Order and disorder in ecological time-series: Introducing normalized spectral entropy

Nicola Zaccarelli a,*, Bai-Lian Li b, Irene Petrosillo a, Giovanni Zurlini a

a Landscape Ecology Laboratory, Department of Biological and Environmental Sciences and Technologies–Ecotekne, University of Salento, Prov.le Lecce-Monteroni, 73100 Lecce, Italy
b Ecological Complexity and Modelling Laboratory, Department of Botany and Plant Sciences, University of California, Riverside, CA 92521, USA

A R T I C L E   I N F O

Article history:
Received 8 February 2011
Received in revised form 30 June 2011
Accepted 10 July 2011

Keywords:
Complexity
Spectral analysis
Information theory
Evenness

A B S T R A C T

Information theory and entropy measures have been extensively applied in ecology in different areas like biodiversity assessment, evolution, species interactions, spatial dynamics or landscape analysis. Ecological applications of entropy measures have been primarily focused on structural and functional complexity of systems and less attention has been paid to temporal evolution and dynamics. The aim of this paper is to present “normalized spectral entropy” (Hs), an entropy related index able to measure part of the structural complexity of an ecological time series. Hs quantifies the degree of order and predictability derived by the series’ power spectrum. The index sensitivity to data attributes is investigated by means of time-series surrogates of known properties (i.e., time-series length, power spectrum shape, and time-series values distribution). A procedure to calculate confidence intervals is outlined as a preliminary statistical tool to assess differences among values. Three examples of possible application are described using time series of meteorological variables, vegetation physiological responses and remote sensing images. Results show how Hs is able to contribute to the ongoing debate on how to estimate spatio-temporal complexity of ecological systems, thus making a step forward in the proposed use of complexity as an ecological orientor.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

A major endeavor of ecology is to understand the complexity of ecological patterns and dynamics at different scales, ultimately leading to the comprehension of properties of natural or social-ecological systems (Levin, 1998; Gunderson and Holling, 2001; Proulx, 2007). Addressing such a task requires the integration of knowledge and theories of different scientific fields, moving from physical to social sciences. This need has stimulated the development of tools to describe and quantify several systems’ properties (like resilience, adaptability or self-organization) and system’s behaviors by integrating hierarchies, non-equilibrium dynamics, feedbacks and cross-scale interactions (Berkes et al., 2003; Solé and Bascompte, 2006; Harris, 2007).

Different aspects of the complexity of a system have been addressed such as structural and functional organization or dynamics types (Adami, 2002). Moreover, a variety of definitions of complexity exist (Anand et al., 2010), demonstrating the extent to which this field is an emerging and quickly evolving discipline. It is not surprising that several mathematical approaches and a variety of indices have been developed so far. Among others, statistical mechanics is applied to evaluate species abundance patterns (Dewar and Porté, 2008), graph theory is used in foodwebs and energy networks (Bascompte, 2009; Ings et al., 2009), thermodynamics allows the quantification of energy fluxes and efficiency in system development (Jørgensen, 2009) or fractal analysis assists in the description of landscape mosaic patterns (Zurlini et al., 2006). Green et al. (2005) and Parrott (2010) provide a far more complete review of spatial and temporal methods for measuring ecological complexity. They acknowledge how, among all the fields of science, the field of the information theory and entropy related indices have provided insights for complex systems analysis.

Entropy and information theory have been extensively applied in ecology (Ulanowicz, 2001) in different areas like biodiversity assessment (Magurran, 2003), evolution (Avery, 2003), species interactions and spatial dynamics (Chen et al., 2005; Parrott, 2005) or landscape description (Ekström, 2003). Since the seminal work of MacArthur (1955) exploring Shannon information entropy (H; Shannon, 1948) to describe species diversity in ecological communities, a wide variety of indices have been used such as, among others, Renyi’s generalized entropy (Pielou, 1975), Tsallis non-extensive entropy (Tsallis, 2002), or quadratic entropy (Izsák and Papp, 2000). The relevance of these indices stems from the argument that provided the entropy of a community is considered to
be a measure of uncertainty in the relative abundances of species, the maximum level of uncertainty is then reached when entropy is at its highest. This idea is one of the cornerstones of the diversity theory (Orlóci, 1991) and it provides the foundation for the ongoing debate on the evaluation of biodiversity at different scales in terms of structural (i.e., α-, β- and γ-diversity; Meffe et al., 2002) and functional complexity (i.e., functional traits of species; Mason et al., 2005).

In the context of landscape ecology, entropy-based landscape indices like Shannon’s H or contagion (Li and Reynolds, 1993) are among the most commonly used metrics to represent landscape composition and configuration diversity (Turner et al., 2001). Changes in landscape diversity indexes are hypothesized to reflect changes in the level of human impacts and disturbance regimes (Johnson et al., 2001; Bogaert et al., 2005), species diversity and habitat use (Wagner et al., 2000; Hrabik et al., 2005), and more recently have been applied as surrogates for direct biodiversity estimates using remotely sensed images (Rocchini et al., 2005).

Ecological applications of entropy measures have been primarily focused on systems structural and functional complexity from a spatial or organizational point of view, while less attention has been paid to temporal evolution and dynamics (Parrott, 2010). Research in time-series analysis of ecological indicators has been focused on different aspects of temporal complexity such as the identification of changes and discontinuities using principal component analysis or scale-dependent correlation analysis (Jassby and Powell, 1990; Rodríguez-Arias and Rodó, 2004), the description of regularity and periodic variability by means of the spectral analysis and autoregressive moving average models (Sabo and Post, 2008; Ives et al., 2010), or the recognition of chaotic behaviors and underlying attractor’s properties by estimating Lyapunov exponents (Ellner and Turchin, 1995). Nonetheless, only a few entropy-related measures have been proposed to exploit the information content and assess the regularity of a series (Parrott, 2010). The mean information gain index was developed to address the problem of accounting for the increase in information obtained by extending a signal (Wackerbauer et al., 1994). The fluctuation complexity index was created to evaluate the patterns and the structure of a series in terms of the ordering of, and the relationship between adjacent values (Bates and Shepard, 1993). Both indices suffer from a loss of information due to pre-treatment and coding of data, the need of long series for meaningful conclusions, and they do not exploit the frequency domain of a series.

The aim of this paper is to introduce an index called “normalized spectral entropy” as a tool to measure part of the structural complexity of an ecological time series. The normalized spectral entropy is an entropy-related index able to describe the degree of order and predictability (i.e., regularity) within a time-series of observations based on its power spectrum. The index sensitivity to data attributes is investigated by means of time-series surrogates of known properties (i.e., time-series length, power spectrum shape, and time-series value distribution). A procedure to calculate confidence intervals is outlined as a preliminary statistical framework to assess differences among values, and three examples of applications are described using real datasets.

2. Materials and methods

2.1. Normalized spectral entropy

Spectral entropy has been used as an entropy-like measure to characterize relative order in time dynamics of a variety of physical and biological systems. It has been applied to investigate coherent structure creation in turbulent fluids, Brownian particles or nonlinear non-equilibrium (Crepeau and Isaacson, 1991), to characterize depth of sedation (Ferenets et al., 2006) or as an indicator of the degree of temporal self-organization of ecosystems in a landscape (Li, 2000). Spectral entropy \( H_S \) is defined as:

\[
H_S = -\sum_{k=1}^{N} p_k \ln(p_k)
\]

where \( k \) represents the frequency \( \lambda_k \) and \( p_k \) is given by:

\[
p_k = \frac{|\lambda_k|^2}{\sum_i{|\lambda_i|^2}}
\]

The term \( |\lambda_k|^2 \) is the Fourier power spectrum of a time-series at frequency \( \lambda_k \) (Legendre and Legendre, 2000). A power spectrum is the description of the importance of the different frequencies necessary to reconstruct a signal based on a series of sine and cosine functions (i.e., a Fourier transformation). \( H_S \) is the Shannon information entropy applied to the power spectrum, where each particular state is described by the relative contribution of the frequency \( \lambda_k \) to the overall power. Usually spectral entropy is normalized (\( H_S \)) to the range of values between 0 and 1, by dividing for \( \ln(N) \), where \( N \) is the number of frequencies derived by the Fourier transform and equal to half the length of the time series. The term \( \ln(N) \) is the theoretical maximum value of entropy when all examined cases have the same value. In this form spectral entropy is mathematically equivalent to the measure of disorder proposed by Landsberg (1986) or the evenness index introduced by Piérou (1975).

Fig. 1 provides an illustrative example of how the index is calculated in two cases of the same dataset generated from a cosine wave function. \( H_S \) can clearly differentiate the two signals having very different levels of what can intuitively be considered order and regularity in the sequence (see the on-line Appendix A for more details and code).

2.2. Surrogates generation

The generation of reference time series of predefined power spectrum and probability density function (PDF) of data values, the so-called surrogates, is widely used in time-series analysis (Schreiber and Schmitz, 2000). We used surrogates to estimate the dependence of normalized spectral entropy on signal properties and length, by using a method which iteratively corrects deviations in spectrum and distribution from a goal set by specific constraints. A detailed description and a mathematical explanation of the algorithm can be found in Schreiber and Schmitz (1996) and Ferenets et al. (2006).

Briefly, the iterative procedure starts by defining a power spectrum \( |\lambda_k|^2 \) of interest and by generating a random signal \( s \) of length \( N_1 \) of a specific PDF, along with a copy \( \mathcal{C} \) of \( s \) sorted in ascending order. At each iteration \( i \) two steps are performed. In step one, the discrete Fourier transform of \( s(0) \), \( \text{FFT}(0) \), is obtained. The amplitude \( |\text{FFT}(0)| \) is replaced by the desired one (i.e., with \( \sqrt{|\lambda_k|^2} \)), and the complex phase of \( \text{FFT}(0) \) is kept. In this way, the inverse Fourier transform is obtained to create a new signal \( s(0) \). In step two, a new sequence \( s^{(i+1)} \) is constructed by replacing each \( s(0) \) with \( j \) from 1 to \( N_1 \) with the value in \( \mathcal{C} \) having the same rank order. Finally, a relative error, \( \varepsilon(0) \), between the desired power spectrum, and the obtained one for \( s(0) \) is computed as:

\[
\varepsilon(0) = \frac{\sum (|\text{FFT}(0)|^2 - |\lambda_k|^2)^2}{\sum (|\lambda_k|^2)^2}
\]

In the first step the desired power spectrum is enforced, but the probability distribution is altered. In the second step the probability
distribution is guaranteed, but the power spectrum is modified. The procedure can be iterated till certain criteria are matched, i.e. a maximum number of cycles is reached or $a^0$ is lower than a certain threshold. We set a maximum number of iterations equal to 100 and an error threshold of $10^{-3}$.

To evaluate the sensitivity of $H_0$, a normal probability distribution with zero mean was chosen as a PDF to generate the starting random signals. The probability standard deviation ($\sigma$) was selected to have the following values: 0.01, 0.05, 0.1, 0.5, 1, 5, and 10. The power spectrum was selected to have a rectangular shape (i.e., equal value for all frequencies), starting from zero, within a rescaled frequency range from 0 to 0.5, corresponding to the average signal power (i.e., the DC component) and the Nyquist frequency, respectively. The frequency interval covered by the different rectangular spectra ($\nu$) was set to have the following values: 0.01, 0.025, 0.05, 0.1, 0.25, 0.5, and 1. For example, the value of 0.05 means that the bandwidth of the corresponding signal covers 5% of the whole bandwidth.

The process of surrogates generation was performed for four different signal lengths (L) of 64, 128, 256, and 1024 data points. No specific sampling frequency or time interval is assumed. Thus, it is possible to evaluate potential scale dependency effects by considering an equal grain (i.e., sampling frequency) among different signal lengths, or alternatively fixing the signal extent (i.e., total time length). For each combination of parameters (i.e., $L$, $\sigma$ and $\nu$) 1000 surrogates were derived. For the first two shortest lengths of signals the lower values of $\nu$ were not calculated due to the lack of frequencies that could be generated. For the surrogates generation we develop specific procedures using the R statistical language (R Development Core Team, 2009).

2.3. Statistical procedure for confidence interval estimate

Statistical analysis of the entropy measures has received considerable attention since the development of the concept using, for instance, random ordering coupled with distribution probability models to estimate the diversity in a large community (for a review consider Magurran, 2003) or bootstrap resampling of error matrix for composition-based landscape indexes (Hess and Bay, 1997). Esteban and Morales (1995) developed a mathematical expression using which most entropy measures can be obtained. They also proposed formulas to test statistical hypotheses or to build confidence intervals starting from simple random samples of independently data. However, the assumption of independently data can break down in practice either because data are not independent (i.e., temporal correlation) or because data are not identically distributed, or both. In order to overcome the aforementioned limitations we propose an algorithm based on the local bootstrap of the power spectral density function and the use of bias-corrected and accelerated procedure (BCa) to calculate confidence intervals which are median unbiased and adjusted for skewness (DiCiccio and Efron, 1996).

Paparoditis and Politis (1999) introduced the idea of a "local" bootstrap as a modification of the ordinary bootstrap scheme of "global" resampling of a time series, i.e. of resampling from the whole set of observations. The basic principle underlying the procedure lies in the fact that given a power spectrum $|\lambda_k|^2$ obtained by a Fourier transform of a time series $X_1,X_2,\ldots,X_T$ over a set of m frequency $\lambda_k$ in (0, 0.5) the random vector $\langle|\lambda_1|^2, |\lambda_2|^2, \ldots, |\lambda_m|^2\rangle$ is asymptotically distributed as a vector of independent exponential variables and the mean value for $|\lambda_k|^2$ is the spectral density value for the frequency. Assuming that the spectral density $f(\lambda)$ is a smooth function of $\lambda$, it is expected that the sampling behavior of the power spectrum at any particular frequency $\lambda_k$ will be very similar for power spectrum ordinates corresponding to frequencies in a small neighborhood of $\lambda_k$. So it is possible to derive replicates at frequency $\lambda_k$ by choosing locally with replacement between ordinates of frequencies "near" to the one of interest. Once a width $m_w$ of a frequency window defining the neighborhood of $\lambda_k$ is selected, a bootstrapped replicate of the original power spectrum can be derived by resampling locally with replacement all along the frequencies obtained by a discrete Fourier transform of the original signal. For a detailed description and mathematical proofs see Paparoditis and Politis (1999).

For the resampling and related calculations we develop a function using the R statistical language (R Development Core Team, 2009) and the TSA package (Cryer and Chan, 2008) available in the on-line Appendix A. All BCa confidence intervals were derived using 5000 replicates, a starting two tails $\alpha$ probability value of 0.05 and setting $m_w$ equal to 3 (Paparoditis and Politis, 1999). Spectral
entropy was measured after power spectra were smoothed using a Daniell filter with a kernel dimension of two and after linear detrending.

2.4. Application to real time series

Three examples of applications of HSn are presented using real data to demonstrate how to use the index for the evaluation and comparison of the degree of order of real time series.

In the first example we compare the cumulative precipitation (CP) and the absolute maximum temperature (Tmax) between two meteorological stations (i.e., “Lecce” 18° 9'E, 40° 13.8'N and “S. Maria di Leuca” 18° 21'E, 39° 49.2'N) belonging to the same climatic region (i.e., Mediterranean climate) but 45 kilometers apart within the Province of Lecce (South Italy). The time series are composed of 2,124 values measured on a ten days basis starting from January 1951 to December 2009 extracted from the Italian National Environmental Information System (http://www.sciainet.apat.it/). The two meteorological variables could be used as the extreme events indicators and may be related to effects of climate change.

In the second example we focus on detecting possible changes in HSn within the same time series. Such changes could be related to a modification in number, strength or role of the controlling set of parameters and constraints acting on a system. Three different chunks of 576 measures of ozone concentration ([O3] in ppb) and total ozone flux (F2O3 in nmol m−2 s−1) were compiled by extracting hourly mean values from the 7 to the 30 of the month of June, July and August from the flux measurements carried out from the 6 June to the 8 December 2004 at the Castelporziano Natural State Reserve (41° 44’ N, 12° 25’ E; Rome, Italy), within a stand of typical Mediterranean vegetation dominated by oak (Quercus ilex). The original data set has been extensively analyzed by Gerosa et al. (2009) showing how ozone flux time patterns change from June to December during the exceptionally dry year of 2004.

In the third example we demonstrate a way to describe spatial patterns of spectral entropy. A map of HSn values is obtained by processing a series of 161 remote sensed images, from January 2003 to December 2005, acquired by the Moderate Imaging Reflectance Spectrometer sensor (MODIS; Parkinson, 2003) downloaded from the FluxNet project web-site (http://www.fluxnet.org/) for the Castelporziano Natural State Reserve site. As an ecological variable we used the Normalized Difference Vegetation Index (NDVI; Pettorelli et al., 2005) provided by images with a spatial grain of 250 meters on a 16 days basis. NDVI has been validated as a robust indicator of vegetation photosynthesis, with built-in relationships to social-ecological processes such as habitat conversion or crop rotation, and related to a number of vegetation indicators and characteristics (see Young and Harris, 2005, and references therein). A land-cover map for the area has been used to support the interpretation of the HSn results.

3. Results

3.1. Surrogates analysis

The surrogate analysis demonstrates three major results (Fig. 2). First, HSn does not depend on the amplitude distribution of signal values (σ) as showed by lines that are parallel to the x-axis. Such lack of dependence is not surprising as the mathematical formula defining the index does not include the phase information of the Fourier transform for the calculations. Second, HSn increases as the frequency interval portion of the rectangular power spectrum (ν) is augmented. The relation is far from linear for values of ν lower than 0.5, to approximately change to linear when the power of the time series is progressively more distributed among Fourier frequencies.

Third, surrogates length (L) is strongly affecting the index behavior by inflating HSn because the signal dimension is increased. Despite its normalized form, such scale dependency is especially evident for the lower ν values. For ν equal to 0.1, spectral entropy starts from 0.446 at L equal to 64 data points, to peak at 0.623 when L is equal to 1024. The difference between the shortest and longest surrogate is of 0.177, with a percentage increase of 39.7. For ν equal to 0.25, HSn is augmented of 0.11, from 0.657 to 0.765 with an increase of the 16.4%. For ν equal to 0.50, HSn changes from 0.824 to 0.874 with a 6.1% variation.

3.2. Meteorological data

Power spectra of the cumulative precipitation (CP) and the absolute maximum temperature (Tmax) show distinctive temporal patterns for the two meteorological stations in the Lecce Province (Fig. 3). CP reflects a more disordered behavior with one clear peak at 0.029 (i.e., a one year cycle) and the remaining signal variance distributed randomly around the mean power of 770 for “Lecce” and of 790 for “S. Maria di Leuca” like in a white noise signal. Tmax demonstrates a strong cyclic pattern with the signal variance concentrated in two significant peaks at 0.029 and 0.056 (i.e., a one year and a half year cycle respectively). Other frequencies contribute for a very small amount.

Spectral entropy helps to synthesize such differences in temporal patterns by showing values near one for CP (i.e., white noise) and around 0.450 for Tmax (Table 1). When comparing Bca confidence intervals it is clear that the two stations show no differences for entropy estimates for absolute maximum temperatures as opposed to cumulative precipitations which do not exhibit overlapping ranges. For all the time series, confidence interval limits reflect bootstrapped index distribution with a pronounced skewed shape, as lower and upper probability bounds are far from the desired 0.025 and 0.975 probability values (Table 1).

Differences in HSn between stations are the result of the actual geographic location of the sites. The “S. Maria di Leuca” station is less than one kilometer from the coast in the most southern point of the Lecce Province. It is more exposed to winds and clouds coming from the sea. The “Lecce” station is more than 20 kilometers inland in a less windy area. In both cases solar irradiance, which is affecting Tmax, is identical.

3.3. Ozone case study

By comparing different subsequent months for the ozone concentration ([O3]) and total ozone flux (F2O3) time series it is possible to highlight differences in frequency patterns and in HSn values. Ozone concentration shows three characteristics peaks for August and July, and two for June, around which the signal variance in

<table>
<thead>
<tr>
<th>Time series</th>
<th>HSn</th>
<th>Bca confidence intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td>CP “Lecce”</td>
<td>0.963</td>
<td>0.953–0.973</td>
</tr>
<tr>
<td>CP “S Maria di Leuca”</td>
<td>0.935</td>
<td>0.912–0.952</td>
</tr>
<tr>
<td>Tmax “Lecce”</td>
<td>0.448</td>
<td>0.443–0.503</td>
</tr>
<tr>
<td>Tmax “S Maria di Leuca”</td>
<td>0.451</td>
<td>0.444–0.480</td>
</tr>
<tr>
<td>Oi1 June</td>
<td>0.609</td>
<td>0.590–0.665</td>
</tr>
<tr>
<td>Oi1 July</td>
<td>0.573</td>
<td>0.555–0.612</td>
</tr>
<tr>
<td>Oi1 August</td>
<td>0.526</td>
<td>0.510–0.550</td>
</tr>
<tr>
<td>FtO2 June</td>
<td>0.660</td>
<td>0.630–0.738</td>
</tr>
<tr>
<td>FtO2 July</td>
<td>0.678</td>
<td>0.643–0.708</td>
</tr>
<tr>
<td>FtO2 August</td>
<td>0.652</td>
<td>0.621–0.723</td>
</tr>
</tbody>
</table>

Fig. 2. Normalized spectral entropy ($H_{sn}$) of time-series surrogates of different lengths ($L$) as function of the frequency interval ($v$; left) and of the standard deviation ($\sigma$; right) of the PDF generating the data values for some frequency intervals.

Fig. 3. Power spectra of cumulative precipitation (CP) and absolute maximum temperature ($T_{\text{max}}$) for the meteorological station of Lecce (black line) and of “S. Maria di Leuca” (grey line) for the ten days based time series from 1951 to 2009.

Fig. 4. Power spectra of ozone concentration ([O$_3$] in ppb) and ozone total flux ($F$\text{O$_3$} in nmol m$^{-2}$ s$^{-1}$) for the 576 data points time series for June, July, and August 2004 for the Natural State Reserve of Castelporziano FluxNet tower site.
concentrated. Such peaks occur at 0.045, 0.085 and 0.130 corresponding to a 22 h, a 12 h and a 7.6 h cycle respectively (Fig. 4). For the remaining portion of the spectrum the monthly signals seem to be more diverse. Such general pattern is mimicked by the total ozone flux with the exception of the last peak, which is far less important and is present only in July.

Normalized spectral entropy shows medium level of order with values ranging from 0.526 to 0.678 (Table 1). The series of F:\O3 show remarkably similar values, with BCa confidence intervals strongly overlapping. The series of ozone concentration are different only for August, while June and July have partially overlapping confidence intervals. Likewise for the meteorological variables, confidence intervals reflect bootstrapped distributions with pronounced skewness and probability bounds far from the starting points of 0.025 and 0.975 (Table 1).

Differences between H\textsubscript{sn} for [O\textsubscript{3}] and for F:\O3 are linked to the interplay of temperature, precipitation trends, and physiological response of plant to water. It can be observed that the ozone concentration variability is more influenced by high mean and maximum temperature, and total precipitation amount of June and July. While more stable weather and dryer conditions in August have shaped ozone fluctuation into a more regular sequence (Gerosa et al., 2009). The ozone flux peaked in July due to higher water availability and it has been controlled by leaf wetness in dry months. This in turn has partially counteracted the plant stomatal conductance and uptake, and the decrease in air turbulence determined by a lower amount of available energy associated to more regular weather conditions (Gerosa et al., 2009).

3.4. Spatial analysis

The map of H\textsubscript{sn} for the MODIS NDVI time series of the Castelporzio site shows distinctive spatial patterns clearly linked to the different land-cover classes (Fig. 5). Higher H\textsubscript{sn} is associated to the coast and urban areas, with the four classes of “beaches and dunes”, “sclerophyllous vegetation”, “industrial and commercial units” and “continuous urban fabric” having a mean value of 0.782, 0.701, 0.687 and 0.676 respectively. Lower mean spectral entropy scores are measured for the class “non-irrigated arable land” (H\textsubscript{sn} of 0.640) and for the class “mixed forest” (H\textsubscript{sn} of 0.631). Such diversity can be explained by three main reasons. First, the main vegetation types of the area are characterized by differences in phenological cycles and abilities to cope with weather constraints such as water availability and maximum temperature (Gerosa et al., 2009). Oak forests have more regular cycles because of their higher buffer capacity compared to Mediterranean maquis or pastures due to their self-regulating mechanisms to cope with drought and higher temperature. Second, degrees of human disturbance and regulation are different among land-cover classes in terms of intensity, temporal patterns and spatial extent. Cycles in arable lands are under complete control of farmers (i.e., watering, ploughing and harvesting), and therefore are more regular, and less dependent on seasonality and weather variability. Third, NDVI is especially suited for vegetation monitoring. Its behavior is strongly influenced by the amount of water, bare soil or concrete and paved surfaces. Along the coast, sea waves contaminate the NDVI signal with an erratic component. In urban areas, NDVI performs purely recording changes in color, solar lighting or material types of buildings and streets (Pettorelli et al., 2005).

The map of H\textsubscript{sn} allows testing the differences among locations beside land-cover types. By comparing distances for each pixel among upper and lower bounds of all the BCa confidence intervals, we were able to identify a minimum significant distance. Two locations, adjacent or not, have different H\textsubscript{sn} values when their difference is bigger than 0.042. This minimum distance helps to assess the spatial degree of homogeneity in the order of the temporal trajectories of monitored processes among portions of the same classified patches. So locations can be clumped in groups of pixels reflecting differences in the entropy of NDVI time series, like within the patch of mixed forest containing the tower site.

4. Discussion

Normalized spectral entropy (H\textsubscript{sn}) has been introduced as an ecological index for the analysis of the temporal structural complexity of a system (i.e., order/disorder degree sensu Landsberg, 1986). When compared to existing temporal entropy-related measures proposed in the ecological literature (Parrott, 2010), H\textsubscript{sn} shows its unique properties in exploiting the frequency domain information to characterize the complexity of the patterns of a signal through the power spectrum. This can be interpreted as a decomposition of the total signal variance into independent frequency components by means of a Fourier transform (Legendre and Legendre, 2000). H\textsubscript{sn} is an example of what Atmanspacher (2007) defines as “type one” measure of complexity, as it is a function increasing monotonously from predictability to uncertainty or from order to disorder in the signal variability over frequencies. H\textsubscript{sn} quantifies the evenness of the diversity of time-series frequencies’ patterns described by the signal power spectrum, and in this respect it is intuitively equivalent to Pielou’s evenness index, so familiar to ecologists (Magurran, 2003). Low values of H\textsubscript{sn} describe a time series with few dominant frequencies, exhibiting periodic cycles and an ordered behavior, like in a community where few species are abundant and all others are rare. Such signal is deemed as more foreseeable or less uncertain because it allows more precise predictions of the near future evolutions. For a time series showing equal variance for all frequencies, like a community with evenly abundant species, H\textsubscript{sn} reaches its maximum of one. In this case any prediction on future results is more difficult because any outcome is equally possible in terms of frequencies.

The three case study presented in this paper exemplify possible applications of H\textsubscript{sn} as an indicator of order/predictability within the temporal patterns of an ecological variable. In the first and second example, the index is able to clearly synthesize and distinguish different time trajectories as well as supporting the identification of changes in regularity within the same signal. Coupled with the power spectrum, H\textsubscript{sn} helps to characterize types and scales of temporal dynamics for abiotic constrains shading lights on the variability of ecological responses, and, thus, possibly leading to a better understanding of disturbance regimes and system evolutions. When ecosystem’s function or process indicators are analyzed, H\textsubscript{sn} could be used as a holistic indicator for system level properties able to characterize heterogeneity in time and pointing to the system’s self-organization strength (Li, 2006). Any ecological system is able to create metastable dynamics through not only internal interactions or the development of dissipative structures, but also by the modification of external constraints and disturbances in a way that they would reinforce (or at least do not disrupt) internally generated order (Nicolis and Prigogine, 1977). Thus, by comparing the values of normalized spectral entropy between environmental inputs and the system’s responses it is possible to derive a first approximation of the strength of the system’s filter ability and self-organization capacity. This approach would be different from the use of mutual information which involves the comparison of the rhythms of two time-series to quantify their degree of dynamic cohesion (Cazelles, 2004).

In the third example we try to show how H\textsubscript{sn} can effectively enhance the analysis of a landscape. A land-cover map, like the one presented in Figure 5, is a static representation of a complex and dynamic system masking relevant ecological differences.
among patches. Some of these differences can be reviled by $H_{sn}$, such as changes in crop rotation schemes in arable lands which may affect productivity, soil resource extraction or pests’ dynamics (Tschamkette et al., 2005). Remote sensing is a primary source of information to study the complexity of social-ecological systems at the landscape level in terms of dynamics at multiple spatial and temporal scales as it provides consistent time series of integrated ecosystem measurements (Pettorelli et al., 2005). We believe that by describing the degree of order of a time series, spectral entropy, can be used in conjunction with Remote Sensing to support monitoring in regional areas for location and distribution of land-cover changes. This will assist in the characterization of the links between land-cover configuration and composition with ecological processes such as primary production or carbon sequestration (Kerr and Ostrovsky, 2003), or to support in the establishment of a posteriori links between policy decisions, regulatory actions and subsequent land-use activities (Lunetta et al., 2006).

The use of time-series surrogates demonstrates that $H_{sn}$ is sensitive to the length of the signal like other entropy related measures of time complexity (Parrott, 2010). One reason for such dependence is the way $H_{sn}$ is calculated. The existence of a correlation among Pielou’s evenness index and the number of species (S) in a community was first identified by DeBenedictis (1973) on a mathematical basis: both terms in the ratio to calculate the evenness are linked to S. Later Smith and Wilson (1996) showed how the aforementioned relation starts to decrease after S exceeds roughly 25. Similarly, our analysis shows that when $v$ is greater than 0.5, differences in $H_{sn}$ tend to progressively decrease as a function of N, because the number of frequencies used in computing the index is gradually bigger than 25. A second cause of dependence is due to the modality we used to derive the power spectrum of the series. The fast Fourier transform (FFT) is dependent on the total length (L) both in terms of the number of frequencies that could be evaluated and in terms of the reliability of the results (cf. chapter 4 in Shumway and Stoffer, 2006). Ecological time series have proven to be exceptionally difficult to analyze because they are generally noisy, show typical non-linear patterns, and have short lengths (Bjørnstad and Grenfell, 2001). All these properties tend to violate the assumptions of the linear time-series analysis tools, partially compromising the modeling strength and reliability of power spectrum estimates through the FFT. By carefully manipulating the time series, like filtering and detrending, it is generally possible to reduce most of the problems, but still the relevance of the length of the signal cannot be avoided. Until a solution to such scale dependency is identified, we suggest that caution has to be paid when comparing entropies estimates of time series of different length in particular when L is lower than 256 points.

Finally, many indices are estimated quantities and, thus, are prone to various sources of uncertainty and/or variability like sampling protocols, error propagation or spatio-temporal variation (cf. Chapter 1 of Legendre and Legendre, 2000). The ability to identify statistically significant changes or to draw statistically meaningful conclusions concerning any cause-effect or correlation with ecologically relevant responses can be impaired by the lack of information concerning their uncertainty (Li and Wu, 2004). Furthermore statistical differences, although not always ecologically significant, may suggest the presence of underlying processes eventually leading to ecological differences over time (Holling and Allen, 2002). We provide a statistical procedure to derive a first approximation of the confidence limits for testing the differences and significance levels for $H_{sn}$. Though we are aware that such a procedure could be improved and a more rigorous statistical framework can be set up, like for instance by using an ad hoc probability distribution or addressing type one error inflation with multiple comparisons (Roback and Askins, 2005), we strongly believe that the availability of a measure of uncertainty of the calculated value of $H_{sn}$ is a major advantage for stimulating and testing working hypothesis.
5. Conclusions

We propose normalized spectral entropy as a valuable tool for ecological applications to quantify and analyze order or predictability of an ecological time series. This index is an answer to the needs of new measures of complexity to better understand the dynamics of adaptive systems and to support the evaluation of system level properties like self-organization. We believe $H_S$ and the related statistical framework can contribute to the ongoing debate on how to estimate spatio-temporal complexity, thus making a step forward in the proposed use of complexity as an ecological orientor (Müller, 2005; Proulx and Parrott, 2008).

Like other complexity measures, $H_S$ is a data demanding index. In physical and medical applications this constraint can, generally, be easily overcome. In many ecological researches the collection of long time series may be a problem for multiple reasons (e.g., costs, project length limits, or sampling scheme). But it is worth noting that in the last decade or so automated ecological monitoring systems have been progressively more installed in research sites (e.g., the FluxNET network). National and international networks for long term ecological research have been established (e.g., the umbrella International Long Term Ecological Research, http://www.ILTERnet.org), the number of public access ecological time-series databases has increased (e.g., the Global Population Dynamics Database, http://cpbdus1.bio.ic.ac.uk/npdd), and the availability and use of remotely sensed data for ecological applications is more common (Kerr and Ostrovsky, 2003). All these advancements should greatly contribute to the application of complexity measure in general, and the use of spectral entropy in particular, especially in the perspective of climate change assessment and ecosystem’s service evaluation.

Finally, we consider normalized spectral entropy ready for use in ecological applications, but further research is necessary not only to cope with its scale dependence from signal’s length. Open issues are how to address effects of rare and extreme events (i.e., spikes or bursts in the signal) or to overcome explicit violations of a Fourier transform, in particular the lack of stationarity in the signal. Currently we are working on such aspects by exploring the use of autoregressive modeling for estimating the power spectrum and extending spectral entropy to wavelet analysis as the global wavelet power spectrum seems a promising tool to address the majority of such issues.

Acknowledgements

We are grateful to the two anonymous reviewers for critical comments and insights that greatly improved the manuscript. Dr. Zaccarelli has been supported for this work by a grant of a PRIN 2008 of the Italian Ministry of University. The Authors thanks Dr. G. Gerosa, University “Cattolica del Sacro Cuore” of Milan (Italy) for providing the data and comments on the ozone concentration and flux time series. The R script to calculate $H_S$ along with an example of the meteorological data set used in the paper can be found in the on-line Appendix A of the journal web-site.

Appendix A. Supplementary data


References
