Catena 77 (2009) 56–64

Contents lists available at ScienceDirect

Catena

journal homepage: www.elsevier.com/locate/catena

Nonlinear spatial series analysis from unidirectional transects of soil physical properties

H. Millán *, I. García-Fornaris, M. González-Posada

Department of Physics and Chemistry, University of Granma, Apdo.21, Bayamo 85100, Granma, Cuba

ARTICLE INFO

Article history:
Received 6 June 2008
Received in revised form 5 December 2008
Accepted 12 December 2008

Keyword:
Soils
Chaos theory
Soil water content
Electrical conductivity

ABSTRACT

Soil is usually presented as a complex dynamical system. Nevertheless, evidences based on the theoretical background of complex system physics are still scarce. The main objective of this work was to search for chaotic parameters using some basic concepts of nonlinear dynamical system theory with spatial series of soil properties. Three spatial series consisting of 1000 data point transects were used. We selected horizontal and vertical electrical conductivity (EC_h and EC_v, respectively) and gravimetric water content from a Vertisol (Typic Hapludert) under rice cropping. Each spatial transect was oriented from South to North with 1-m spacing. It was used the TISEAN Software Package (a public domain software available at http://www.mpipks-dresdren.mpg.de/~tisean) for deriving nonlinear parameters from spatial series. We found interesting evidences of chaotic behaviour as maximal Lyapunov exponents were all positive. They ranged from λ_m=0.129 for water content to λ_m=0.219 for EC_v (filtered series in each case). Original (unfiltered), filtered, and surrogate spatial series confirmed these findings as they also showed positive Lyapunov exponents. All the spatial series showed a higher deterministic component ($|\kappa|$>0.9 in most cases). The Lyapunov range of correlation (ρ) was within the limits 4.56 m (EC_v) to 7.75 m (gravimetric water content) as usually reported from geostatistical investigations. Future works based on dynamical system theory will advance our knowledge on spatial variability of important soil properties and the emergence of deterministic and/or stochastic components.

1. Introduction

The interpretation of soil as a complex, nonlinear system has been a key issue for understanding spatio/temporal soil evolution. The term complex system is very difficult for fitting to open media (e.g. soils) where many variables and state factors are present. As open systems they are probably far from equilibrium. Soils are basic components of ecological systems. However, analysing soils as nonlinear physical systems would be always limited to few variables collected within a narrow range of spatial or temporal scales. Many papers have been published discussing how a novel physical theory (Nonlinear Dynamical System Theory) can explain and/or predict the nonlinear evolution of soil system (Culling, 1988), ecological data (Turching and Taylor, 1993) or soil formation (Phillips, 1998). The extrapolation of time series methods to investigate spatial series is not new in soil and related sciences (Pike and Rozema, 1975). Spectral analyses (e.g. Fast Fourier Transform) conducted on spatial series of soil properties have been used for computing fractal dimensions. This assumes an equivalence between time and spatial series. This sort of equivalence has been used by Ohimya (1991), Kemblowski and Chang (1993) and Pan and Lu (1994) for estimating fractal dimensions from unidirectional transects. However, the explicit use of chaos theory methodology is less reported. In practical terms, long time series are easier to collect than spatial series. Xie and Chen (2004) have found positive Lyapunov exponents from spatial series of gold grade and concluded that the evolution of ore-forming is a dynamical process with chaotic evolution.

The search for nonlinear dynamics usually involves deterministic forms of nonlinear differential equations. Phillips (2002, 2008) has called the attention on the need to develop hypotheses from nonlinear theory which are testable with field observations. Some recent projects such as TISEAN (Hegger et al., 1999) allow nonlinear analyses directly from spatial and/or temporal series of empirical data. This could be a valuable tool for fresh soil studies where medium to large spatial and/or time series of soil data are available. Like a time series, a spatial series is a sequence of scalars representing soil property values as a function of its spatial position (independent variable). The variability pattern of a soil property along a given direction could be a stochastic or a deterministic function of the measurement location. That pattern variation could be associated to the nonlinear response of many interacting variables within soil system. Kantz and Schriber (2003) consider that theoretical arguments on chaotic dynamics also hold for physical quantities measured as a function of their spatial distance. However, one has to consider three situations. First, each physical measurement reveals the

* Corresponding author. Tel.: +53 23 427392.
E-mail address: hmillanv@udg.co.cu (H. Millán).

0341-8162/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.catena.2008.12.009

effect of an underlying dynamics, second, each univariate spatial series represents a dynamical sub-system (e.g. near soil surface salt distribution) and third, the spatial dynamics of soil properties occurs within uniformly managed fields (Timlin et al., 1998). Under these considerations, one could hypothesize that dynamical system concepts allow to search for chaotic invariants from spatial series of soil data. The main objective of this work was to search for chaotic parameters using some basic concepts of nonlinear dynamical system theory with spatial series of soil properties.

2. Theoretical background

There are two basic ways for chaos studies: first, solving nonlinear partial or ordinary differential equations for deriving chaotic parameters or attractors (e.g. Lyapunov exponents or fractals) (Lorenz, 1963) and second detecting chaotic dynamics directly from time series (Abarbanel, 1996; Kodba et al., 2005). A dynamical system is usually observed as a function of time. However, at a first glance there is no restriction for exploring it as a function of spatial coordinates. Xie and Chen (2004) have used this theoretical approach with spatial series of gold grade. That is, there is no mathematical constrain between a time series (e.g. meteorological time series) and a spatial one (for instance, a unidirectional transect of soil data collected at equal spacing intervals).

In the present study we managed three relevant concepts commonly used within the rationale of nonlinear analysis. They were surrogate data, minimal embedding delay (spatial lag delay) determination after averaged mutual information estimation, embedding dimension (m) determination after fraction of false nearest neighbors calculation and estimation of maximal Lyapunov exponent. In addition, each considered spatial series was subjected to a determinism test. The method of surrogate data has been previously used for nonlinearity tests (Theiler et al., 1992; Schreiber and Schmitz, 1996; Chakraborty and Roy, 2006). The basic point is to formulate a null hypothesis (e.g. the spatial series of a soil property was generated by a Gaussian stochastic process) and then trying to reject that hypothesis by comparing results for the data set to corresponding realizations of the null hypothesis (Hegger et al., 1999). As an alternative, the method of surrogates rescales to the exact values of the data and the Fourier Transform is taken to the amplitudes derived from the original data set. The correspondence between surrogates and the original data either tends to zero as the number of iterations increases or converges to a finite value acting as an uncertainty threshold. That is, surrogates are treated as random data sets resembling the same autocorrelation structure as the original series.

The concept of mutual information (M_I hereafter) was introduced by Fraser and Swinney (1986) for estimating an appropriate time lag value. In our case the mutual information between y_i and $y_{(i+x)}$ quantifies the information at state $y_{(i+x)}$ under the assumption that state y_i is known. The rationale presented by Fraser and Swinney (1986), adapted to the present study, is as follows. Given a spatial series $\{y_0,y_1,y_2,\ldots,y_i,\ldots,y_n\}$ with minimum ($y_{\min}$) and maximum ($y_{\max}$) values one computes the absolute difference $|y_{\max}-y_{\min}|$, then this difference is partitioned into k equally sized intervals, with j as large as a possible integer number, then:

$$M_I = -\sum_{h=1}^{j}\sum_{k=1}^{j} P_{h\cdot k}(x)\ln\frac{P_{h,k}(x)}{P_h P_k} \quad (1)$$

where P_h and P_k are the probabilities that the variable takes a value within the h-th and k-th grids, respectively, and $P_{h.k}(x)$ would be the joint probability that y_i is in the h grid and y_{i+x} is in the k grid. The case $P_{h,k}(x)=P_hP_k$ implies no correlation between y_i and y_{i+x} ($M_I(x)\rightarrow 0$). In general, the first minimum of $M_I(x)$ versus x defines a suitable value for the spatial lag (x). In addition, we also considered the structure of the autocorrelation function corresponding to each spatial series.

The false nearest neighbor is a geometrical method allowing a determination of the minimal embedding dimension (m) or the number of active degrees of freedom of the system (Kennel et al., 1993). Its implementation allows one to find a minimal spatial separation of valid neighbors (Hegger et al., 1999). One needs to contruct a vector sequence $\vec{p}(i)=\{y_i,y_{i+x},y_{i+2x},\ldots,y_{i+(m-1)x}\}$, where x is the embedding delay (spatial lag) (Takens, 1981) within the m-dimensional embedding space, after that a neighbour $\vec{p}(j)$ is found such that $|p(i)-p(j)|<\delta$, where δ is a small constant of the order of the standard deviation of the data. A normalized distance Γ_i between the $(m+1)$th embedding coordinate of points $p(i)$ and $p(j)$ could be computed:

$$\Gamma(i) = \frac{|y_{i+m\tau} - y_{j+m\tau}|}{|p(i)-p(j)|} \quad (2)$$

If the distance of the iteration to the nearest neighbor ratio exceeds a defined threshold (e.g. data standard deviation in this case), the point is considered as a false neighbor. The final result is a fraction (e.g. percentage) of false neighbors for each embedding dimension.

Maximal Lyapunov exponents are strong indicators of chaotic behaviour of a system. This is a parameter characterizing the separation rate of closer trajectories within a phase space (Sprott, 2003). Even when most investigations use the maximal Lyapunov exponent, we worked with the Lyapunov spectra (Sano and Sawada, 1985) as they also can be computed from the spatial series without an explicit mathematical model (Kantz and Schriber, 2003). In addition, the Kaplan–Yorke dimension (D_{KY}) is computed from the spectrum of Lyapunov exponents (Kaplan and Yorke, 1987). Hereafter, we fitted nonlinear time series concepts to nonlinear spatial series. This implies a spatial sequence $\{y_0,y_1,y_2,\ldots,y_i,\ldots,y_n\}$, where y_i represents the soil property measured at the point $i=0,1,2,\ldots,n$. Here, we adjust the theoretical rationale for computing the largest Lyapunov exponent ($\lambda_{\max}$). This is the same as that presented within the TISEAN project. In fact, there is an ample literature presenting many variants for $\lambda_{\max}$ estimations. Let us consider the spatial series data (point soil property in this case) as a trajectory and assume a Euclidean distance $\Delta=y_n-y_{n'}$ representing a small perturbation due to nearest neighbors to the data point y_n. The future of Δ can be estimated from the spatial series according to the algorithm of Wolf et al. (1985):

$$S(\varepsilon,m,x) \propto \log\sum|y_{n+x}-y_{n'+x}| \propto \lambda_{\max}x \quad (3)$$

where ε is the spacing of a two dimensional grid constructed for defining nearest neighbours. In general, the slope of a semi-log plot of $S(\varepsilon,m,x)$ versus x would estimate numerically the maximal Lyapunov exponent. Based on $\lambda_{\max}$ values, three cases are distinguished from physical systems:

i) $\lambda_{\max}<0$ represents a dissipative dynamical system with asymptotic stability.
ii) $\lambda_{\max}=0$ is characteristic of conservative dynamical systems. This could be a very rare case for open, real systems like soils.
iii) $\lambda_{\max}>0$ is the exponent of unstable and chaotic systems. This situation does not reject the existence of some type of organization and/or pattern emergence (e.g. fractal or multifractal structures).

The Kaplan–Yorke dimension (D_{KY}) is an interesting nonlinear parameter as it is usually associated to the complex structure of the attractor confining the dynamics of the system. It was originally stated as a conjecture (Kaplan and Yorke, 1987) but it has proved to be true. Many authors consider D_{KY} as a measure of complexity, strangeness and fractal dimension of the attractor (Hegger et al., 1999; Sprot, 2007). In addition, Frederickson et al. (1983) also identified D_{KY} with the information dimension of the system. The practical computation of D_{KY} requires the spectrum of Lyapunov exponents. Consider a chaotic system with m degree of freedom (m embedding dimensions