages with Phase Space Methods

2.4.1. Reinterpreting Entropies As Phase Space Characteristics

In their most common formulation, many of the aforementioned characteristics are explicitly based on a discretization of the system under study (either via coarse-graining and subsequent symbolization or using order patterns or related approaches), giving rise to an information-theoretic interpretation of entropies and related complexity measures. Consequently, these measures are often considered as providing one major field of nonlinear time series analysis methods, which is opposed to another class of well-developed theoretical frameworks making use of the notion of the phase space of a dynamical system. In the latter case, one commonly takes sets of variables (or time-lagged replications of one variable) into account, which are assumed to span a metric space in which the system's behavior can be fully described by means of distances between sets of "state vectors" sampled from the observed trajectory.

Reconsidering the entropic measures described above in more detail, one recognizes that there are also multiple quantifiers that make explicit use of distances (*i.e.*, approximate, sample and fuzzy entropy), which implies that the latter methods can also be considered as phase space-based approaches if one identifies subsequent observations as individual "coordinates" of some abstract space (*i.e.*, a "naive" embedding space of variable dimension). In fact, as already stated above, the mentioned class of entropy characteristics is based on ideas formulated by Grassberger and Procaccia about 30 years ago, which are among the pioneering works utilizing the concept of phase space for nonlinear time series analysis.

In a similar spirit, state-of-the-art estimators of mutual information and causality concepts based on this framework make use of the phase space concept by replacing the traditional histogram-based symbolizations or order patterns by nearest-neighbor relationships between sampled state vectors.

Specifically, k-nearest neighbor estimators of entropies [49], MI [47] and CMI [66] have been found practically useful and have desirable numerical properties not shared by other estimators. Here, we just briefly illustrate the basic idea of k-nearest neighbor estimates for CMI.

As the first step, the number of neighbors, k, to be considered around each sampled state vector in the joint space of (X, Y, Z) is chosen as a free parameter. For each sample corresponding to observations at time t, the supremum norm distance, ε_t, to the k-th nearest neighbor defines a cube of a length of $2\varepsilon_t$ around the corresponding state vector. Next, the numbers of points in the respective subspaces with a distance less than ϵ_t are counted, yielding $k_{z,t}$, $k_{xz,t}$ and $k_{yz,t}$. Then, the CMI is estimated as:

$$\widehat{I}(X;Y|Z) = \psi(k) + \frac{1}{T}\sum_{t=1}^{T}\left[\psi(k_{z,t}) - \psi(k_{xz,t}) - \psi(k_{yz,t})\right] \tag{43}$$

where ψ is the Digamma function. Smaller k result in smaller cubes, and since (as assumed in the estimator's derivation) the density is approximately constant in these cubes, the estimator has a low bias. Conversely, for large k, the bias is stronger, but the variance gets smaller. Note, however, that for independent processes, the bias is approximately zero, *i.e.*, $\widehat{I}(X;Y|Z) \approx 0$; large k are therefore better suited for independence tests, as needed in the algorithm to infer the time series graph.

2.4.2. Other Entropy Measures from Phase Space Methods

Besides the above-mentioned examples, entropic quantities frequently occur in various phase space-based methods of nonlinear time series analysis. One prominent example is the Grassberger-Procaccia algorithm for estimating the correlation dimension of chaotic attractors [67,68]. As in the calculation of $ApEn$, $SampEn$ and $FuzzyEn$, let us consider the m-dimensional delay vectors, $\mathbf{X}_i^m$, where the embedding dimension, m, and the mutual temporal offset between the components of $\mathbf{X}_i^m$ have been optimized, so as to yield coordinates, which are as independent as possible and represent the "true" dimensionality of the recorded system appropriately. Computing the pairwise distance of the resulting delay vectors, we may distinguish between mutually "close" and "distant" vectors by comparing the distance with some predefined global threshold value, r, which is summarized in the *recurrence matrix* [69]:

$$\mathbf{R}_{ij}(r) = \Theta(r - \|\mathbf{X}_i^m - \mathbf{X}_j^m\|) \tag{44}$$

where $\Theta(\cdot)$ is the Heaviside "function". The fraction of mutually close pairs of vectors:

$$C^m(r) = \frac{2}{N'(N'-1)}\sum_{i<j}\mathbf{R}_{ij}(r) \tag{45}$$

(with $N' = N - m + 1$) is called the *correlation sum* or *recurrence rate* of the observed trajectory. For a dissipative chaotic system with attractor dimension $< m$, the latter obeys a characteristic scaling behavior as:

$$C^m(r) \sim r^{D_2}\exp(-m\,H_2\,\Delta t) \tag{46}$$

with Δt denoting the sampling interval of the time series data and D_2 the attractor's correlation dimension. By varying r and m, it is thus possible to estimate the second-order Rényi entropy, H_2, together with the correlation dimension from the correlation sum.

Alternatively, other properties of the recurrence matrix may be exploited for estimating different entropic quantities. Among others, the length distribution, $p(l)$, of "diagonal lines" formed by non-zero entries in $\mathbf{R}$ provides a versatile tool for estimating H_2 from time series data [70]. A similar approach can be used for estimating the generalized MI of second order. The Shannon entropy, $ENTR$, of the discrete distribution, $p(l)$, has been used as a heuristic complexity measure in the framework of the so-called recurrence quantification analysis [69]. However, it has been shown that this definition does not allow for properly distinguishing between different "degrees" of chaotic dynamics in the desired way. In turn, the Shannon entropy associated with the length distribution, $p(\bar{l})$, of diagonal lines in $\mathbf{R}$ formed exclusively by zero entries provides a measure of dynamical disorder that increases with the complexity of chaotic dynamics as measured by the system's largest Lyapunov exponent [71]. Even more, the latter quantity shows an excellent agreement with the classical Shannon entropy estimate based on symbolic dynamics, but does not share some of the disadvantages, due to the coarse-graining, which arise in the latter case.

It is worth noting that, recently, there have been the first successful attempts to transfer the idea of "recurrences" in phase space to symbolic data [11,72,73]. A special case is order pattern recurrence plots [74,75], which have been successfully applied in the analysis of neurophysiological time series.

2.4.3. Causality from Phase Space Methods

For the sake of completeness, we briefly mention that similar to the characterization of dynamical complexity with entropic characteristics, also the complementary bi- and multi-variate problem of detecting causal interrelationships can be complementarily tackled by information-theoretic and phase space methods. Distance-based characteristics allowing one to reliably detect directional couplings between different variables have been developed and applied by various authors, e.g., [76–82]. A specific class of approaches is based on the recurrence matrix, making use of conditional recurrence probabilities [83,84] or asymmetries in coupled network statistics [85]. Since these methods are not within the focus of this review, we refer to the aforementioned references for further details.

3. Applications

3.1. Space Weather and Magnetosphere

During the past two decades, it has become increasingly evident that the magnetosphere cannot be viewed as an autonomous system, but should rather be examined in terms of driven nonlinear dynamical systems [86]. As early as 1990, Tsurutani *et al.* [87] provided one of the first indications of a nonlinear process in the magnetosphere, claiming that there is a nonlinear response of the AE index to the southward component of the interplanetary magnetic field (IMF). At the same time, Baker *et al.* were able to provide a mechanical analogue to model the driving of the magnetosphere, concluding that using tools of chaos theory, we may gain considerable insight into the storm-time behavior of the magnetosphere [88]. The discussion on low-dimensional chaos in magnetospheric activity was continued in many important works [89–92]. On the other hand, Chang and co-workers [93–96] suggested that the magnetospheric dynamics is that of an infinite-dimensional nonlinear system near criticality, a hypothesis

that was later supported by statistical analyses of the AE index [97,98] and auroral observations obtained by the Ultraviolet Imager (UVI) on the Polar spacecraft [99,100].

In situ observations revealed that scale-invariance, turbulence, non-Gaussian fluctuations and intermittency are essential elements of plasma sheet dynamics [101–103], while both *in situ* and global observations pointed out the possible emergence of dynamical phase transitions and symmetry breaking phenomena in coincidence with the occurrence of local and global relaxation processes [104–107].

Due to the above, it becomes obvious that the classical magnetohydrodynamic (MHD) approach is not sufficient to describe the more dynamical aspects of many magnetospheric phenomena, since the emergence of long-range correlations, as well as the coupling between various scales demand the utilization of novel concepts that were developed with respect to the multi-scale dynamics of out-of-equilibrium systems [28,29,108–110].

In this spirit, multi-fractal models have been employed for the description of space weather [111]. In addition, the emergence of self-organized criticality (SOC) has been examined in both AE index and solar wind data [112,113], as a means of understanding the global energy storage and release in the coupled solar wind–magnetosphere system [114,115] and as a paradigm for the magnetotail dynamics [116]. Moreover, the probability distributions of physical variables in turbulent space plasmas are scale-dependent and perfectly well modeled by non-extensive one-parameter (Tsallis) distributions, where q measures the degree of correlations (long-range interactions) in the system [117].

Magnetic storms are the most prominent global phenomenon of geospace dynamics, interlinking the solar wind, magnetosphere, ionosphere, atmosphere and, occasionally, the Earth's surface [118–120]. Magnetic storms produce a number of distinct physical effects in the near-Earth space environment: acceleration of charged particles in space, intensification of electric currents in space and on the ground, impressive aurora displays and global magnetic disturbances on the Earth's surface [121]. The latter serve as the basis for storm monitoring via the hourly disturbance storm time (Dst) index, which is computed from an average over four mid-latitude magnetic observatories [122].

The Dst data used here for illustrative purposes (Figure 2) include two intense magnetic storms, which occurred on March 31, 2001, and November 6, 2001, with minimum Dst values of -387 nT (nanoTesla) and -292 nT, respectively, as well as a number of smaller events (e.g., May and August, 2001, with Dst <-100 nT in both cases).

All entropy calculations were performed on 256 sample long windows sliding along the time series with a step of one sample. For the case of symbolic Tsallis entropy, the value $q = 1.84$ [123] was found for the non-extensivity parameter, q. For the $FuzzyEn$ calculations, the exponential function $\mu\left(d_{ij}^m, r\right) = \exp\left(-\left(d_{ij}^m/r\right)^n\right)$ with $n = 2$ has been used as the fuzzy membership function, using $m = 2$ and $r = 0.65 \cdot STD$, where STD is the standard deviation of the analyzed time series fragment, allowing fragments with different amplitudes to be compared quantitatively. The same values for m and r were also used in the case of $ApEn$ and $SampEn$ calculations. The results of the corresponding analyses are shown in Figure 2.

Approximate entropy, sample entropy, Fuzzy entropy and non-extensive Tsallis entropy sensitively trace the complexity dissimilarity among different "physiological" (quiet times) and "pathological" (intense magnetic storms) states of the magnetosphere. They imply the emergence of two distinct patterns: (1) a pattern associated with the intense magnetic storms, which is characterized by a higher

degree of organization; and (2) a pattern associated with normal periods, which is characterized by a lower degree of organization. In general, all four entropy measures clearly distinguish between the different complexity regimes in the Dst time series (see the red part of the corresponding plots in Figure 2).

Figure 2. From top to bottom: Dst time series from January 1, 2001, to December 31, 2001, associated with two intense magnetic storms (marked in red) and the corresponding $ApEn$, $SampEn$, $FuzzyEn$ and non-extensive Tsallis entropy, S_q, for sliding windows in time. The common horizontal axis is the time (in days), denoting the relative time position from the beginning of the Dst data. The triangles denote five time intervals in which the first, third and fifth intervals correspond to higher entropies, whereas the second and fourth time windows exhibit lower entropies (*cf.* the parts of the entropy plots shown in red).

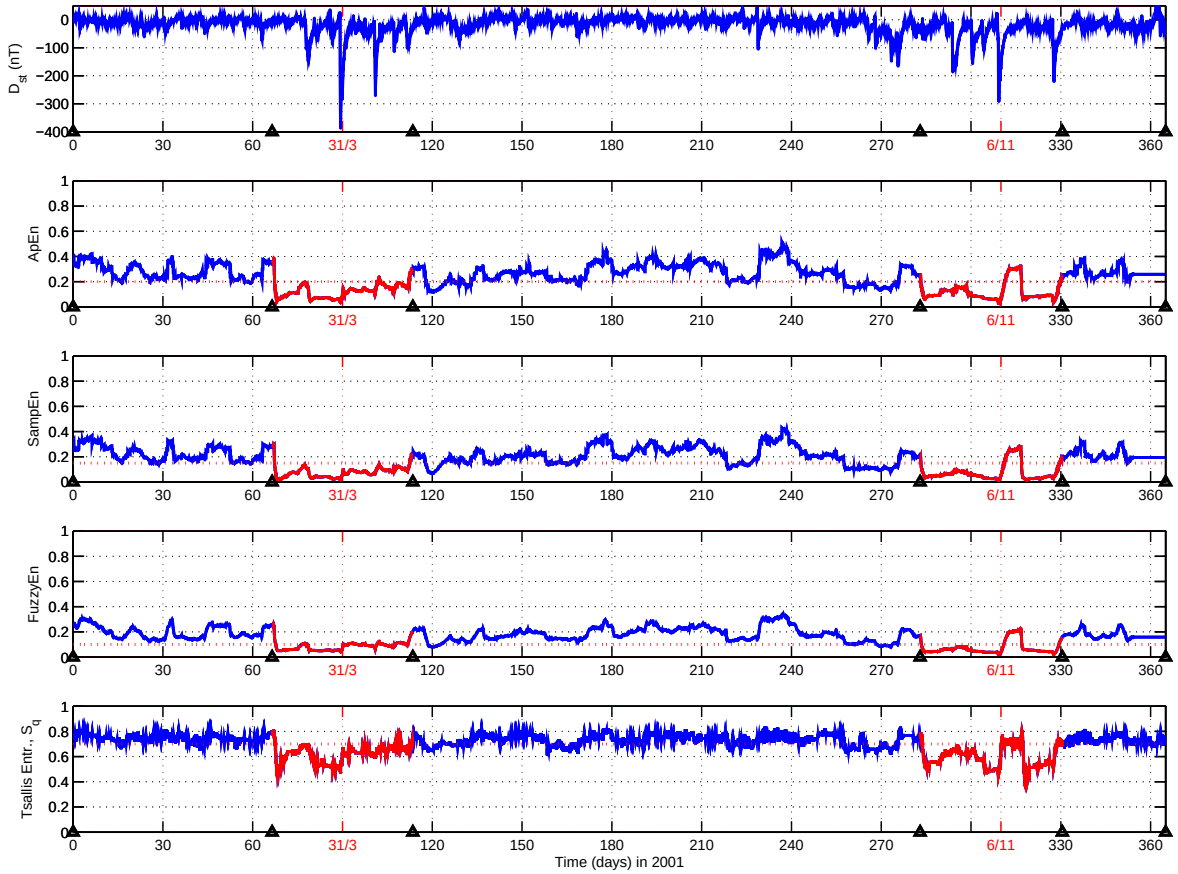


Figure 2 indicates that $ApEn$ and S_q yield superior results in comparison with the other entropy measures regarding the detection of dynamical complexity in the Earth's magnetosphere (*i.e.*, they offer a clearer picture of the transition). A possible explanation for this is that, on the one hand, S_q is an entropy obeying a non-extensive statistical theory, which is different from the classical Boltzmann-Gibbs statistical mechanics. Therefore, it is expected to better describe the dynamics of the magnetosphere, which is a non-equilibrium physical system with large variability. On the other hand, $ApEn$ is more stable when dealing with non-stationary signals from dynamical systems (such as the magnetospheric signal) than the other entropy measures presented here.

Figure 2 could also serve for placing boundary values or thresholds for $ApEn$ and Tsallis entropy in order to distinguish the different magnetospheric states. Along these lines, we suggest the limits of 0.2 and 0.7 for $ApEn$ and Tsallis entropy, respectively. Similarly, the values of 0.15 and 0.1 could play the role of indicative limits for the transition between the quiet-time and stormy magnetosphere in the cases of $SampEn$ and $FuzzyEn$, respectively.

3.2. Preseismic Electromagnetic Emissions

Under specific conditions, the Earth's heterogeneous crust experiences sudden releases of energy following large-scale mechanical failures. These phenomena of a potentially catastrophic nature are known as earthquakes (EQs). In a simplified view, a significant EQ is what happens when the two surfaces of a major fault slip, one over the other, under the stresses rooted in the motion of tectonic plates. Due to the catastrophic nature of some significant EQs and their direct impact on human life, there is a great interest in all related phenomena, especially those that are systematically observed prior to EQ occurrence and could therefore be regarded as potential EQ precursors.

Electromagnetic (EM) phenomena associated with EQs have been repeatedly recognized as soon as human history began to be recorded. The EQ light phenomenon has been reported all around the world already since ancient times [124], given that no specific instruments, but only eye-witnesses were necessary for the observations. However, since the existence of appropriate instrumentation for their measurement, a wide variety of EM phenomena have been observed [125]. For example, the existence of magnetic effects associated with earthquakes or volcanic activity was first suggested by Kalashnikov in 1954 [126].

Focused research in the field, which now is usually referred to as "Seismo-Electromagnetics", practically started in the 1980s, extending from direct current (DC) to very high frequency (VHF) bands, while it concerns the theoretical and experimental study of either EM signals supposedly emitted from the focal zone or anomalous transmission of EM waves over the epicentral region [124]. From that time onwards, a large number of telluric current anomalies (e.g., [127,128]), ultra-low frequency (ULF) anomalies (e.g., [129–133]), very low frequency (VLF) to VHF anomalies (e.g., [133–138] and references therein), as well as a variety of atmospheric-ionospheric anomalies attributed to the phenomenon known as lithosphere-atmosphere-ionosphere (LAI) coupling (e.g., [124,139–142] and the references therein) have been reported, analyzed and interpreted. More information, historical data and reviews can be found in [124,125,143–145] and their references.

The identification of an observed EM anomaly, such as an EQ-related one, is a challenging research topic. The time-lead of the EM anomaly relative to the EQ occurrence *per se* is not credible evidence. Due to the multi-facet nature of EQ preparation processes, a multidisciplinary analysis of the EM anomalies is needed before its classification as EQ-related for the identification of features that characterize the EQ preparation processes.

Since the early work of Mandelbrot [146], the self-affine nature of faulting and fracture is widely documented based on the analysis of data from both field observations and laboratory experiments; *cf.* [147] and the references therein. Moreover, several characteristics of seismicity and fracture (like criticality, complexity, frequency-magnitude laws, *etc.*) have been investigated both in the laboratory

and at the geophysical scale (e.g., [148–153] and the references therein). These findings led to the hypothesis that any EM emissions related to EQ preparation should present specific characteristics, like fractal scaling, criticality, phase transition characteristics, symmetry breaking, frequency-size laws, universal structural patterns of fracture and faulting, complexity characteristics, long-range correlations, memory, *etc.*, depending on the phase of the EQ preparation process to which they are related. A number of analyses in field-recorded preseismic EM signals have verified this hypothesis (e.g., [18,136–138,148,153–159] and references therein).

As previously stated, for systems exhibiting long-range correlations, memory or fractal properties, non-extensive Tsallis entropy is considered an appropriate mathematical tool [1,25,26,160]. A central property of the EQ preparation process is the occurrence of coherent large-scale collective behavior with a very rich structure, resulting from repeated nonlinear interactions among the constituents of the system [147,161]. Consequently, Tsallis entropy is also a physically meaningful tool for investigating the launch of a fracture-induced EM precursor.

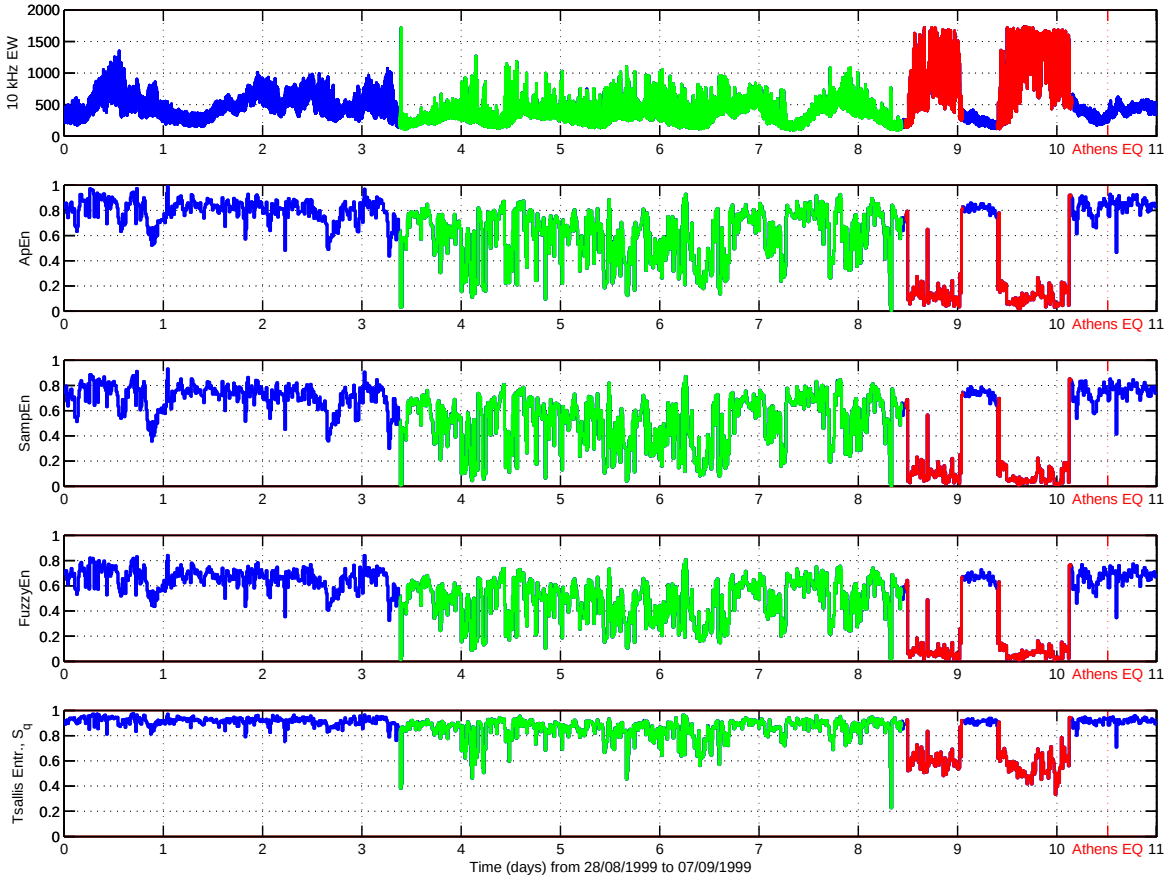
A way to examine transient phenomena like preseismic EM emissions is to analyze the corresponding time series into a sequence of distinct time windows. The aim is to discover a clear difference in the dynamical characteristics as the catastrophic event is approaching. As already pointed out, $ApEn$, $SampEn$ and $FuzzyEn$ are especially useful for the analysis of short time series contaminated by noise. Since the recorded EM time-series are usually strongly contaminated by the background noise at the location of the field measurement station, $ApEn$, $SampEn$ and $FuzzyEn$ are particularly appropriate for their analysis in the form of distinct (short) time windows.

Seismicity is a critical phenomenon [147,148,155]. Thus, it is expected that a significant change in the statistical pattern, namely, the appearance of entropy "drops", represents a deviation from the normal behavior of the EM recordings, revealing the presence of an EM anomaly. Specifically, we aim to reveal any increase of the organization level (entropy drop) in the observed EM time series as approaching the time of the EQ occurrence, by applying the different entropic metrics of symbolic Tsallis entropy, $ApEn$, $SampEn$ and $FuzzyEn$.

As an illustrative example, we consider here the preseismic kilohertz (kHz) EM activities associated with the Athens (Greece) earthquake ($M = 5.9$) that occurred on 7 September 1999 [137,138,156,157, 162–166]. All entropy calculations were performed on 1,024 sample long windows sliding along the time series with 25% mutual overlap, *i.e.*, with a step of 768 samples. For the symbolic Tsallis entropy, the value $q = 1.8$ [137] was found for the non-extensivity parameter, q. Here, the results for words of a length of $n = 3$ are presented. $ApEn$, $SampEn$ and $FuzzyEn$ have been estimated using the same setting as for the Dst time series in Section 3.1.

Figure 3 shows the same analyses as those presented for the Dst index (Figure 2) for the 10 kHz preseismic EM time series associated with the Athens earthquake. It can be observed that lower entropy values, compared to that of the background noise, can be identi?ed between the time points $\sim$ 292,600 s and $\sim$729,000 s (green-colored part in Figure 3), although sparsely distributed in time. In turn, the entropy values suddenly drop during the two strong EM pulses at the tail of the analyzed excerpt of the Athens kHz signal, signifying a different behavior. This has been attributed to a new distinct phase in the tail of the EQ preparation process, which is characterized by a significantly higher degree of organization and lower complexity in comparison to that of the preceding phase (e.g., [153,157]).

Figure 3. From top to bottom: Part of the recorded time series of the 10 kHz (east-west) magnetic field strength (in arbitrary units) covering an 11-days period from August 28, 1999, 00:00:00 UT (Universal Time) , to September 7. 1999, 23:59:59 UT, associated with the Athens EQ, and the corresponding $ApEn$, $SampEn$, $FuzzyEn$ and non-extensive Tsallis entropy for sliding windows in time (see text). The common horizontal axis is the time (in days), denoting the relative time position from the beginning of the analyzed part of the EM recording. The preseismic EM emissions exhibit lower entropy values, which are observed in two distinct phases: the first phase characterized by lower entropy values sparsely distributed in time (*cf.* the part of the time series and entropy plots displayed in green), followed by a phase characterized by continuous significantly lower entropy values (time series and entropy plotted in red).



From the obtained results, it seems that all employed entropic metrics display a comparable performance. However, $ApEn$, $SampEn$ and $FuzzyEn$ generally present stronger fluctuations with time, providing a higher resolution and a clearer difference of the EM anomalies, especially the two strong EM pulses, from the background noise baseline. It should be noted that the Tsallis entropy calculations presented in Figure 3 were performed without an explicit check of stationarity, considering that the window length of 1,024 points is small enough to provide pseudo-stationarity conditions.

3.3. Climate and Related Fields

Due to the relatively large amount of data on climate variability and the closely related aspects of Earth system dynamics (for example, hydrology and biogeochemistry), there have been numerous applications of statistical mechanics and information-theoretic concepts, including various entropic characteristics and complexity measures, in this field. Due to the large number of potential applications in this area, we restrict ourselves in the following to summarizing some examples of recent efforts related to certain specific climate-related studies, which can be considered representative for a great variety of potential research questions.

3.3.1. Complexity of Present-Day Climate Variability

The quantitative characterization of dynamical complexity in today's Earth climate has been the subject of many studies. Notably, entropic characteristics and associated measures of complexity are particularly well suited for this purpose. For example, Paluš and Novotná [167] demonstrated the use of redundancies (*i.e.*, multivariate generalizations of MI), in combination with surrogate data preserving the linear correlation properties of the original time series, for testing for the nonlinearity of different climatic variables (surface air pressure and temperature, geopotential height data), a necessary prerequisite for the emergence of chaotic dynamics.

On the global scale, von Bloh *et al.* [168] studied surface air temperature observations, as well as the corresponding outcomes of a global circulation model. In order to reveal the spatial characteristics associated with the nonlinear dynamics of the climate system, they considered the spatial patterns of the second-order Rényi entropy, H_2, estimated from recurrence plots (A similar analysis has been recently presented by Paluš *et al.* [169] in terms of a Gaussian process entropy rate computed for globally distributed surface air temperature anomalies). Their results suggested that state-of-the-art climate models (at the time of this study) have not been able to trace the fundamental spatial signatures of nonlinearity in observed climate dynamics, such as marked latitudinal, as well as land-ocean entropy gradients. The latter finding can be considered a well-established fact, which has been described for other nonlinear characteristics, such as the Hurst exponent describing the degree of long-range dependence in temperature fluctuations (see, e.g., [170–173]).

One reason for the observed complexity of temperature fluctuations is that the climate exhibits relevant variability on a multitude of different time-scales, from diurnal to seasonal and inter-annual scales. The longer time-scales are particularly highlighted by nonlinear large-scale oscillatory (dipole) patterns, such as the El Niño/Southern Oscillation (ENSO). Investigating the spatial profile and temporal variability of such patterns and characterizing their degree of dynamical disorder can provide a vast amount of complementary information not provided by classical linear analysis techniques. Among other entropic characteristics, Tsallis' non-extensive entropy has been found particularly useful for this purpose, once more making use of the fact that the climate system is typically far from equilibrium.

Ausloos and Petroni [174] studied the normalized variability of the southern oscillation index (SOI) on different time-scales. Their results indicated a qualitative change in the non-extensivity as the considered scales vary, leading to a classical Boltzmann-Gibbs scaling ($q \to 1$) at larger scales. Beyond exclusive consideration of S_q, in the non-extensive statistical mechanics framework, characteristic

scaling exponents can also be found for other quantities, which is formalized by Tsallis' q-triplet [175], which has been obtained for daily ENSO fluctuations [176], but also other geoscientific variables, such as the depth of the stratospheric ozone layer [177], various meteorological variables, earthquake inter-occurrence times, magnetospheric dynamics, *etc.* [178].

In addition to non-extensive entropy, also other information-theoretic entropy concepts have great potentials for discriminating between climatological records regarding their respective dynamical complexity. As a recent example, we mention here the consideration of the complexity-entropy causality plane for the characterization of daily streamflow time series [36], vegetation dynamics based on remote sensing data [179] or the degree distribution of a complex network obtained from correlations among surface air temperature records in the ENSO region in the equatorial Pacific [180].

3.3.2. Dynamical Transitions in Paleoclimate Variability

In order to obtain reliable estimates of the dynamical complexity associated with the variability of large-scale climate modes, such as ENSO, on long time-scales, instrumental records may provide insufficient time coverage. Instead, it is feasible to utilize paleoclimate proxy data for obtaining information on nonlinear variability during the last few centuries, millennia or even the deeper past. In such situations, entropic characteristics can again provide interesting insights into qualitative changes in historical climate variability.

Saco *et al.* [181] studied the histogram-based Shannon entropy, as well as permutation entropy for a proxy record obtained from a sedimentary sequence from Laguna Pallcacocha, Ecuador. Their study revealed a distinct time interval with markedly reduced permutation entropy corresponding to a known period of large-scale rapid climate change (RCC) between 9,000 and 8,000 years before present (BP). The latter period includes the globally observed 8.2 kyr (kiloyear) event, which is believed to have originated from a strong meltwater pulse into the North Atlantic [182]. In terms of dynamical entropy, the obtained results indicate a considerable degree of regularity in the variability of environmental conditions. We hypothesize that this regularity indeed reflects the marked variability profile during the 8.2 kyr event, which exceeds the typical magnitude of "background fluctuations". Notably, in the case of other, weaker RCC episodes during the last 8,000 years [182], there is no such clear decrease in permutation entropy. Instead, the corresponding entropy values are found to display a marked low-frequency variability, where even some of the temporary maxima of entropy coincide with known RCC episodes.

The special characteristics of the time interval between 9,000 and 8,000 years BP are supported by complementary analyses of Ferri *et al.* [183] using Tsallis' q-triplet [175], who found that the three characteristic scaling exponents from the triplet change their relative ordering during this time interval. The aforementioned results once more underline the relevance of entropic characteristics (especially in the non-extensive framework) as indicators of approaching critical transitions, which is additionally supported by the results of another study on some Antarctic ice core records [184].

The previous results highlight the great potentials of nonlinear methods of time series analysis for the investigation of paleoclimate variability, which are not restricted to statistical mechanics and information-theoretic approaches, but have also been demonstrated for phase space-based characteristics, such as different quantifiers based on recurrence plots [85,185–189]. As an important cautionary note, we

need to mention the presence of severe methodological problems, due to the non-uniform sampling and time-scale uncertainty of paleoclimate proxy data [190–192], which often do not permit a straightforward use of existing (linear and nonlinear) methods of time series analysis, but call for the development of specifically tailored estimators. Notable exceptions are tree rings, annually laminated (varved) sediments and other layer-forming geological archives (e.g., some ice cores, corals, *etc.*), where a reliable age-depth model can eventually be inferred based on layer counting.

3.3.3. Hydro-Meteorology and Land-Atmosphere Exchanges

Another climate-related recent field of application of information-theoretic concepts is the investigation of cross-scale processes in biogeochemistry and ecohydrology, *i.e.*, the investigation of interdependencies between atmospheric dynamics and hydrology, on the one hand, and the biosphere (especially vegetation), on the other hand. In general, the biosphere is considered an important part of the Earth system, which exhibits mutual interactions with other components on, or even between, various spatial and temporal scales. The identification of such interdependencies, along with the associated relevant scales, is a challenging task for which bi- and multi-variate characteristics, based on information theory, appear particularly well suited.

Realizing the latter fact, the application of methods, such as mutual information or transfer entropy, has recently gained increasing importance. Brunsell and co-workers [193,194] applied entropy concepts for quantifying the information gained by surface vegetation as a function of the time-scale covered by the input precipitation field, revealing a clear relationship between information content and data resolution. In addition, they were able to characterize the sensitivity of evapotranspiration estimates to the spatial scales of vegetation and soil moisture dynamics obtained from satellite measurements that have been used for deriving these estimates. Stoy *et al.* [195] utilized the Shannon entropy, as well as relative entropy (Kullback-Leibler divergence) for assessing the amount of information contributed by the aggregation of scales in remote sensing data. Based on the inferred loss of information, they were able to define an "optimum" pixel size for working with such data. Subsequently, Brunsell [196] applied Shannon entropy, relative entropy and mutual information content for investigating the spatial variation in the temporal scaling of daily precipitation data, revealing the existence of distinct scaling regions in precipitation not discriminated before by traditional analysis techniques.

While in the aforementioned examples, information content and information transfer have been considered for different datasets with the same spatial and/or temporal resolution, a more general framework was put forward recently [197]. In this context, Shannon entropy and relative entropy have been used for identifying the dominant spatial scales in land-atmosphere H_2O fluxes between surface and atmosphere obtained from remote sensing data and investigating the interactions of corresponding processes for different stages of the phenological cycle (*i.e.*, the developmental stages of vegetation). The obtained results provided a deep and unique look into the multi-scale spatial structure of evapotranspiration, one of the most important climate variables determining vegetation growth. In a similar spirit, Cochran *et al.* [198] used relative entropy together with a wavelet-based multi-resolution analysis of ground-based eddy covariance measurements from so-called flux towers for studying the CO_2 exchange between the atmospheric boundary layer and the free troposphere. The aforementioned studies exemplify the relevance of entropic characteristics for studying various aspects of climate-vegetation

feedback on local to global scales, thereby providing previously unrecognized information on the terrestrial carbon cycle and, hence, the Earth's biogeochemistry.

As a complementary approach to studying interdependencies between the biosphere and climate system, Ruddell and Kumar put forward the idea of information process networks in ecohydrology, characterizing different types of coupling corresponding to synchronization, feedback and forcings [199,200]. Here, the quantification of information flows between simultaneously measured variables at various time lags by means of transfer entropy provides weighted network representations of the corresponding interaction structures, which can be further quantified using proper statistical measures. Kumar and Ruddell [201] successfully applied their formalism for inferring feedback loops from flux tower measurements in seven ecosystems in different climatic regions. As a particularly relevant result, they provided evidence that feedbacks tend to maximize information production, since variables participating in feedback loops exhibit "moderated" variability. The amount of information production was found to increase with gross ecosystem productivity and revealed a clear dependence of the corresponding processes on the specific ecological and climate setting, seasonal patterns, *etc.* Specifically, drier and colder ecosystems were found to respond more intensely and immediately to small changes in water and energy supply, which is in accordance with the general experience in ecology.

3.3.4. Interdependencies and Causality between Atmospheric Variability Patterns

One important recent field of application of mutual information is the construction of *climate networks* from spatially distributed climate records. Here, individual time series corresponding to data from distinct measurement stations or grid points in a climate model are considered as variables, the strength of statistical associations between which can then be characterized by a variety of linear, as well as nonlinear measures. Selecting the strongest and/or statistically most significant pairwise associations provides a discrete set of spatial interdependencies, which can be further analyzed by network-theoretic concepts. Besides classical linear correlations and concepts referring to different notions of synchronization [202–205] (which have proven their potentials for bivariate analyses of climate data [206,207] before applications to network construction), application of mutual information based on symbolic dynamics [208,209] or order patterns [210,211] has attracted considerable interest and leads to a multi-scale description of the backbone of relevant spatial interdependencies in the climate system. Recently, the corresponding framework has been successfully generalized to studying directed interrelationships by making use of conditional mutual information [60]. We need to mention that the climate network approach based on information-theoretic characteristics is a subject of ongoing methodological discussions, referring to effects, such as the intrinsic dynamic complexity (which can be characterized itself by concepts, such as Gaussian process entropy rates [169]) or data artifacts resulting in spurious nonlinearities [212].

In addition to climate network construction and analysis, there have been first successful case studies applying information-theoretic concepts for coupling and causality analysis associated with some more specific research questions from climatology. Pompe and Runge [65] used a technique based on permutation entropy for studying the mutual interdependency between sea-surface temperature (SST) anomalies at two specific sites in the North Atlantic, motivated by the path of the North Atlantic Current. Their analysis revealed two distinct, relevant time-scales of oceanic coupling in addition to

an instantaneous interrelationship; the latter is most likely mediated by atmospheric processes, namely the North Atlantic Oscillation (NAO) dipole pattern. These non-zero coupling delays were not as clearly detectable using the traditional mutual information or linear cross-correlation. In a similar study, graphical models extending the established transfer entropy framework have proven their skills in obtaining climatologically meaningful coupling patterns from the station data of daily sea-level pressure over Central and Eastern Europe during winter [59]. Another recent study making use of the same formalism investigated the vertical heat transport between the sea surface and upper troposphere based on monthly air temperature anomalies at different altitudes, having distinct implications for our mechanistic picture of the functioning of the atmospheric Walker circulation [64].

As an illustrative case study, we recall here the example originally provided by [213], who demonstrated the estimation algorithm of MIT using the linear partial correlation and its application to studying the Pacific Walker circulation based on reanalysis data with monthly resolution. In the following, we re-examine this example using the nonlinear conditional mutual information estimated with the k-nearest neighbor approach described in Section 2.4.1. .

The basic mechanism of the Walker circulation [214–217] suggests that heating of the western equatorial Pacific, which is accompanied by low surface air pressure, promotes upwelling of waters in the eastern Pacific. This leads to easterly trade winds along the surface of the equatorial Pacific. The moist air then rises above the western Pacific, and the circulation is closed by the dry air masses descending along the central and eastern Pacific.

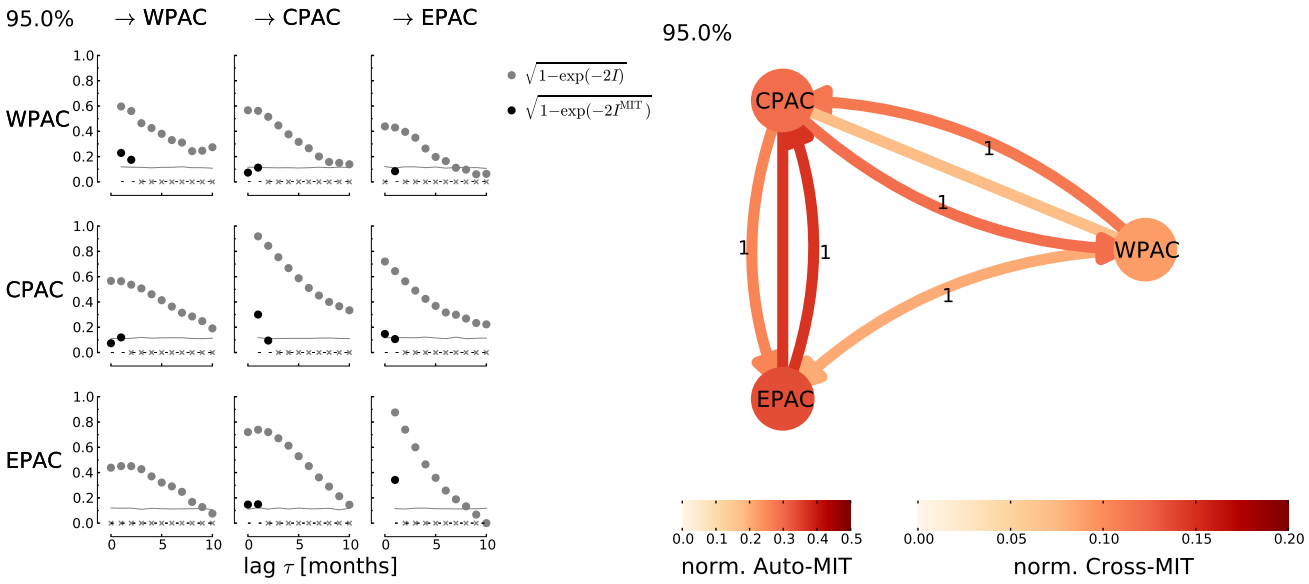
To test whether the feedback between the eastern Pacific (EPAC, represented here by the average SST over 80°–100° W, 5°S–5°N) and western Pacific (WPAC, average sea-level pressure anomalies over the region 130°–150° E, 5°S–5°N) was mediated via the surface of the central equatorial Pacific (CPAC, average SST over 150°–120° W, 5°S–5°N), we study the time series graph and MIT of the three-variable process (EPAC, CPAC, WPAC).

Figure 4 shows the lagged MI and the significant values of MIT after the time series graph was estimated using a significance level of $\alpha = 95\%$. The MI lag function is significant between all pairs of variables, making a precise assessment of the causal structure impossible. For example, we find broad peaks between EPAC $\to$ CPAC and CPAC $\to$ WPAC, while the MIT values are significant only at lags 1 and 0 (indicating a contemporaneous dependency). Very interestingly, while MI features a broad peak in EPAC $\to$ WPAC, the time series graph shows no link (even for a lower significance level of 90%). Indeed, this link is obviously mediated via the surface of the central equatorial Pacific. In turn, the MIT of the link WPAC $\to$ EPAC at lag 1 is still significant, which demonstrates that the link back takes a different path, not via the surface of the equatorial central Pacific, which is in line with the climatological theory that the air masses return into the high troposphere.

In summary, the time series graph approach confirms the Walker circulation pattern, and the MIT further gives an estimate of the strengths of the underlying causal dependencies. We find generally very strong auto-MIT values, especially for the temperature variables, CPAC and EPAC, which indicates their strong inertia due to the large specific heat capacity of the ocean. Furthermore, the MIT values of the couplings EPAC $\to$ CPAC and CPAC $\to$ WPAC are comparable, whereas the MI peak of the former is significantly larger. This effect can be explained by the strong auto-dependencies within EPAC and CPAC, which increase MI, but are 'filtered out' in MIT. Another interesting point is that the link

WPAC → CPAC that describes the descending mechanism is recovered here, while it was missing in the linear analysis in [213]. This finding suggests a nonlinear character of this dependency. However, in the simple Walker circulation picture, one would not expect the link CPAC → EPAC, which could be explained by the fact that the above described Walker circulation pattern is different during El Niño events, where the region of atmospheric updrafts shifts towards the central Pacific and also broadens markedly. Here, we used the whole time sample to test the hypothesis that the "average" influence is mediated via the central Pacific; a time-resolved analysis will be the subject of future research.

Figure 4. (**Left**) Lag functions of mutual information (gray) and significant momentary information transfer (MIT) values (black) for all pairs of variables (western Pacific (WPAC), central equatorial Pacific (CPAC), eastern Pacific (EPAC)). Note that the (conditional) mutual information ((C)MI) values have been normalized to the scale of (partial) correlations according to Equation (32). The gray horizontal line denotes the significance threshold for MI only. (**Right**) Process graph as in Figure 1b, where the node color represents the auto-MIT strength at lag 1 and the link color the cross-MIT strength, with the label denoting the lag. Straight lines mark contemporaneous links. The most important finding is the vanishing link EPAC → WPAC, which shows that the influence of the east on the west Pacific is mediated via the surface of the equatorial central Pacific. In turn, the link WPAC → EPAC is conserved, implying that this influence was not mediated via this region.



4. Discussion

The examples discussed in Section 3 demonstrate that entropic characteristics provide versatile means to detect not only spatial patterns, but also temporal changes of dynamical complexity, which can be particularly relevant for anticipating approaching qualitative changes in the dynamical regime of the system under study. From a statistical point of view, such critical phenomena (extreme events) should be characterized (among other features, [218–222]) by the following crucial symptoms: (i) a high degree of organization or high information content; (ii) strong persistency, indicating the presence of a positive

feedback mechanism in the underlying phenomenon that drives the systems out of equilibrium; (iii) the existence of a clear preferred direction of the underlying activities; and (iv) the absence of any footprint of a second-order transition in equilibrium.

In this review, we have focused on the first of the aforementioned symptoms in terms of entropies of different kinds: on the one hand, non-extensive Tsallis entropy using symbolic dynamics techniques and, on the other hand, approximate entropy, sample entropy and fuzzy entropy. We note that it has been established that physical systems exhibiting critical phenomena, such as magnetic storms, earthquakes or climate shifts, which are characterized by long-range interactions or long-term memories, or are of a multi-fractal nature, are best described by the generalized statistical mechanics formalism proposed by Tsallis [1,25]. Therefore, although the various introduced measures of dynamical organization or information content can reveal crucial features associated with the launch of an extreme event, namely, low complexity, high degree of organization or high information content, the use of Tsallis entropy seems to be a particularly appropriate measure from a physical point of view.

We clarify that the use of Tsallis entropy leads to double information. The entropic index, q, describes the deviation of Tsallis entropy from the standard Boltzmann-Gibbs entropy, namely, the appearance of long-range interactions, long-term memory and/or multi-fractal behavior. However, the index, q, itself is not a measure of the complexity or organization of the system. In turn, the time variation of the Tsallis entropy, S_q, for a given q quantifies the dynamic changes of the complexity or organization of the system. Lower S_q values characterize the portions of the signal with lower complexity or higher organization.

In contrast to the non-extensive statistical mechanics framework, $ApEn$, $SampEn$ and $FuzzyEn$ are especially useful for the analysis of short time series contaminated by noise. For example, since the recorded preseismic EM emissions time series are usually strongly contaminated by such background noise, $ApEn$, $SampEn$ and $FuzzyEn$ are particularly appropriate for their analysis in the form of distinct (short) time windows. We expect similar benefits for the analysis of certain time series in a climatic context, as well.

Mounting empirical evidence has been supporting the possibility that a number of systems arising in disciplines as diverse as physics, biology, engineering and economics may have certain quantitative features that are intriguingly similar. These properties can be conveniently grouped under the headings of scale invariance and universality [223–228]. For instance, de Arcangelis *et al.* [228] have shown that the stochastic processes underlying apparently different phenomena, such as solar flares and earthquakes, have universal properties.

Balasis *et al.* [229] have recently established the principle of universality in solar flares, magnetic storms and earthquake dynamics. The aforementioned similarity is quantitatively supported by the observation of power-laws in the distributions of solar flare and magnetic storm energy related to a non-extensive Tsallis formalism that gives the Gutenberg-Richter law for the earthquake magnitude distribution as a special case. In a later study [230], Balasis *et al.* extended their previous work [229] to include preseismic EM emissions. They, thus, provide evidence for universal behavior in the Earth's magnetosphere, solar corona and in the coupled Earth's ionosphere, atmosphere and lithosphere system, where magnetic storms, solar flares and preseismic kHz EM anomalies occur, respectively.

Another interesting observation that witnesses universality among different systems is the emergence of discrete scale invariance. Discrete scale invariance has been documented so far in ruptures, seismicity,

financial systems (for a review see [147]) and, also, quite recently, for Dst time series that represent magnetospheric activity [231].

The evidence for universal statistical behavior suggests the possibility of a common approach to the forecasting of space weather and earthquakes. In any case, the transfer of ideas and methods of seismic forecasting to the prediction of solar flares and magnetic storms could improve space weather forecasting. Similar concepts could be applied to the treatment of space weather and earthquakes along with extreme events arising from meteorological hazards. For instance, weather forecasting techniques could be applied and tested to predict the evolution of space weather or earthquakes and *vice versa*.

Besides the quantitative characterization of dynamical complexity, various recent examples summarized in this review point to a considerable importance of entropy-based characteristics for identifying and quantifying (possibly nonlinear) interdependencies between different (geophysical or other) variables, variability at different scales, *etc.* Besides obtaining estimates of the strengths of bivariate interrelationships between two variables and assessing the corresponding significance by means of suitable statistical tests, unveiling and characterizing asymmetric interdependencies is a crucial point for developing a better "mechanistic" understanding of the governing processes, which can exhibit symmetric interactions between different quantities, but also distinct "driver-response" relationships. The latter are of particular relevance when thinking about causality issues, e.g., trying to understand how certain dynamical patterns in one variable affect the behavior of another one. As a specific geoscientific example for this kind of research question, we mention the not yet systematically understood effect of extreme climatic conditions on vegetation dynamics in different types of ecosystems [232].

Among other complications for appropriate analysis and modeling of geoscientific data, cross-scale variability is a common problem in Earth sciences [3] when it comes to the detection, quantification and attribution of physical processes interlinking such "components". Notably, such dynamical interdependencies foster complex nonlinear dynamics in terms of feedbacks or various kinds of synchronization. Beyond certain climatic examples mentioned in this review, we envision the potential of modern approaches related to transfer entropies, momentary information transfer and graphical models for tackling contemporary research questions in various fields of Earth system science. Among others, these information-theoretic functionals have great potentials for applications to problems, such as the time-dependent coupling between solar wind and the Earth's magnetosphere and the geospace (storm–substorm relationship and interdependency [233]).

Finally, we envision the systematic application of information-theoretic approaches for the explicit study of spatio-temporal dynamics. Specifically, combinations with complex network analysis, where concepts like ordinary or conditional mutual information have already been successfully utilized for inferring the most relevant couplings among sets of variables in a climate context [60,208–211], have great potentials for unveiling and quantifying possibly unknown dynamical interrelationships, responses and feedbacks. Besides their application in the context of climate networks, the latter concepts have also been successfully utilized for studying the spatio-temporal backbone of seismicity [234], suggesting a much wider range of potential research questions to be addressed with similar approaches.

5. Conclusion

Entropy characteristics based on statistical mechanics and information-theoretic considerations exhibit a large degree of conceptual similarity. For example, the classical Boltzmann-Gibbs entropy is known to have its analog in the Shannon entropy originally introduced in a telecommunications science context. This duality, which is further elaborated on in terms of the recently developed non-extensive statistical mechanics framework, provides the basic foundation and rationale for the systematic application of information-theoretic notions of entropies and associated complexity measures and, also, complementary phase space-based entropy characteristics to investigating dynamical complexity in "abstract" (mathematical), as well as observational (physical) systems. More specifically, the idea of quantifying the disorder of a configuration in a microscopic thermodynamic framework is transferred to "configurations" representing observable dynamics, *i.e.*, patterns arising due to the evolution of the system under study, which allows quantifying dynamical complexity by various types of entropies and associated complexity measures.

In this review, we have identified four main fields of application of such characteristics in the Earth sciences:

- the measurement of complexity of stationary records (e.g., for identifying spatial patterns of complexity in climate data);
- the identification of dynamical transitions and their preparatory phases from non-stationary time series (*i.e.*, critical phenomena associated with approaching "singular" extreme events, like earthquakes, magnetic storms or climatic regime shifts known from paleoclimatology, but expected to be possible in the future climate, associated with the presence of climatic tipping points [235,236]);
- the characterization of complexity and information transfer between variables, subsystems or different spatial and/or temporal scales (e.g., couplings between solar wind and the magnetosphere or the atmosphere and vegetation);
- the identification of directed interdependencies between variables related to causal relationships between certain geoscientific processes, which are necessary for an improved process-based understanding of the coupling between different variables or even systems, a necessary prerequisite for the development of appropriate numerical simulation models.

For tackling the aforementioned, as well as conceptually similar research questions, different classes of entropic characteristics have been discussed, illustrating their respective strengths for different kinds of data and problems. Notably, methodological potentials and limitations of the considered approaches are closely associated with the specific properties of the available time series:

- Symbolic dynamics approaches (e.g., block entropies, permutation entropy, Tsallis' non-extensive entropy) typically require a considerable amount of data for their accurate estimation, which is mainly due to the systematic loss of information detail in the discretization process. Consequently, we suggest that such approaches are particularly well suited for studying long time series of

approximately stationary processes (e.g., climate or hydrological time series) and high-resolution data from non-stationary and/or non-equilibrium systems (e.g., electromagnetic recordings associated with seismic activity).

- Tsallis' non-extensive entropy is specifically tailored for describing non-equilibrium phenomena associated with changes in the dynamical complexity of recorded fluctuations, as arising in the context of critical phenomena, certain extreme events or dynamical regime shifts.

- Distance-based entropies ($ApEn$, $SampEn$ and $FuzzyEn$) provide a higher degree of robustness to short and possibly noisy data than other entropy characteristics based on any kind of symbolic discretization. This suggests their specific usefulness for studying changes of dynamical complexity across dynamical transitions in a sliding windows framework.

- Directional bivariate measures and graphical models allow for a statistical evaluation of causality between variables. In this spirit, this class of approaches provides a versatile and widely applicable toolbox, where the graphical model idea gives rise to a multitude of bivariate characteristics (including conditional mutual information and transfer entropy as special cases), but may exhibit a considerable degree of algorithmic complexity in practical estimation.

We emphasize that the research questions and data properties mentioned above are relatively generic and can be found in various fields of the Earth sciences, but also other scientific disciplines. Learning from successful applications of entropy concepts originating in statistical mechanics or information theory in one field can therefore provide important insights or conceptual ideas for other areas (in the Earth sciences or beyond) or even stimulate new research questions and approaches. Some corresponding examples for prospective problems to be studied by such means have been mentioned in Section 4. In this spirit, this review is intended to provide such a stimulation, without aiming to provide a complete summary of all available approaches or recent applications.

Acknowledgments

This work has been financially supported by the joint Greek-German IKYDA 2013 project Transdisciplinary assessment of dynamical complexity in magnetosphere and climate: A unified description of the nonlinear dynamics across extreme events funded by Greek State Scholarships Foundation (IKY) and Deutscher Akademischer Austausch Dienst (DAAD). The Dst data are provided by the World Data Center for Geomagnetism, Kyoto [122]. We also would like to acknowledge the support of the "Learning about Interacting Networks in Climate" (LINC) project (no. 289447) funded by EC's Marie-Curie Initial Training Networks (ITN) program (FP7-PEOPLE-2011-ITN).

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Tsallis, C. *Introduction to Nonextensive Statistical Mechanics, Approaching a Complex World*; Springer: New York, NY, USA, 2009.
2. Turcotte, D.L. Modeling Geocomplexity: "A New Kind of Science". In *Earth & Mind: How Geoscientists Think and Learn about the Earth. GSA Special Papers volume 413, 2006*; Manduca, C., Mogk, D., Eds.; Geological Society of America: Boulder, CO, USA, 2007; pp. 39–50.
3. Donner, R.V.; Barbosa, S.M.; Kurths, J.; Marwan, N. Understanding the earth as a complex system—Recent advances in data analysis and modelling in earth sciences. *Eur. Phys. J. Spec. Top.* **2009**, *174*, 1–9.
4. Donner, R.V.; Barbosa, S.M. (Eds.) *Nonlinear Time Series Analysis in the Geosciences —Applications in Climatology, Geodynamics, and Solar-Terrestrial Physics*; Springer: Berlin, Germany, 2008.
5. Kurths, J.; Voss, A.; Saparin, P.; Witt, A.; Kleiner, H.; Wessel, N. Quantitative analysis of heart rate variability. *Chaos* **1995**, *5*, 88–94.
6. Kantz, H.; Schreiber, T. *Nonlinear Time Series Analysis*; Cambridge University Press: Cambridge, UK, 1997.
7. Sprott, J.C. *Chaos and Time Series Analysis*; Oxford University Press: Oxford, UK, 2003.
8. Hao, B.-L. *Elementary Symbolic Dynamics and Chaos in Dissipative Systems*; World Scientific: Singapore, Singapore, 1989.
9. Karamanos, K.; Nicolis, G. Symbolic dynamics and entropy analysis of Feigenbaum limit sets. *Chaos Solitons Fractals* **1999**, *10*, 1135–1150.
10. Daw, C.S.; Finney, C.E.A.; Tracy, E.R. A review of symbolic analysis of experimental data. *Rev. Scient. Instrum.* **2003**, *74*, 915–930.
11. Donner, R.; Hinrichs, U.; Scholz-Reiter, B. Symbolic recurrence plots: A new quantitative framework for performance analysis of manufacturing networks. *Eur. Phys. J. Spec. Top.* **2008**, *164*, 85–104.
12. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423, 623–656.
13. Saparin, P.; Witt, A.; Kurths, J.; Anishchenko, V. The renormalized entropy—An appropriate complexity measure? *Chaos, Solitons Fractals* **1994**, *4*, 1907–1916
14. Rényi, A. On Measures of Entropy and Information. In Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, CA, USA, 20–30 June 1960; University of California Press: Berkeley, CA, USA, 1961; volume I, pp. 547–561.
15. Hartley, R.V.L. Transmission of Information. *Bell Syst. Tech. J.* **1928**, *7*, 535–563.
16. Graben, P.; Kurths, J. Detecting subthreshold events in noisy data by symbolic dynamics. *Phys. Rev. Lett.* **2003**, *90*, doi:10.1103/PhysRevLett.90.100602.

17. Eftaxias, K.; Athanasopoulou, L.; Balasis, G.; Kalimeri, M.; Nikolopoulos, S.; Contoyiannis, Y.; Kopanas, J.; Antonopoulos, G.; Nomicos, C. Unfolding the procedure of characterizing recorded ultra low frequency, kHz and MHz electromagnetic anomalies prior to the L'Aquila earthquake as pre-seismic ones. Part I. *Nat. Hazards Earth Syst. Sci.* **2009**, *9*, 1953–1971.
18. Potirakis, S.M.; Minadakis, G.; Eftaxias, K. Analysis of electromagnetic pre-seismic emissions using Fisher Information and Tsallis entropy. *Physica A* **2012**, *391*, 300–306.
19. Nicolis, G.; Gaspard, P. Toward a probabilistic approach to complex systems. *Chaos, Solitons Fractals* **1994**, *4*, 41–57.
20. Ebeling, W.; Nicolis, G. Word frequency and entropy of symbolic sequences: A dynamical perspective. *Chaos, Solitons Fractals* **1992**, *2*, 635–650.
21. Khinchin, A.I. *Mathematical Foundations of Information Theory*; Dover: New York, NY, USA, 1957.
22. McMillan, B. The basic theorems of information theory. *Ann. Math. Stat.* **1953**, *24*, 196–219.
23. Grassberger, P. Toward a quantitative theory of self-generated complexity. *Int. J. Theor. Phys.* **1986**, *25*, 907–938.
24. Wackerbauer, R.; Witt, A.; Atmanspacher, H.; Kurths, J.; Scheingraber, H. A comparative classification of complexity measures. *Chaos Solitons Fractals* **1994**, *4*, 133–173.
25. Tsallis, C. Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Phys.* **1988**, *52*, 479–487.
26. Tsallis, C. Generalized entropy-based criterion for consistent testing. *Phys. Rev. E.* **1998**, *58,* 1442–1445.
27. Zunino L.; Perez D.; Kowalski A.; Martin M.; Garavaglia M.; Plastino A.; Rosso O. Fractional Brownian motion, fractional Gaussian noise and Tsallis permutation entropy. *Physica A* **2008**, *387*, 6057–6068.
28. Balasis, G.; Daglis, I.A.; Papadimitriou, C.; Kalimeri, M.; Anastasiadis, A.; Eftaxias, K. Dynamical complexity in Dst time series using nonextensive Tsallis entropy. *Geophys. Res. Lett.* **2008**, *35*, doi:10.1029/2008GL034743.
29. Balasis, G.; Daglis, I.A.; Papadimitriou, C.; Kalimeri, M.; Anastasiadis, A.; Eftaxias, K. Investigating dynamical complexity in the magnetosphere using various entropy measures. *J. Geophys. Res.* **2009**, *114*, doi:10.1029/2008JA014035.
30. Bandt, C. Ordinal time series analysis. *Ecol. Model.* **2005**, *182*, 229–238.
31. Bandt, C.; Pompe, B. Permutation entropy: A natural complexity measure for time series. *Phys. Rev. Lett.* **2005**, *88*, doi:10.1103/PhysRevLett.88.174102.
32. Riedl, M.; Müller, A.; Wessel, N. Practical considerations of permutation entropy—A tutoral review. *Eur. Phys. J. Spec. Top.* **2013**, *222*, 249–262.
33. Amigó, J.; Keller, K. Permutation entropy: One concept, two approaches. *Eur. Phys. J. Spec. Top.* **2013**, *222*, 263–273.
34. Martin, M.T.; Plastino, A.; Rosso, O.A. Generalized statistical complexity measures: Geometrical and analytical properties. *Physica A* **2006**, *369*, 439–462.
35. Rosso, O.A.; Larrondo, H.A.; Martin, M.T.; Plastino, A.; Fuentes, M.A. Distinguishing noise from chaos. *Phys. Rev. Lett.* **2007**, *99*, doi:10.1103/PhysRevLett.99.154102.

36. Lange, H.; Rosso, O.A.; Hauhs, M. Ordinal pattern and statistical complexity analysis of daily stream flow time series. *Eur. Phys. J. Spec. Top.* **2013**, *222*, 535–552.
37. Pincus, S.M. Approximate entropy as a measure of system complexity. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 2297–2301.
38. Pincus, S.M.; Goldberger, A.L. Physiological time-series analysis: What does regularity quantify? *Am. J. Physiol.* **1994**, *266*, H1643–H1656.
39. Pincus, S.M.; Singer B.H. Randomness and degrees of irregularity. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 2083–2088.
40. Grassberger, P.; Procaccia, I. Estimation of the Kolmogorov entropy from a chaotic signal. *Phys. Rev. A* **1983**, *28*, 2591–2593.
41. Pincus, S.M.; Keefe, D. Quantification of hormone pulsatility via an approximate entropy algorithm. *Am. J. Physiol. Endocrinol. Metab.* **1992**, *262*, E741–E754.
42. Richman, J.S.; Moorman, J.R. Physiological time-series analysis using approximate and sample entropy. *Am. J. Physiol. Heart Circ. Physiol.* **2000**, *278*, H2039–2049.
43. Chen, W.; Wang, Z.; Xie H.; Yu, W. Characterization of surface EMG signal based on fuzzy entropy. *IEEE Trans. Neural Syst. Rehab. Eng.* **2007**, *15*, 266–272.
44. Chen, W.; Zhuang, J.; Yu, W.; Wang, Z. Measuring complexity using FuzzyEn, ApEn, and SampEn. *Med. Eng. Phys.* **2009**, *31*, 61–68.
45. Reshef, D.N.; Reshef, Y.A.; Finucane, H.K.; Grossman, S.R.; McVean, G.; Turnbaugh, P.J.; Lander, E.S.; Mitzenmacher, M.; Sabeti, P.C. Detecting novel associations in large data sets. *Science* **2011**, *334*, 1518–1524.
46. Fraser, A.M.; Swinney, H.L. Independent coordinates for strange attractors from mutual information. *Phys. Rev. A* **1986**, *33*, 1134–1140.
47. Kraskov, A.; Stögbauer, H.; Grassberger, P. Estimating mutual information. *Phys. Rev. E* **2004**, *69*, doi:10.1103/PhysRevE.69.066138.
48. Cellucci, C.J.; Albano, A.M.; Rapp, P.E. Statistical validation of mutual information calculations: Comparison of alternative numerical algorithms. *Phys. Rev. E* **2005**, *71*, doi:10.1103/PhysRevE.71.066208.
49. Kozachenko, L.F.; Leonenko, N.N. Sample estimate of the entropy of a random vector. *Probl. Peredachi Inf.* **1987**, *23*, 95–101.
50. Pompe, B. On some entropy methods in data analysis. *Chaos Solitons Fractals* **1994**, *4*, 83–96.
51. Prichard, D.; Theiler, J. Generalized redundancies for time series analysis. *Physica D* **1995**, *84*, 476–493.
52. Paluš, M. Testing for nonlinearity using redundancies: Quantitative and qualitative aspects. *Physica D* **1995**, *80*, 186–205.
53. Paluš, M. Detecting nonlinearity in multivariate time series. *Phys. Lett. A* **1996**, *213*, 138–147.
54. Schreiber, T. Measuring information transfer. *Phys. Rev. Lett.* **2000**, *85*, 461–464.
55. Granger, C. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* **1969**, *37*, 424–438.
56. Chen, Y.; Rangarajan, G.; Feng, J.; Ding, M. Analyzing multiple nonlinear time series with extended Granger causality. *Phys. Lett. A* **2004**, *324*, 26–35.

57. Ancona, N.; Marinazzo, D.; Stramaglia, S. Radial basis function approach to nonlinear Granger causality of time series. *Phys. Rev. E* **2004**, *70*, doi:10.1103/PhysRevE.70.056221.
58. Barnett, L.; Barrett, A.; Seth, A. Granger causality and transfer entropy are equivalent for Gaussian variables. *Phys. Rev. Lett.* **2009**, *103*, doi: 10.1103/PhysRevLett.103.238701.
59. Runge, J.; Heitzig, J.; Petoukhov, V.; Kurths, J. Escaping the curse of dimensionality in estimating multivariate transfer entropy. *Phys. Rev. Lett.* **2012**, *108*, doi:10.1103/PhysRevLett.108.258701.
60. Hlinka, J.; Hartman, D.; Vejmelka, M.; Runge, J.; Marwan, N.; Kurths, J.; Paluš, M. Reliability of inference of directed climate networks using conditional mutual information. *Entropy* **2013**, *15*, 2023–2045.
61. Lauritzen, S.L. *Graphical Models*; Oxford University Press: Oxford, UK, 1996.
62. Dahlhaus, R. Graphical interaction models for multivariate time series. *Metrika* **2000**, *51*, 157–172.
63. Eichler, M. Graphical modelling of multivariate time series. *Probab. Theor. Rel. Fields* **2012**, *153*, 233–268.
64. Runge, J.; Heitzig, J.; Marwan, N.; Kurths, J. Quantifying causal coupling strength: A lag-specific measure for multivariate time series related to transfer entropy. *Phys. Rev. E* **2012**, *86*, doi:10.1103/PhysRevE.86.061121.
65. Pompe, B.; Runge, J. Momentary information transfer as a coupling measure of time series. *Phys. Rev. E* **2011**, *83*, doi:10.1103/PhysRevE.83.051122.
66. Frenzel, S.; Pompe, B. Partial mutual information for coupling analysis of multivariate time series. *Phys. Rev. Lett.* **2007**, *99*, doi:10.1103/PhysRevLett.99.204101.
67. Grassberger, P.; Procaccia, I. Measuring the strangeness of strange attractors. *Physica D* **1983**, *9*, 189–208.
68. Grassberger, P.; Procaccia, I. Characterization of strange attractors. *Phys. Rev. Lett.* **1983**, *50*, 346–349.
69. Marwan, N.; Romano, M.C.; Thiel, M.; Kurths, J. Recurrence plots for the analysis of complex systems. *Phys. Rep.* **2007**, *438*, 237–329.
70. Thiel, M.; Romano, M.C.; Read, P.L.; Kurths, J. Estimation of dynamical invariants without embedding by recurrence plots. *Chaos* **2004**, *14*, 234–243.
71. Letellier, C. Estimating the Shannon entropy: Recurrence plots *versus* symbolic dynamics. *Phys. Rev. Lett.* **2006**, *96*, doi:10.1103/PhysRevLett.96.254102.
72. Faure, P.; Lesne, A. Recurrence plots for symbolic sequences. *Int. J. Bifurcation Chaos* **2011**, *20*, 1731–1749.
73. Beim Graben, P.; Hutt, A. Detecting recurrence domains of dynamical systems by symbolic dynamics. *Phys. Rev. Lett.* **2013**, *110*, doi:10.1103/PhysRevLett.110.154101.
74. Groth, A. Visualization of coupling in time series by order recurrence plots. *Phys. Rev. E.* **2005**, *72*, doi:10.1103/PhysRevE.72.046220.
75. Schinkel, S.; Marwan, N.; Kurths, J. Order patterns recurrence plots in the analysis of ERP data. *Cogn. Neurodyn.* **2007**, *1*, 317–325.
76. Arnhold, J.; Grassberger, P.; Lehnertz, K.; Elger, C.E. A robust method for detecting interdependencies: Application to intracranially recorded EEG. *Physica D* **1999**, *134*, 419–430.

77. Quian Quiroga, R.; Kraskov, A.; Kreuz, T.; Grassberger, P. Performance of different synchronization measures in real data: A case study on electroencephalographic signals. *Phys. Rev. E* **2002**, *65*, doi:10.1103/PhysRevE.65.041903.
78. Andrzejak, R.G.; Kraskov, A.; Stögbauer, H.; Mormann, F.; Kreuz, T. Bivariate surrogate techniques: Necessity, strengths, and caveats. *Phys. Rev. E* **2003**, *68*, doi:10.1103/PhysRevE.68.066202.
79. Smirnov, D.A.; Andrzejak, R.G. Detection of weak directional coupling: Phase-dynamics approach *versus* state-space approach. *Phys. Rev. E* **2005**, *71*, doi:10.1103/PhysRevE.71.036207.
80. Andrzejak, R.G.; Ledberg, A.; Deco, G. Detecting event-related time-dependent directional couplings. *New J. Phys.* **2006**, *8*, doi:10.1088/1367-2630/8/1/006.
81. Chicharro, D.; Andrzejak, R.G. Reliable detection of directional couplings using rank statistics. *Phys. Rev. E* **2009**, *80*, doi:10.1103/PhysRevE.80.026217.
82. Sugihara, G.; May, R.; Ye, H.; Hsieh, C.-H.; Deyle, E.; Fogarty, M.; Munch, S. Detecting causality in complex ecosystems. *Science* **2012**, *338*, 496–500.
83. Romano, M.C.; Thiel, M.; Kurths, J.; Grebogi, C. Estimation of the direction of the coupling by conditional probabilities of recurrence. *Phys. Rev. E* **2007**, *76*, doi: 10.1103/PhysRevE.76.036211.
84. Zou, Y.; Romano, M.C.; Thiel, M.; Marwan, N.; Kurths, J. Inferring indirect coupling by means of recurrences. *Int. J. Bifurcation Chaos* **2011**, *21*, 1009–1111.
85. Feldhoff, J.H.; Donner, R.V.; Donges, J.F.; Marwan, N.; Kurths, J. Geometric detection of coupling directions by means of inter-system recurrence networks. *Phys. Lett. A* **2012**, *376*, 3504–3513.
86. Klimas, A.J.; Vassiliadis, D.; Baker, D.N.; and Roberts, D.A. The organized nonlinear dynamics of the magnetosphere, *J. Geophys. Res.* **1996**, *101*, 13089–13113.
87. Tsurutani, B.; Sugiura, M.; Iyemori, T.; Goldstein, B.; Gonzalez, W.; Akasofu, A.; Smith, E. The nonlinear response of AE to the IMF BS driver: A spectral break at 5 hours. *Geophys. Res. Lett.* **1990**, *17*, 279–282.
88. Baker, D.N.; Klimas, A.J.; McPherron, R.L.; Buchner, J. The evolution from weak to strong geomagnetic activity: An interpretation in terms of deterministic chaos. *Geophys. Res. Lett.* **1990**, *17*, 41–44.
89. Vassiliadis, D.V.; Sharma, A.S.; Eastman, T.E.; Papadopoulos, K. Low-dimensional chaos in magnetospheric activity from AE time series. *Geophys. Res. Lett.* **1990**, *17*, 1841–1844.
90. Sharma, A.S.; Vassiliadis, D.; Papadopoulos, K. Reconstruction of low-dimensional magnetospheric dynamics by singular spectrum analysis. *Geophys. Res. Lett.* **1993**, *20*, 335–338.
91. Pavlos, G.P.; Athanasiu, M.A.; Rigas, A.G.; Sarafopoulos, D.V.; Sarris, E.T. Geometrical characteristics of magnetospheric energetic ion time series: Evidence for low dimensional chaos. *Ann. Geophys.* **2003**, *21*, 1975–1993.
92. Vörös, Z.; Baumjohann, W.; Nakamura, R.; Runov, A.; Zhang, T.L.; Volwerk, M.; Eichelberger, H.U.; Balogh, A.; Horbury, T.S.; Glassmeier, K.-H.; *et al*. Multi-scale magnetic field intermittence in the plasma sheet. *Ann. Geophys.* **2003**, *21*, 1955–1964.

93. Chang, T. Low-dimensional behavior and symmetry breaking of stochastic systems near criticality-can these effects be observed in space and in the laboratory? *IEEE Trans. Plasma Sci.* **1992**, *20*, 691–694.
94. Chang, T. Intermittent turbulence in the magnetotail. *EOS Trans. Suppl.* **1998**, *79*, S328.
95. Chang, T. Self-organized criticality, multi-fractal spectra, sporadic localized reconnections and intermittent turbulence in the magnetotail. *Phys. Plasmas* **1999**, *6*, 4137–4145.
96. Chang, T.; Wu, C.C.; Angelopoulos, V. Preferential acceleration of coherent magnetic structures and bursty bulk flows in Earth's magnetotail. *Phys. Scr.* **2002**, *T98*, 48–51.
97. Consolini, G. Sandpile Cellular Automata and Magnetospheric Dynamics. In *Cosmic Physics in the Year 2000*; Aiello, S., Ed.; SIF: Bologna, Italy, 1997; pp. 123–126.
98. Consolini, G.; Chang, T. Complexity, magnetic field topology, criticality, and metastability in magnetotail dynamics. *J. Atmos. Sol. Terr. Phys.* **2002**, *64*, 541–549.
99. Lui, A.T.Y.; Chapman, S.C.; Liou, K.; Newell, P.T.; Meng, C.I.; Brittnacher, M.; Parks. G.K. Is the dynamic magnetosphere an avalanching system? *Geophys. Res. Lett.* **2000**, *27*, 911–914.
100. Uritsky, V.M.; Klimas, A.J.; Vassiliadis, D.; Chua, D.; Parks, G.D. Scalefree statistics of spatiotemporal auroral emissions as depicted by POLAR UVI images: The dynamic magnetosphere is an avalanching system. *J. Geophys. Res.* **2002**, *107*, doi:10.1029/2001JA000281.
101. Angelopoulos, V.; Mukai, T.; Kokubun, S. Evidence for intermittency in Earth's plasma sheet and implications for self-organized criticality. *Phys. Plasmas* **1999**, *6*, 4161–4168.
102. Borovsky, J.E.; Funsten, H.O. MHD turbulence in the Earth's plasma sheet: Dynamics, dissipation, and driving. *J. Geophys. Res.* **2003**, *108*, doi:10.1029/2002JA009625.
103. Weygand, J.M.; Kivelson, M.G.; Khurana, K.K.; Schwarzl, H.K.; Thompson, S.M.; McPherron, R.L.; Balogh, A.; Kistler, L.M.; Goldstein, M.L.; Borovsky, J.; *et al.* Plasma sheet turbulence observed by Cluster II. *J. Geophys. Res.* **2005**, *110*, doi: 10.1029/2004JA010581.
104. Consolini, G.; Lui, A.T.Y. Sign-singularity analysis of current disruption. *Geophys. Res. Lett.* **1999**, *26*, 1673–1676.
105. Consolini, G.; Lui, A.T.Y. Symmetry Breaking and Nonlinear Wave-Wave Interaction in Current Disruption: Possible Evidence for a Phase Transition. In *Magnetospheric Current Systems—Geophysical Monograph*; Ohtani, S., Fujii, R., Hesse, M., Lysak, R.L. Eds.; **2000**, *118*, 395–401.
106. Sitnov, M.I.; Sharma, A.S.; Papadopoulos, K.; Vassiliadis, D.; Valdivia, J.A.; Klimas, A.J.; Baker, D.N. Phase transition-like behavior of the magnetosphere during substorms. *J. Geophys. Res.* **2000**, *105*, 12955–12974.
107. Sitnov, M.I.; Sharma, A.S.; Papadopoulos, K.; Vassiliadis, D. Modeling substorm dynamics of the magnetosphere: From self-organization and self-organized criticality to non-equilibrium phase transitions. *Phys. Rev. E* **2001**, *65*, doi:10.1103/PhysRevE.65.016116.
108. Consolini, G.; Chang, T. Magnetic field topology and criticality in geotail dynamics: Relevance to substorm phenomena. *Space Sci. Rev.* **2001**, *95*, 309–321.

109. Balasis, G.; Daglis, I.A.; Kapiris, P.; Mandea, M.; Vassiliadis, D.; Eftaxias, K. From pre-storm activity to magnetic storms: A transition described in terms of fractal dynamics. *Ann. Geophys.* **2006**, *24*, 3557–3567.
110. De Michelis, P.; Consolini, G.; Tozzi, R. On the multi-scale nature of large geomagnetic storms: An empirical mode decomposition analysis. *Nonlinear Processes Geophys.* **2012**, *19*, 667–673.
111. Wanliss, J.A.; Anh, V.V.; Yu, Z.-G.; Watson, S. Multifractal modeling of magnetic storms via symbolic dynamics analysis. *J. Geophys. Res.* **2005**, *110*, doi:10.1029/2004JA010996.
112. Uritsky, V.M.; Pudovkin, M.I. Low frequency $1/f$-like fluctuations of the AE-index as a possible manifestation of self-organized criticality in the magnetosphere. *Ann. Geophys.* **1998**, *16*, 1580–1588.
113. Freeman, M.P.; Watkins, N.W.; Riley, D. J. Power law distributions of burst duration and interburst interval in the solar wind: Turbulence or dissipative self-organized criticality? *Phys. Rev. E* **2000**, *62*, 8794–8797.
114. Chapman, S. C.; Watkins, N. W.; Dendy, R. O.; Helander P.; Rowlands, G. A simple avalanche model as an analogue for the magnetospheric activity. *Geophys. Res. Lett.* **1998**, *25*, 2397–2400.
115. Chapman, S.; Watkins, N. Avalanching and self organised criticality. A paradigm for geomagnetic activity? *Space Sci. Rev.* **2001**, *95*, 293–307.
116. Consolini, G. Self-organized criticality: A new paradigm for the magnetotail dynamics. *Fractals* **2002**, *10,* 275–283.
117. Leubner, M.P.; Vörös, Z. A nonextensive entropy approach to solar wind intermittency. *Astrophys. J.* **2005a**, *618*, 547–555.
118. Daglis, I.A.; Baker, D.N.; Galperin, Y.; Kappenman, J.G.; Lanzerotti, L.J. Technological impacts of space storms: Outstanding issues. *EOS Trans. AGU* **2001**, *82*, 585–592.
119. Daglis, I.A.; Kozyra, J.; Kamide, Y.; Vassiliadis, D.; Sharma, A.; Liemohn, M.; Gonzalez, W.; Tsurutani, B.; Lu, G. Intense space storms: Critical issues and open disputes. *J. Geophys. Res.* **2003**, *108*, doi:10.1029/2002JA009722.
120. Daglis, I.A.; Balasis, G.; Ganushkina, N.; Metallinou, F.-A.; Palmroth, M.; Pirjola, R.; Tsagouri, I.A. Investigating dynamic coupling in geospace through the combined use of modeling, simulations and data analysis. *Acta Geophys.* **2009**, *57*, 141–157.
121. Baker, D.N.; Daglis, I.A. Radiation Belts and Ring Current. In *Space Weather: Physics and Effects*; Bothmer, V., Daglis, I.A., Eds.; Springer: Berlin, Germany, 2007; pp. 173–202.
122. Geomagnetic Equatorial Dst index Home Page. available online: http://wdc.kugi.kyoto-u.ac.jp/dst-dir/index.html (accessed on 1 August 2013).
123. Balasis, G.; Daglis, I.A.; Papadimitriou, C.; Anastasiadis, A.; Sandberg, I.; Eftaxias, K. Quantifying dynamical complexity of magnetic storms and solar flares via nonextensive tsallis entropy. *Entropy* **2011**, *13*, 1865–1881.
124. Uyeda, S.; Kamogawa, M.; Nagao, T. Earthquakes, Electromagnetic Signals of. In *Encyclopedia of Complexity and Systems Science*; Meyers, R.A., Ed.; Springer–Verlag: New York, NY, USA, 2009; pp. 2621–2635.
125. Uyeda, S.; Nagao, T.; Kamogawa, M. Short-term earthquake prediction: Current status of seismo-electromagnetics. *Tectonophysics* **2009**, *470*, 205–213.

126. Kalashnikov, A.D. Potentialities of magnetometric methods for the problem of earthquake forerunners. *Tr. Geofiz. Inst. Akad. Nauk. SSSR.* **1954**, *25*, 180–182.
127. Varotsos, P.; Kulhanek, O. (Eds.) Measurements and theoretical models of the Earth's electric field variations related to earthquakes. *Tectonophysics* **1993**, *224*, 1–88.
128. Varotsos, P.A. *The Physics of Seismic Electric Signals*; TerraPub: Tokyo, Japan, 2005.
129. Park, S.; Johnston, M.; Madden, T.: Morgan, F.; Morrison, H. Electromagnetic precursors to earthquakes in the ULF band: A review of observations and mechanisms. *Rev. Geophys.* **1993**, *31*, 117–132.
130. Hayakawa, M.; Ito, T.; Smirnova, N. Fractal analysis of ULF geomagnetic data associated with the Guam eartquake on 8 August 1993. *Geophys. Res. Lett.* **1999**, *26*, 2797–2800.
131. Hayakawa, M.; Itoh, T.; Hattori, K.; Yumoto, K. ULF electromagnetic precursors for an earthquake at Biak, Indonesia in 17 February 1996. *Geophys. Res. Lett.* **2000**, *27*, 1531–1534.
132. Hattori, K. ULF geomagnetic changes associated with large earthquakes. *Terr. Atmos. Ocean Sci.* **2004**, *15*, 329–360.
133. Eftaxias, K.; Kapiris, P.; Polygiannakis, J.; Peratzakis, A.; Kopanas, J.; Antonopoulos, G.; Rigas, D. Experience of short term earthquake precursors with VLF-VHF electromagnetic emissions. *Nat. Hazards Earth Syst. Sci.* **2003**, *3*, 217–228.
134. Gokhberg M.B.; Morgounov, V.A.; Yoshino, T.; Tomizawa, I. Experimental measurement of electromagnetic emissions possibly related to earthquakes in Japan. *J. Geophys. Res.* **1982**, *87*, 7824–7828.
135. Asada, T.; Baba, H.; Kawazoe, K.; Sugiura, M. An attempt to delineate very low frequency electromagnetic signals associated with earthquakes. *Earth Planets Space* **2001**, *53*, 55–62.
136. Contoyiannis, Y.F.; Potirakis, S.M.; Eftaxias, K. The Earth as a living planet: human-type diseases in the earthquake preparation process. *Nat. Hazards Earth Syst. Sci.* **2013**, *13*, 125–139.
137. Papadimitriou, C.; Kalimeri, M.; Eftaxias, K. Nonextensivity and universality in the earthquake preparation process. *Phys. Rev. E* **2008**, *77*, doi:10.1103/PhysRevE.77.036101.
138. Potirakis, S.M.; Minadakis, G.; Eftaxias, K. Relation between seismicity and pre-earthquake electromagnetic emissions in terms of energy, information and entropy content. *Nat. Hazards Earth Syst. Sci.* **2012**, *12*, 1179–1183.
139. *Seismo Electromagnetics: Lithosphere-Atmosphere-Ionosphere Coupling*; Hayakawa, M., Molchanov, O., Eds.; TerraPub: Tokyo, Japan, 2002.
140. Pulinets, S.; Boyarchuk, K. *Ionospheric Precursors of Earthquakes*; Springer: Berlin, Germany, 2005.
141. Kamagowa, M. Preseismic lithosphere-atmosphere-ionosphere coupling. *EOS Trans. AGU* **2006**, *87*, 417–424.
142. Shen, X.; Zhang, X.; Wang, L.; Chen, H.; Wu, Y.; Yuan, S.; Shen, J.; Zhao, S.; Qian, J.; Ding, J. The earthquake-related disturbances in ionosphere and project of the first China seismo-electromagnetic satellite. *Earthq. Sci.* **2011**, *24*, 639–650.
143. Parrot, M.; Achache, J.; Berthelier, J.J.; Blanc, E.; Deschamps, A.; Lefeuvre, F.; Menvielle, M.; Plantet, J.L.; Tarits, P.; Villain, J.P. High-frequency seismo-electromagnetic effects. *Phys. Earth Planet. Inter.* **1993**, *77*, 65–83.

144. Molchanov, O.A.; Hayakawa, M. *Seismo-Electromagnetics and Related Phenomena: History and Latest Results*; TerraPub: Tokyo, Japan, 2008.
145. Hayakawa, M.; Hobara, Y. Current status of seismo-electromagnetics for short-term earthquake prediction. *Geomat. Nat. Hazards Risk* **2010**, *1*, 115–155.
146. Mandelbrot, B.B. *The Fractal Geometry of Nature*; Freeman: New York, NY, USA, 1983.
147. Sornette, D. *Critical Phenomena in Natural Sciences: Chaos, Fractals, Selforganization and Disorder: Concepts and Tools*; Springer: Berlin, Germany, 2000.
148. Varotsos, P.; Sarlis, N.; Skordas, E.; Uyeda, S.; Kamogawa, M. Natural time analysis of critical phenomena. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 11361–11364.
149. Bak, P. *How Nature Works*; Springer: New York, NY, USA, 1996.
150. Rundle, J.B.; Holliday, J.R.; Turcotte, D. L.; Klein, W. Self-Organized Earthquakes. In *American Geophysical Union Fall Meeting Abstracts*; AGU: San Francisco, CA, USA, 2011; volume 1, p. 01.
151. Yoder, M.R.; Turcotte, D.L.; Rundle, J.B. Earthquakes: Complexity and Extreme Events. In *Extreme Events and Natural Hazards: The Complexity Perspective, Geophysical Monograph*; Sharma, A.S., Bunde, A., Dimri, V.P., Baker, D.N., Eds.; American Geophysical Union: Washington, D.C., WA, USA, 2012; Volume 196, pp. 17–26.
152. Main, I.G.; Naylor, M. Entropy production and self-organized (sub)criticality in earthquake dynamics, *Phil. Trans. R. Soc. A* **2010**, *368*, 131–144.
153. Potirakis, S.M.; Minadakis, G.; Nomicos, C.; Eftaxias, K. A multidisciplinary analysis for traces of the last state of earthquake generation in preseismic electromagnetic emissions. *Nat. Hazards Earth Syst. Sci.* **2011**, *11*, 2859–2879.
154. Eftaxias, K. Are There Pre-Seismic Electromagnetic Precursors? A Multidisciplinary Approach. In *Earthquake Research and Analysis—Statistical Studies, Observations and Planning*; D'Amico, S., Ed.; InTech: Rijeka, Croatia, 2012; doi: 10.5772/28069.
155. Potirakis, S.M.; Karadimitrakis, A.; Eftaxias, K. Natural time analysis of critical phenomena: The case of pre-fracture electromagnetic emissions. *Chaos* **2013**, *23*, doi: 10.1063/1.4807908.
156. Potirakis, S.M.; Minadakis, G.; Eftaxias, K. Sudden drop of fractal dimension of electromagnetic emissions recorded prior to significant earthquake. *Nat. Hazards* **2012**, *64*, 641–650.
157. Minadakis, G.; Potirakis, S.M.; Nomicos, C.; Eftaxias, K. Linking electromagnetic precursors with earthquake dynamics: An approach based on nonextensive fragment and self-affine asperity models. *Physica A* **2012**, *391*, 2232–2244.
158. Sarlis, N.V.; Skordas, E.S.; Varotsos, P.A.; Nagao, T.; Kamogawa, M.; Tanaka, H.; Uyeda, S. Minimum of the order parameter fluctuations of seismicity before major earthquakes in Japan. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 13734–13738.
159. Ida, Y.; Yang, D.; Li, Q.; Sun, H.; Hayakawa, M. Fractal analysis of ULF electromagnetic emissions in possible association with earthquakes in China. *Nonlinear Processes Geophys.* **2012**, *19*, 577–583.
160. Abe, S.; Herrmann, H.; Quarati, P.; Rapisarda, A.; Tsallis, C.. Complexity, Metastability, and Nonextensivity. In Proceedings of the AIP Conference, Zaragoza, Spain, 18–20 April 2007; Volume 95, American Institute of Physics: College Park, MD, USA, 2007.

161. Rundle, J.; Turcotte, D.; Shcherbakov, R.; Klein, W.; Sammis, C. Statistical physics approach to understanding the multiscale dynamics of earthquake fault systems. *Rev. Geophys.* **2003**, *41*, doi:10.1029/2003RG000135.
162. Eftaxias, K.; Kapiris, P.; Polygiannakis, J.; Bogris, N.; Kopanas, J.; Antonopoulos, G.; Peratzakis, A.; Hadjicontis, V. Signatures of pending earthquake from electromagnetic anomalies. *Geophys. Res. Lett.* **2001**, *28*, 3321–3324.
163. Kapiris, P.; Balasis, G.; Kopanas, J.; Antonopoulos, G.; Peratzakis, A.; Eftaxias, K. Scaling similarities of multiple fracturing of solid materials. *Nonlinear Processes Geophys.* **2004**, *11*, 137–151.
164. Contoyiannis, Y.; Kapiris, P.; Eftaxias, K. Monitoring of a preseismic phase from its electromagnetic precursors. *Phys. Rev. E* **2005**, *71*, doi:10.1103/PhysRevE.71.066123.
165. Karamanos, K.; Dakopoulos, D.; Aloupis, K.; Peratzakis, A.; Athanasopoulou, L.; Nikolopoulos, S.; Kapiris, P.; Eftaxias, K. Pre-seismic electromagnetic signals in terms of complexity. *Phys. Rev. E* **2006**, *74*, doi:10.1103/PhysRevE.74.016104.
166. Eftaxias, K.; Sgrigna, V.; Chelidze, T. Mechanical and electromagnetic phenomena accompanying preseismic deformation: From laboratory to geophysical scale. *Tectonophysics* **2007**, *431*, 1–301.
167. Paluš, M.; Novotná, D. Testing for nonlinearity in weather records. *Phys. Lett. A* **1994**, *193*, 67–74.
168. Von Bloh, W.; Romano, M.C.; Thiel, M. Long-term predictability of mean daily temperature data. *Nonlinear Processes Geophys.* **2005**, *12*, 471–479.
169. Paluš, M.; Hartman, D.; Hlinka, J.; Vejmelka, M. Discerning connectivity from dynamics in climate networks. *Nonlinear Processes Geophys.* **2011**, *18*, 751–763.
170. Bunde, A.; Havlin, S. Power-law persistence in the atmosphere and the oceans. *Physica A* **2002**, *314*, 15–24.
171. Vjushin, D.; Govindan, R.B.; Brenner, S.; Bunde, A.; Havlin, S.; Schellnhuber, H.-J. Lack of scaling in global climate models. *J. Phys. Condens. Matter* **2002**, *14*, 2275–2282.
172. Govindan, R.B.; Vjushin, D.; Bunde, A.; Brenner, S.; Havlin, S.; Schellnhuber, H.-J. Global climate models violate scaling of the observed atmospheric variability. *Phys. Rev. Lett.* **2002**, *89*, doi:10.1103/PhysRevLett.89.028501.
173. Fraedrich, K.; Blender, R. Scaling of atmosphere and ocean temperature correlations in observations and climate models. *Phys. Rev. Lett.* **2003**, *90*, doi:10.1103/PhysRevLett.90.108501.
174. Ausloos, M.; Petroni, F. Tsallis non-extensive statistical mechanics of El Niño southern oscillation index. *Physica A* **2007**, *373*, 721–736.
175. Tsallis, C. Dynamical scenario for nonextensive statistical mechanics. *Physica A* **2004**, *340*, 1–10.
176. Petroni, F.; Ausloos, M. High frequency (daily) data analysis of the Southern Oscillation Index: Tsallis nonextensive statistical mechanics approach. *Eur. Phys. J. Spec. Top.* **2011**, *143*, 201–208.
177. Ferri, G.L.; Reynoso Savio, M.F.; Plastino, A. Tsallis' q-triplet and the ozone layer. *Physica A* **2010**, *389*, 1829–1833.
178. Pavlos, G.P.; Xenakis, M.N.; Karakatsanis, L.P.; Iliopoulos, A.C.; Pavlos, A.E.G.; Sarafopoulos, D.V. University of Tsallis non-extensive statistics and fractal dynamics for complex systems. *Chaotic Model. Simul.* **2012**, *2*, 395–447.

179. Zemp, D. The Complexity of the Fraction of Absorbed Photosynthetically Active Radiation on a Global Scale. Diploma Thesis, Norwegian Forest and Landscape Institute, Ås and French National College of Agricultural Sciences and Engineering, Toulouse, France, 2012.
180. Carpi, L.C.; Rosso, O.A.; Saco, P.M.; Gómez Ravetti, M. Analyzing complex networks evolution through Information Theory quantifiers. *Phys. Lett. A* **2011**, *375*, 801–804.
181. Saco, P.M.; Carpi, L.C.; Figliola, A.; Serrano, E.; Rosso, O.A. Entropy analysis of the dynamics of El Niño/Southern Oscillation during the Holocene. *Physica A* **2010**, *389*, 5022–5027.
182. Mayewski, P.A.; Rohling, E.E.; Stager, J.C.; Karlén, W.; Maasch, K.A.; Meeker, L.D.; Meyerson, E.A.; Gasse, F.; van Kreveld, S.; Holmgren, K.; *et al*. Holocene climate variability. *Quatern. Res.* **2004**, *63*, 243–255.
183. Ferri, G.L.; Figliola, A.; Rosso, O.A. Tsallis' statistics in the variability of El Niño/Southern Oscillation during the Holocene epoch. *Physica A* **2012**, *391*, 2154–2162.
184. Gonzalez, J.J.; de Faria, E.L.; Albuquerque, M.P.; Albuquerque, M.P. Nonadditive Tsallis entropy applied to the Earth's climate. *Physica A* **2011**, *390*, 587–594.
185. Marwan, N.; Trauth, M.H.; Vuille, M,; Kurths, J. Comparing modern and Pleistocene ENSO-like influences in NW Argentina using nonlinear time series analysis methods. *Clim. Dyn.* **2003**, *21*, 317–326.
186. Trauth, M.H.; Bookhagen, B.; Marwan, N.; Strecker, M.R. Multiple landslide clusters record Quaternary climate changes in the northwestern Argentine Andes. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **2003**, *194*, 109–121.
187. Donges, J.F.; Donner, R.V.; Rehfeld, K.; Marwan, N.; Trauth, M.H.; Kurths, J. Identification of dynamical transitions in marine palaeoclimate records by recurrence network analysis. *Nonlinear Processes Geophys.* **2011**, *18*, 545–562.
188. Donges, J.F.; Donner, R.V.; Trauth, M.H.; Marwan, N.; Schellnhuber, H.-J.; Kurths, J. Nonlinear detection of paleoclimate-variability transitions possibly related to human evolution. *Proc. Natl. Acad. Sci. USA* **2011**, 20422–20427.
189. Malik, N.; Zou, Y.; Marwan, N.; Kurths, J. Dynamical regimes and transitions in Plio-Pleistocene Asian monsoon. *Europhys. Lett.* **2012**, *97*, doi:10.1209/0295-5075/97/40009.
190. Witt, A.; Schumann, A.Y. Holocene climate variability on millennial scales recorded in Greenland ice cores. *Nonlinear Processes Geophys.* **2005**, *12*, 345–352.
191. Telford, R.J.; Heegaard, E.; Birks, H.J.B. All age-depth models are wrong: But how badly? *Quat. Sci. Rev.* **2004**, *23*, 1–5.
192. Rehfeld, K.; Marwan, N.; Heitzig, J.; Kurths, J. Comparison of correlation analysis techniques for irregularly sampled time series. *Nonlinear Processes Geophys.* **2011**, *18*, 389–404.
193. Brunsell, N.A.; Young, C.B. Land surface response to precipitation events using MODIS and NEXRAD data. *Int. J. Remote Sens.* **2008**, *29*, 1965–1982.
194. Brunsell, N.; Ham, J.; Owensby, C. Assessing the multi-resolution information content of remotely sensed variables and elevation for evapotranspiration in a tall-grass prairie environment. *Remote Sens. Environ.* **2008**, *112*, 2977–2987.

195. Stoy, P.C.; Williams, M.; Spadavecchia, L.; Bell, R.A.; Prieto-Blanco, A.; Evans, J.G.; Wijk, M.T. Using information theory to determine optimum pixel size and shape for ecological studies: Aggregating land surface characteristics in Arctic ecosystems. *Ecosystems* **2009**, *12*, 574–589.
196. Brunsell, N.A. A multiscale information theory approach to assess spatial-temporal variability of daily precipitation. *J. Hydrol.* **2010**, *385*, 165–172.
197. Brunsell, N.A.; Anderson, M.C. Characterizing the multi-scale spatial structure of remotely sensed evapotranspiration with information theory. *Biogeosciences* **2011**, *8*, 2269–2280.
198. Cochran, F.V.; Brunsell, N.A.; Mechem, D.B. Comparing surface and mid-tropospheric CO_2 concentraltions from Central U.S. grasslands. *Entropy* **2013**, *15*, 606–623.
199. Ruddell, B.L.; Kumar, P. Ecohydrological process networks: I. Identification. *Water Resour. Res.* **2009**, *45*, doi:10.1029/2008WR007279.
200. Ruddell, B.L.; Kumar, P. Ecohydrological process networks: II. Analysis and characterization. *Water Resour. Res.* **2009**, *45*, doi:10.1029/2008WR007280.
201. Kumar, P.; Ruddell, B.L. Information driven ecohydrological self-organization. *Entropy* **2010**, *12*, 2085–2096.
202. Rybski, D. Untersuchung von Korrelationen, Trends und synchronem Verhalten in Klimazeitreihen. Ph.D. Thesis, Justus Liebig University of Gießen, Gießen, Germany, 2006.
203. Yamasaki, K.; Gozolchiani, A.; Havlin, S. Climate networks based on phase synchronization analysis track El Niño. *Prog. Theor. Phys. Suppl.* **2009**, *179*, 178–188.
204. Malik, N.; Marwan, N.; Kurths, J. Spatial structures and directionalities in monsoonal precipitation over south Asia. *Nonlinear Processes Geophys.* **2010**, *17*, 371–381.
205. Malik, N.; Bookhagen, B.; Marwan, N.; Kurths, J. Analysis of spatial and temporal extreme monsoonal rainfall over South Asia using complex networks. *Clim. Dyn.* **2012**, *39*, 971–987.
206. Rybski, D.; Havlin, S.; Bunde, A. Phase synchronization in temperature and precipitation records. *Physica A* **2003**, *320*, 601–610.
207. Maraun, D.; Kurths, J. Epochs of phase coherence between El Niño/Southern Oscillation and Indian monsoon. *Geophys. Res. Lett.* **2005**, *32*, doi:10.1029/2005GL023225.
208. Donges, J.F.; Zou, Y.; Marwan, N.; Kurths, J. Complex networks in climate dynamics. Comparing linear and nonlinear network construction methods. *Eur. Phys. J. Spec. Top.* **2009**, *174*, 157–179.
209. Donges, J.F.; Zou, Y.; Marwan, N.; Kurths, J. The backbone of the climate network. *Europhys. Lett.* **2009**, *87*, doi:10.1209/0295-5075/87/48007.
210. Barreiro, M.; Marti, A.C.; Masoller, C. Inferring long memory processes in the climate network via ordinal pattern analysis. *Chaos* **2011**, *21*, doi:10.1063/1.3545273.
211. Deza, J.I.; Barreiro, M.; Masoller, C. Inferring interdependencies in climate networks constructed at inter-annual, intra-season and longer time scales. *Eur. Phys. J. Spec. Top.* **2013**, *222*, 511–523.
212. Hlinka, J.; Hartman, D.; Vejmelka, M.; Novotná, D.; Paluš, M. Non-linear dependence and teleconnections in climate data: Sources, relevance, nonstationarity. *Clim. Dyn.* **2013**, doi:10.1007/s00382-013-1780-2.
213. Runge, J.; Petoukhov, V.; Kurths, J. Quantifying the strength and delay of climatic interactions: The ambiguities of cross correlation and a novel measure based on graphical models. *J. Clim.* **2013**, doi: 10.1175/JCLI-D-13-00159.1.

214. Bjerknes, J. Atmospheric teleconnections from the equatorial pacific. *Mon. Weather Rev.* **1969**, *97*, 163–172.
215. Walker, G.T. Correlation in seasonal variations of weather, VIII: A preliminary study of world weather. *Mem. Ind. Meteorol. Dep.* **1923**, *24*, 75–131.
216. Walker, G.T. Correlation in seasonal variations of weather, IX: A further study of world weather. *Mem. Ind. Meteorol. Dep.* **1924**, *24*, 275–332.
217. Wang, C. Atmospheric circulation cells associated with the El-Niño-Southern Oscillation. *J. Clim.* **2002**, *15*, 399–419.
218. Held, H.; Kleinen, T. Detection of climate system bifurcations by degenerate fingerprinting. *Geophys. Res. Lett.* **2004**, *31*, doi:10.1029/2004GL020972.
219. Livina, V.N.; Lenton, T.M. A modified method for detecting incipient bifurcations in a dynamical system. *Geophys. Res. Lett.* **2007**, *34*, doi:10.1029/2006GL028672.
220. Dakos, V.; Scheffer, M.; van Nes, E.H.; Brovkin, V.; Petoukhov, V.; Held, H. Slowing down as an early warning signal for abrupt climate change. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 14308–14312.
221. Scheffer, M.; Bascompte, J.; Brock, W.A.; Brovkin, V.; Carpenter, S.R.; Dakos, V.; Held, H.; van Nes, E.H.; Rietkerk, M.; Sugihara, G. Early-warning signals for critical transitions. *Nature* **2009**, *461*, 53–59.
222. Kühn, C. A mathematical framework for critical transitions: bifurcations, fast-slow systems and stochastic dynamics. *Physica D* **2011**, *240*, 1020–1035.
223. Stanley, H.E. Scaling, universality, and renormalization: Three pillars of modern critical phenomena. *Rev. Mod. Phys.* **1999**, *71*, S358–S366.
224. Stanley, H.E.; Amaral, L.; Gopikrishnan, P.; Ivanov, P.; Keitt, T.; Plerou, V. Scale invariance and universality: Organizing principles in complex systems. *Physica A* **2000**, *281*, 60–68.
225. Sornette, D.; Helmstetter, A. Occurrence of finite-time singularities in epidemic models of rupture, earthquakes, and starquakes. *Phys. Rev. Lett.* **2002**, *89*, doi: 10.1103/PhysRevLett.89.158501.
226. Sornette, D. Predictability of catastrophic events: Material rupture, earthquakes, turbulence, financial crashes and human birth. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 2522–2529.
227. Kossobokov, V.; Keillis-Borok, V.; Cheng, B. Similarities of multiple fracturing on a neutron star and on Earth. *Phys. Rev. E* **2000**, *61*, 3529–3533.
228. De Arcangelis, L.; Godano, C.; Lippiello, E.; Nicodemi, M. Universality in solar flare and earthquake occurrence. *Phys. Rev. Lett.* **2006**, *96*, doi:10.1103/PhysRevLett.96.051102.
229. Balasis, G.; Daglis, I.A.; Anastasiadis, A.; Papadimitriou, C.; Mandea, M.; Eftaxias, K. Universality in solar flare, magnetic storm and earthquake dynamics using Tsallis statistical mechanics. *Physica A* **2011**, *390*, 341–346.
230. Balasis, G.; Papadimitriou, C.; Daglis, I.A.; Anastasiadis, A.; Sandberg I.; Eftaxias, K. Similarities between extreme events in the solar-terrestrial system by means of nonextensivity. *Nonlinear Processes Geophys.* **2011**, *18*, 563–572.
231. Balasis, G.; Papadimitriou, C.; Daglis, I.A.; Anastasiadis, A.; Athanasopoulou, L.; Eftaxias, K. Signatures of discrete scale invariance in Dst time series. *Geophys. Res. Lett.* **2011**, *38*, doi:10.1029/2011GL048019.

232. Zscheischler, J.; Mahecha, M.D.; Harmeling, S.; Reichstein, M. Detecting and attribution of large spatiotemporal extreme events in Earth observation data. *Ecol. Inform.* **2013**, *15*, 66–73.
233. De Michelis, P.; Consolini, G.; Materassi, M.; Tozzi, R. An information theory approach to the storm-substorm relationship. *J. Geophys. Res.* **2011**, *116*, doi:10.1029/2011JA016535.
234. Jimenez, A. A complex network model for seismicity based on mutual information. *Physica A* **2013**, *10*, 2498–2506, doi: 10.1016/j.physa.2013.01.062.
235. Lenton, T.M.; Held, H.; Kriegler, E.; Hall, J.W.; Lucht, W.; Rahmstorf, S.; Schellnhuber, H.J. Tipping elements in the Earth's climate system. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 1786–1793.
236. Lenton, T.M. Early warning of climate tipping points. *Nat. Clim. Chang.* **2011**, *1*, 201–209.