

Original Research Article

Interaction network based early-warning indicators of vegetation transitions

G. Tirabassi^{a,*}, J. Viebahn^b, V. Dakos^c, H.A. Dijkstra^b, C. Masoller^a, M. Rietkerk^d, S.C. Dekker^d^a Universitat Politècnica de Catalunya, Barcelona, Spain^b Institute for Marine and Atmospheric Research Utrecht, Utrecht University, Utrecht, the Netherlands^c Estación Biológica de Doñana, CSIC, Sevilla, Spain^d Department of Environmental Sciences, Utrecht University, Utrecht, the Netherlands

ARTICLE INFO

Article history:

Received 9 February 2014

Received in revised form 2 June 2014

Accepted 16 June 2014

Available online

Keywords:

Desertification

Early-warning indicators

Networks

Catastrophic transitions

ABSTRACT

Changes in vegetation patterns in semi-arid regions can precede the abrupt transition to bare soil. Here, complex network techniques are used to develop novel early-warning indicators for these desertification transitions. These indicators are applied to results from a local positive feedback vegetation model and are compared to classical indicators, such as the autocorrelation and variance of biomass time series. A quantitative measure is also introduced to evaluate the quality of the early-warning indicators. Based on this measure, the network-based indicators are superior to the classical ones, being more sensitive to the presence of the transition point.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

There are now numerous indications, from observations and models, that transitions between different vegetation states can occur due to the existence of multiple equilibrium states (Scheffer et al., 2001; Rietkerk et al., 2004; Kéfi et al., 2007, 2013). One of these transitions is between vegetated and bare soil states in semi-arid regions and is associated with desertification. Other transitions are, for example, those between savannah and forest states as found in observations of woody cover (Hirota et al., 2011) and above ground biomass (Yin et al., 2014). In the latter case, the equilibrium states appear as different maxima in the probability density function (PDF) of these quantities. Although transient growth effects in the relatively short length of the time series obscure the interpretation of the PDFs, it is plausible that multiple equilibrium states exist (Yin et al., 2014).

The existence of multiple equilibria is supported by models of vegetation dynamics of various levels of complexity. These models can be mainly classified into spatial (PDEs) and non-spatial (ODEs) models, and both types can display multiple equilibria and

catastrophic transitions to desertification. Looking at the spatial models, most of them are of reaction-diffusion type, contain positive feedbacks at different scales, represent the dissipation mechanism as diffusion, and use precipitation as stressor. Analysis of three basic spatial models which describe desertification due to decreasing precipitation (Dakos et al., 2011) indicates that the route to desertification occurs through a saddle-node bifurcation which gives rise to hysteresis.

From a land management point of view, these theoretical results suggest that if transitions to desert states occur in a sudden and unexpected way, it would be crucial to develop early-warning signals in order to prevent or at least prepare for such transitions. Previous studies (Scheffer et al., 2009) proposed that such indicators could be developed based on critical slowing down (CSD), that is, the slow recovery of a system to small perturbations which appears to be a generic phenomenon in the vicinity of bifurcation points (Wissel, 1984). However, CSD can only be detected indirectly by specific statistical properties of the dynamics of a system such as an increase in spatial and temporal correlation as well as variance (Scheffer et al., 2009; Guttal and Jayaprakash, 2009; Dakos et al., 2010). These CSD-based statistical properties have been suggested to act as early-warning signals for critical transitions (Scheffer et al., 2009), and they have been experimentally demonstrated to exist in various living systems (Drake and Griffen, 2010; Carpenter et al., 2011; Dai et al., 2012).

* Corresponding author at: Rambla de St. Nebridi 22, Gaia building, 08222 Terrassa, Spain. Tel.: +34 937398556.

E-mail address: giulio.tirabassi@upc.edu (G. Tirabassi).

Recently, complex network based measures were proposed as leading indicators of critical transitions in the ocean circulation – also known to be associated with saddle-node bifurcations (Mheen et al., 2013; Viebahn and Dijkstra, 2014). It was shown that an upcoming saddle-node bifurcation can be adequately detected by network measures which are constructed via coarse-graining of the cross-correlation matrix of the spatially extended system.

In this paper, we apply network-based indicators to time series stemming from a vegetation model which exhibits a critical transition to desertification. Both the vegetation model and the network techniques are presented in Section 2. Our aim is to test whether novel network-based measures are able to indicate the upcoming collapse of vegetation to a desert state (Section 3) and to assess their potential as tools for anticipating critical transitions in a wide range of ecological systems (Section 4). The paper is closed with a summary and discussion (Section 5).

2. Model and methods

In this section, we first present the spatial vegetation model used to produce biomass time series, and then present the methods of network construction and analysis.

2.1. A spatial model of vegetation dynamics with a local positive feedback

The local positive feedback model (LPF) is described by the following set of stochastic differential equations (Shnerb et al., 2003; Guttal and Jayaprakash, 2007; Dakos et al., 2011),

$$\frac{\partial w}{\partial t} = R - \frac{w}{\tau_w} - \Lambda w B + D \nabla^2 w + \sigma_w w_0 \xi^w(t), \tag{1}$$

$$\frac{\partial B}{\partial t} = \rho B \left(\frac{w}{w_0} - \frac{B}{B_c} \right) - \mu \frac{B}{B + B_0} + D \nabla^2 B + \sigma_B B_0 \xi^B(t), \tag{2}$$

where w (in mm) is the soil water amount and B (in g/m^2) is the vegetation biomass. The quantity D is the diffusivity and $\tau_w, \mu, \rho, \Lambda, w_0, B_0, B_c$ are additional constants explained in Table 1. Finally, R is the amount of rainfall which is used as the bifurcation parameter of the system. Additive Gaussian white noise, ξ , for which

$$\langle \xi(t) \xi(t') \rangle = \delta(t - t'), \tag{3}$$

is prescribed with amplitudes σ_w and σ_B for soil water and biomass, respectively. A characteristic spatial pattern of biomass as well as time series for different values of R are shown in Fig. 1.

The important mechanism in this model is a positive feedback that causes each patch to have alternative stable states. This is demonstrated in the bifurcation diagram of the LPF model which is sketched in Fig. 2. The deterministic homogeneous solutions of the LPF model and their linear stability can be determined analytically.

Table 1
Parameters of the local positive feedback model (LPF) given by Eq. (2) and same values as in Dakos et al. (2011).

Parameter	Meaning	Value
D	Exchange rate (m^2/day)	0.5
Λ	Water consumption rate by vegetation ($m^2/(g \cdot day)$)	0.12
ρ	Maximum vegetation growth rate (day^{-1})	1
B_c	Vegetation carrying capacity (g/m^2)	10
μ	Maximum grazing rate ($g/(day \cdot m^2)$)	2
B_0	Half-saturation constant of vegetation consumption (g/m^2)	1
σ_w	Standard deviation of white noise in water moisture	0.1
σ_B	Standard deviation of white noise in vegetation biomass	0.25
w_0	Water moisture scale value (mm)	1
B_0	Biomass density scale value (g/m^2)	1
τ_w	Water moisture scale time (day)	1

For all values of R , the trivial solution ($B = 0, w = \tau_w R$) exists. For the standard parameter values shown Table 1, the trivial solution is linearly stable for $R < 2$ mm/day and unstable for $R > 2$ mm/day (see Fig. 2). At $R = 2$ mm/day, a transcritical bifurcation occurs and two additional branches of steady solutions emerge. Solutions on the lower branch are not considered here because they have $B < 0$, i.e., they are physically non-realistic. Solutions on the upper branch are unstable for values of R down to $R_c = 1.067$ mm/day. At this R -value a saddle-node bifurcation occurs which provides a linearly stable upper branch of solutions for $R > 1.067$ mm/day. Finally, a fourth homogeneous solution exists but it has also values of $B < 0$ for every R value and hence is not further considered in this study.

In order to determine inhomogeneous vegetation patterns in the stochastic case, the model equations (2) are numerically solved on a periodic square grid composed of $100 \times 100 = 10^4$ grid cells on a regular lattice with dimension $L = 100$ m. The evaluated model data consists of a set of time series (500 time steps with $\Delta t = 0.01$ days) of statistically equilibrated biomass fields B for different fixed rainfall parameters R . Time series related to 10 different values of R with 1.1 mm/day $\leq R \leq 1.8$ mm/day are analysed. For $R < R_c$ only the desert-like solution, with $B = 0$ over the whole domain, is found.

The temporal and spatial mean values of the biomass distribution of each of these spatially inhomogeneous solutions are plotted as the red dots in Fig. 2. Obviously, the average values of the inhomogeneous solutions are similar to those of the homogeneous solutions, hence, diffusion and noise do not impact on the average state of the system.

2.2. Network approach and analysis

A network is defined by two sets of objects: the so-called nodes, and the set of their mutual connections, namely their links (Caldarelli, 2007). In the following, we will associate a network to the simulated field of biomass. The nodes are defined as the $N = 100 \times 100 = 10^4$ grid cells of the discretised LPF model. In order to define the links between the nodes, the zero-lag temporal correlations between the biomass time series at the different nodes are considered. More precisely, two nodes i and j are linked if the temporal cross-correlation $\mathcal{C}(B_i, B_j)$ of the time series B_i and B_j is statistically significant. A sketch of the network formation is depicted in Fig. 3. Note that in this way the number of nodes of the network is kept fixed, and changes in the network structure due to varying R are solely related to the links, that is, to changes in the values of the cross-correlation matrix $\mathcal{C}(B_i, B_j)$.

A compact way to describe a network composed of N nodes is to consider its adjacency matrix \mathbf{A} , a symmetric $N \times N$ matrix with $A_{ij} = 1$ if node i and node j of the network are linked, and $A_{ij} = 0$ otherwise. Thus, the correlation network of the biomass data is given by the following adjacency matrix,

$$A_{ij} = H(|\mathcal{C}(B_i, B_j)| - \theta), \tag{4}$$

where H is the Heaviside step function, and θ is a constant threshold indicating statistical significance of the cross-correlation $\mathcal{C}(B_i, B_j)$.

To determine the value of θ we build the following test variable for the Student's t -test, i.e.,

$$t = \frac{\theta}{\sqrt{1 - \theta^2}} \sqrt{N_{\text{time-steps}} \frac{1 - r}{1 + r}}, \tag{5}$$

with the null hypothesis $\theta = 0$. Here $r = r(R)$ and $N_{\text{time-steps}}$ are the autocorrelation and the length of the time series, respectively. The test variable takes the effective number of degrees of freedom of the time series into account. From this we can compute the value of θ which ensures statistical significance of correlations larger than

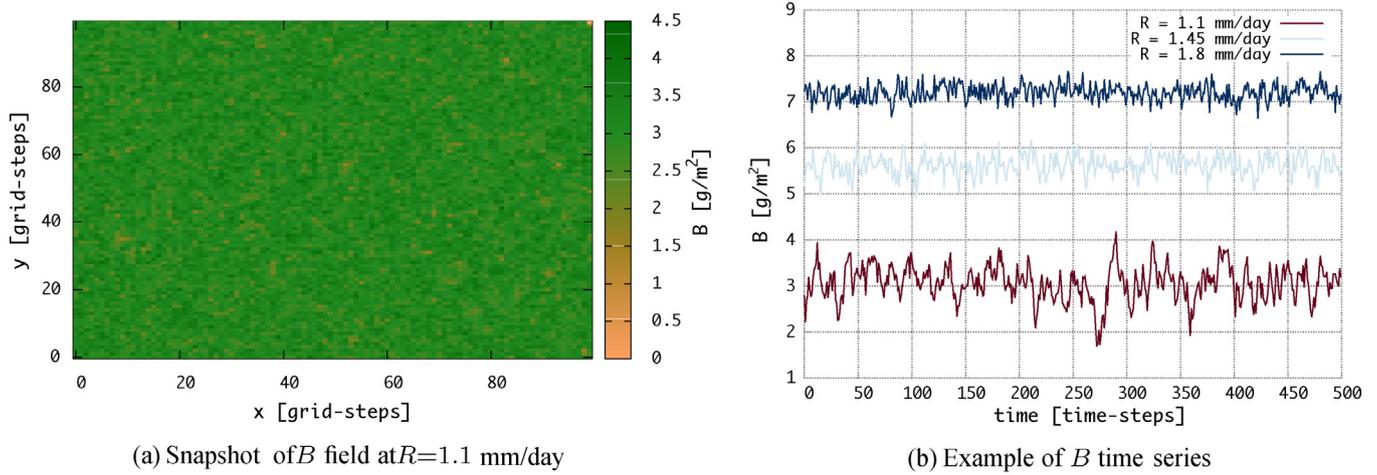


Fig. 1. (a) A snapshot at $t = 500$ time-steps of the biomass field B at $R = 1.1$ mm/day. (b) Example of biomass time series for a node (single grid-cell) i and different values of R .

θ . A value of $\theta = 0.2$ guarantees that, for each value of R , the zero-lag correlation between linked nodes is statistically significant with a p -value smaller than 0.05 and this value is taken in all results below.

A central advantage of this approach is that by taking into account only the statistically significant correlations to establish links between nodes, we can remove part of the noise and extract only the relevant information. This coarse-graining results in a more precise measure and identification of the essential properties of the system under study.

It is important to note that using the temporal cross-correlation among nodes does not directly take their spatial distribution into account. It is only the temporal part of the data which is used to build the network, but the spatial information is kept because the nodes have a definite spatial location in physical space (that is, on a grid). Thus, within this network approach we exploit both the systems temporal information and spatial information which are treated separately in the classical early-warning approaches. After

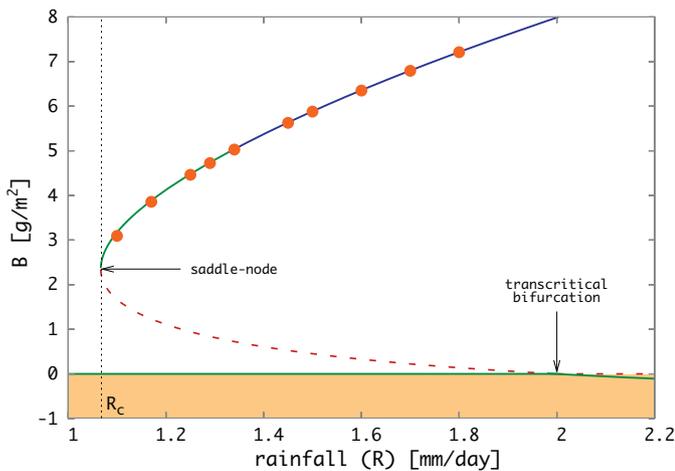


Fig. 2. Bifurcation diagram of the local positive feedback model (LPF) given by Eq. (2). Curves depict steady homogeneous states, that is, determined under vanishing diffusion and noise. Linearly stable branches are denoted by solid lines, whereas linearly unstable branches are indicated by dashed lines. The blue lines mark states for which the Jacobian matrix has eigenvalues with a non-zero imaginary part. The shaded region indicates to the non-physical negative B values. The average value of B for each simulation of biomass evolution via the full Eq. (2) is depicted by dots. Obviously, the average values of the inhomogeneous solutions are similar to those of the homogeneous solutions, hence, diffusion and noise do not impact on the average state of the system. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

the construction of the interaction network of the biomass data, we can study changes in the topology of the network due to varying R . The most basic characteristic of a network is its degree distribution. The degree k_i of a node i is defined as the number of links of the node i . For example, node i in Fig. 3 has degree $k_i = 3$. In general, the degree of a node can be computed from the adjacency matrix by

$$k_i \equiv \sum_{j=1}^N A_{ij}. \tag{6}$$

In our network construction k_i can range from 0 to $N - 1$, because we exclude self-connections.

Another basic network measure is the assortativity a_i of a node i , which is the average degree of its neighbours, that is, of all the nodes to which node i is linked to. In Fig. 3, all three neighbours of node i have degree 2 and hence its assortativity $a_i = 2$. In general, the assortativity coefficient of a node can be computed from the adjacency matrix via

$$a_i \equiv \frac{1}{k_i} \sum_{j=1}^N A_{ij} k_j. \tag{7}$$

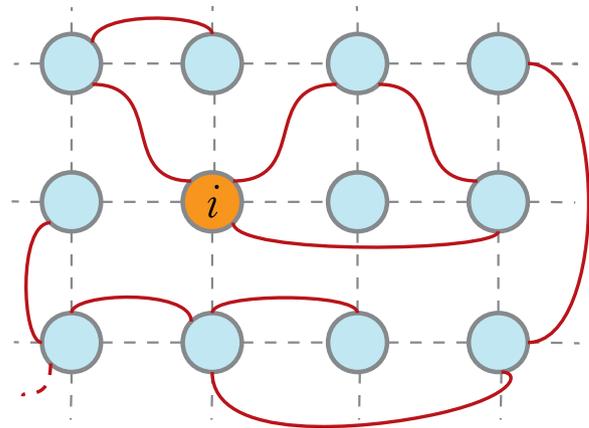


Fig. 3. Sketch of the network associated to the local positive feedback model (LPF) given by Eq. (2). The circles represent the biomass grid-cells, that is, the nodes of the network. The dashed lines represent the computational grid, whereas the solid lines represent the network connections, that is, the links. The node i (marked orange) has degree $k_i = 3$, clustering coefficient $c_i = 0.33$, and assortativity $a_i = 2$. The link length of its three links is $\sqrt{2}$, $\sqrt{2}$, and 2 grid-steps (see text for the definition of these quantities). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

The assortativity characterizes the tendency of a node to be connected to nodes with high degree. Note that a node can have low degree but at the same time high assortativity. As the degree, the assortativity coefficient can range from 0 to $N - 1$.

As a third basic network measure we consider the clustering coefficient c_i that measures the number of links among the neighbours of a node i , weighted by the possible number of links among its neighbours. In the example shown in Fig. 3, there is only one link between the neighbours of node i and there are three possible links among these neighbours and hence the clustering coefficient $c_i = 1/3$. In general, the clustering coefficient of a node can be computed from the adjacency matrix via

$$c_i \equiv \frac{1}{k_i(k_i - 1)} \sum_{j=1}^N \sum_{l=1}^N A_{ij} A_{jl} A_{li} \quad (8)$$

which can take values between 0 and 1. The average clustering coefficient of a network quantifies the presence of strongly connected groups (clusters) in the network.

3. Results

In Section 3.1, the time series of the LPF model related to different values of R are analysed by computing the classical early-warning indicators and focussing on the behaviour of these measures when R approaches the critical value R_c . In Section 3.2, we investigate whether changes in the network topology (diagnosed via the distributions of degree k_i , assortativity a_i and clustering c_i) can be used as alternative network-based indicators of R approaching R_c .

3.1. Classical indicators

The most prominent statistical measure to infer CSD from data is the lag-1 autocorrelation. The spatially averaged autocorrelation of the LPF model data is shown in Fig. 4 for each value of R . As R decreases and approaches R_c , the system indeed experiences an increase in autocorrelation, a distinct fingerprint of CSD.

It has also been suggested that spatial statistics may be used to detect CSD (Dakos et al., 2011). In particular, an increase in spatial correlation of the system is expected when the system experiences CSD close to the transition. A typical measure of spatial correlation that has been used is Moran's coefficient, I , defined here through

the biomass mass field B as

$$I \equiv \frac{N}{\sum_{ij} g_{ij}} \frac{\sum_{ij} g_{ij} (B_i - \bar{B})(B_j - \bar{B})}{\sum_i (B_i - \bar{B})^2}, \quad (9)$$

where $g_{ij} = 1$ if i and j are two adjacent grid cells and $g_{ij} = 0$ otherwise. The spatially averaged biomass is indicated by \bar{B} . Moran's coefficient I is shown in Fig. 4 and demonstrates that the spatial correlation increases in a similar way – although stronger – to the temporal correlation as R decreases and approaches R_c .

However, despite the fact that the classical indicators are able to reflect CSD, they change in a very smooth, gradual and monotonic way. From a strictly local point of view, that is, based on a few closely spaced R values, it is not possible to estimate the proximity of the system to R_c . In other words, the correlation coefficients suffer from a lack of distinct features necessary to provide a pronounced early-warning signal of desertification.

3.2. Interaction network based indicators

In order to give a general characterization of the interaction network of the LPF model biomass data, we briefly discuss its link length distribution before studying the changes in degree, assortativity and clustering due to varying R .

The link length measures how close two linked nodes are in physical space, that is, the link length is the Euclidean distance between two linked nodes (for an example see again Fig. 3), where the grid spacing is taken as a unit length. The distribution of link lengths for different R is shown in Fig. 5. Apparently, the link length distribution does not depend on R . The smaller spike at link length of one grid spacing represents the underlying grid structure and is related to local correlations due to diffusion. The linear behaviour at larger distances can be attributed to the random structure of the network: If the links are randomly distributed in physical space, the probability of two nodes i and j to be connected is independent of their distance. In fact, the number of links of length d of any node i is solely proportional to the number of nodes at distance d from node i . This number scales with the circular area $2\pi d$. Hence, the probability of a certain link length is directly proportional to the link length itself, until the appearance of the boundary reduces the possible number of links again (here at a distance of 50 m, that is $L/2$). Consequently, the interaction network of the simulated biomass field appears to be composed of a random architecture superimposed on the grid structure.

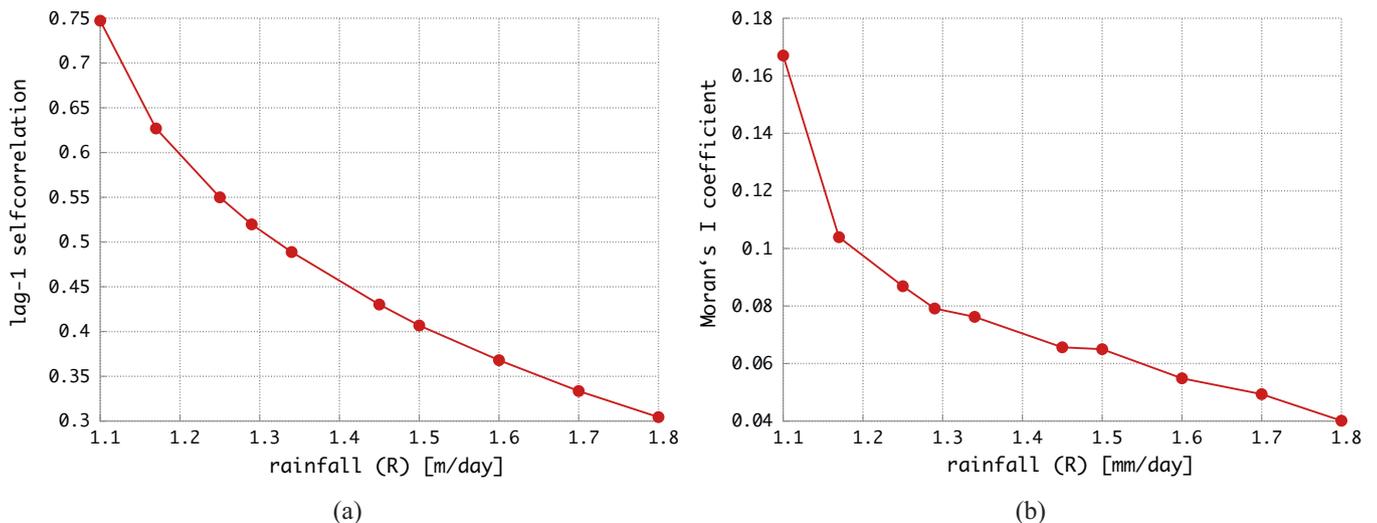


Fig. 4. (a) Spatially averaged lag-1 autocorrelation and (b) Moran's coefficient I computed using the spatial biomass distribution at the last time-step of the simulation for different values of the rainfall parameter R .

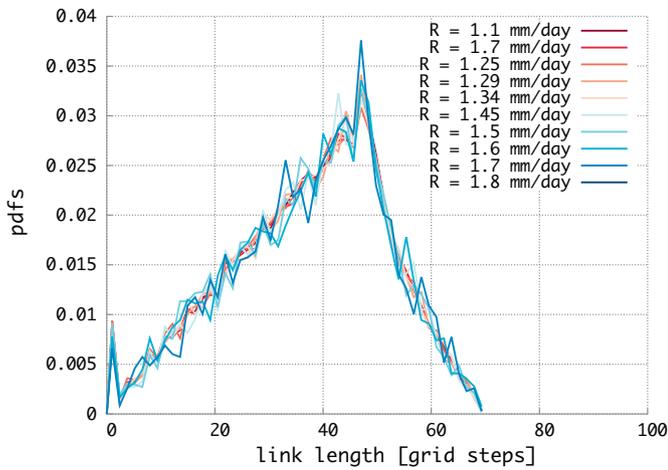


Fig. 5. Link length PDFs for various values of rainfall R . Link length is computed as the Euclidean distance between two linked grid-cells, taking into account periodic boundaries.

Fig. 6 shows the spatial field of the degree distribution for two values of R (at the transition and far away from the transition), together with the degree density distribution (i.e., the PDF of k_i) for different R . With decreasing rainfall R the mean number of links per node in the network increases. Consequently, the degree PDF shifts to the higher degrees when the system approaches the critical transition. The increase of network connectivity is related to higher

values of the cross-correlation among the nodal time series which occurs near the saddle-node bifurcation (Mheen et al., 2013).

This feature can also be seen in the spatial patterns of the degree field. For $R = 1.8$ mm/day (Fig. 6b), the network is almost disconnected and the majority of the nodes has zero or one link. With decreasing rainfall the number of connections increases, and the disconnected nodes eventually join the network. At the transition, the spatial pattern of the degree distribution shows a granular structure, with patches of highly connected nodes (Fig. 6a).

The shift of the degree PDF to higher degrees (Fig. 6c) can be used to develop an indicator of the upcoming transition. Fig. 7a shows the mean of the degree distribution as a function of R . This network measure is highly sensitive to R near criticality, showing a steep increase close to the transition. Additionally, Fig. 7b shows the variance of the degree distribution as a function of R . The increase in variance when approaching the transition point is even more abrupt. The behaviour of the average degree is directly related to high degree values occurring near the saddle-node bifurcation. Close to the transition, the vegetation variability synchronizes over the domain, which can be seen as the spatial expression of CSD. This apparent synchronization produces an increasing number of connections approaching the tipping point. As it can be seen, the panels in Fig. 7 allow quantifying the behaviour displayed by the spatial degree field in Fig. 6. In this way we capture the bidimensional information of the size of the patches and their degree values in a clear way, which monitors quantitatively the presence of the upcoming transition.

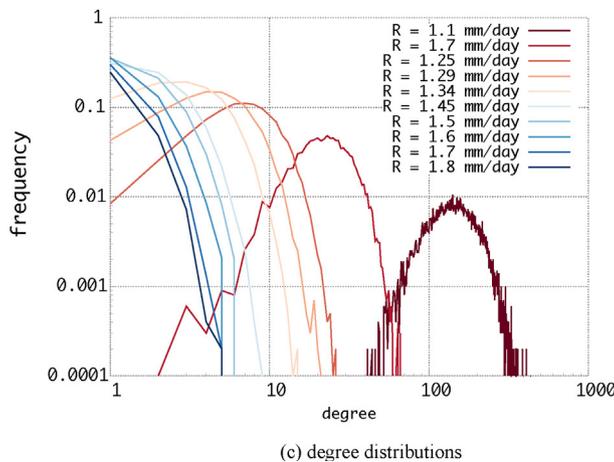
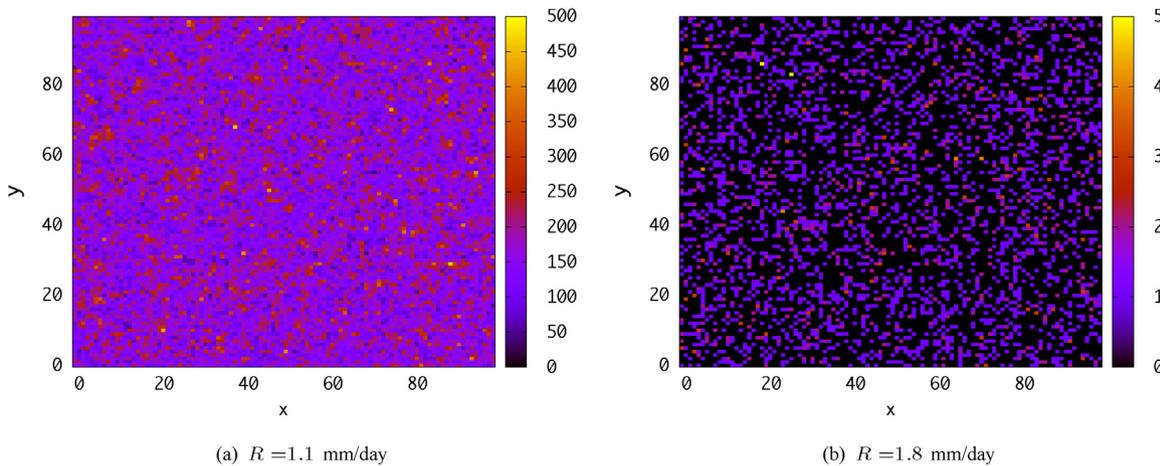


Fig. 6. (a and b) Spatial field and (c) PDF of node degree for different values of R .

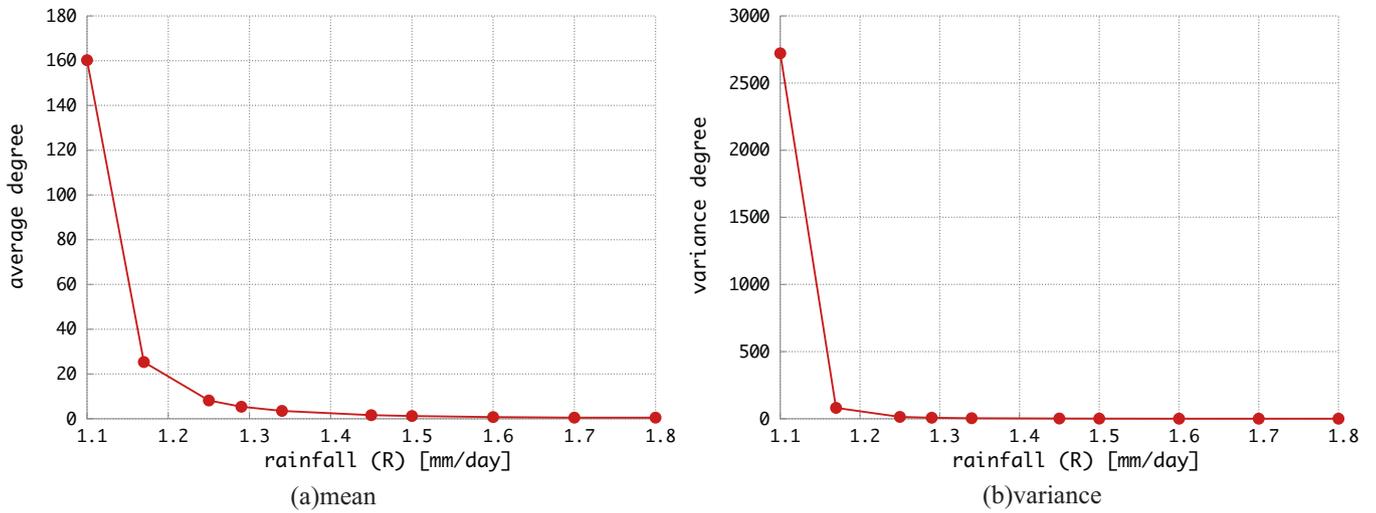


Fig. 7. (a) Mean and (b) variance of network degree distribution as function of the bifurcation parameter R .

Fig. 8 shows the assortativity field of the interaction network and the corresponding PDF for different rainfalls R . The spatial pattern of assortativity is more homogeneous than that of the degree field and hence its PDFs are narrower. Similar to the degree field, there is a substantial shift to larger values of the PDF when R

approaches R_c . The spatial patterns of assortativity display a very noisy structure, compatible with the random network architecture previously inferred from the link length distributions. Furthermore, the average assortativity (not shown) increases sharply when approaching the transition as can also be seen in the PDFs.

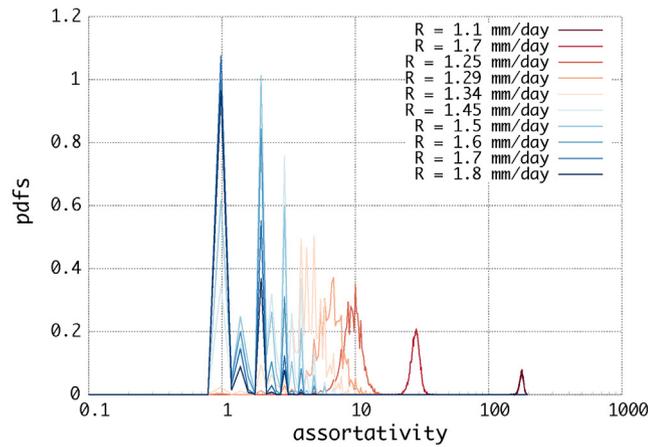
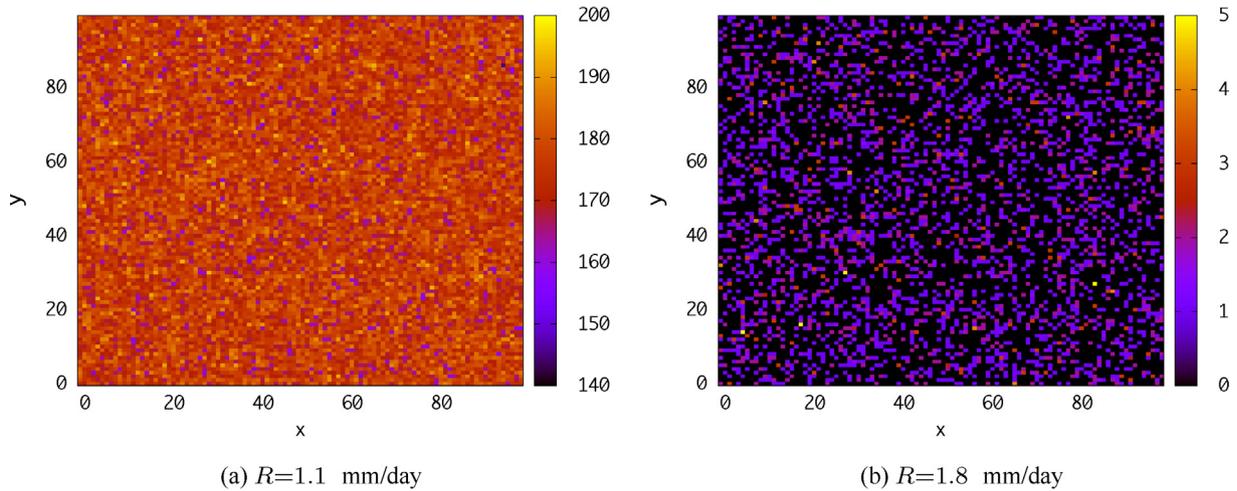


Fig. 8. (a and b) Spatial field and (c) PDF of the assortativity for different values of R . The sharp increase of the assortativity approaching the tipping point is directly related at the behaviour of the corresponding degree field.

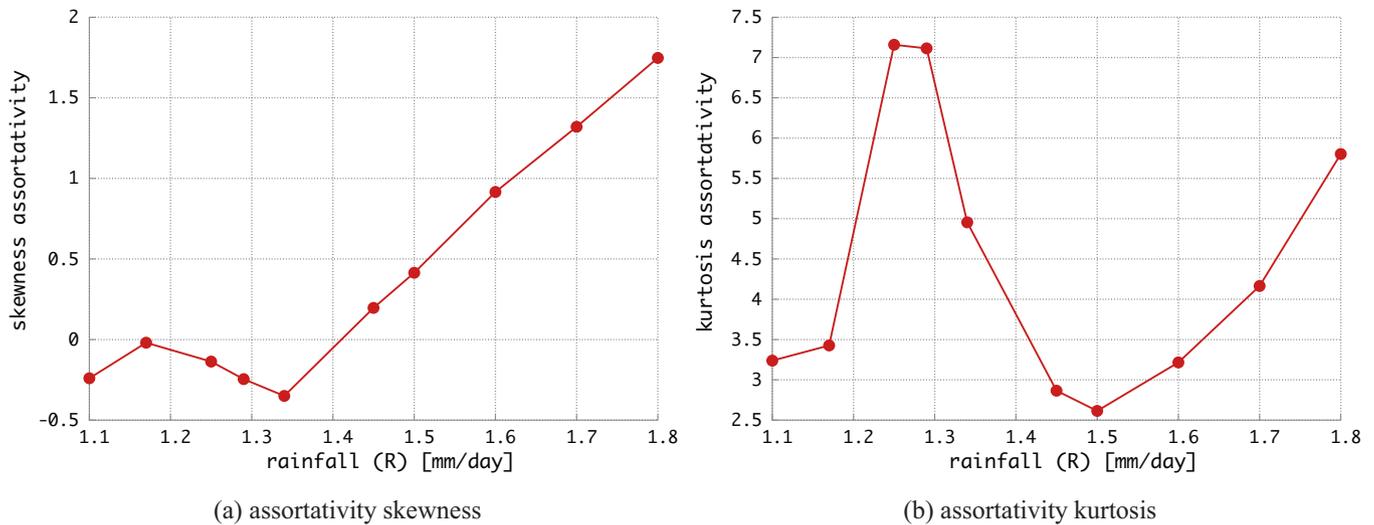


Fig. 9. (a) Skewness and (b) kurtosis of the assortativity distribution for different R .

Fig. 9a–b shows the skewness and kurtosis of the assortativity distribution¹ for different R . The switch in sign of the skewness (Fig. 9a) is a feature that could be related to the forthcoming transition. This switch in sign is unique (a distinct qualitative feature), and it is not prone to false alarms. Furthermore, the kurtosis (Fig. 9b) shows a huge and quick drop just before the transition, providing a distinct warning related to the upcoming transition point, especially if considered together with the behaviour of the skewness. However, the combined analysis of these two quantities reveals a more pronounced change in the assortativity distribution: Just before the transition the skewness approaches 0, while the kurtosis is close to 3; thus the assortativity distribution is close to a Gaussian distribution near R_c . This “Gaussianisation” may therefore be used as an early-warning signal of the transition point.

Finally, Fig. 10 shows the spatial field of the node clustering coefficient and the corresponding PDF for different rainfall parameter R . At high R , the clustering coefficient is zero everywhere. The average node degree is very low such that there are not sufficient links to create clustered structures. With decreasing R the first small clusters start to appear and a non-vanishing average value of the clustering coefficient emerges. The more links are being added to the network with decreasing R , the more clusters are formed and mean clustering coefficient continues to increase. Close to the transition, a richer network structure provides the noisy pattern displayed in Fig. 10a.

Regarding the transition, the mean clustering (not shown) does not provide more information than the mean degree because these two quantities are related: In a completely random network the clustering scales linearly with the number of links. The variance of the clustering distributions (not shown) provides an early-warning indicator of the transition by drastically dropping just before R_c . The combined increase in mean clustering and reduction in variance can be explained as an effect of the increase in spatial coherence displayed by the system. Approaching the transition, patches of vegetation increase their synchronization forming well defined groups of strongly connected nodes. However, these indicators do not perform better than the corresponding degree measures.

In contrast, the skewness and kurtosis of the clustering distribution, shown in Fig. 11c and d, can be used as an early-warning indicator

if the two quantities are monitored together. They are displaying a “Gaussianisation” of the clustering distribution, similar to the assortativity distribution when approaching the tipping point. However, up to this point the “Gaussianisation” of the assortativity and clustering PDFs is only a qualitative feature. In the following section we give a more quantitative measure of this class of early-warning indicators.

4. Quality assessment of early-warning indicators

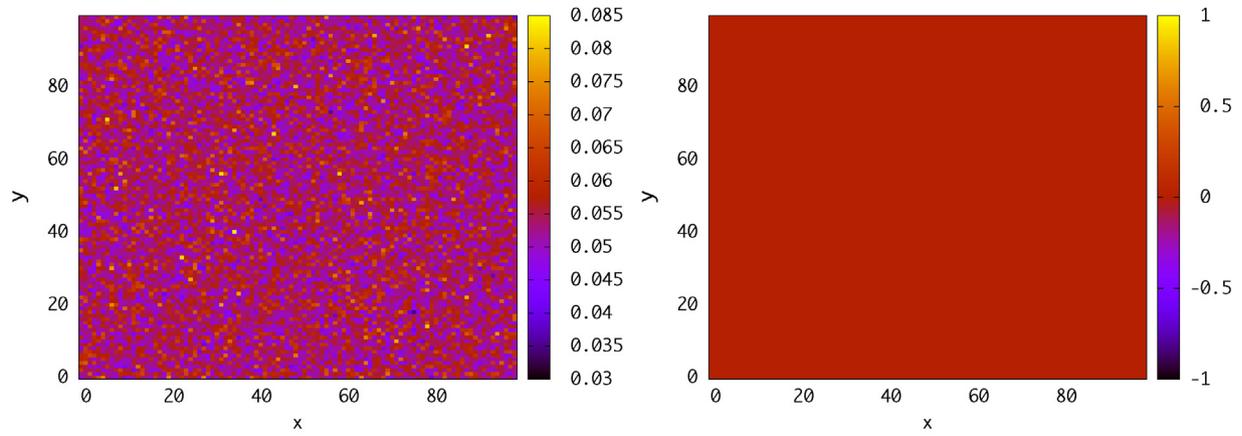
Based on the results of the previous section, we can divide the early-warning indicators considered into two main classes. The first class consists of *scalar-based* indicators, that is, scalar measures which generally change monotonically when the transition point is approached. For example, the mean and variance of degree, assortativity and clustering belong to this class. For these indicators it is generally important to monitor their derivative with respect to the bifurcation parameter R in order to sign proximity of the system to the transition point. The absolute value of the indicator itself gives less information than its abrupt change close to the transition (Van Nes and Scheffer, 2007; Scheffer et al., 2009). The classical indicators are also included in the scalar-based class.

The second class consists of *distribution-based* indicators. These indicators monitor changes in the distribution of key quantities when the system approaches the tipping point. In our case, the assortativity and the clustering belong to this class because their distributions approach Gaussian distributions close to R_c . The degree distribution, instead, does not approach a Gaussian (not shown). The critical normalization of clustering and assortativity can be quantified numerically through the Kullback–Leibler Distance (KLD), also called relative entropy, which measures the distance between two PDFs. Given two one-dimensional distributions $P(x)$ and $Z(x)$, their relative entropy is defined as

$$\text{KLD} \equiv \int_{-\infty}^{\infty} \ln \left(\frac{P(x)}{Z(x)} \right) P(x) dx. \quad (10)$$

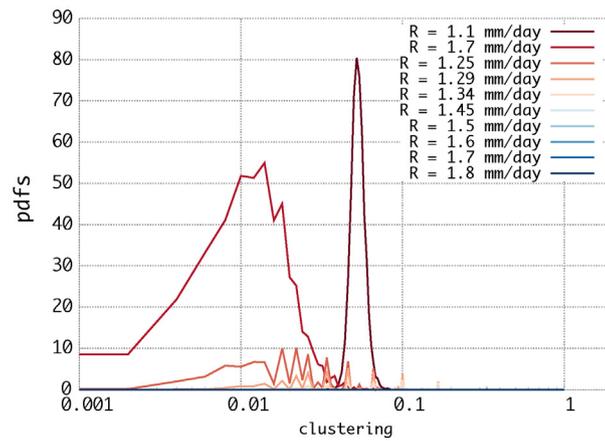
Measuring KLD between assortativity/clustering distributions and Gaussians with the same mean and variance allows to quantify the “Gaussianisation” of these PDFs when the tipping point is approached. We note that the Gaussianisation might a model-specific feature and further analysis is required to assess the generality of this result.

¹ We exclude the mean and variance of the assortativity distribution from the presentation because they are directly connected to the mean and variance of the degree distribution.



(a) $R = 0.1$ mm/day

(b) $R = 0.8$ mm/day



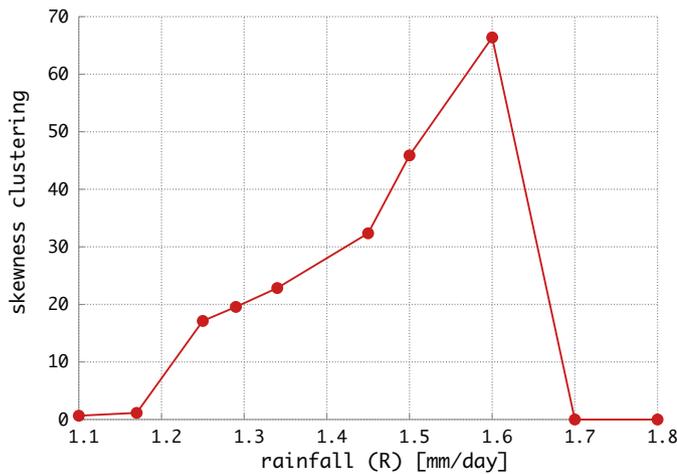
(c) clusteringPDFs

Fig. 10. (a and b) Spatial field and (c) PDF of node clustering for different values of R .

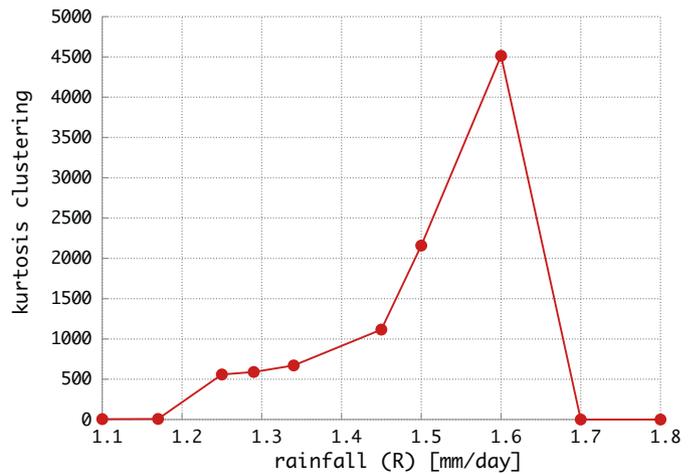
The KLD measure for both assortativity and clustering distributions is plotted in Fig. 12. Obviously, KLD quickly drops to zero when approaching the transition point for both the assortativity and clustering distributions. Small KLD means that the distribution under examination is comparable to a Gaussian,

whereas for large values it deviates from Gaussianity. Unlike the scalar-based case, the absolute value of the indicator is more important than its derivative.

For both indicator classes a quality measure can be defined. First, we define an ϵ -environment around the bifurcation point by



(a) clustering skewness



(b) clustering kurtosis

Fig. 11. (a) Skewness and (b) kurtosis of the clustering distribution for different R .

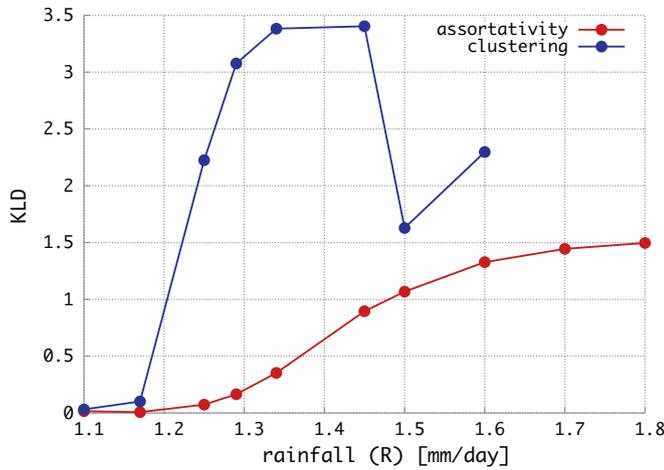


Fig. 12. KLD values for clustering and assortativity PDFs with respect to Gaussian distributions with same mean and variance.

Table 2
Quality values of different early-warning indicators.

Indicator J	Class	Type	Q^J
Average degree	Scalar	Network	0.963
Variance of degree	Scalar	Network	0.996
Variance of assortativity	Scalar	Network	0.971
Average clustering	Scalar	Network	0.904
Assortativity Gaussianity	Distribution	Network	0.973
Clustering Gaussianity	Distribution	Network	0.951
Lag – 1 autocorrelation	Scalar	Classical	0.468
Spatial correlation	Scalar	Classical	0.730

all the R values for which $(R - R_c)/R_c < 0.1$. For a scalar-based indicator J , we then define the normalized quality measure Q_s^J by

$$Q_s^J \equiv \frac{\langle \partial J / \partial R \rangle_{R < \epsilon} - \langle \partial J / \partial R \rangle_{R > \epsilon}}{\langle \partial J / \partial R \rangle_{R < \epsilon} + \langle \partial J / \partial R \rangle_{R > \epsilon}}. \quad (11)$$

where the brackets indicate the mean over the interval indicated. In this way we achieve that if J shows an abrupt change in its derivative close to the transition then we have $Q_s^J \approx 1$. In contrast, if the change of J is merely linear when approaching the tipping point then we obtain $Q_s^J \approx 0$.

In case of distribution-based indicators we can define a similar quality measure Q_d^J by taking into account J itself instead of its derivative, that is,

$$Q_d^J \equiv \frac{\langle J \rangle_{R > \epsilon} - \langle J \rangle_{R < \epsilon}}{\langle J \rangle_{R < \epsilon} + \langle J \rangle_{R > \epsilon}}. \quad (12)$$

In Table 2 the quality values for both network-based and classical indicators are shown. Obviously, the network-based measures have significantly higher early-warning quality than the classical measures.

5. Summary and discussion

In this study the possibility of anticipating desertification transitions in a simple vegetation model using novel interaction network techniques is investigated. Interaction networks are constructed from time series of biomass fields and the topological changes in the resulting networks are studied along a gradient of decreasing rainfall. We find that network measures like degree, assortativity and clustering may offer novel indicators for identifying an upcoming desertification in semi-arid ecosystems.

Our results are consistent with previous studies that have used network measures as early-warning indicators of critical transitions in models. For example, Mheen et al. (2013) used an interaction network approach to obtain an early-warning signal of the Atlantic Meridional Overturning Circulation (AMOC) collapse. In that study the network is built using temperature time series, and the behaviour of the average degree is monitored as function of freshwater input. Similar to the results in this study, the average degree increases sharply approaching the transition. In contrast, Viebahn and Dijkstra (2014) analysed the flow field of the wind-driven ocean circulation introducing a flux-based network approach. Also in that context, the degree of the network increases while the system approaches the transition, but a more precise early-warning indicator is given by the network's closeness which shows a big drop near the tipping point due to a local regime change in the flow field.

In this study we have introduced measures to assess the quality of different early-warning indicators. Using these quality measures we compared the performance of the novel network-based indicators with the classical indicators based on variance and autocorrelation. We find that the scalar network-based indicators have a higher quality value than the classical indicators. Moreover, distribution-based indicators, here calculated from the assortativity and clustering distributions, have also a high quality value. When these distributions become close to Gaussians, there is an early-warning indication of the upcoming transition. Although these observations hint that the indicators we developed here may offer a strong indirect measure of proximity to critical transitions, they may still be prone to similar limitations that classical indicators face, like producing false positives (Boettiger and Hastings, 2012), or requiring a lot of information for their practical application (Dakos et al., 2012).

Regarding the possibility of false positives, it is well known that CSD is a characteristic feature of transitions related to an eigenvalue going to zero. However, not all the transitions accompanied by an eigenvalue going to zero are also catastrophic – that is implying an abrupt discontinuity in the stable branches (Kéfi et al., 2013; Dijkstra, 2011). Thus, CSD can be prone to false alarms, and it will be interesting to test, in future work, the new network indicators against this possible pitfall.

Regarding the sample size of the data, instead, it is clear that the network construction requires a sufficient spatial sampling which temporally based classical indicators do not need. As a test, we computed all the indicators presented in this paper also for other two networks, obtained by spatially down-sampling the original 100×100 dataset considered here. In particular we considered 10×10 and 50×50 sub-sets. We find that the scalar-based indicators still perform quite well, whereas the distribution-based indicators do not. In particular, the assortativity Gaussianity still anticipates the transition in the 50×50 data set, whereas the clustering measure is unable to detect the transition in both data sets. This is a reasonable feature of the proposed distribution-based indicators; in fact a good estimation of a PDF requires a big amount of data, much more than simple low-order moments estimators, as the one used here to build the scalar-based indicators.

While the average degree behaviour as a function of R can be easily linked to the spatial expression of the enhancement of correlations of the time series between the nodes, the interpretation of the other indicators, especially that of the distribution-based indicators, is not straightforward. Nevertheless, the network-based indicators seem to offer a better measure of proximity to the tipping point than the classical ones. This may be attributed to the thresholding of the correlation matrix when constructing the interaction network. This coarse-graining of the information eliminates most of the noise, producing a better signal to noise ratio of the spatial signature of CSD.

The network based indicators hence offer a promising alternative to detect critical transitions. An interesting next step is to analyse the utility of these indicators in a scale dependent feedback model (Dekker et al., 2007; Dakos et al., 2011), for which basic bifurcation diagrams were presented in Dijkstra (2011). Also in such a model, the desertification transition is a saddle-node bifurcation but the approach to this saddle-node with decreasing rainfall is not as smooth as in Fig. 2 due to presence of multiple branches of stable steady states.

Acknowledgements

The Authors would like to acknowledge the support of the LINC project (no. 289447) funded by EC's Marie-Curie ITN program (FP7-PEOPLE-2011-ITN).

CM gratefully acknowledges partial support from the ICREA Academia programme and grant FIS2012-37655-C02-01 from the Spanish MCI.

The research of MR is also supported by the project CASCADE (Seventh Framework Programme FP7/2007-2013 agreement 283068).

VD acknowledges support from a EU Marie Curie Intra-European and a Netherlands Science Foundation Rubicon fellowships.

References

- Boettiger, C., Hastings, A., 2012. Early warning signals and the prosecutor's fallacy. *Proc. R. Soc. B: Biol. Sci.* 279, 4734–4739.
- Caldarelli, G., 2007. *Scale-Free Networks: Complex Webs in Nature and Technology*. Oxford University Press, Oxford.
- Carpenter, S., Cole, J., Pace, M., Batt, R., Brock, W., Cline, T., Coloso, J., Hodgson, J., Kitchell, J., Seekell, D., et al., 2011. Early warnings of regime shifts: a whole-ecosystem experiment. *Science* 332, 1079–1082.
- Dai, L., Vorselen, D., Korolev, K.S., Gore, J., 2012. Generic indicators for loss of resilience before a tipping point leading to population collapse. *Science* 336, 1175–1177.
- Dakos, V., van Nes, E.H., Donangelo, R., Fort, H., Scheffer, M., 2010. Spatial correlation as leading indicator of catastrophic shifts. *Theor. Ecol.* 3, 163–174.
- Dakos, V., Kéfi, S., Rietkerk, M., van Nes, E.H., Scheffer, M., 2011. Slowing down in spatially patterned ecosystems at the brink of collapse. *Am. Nat.* 177, 154–166.
- Dakos, V., Carpenter, S.R., Brock, W.A., Ellison, A.M., Guttal, V., Ives, A.R., Kéfi, S., Livina, V., Seekell, D.A., Van Nes, E.H., et al., 2012. Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. *PLoS One* 7, e41010.
- Dekker, S.C., Rietkerk, M., Bierkens, M., 2007. Coupling microscale vegetation-soil water and macroscale vegetation-precipitation feedbacks in semiarid ecosystems. *Glob. Change Biol.* 13, 671–678.
- Dijkstra, H.A., 2011. Vegetation pattern formation in a semi-arid climate. *Int. J. Bifurc. Chaos* 21, 3497–3509.
- Drake, J.M., Griffen, B.D., 2010. Early warning signals of extinction in deteriorating environments. *Nature* 467, 456–459.
- Guttal, V., Jayaprakash, C., 2007. Impact of noise in bistable ecological systems. *Ecol. Model.* 201, 420–428.
- Guttal, V., Jayaprakash, C., 2009. Spatial variance and spatial skewness: leading indicators of regime shifts in spatial ecological systems. *Theor. Ecol.* 2, 3–12.
- Hirota, M., Holmgren, M., Van Nes, E.H., Scheffer, M., 2011. Global resilience of tropical forest and savanna to critical transitions. *Science* 334, 232–235.
- Kéfi, S., Rietkerk, M., Alados, C., Pueyo, Y., Papanastasis, V., ElAich, A., De Ruiter, P., 2007. Spatial vegetation patterns and imminent desertification in Mediterranean arid ecosystems. *Nature* 449, 213–217.
- Kéfi, S., Dakos, V., Scheffer, M., Van Nes, E.H., Rietkerk, M., 2013. Early warning signals also precede non-catastrophic transitions. *Oikos* 122, 641–648.
- Mheen, M., Dijkstra, H.A., Gozolchiani, A., den Toom, M., Feng, Q., Kurths, J., Hernandez-Garcia, E., 2013. Interaction network based early warning indicators for the Atlantic MOC collapse. *Geophys. Res. Lett.* 40, 2714–2719.
- Rietkerk, M., Dekker, S.C., de Ruiter, P., van de Koppel, J., 2004. Self-organized patchiness and catastrophic shifts in ecosystems. *Science* 305, 1926–1929.
- Scheffer, M., Carpenter, S., Foley, J.A., Folke, C., Walker, B., 2001. Catastrophic shifts in ecosystems. *Nature* 413, 591–596.
- Scheffer, M., Bascompte, J., Brock, W.A., Brovkin, V., Carpenter, S.R., Dakos, V., Held, H., Van Nes, E.H., Rietkerk, M., Sugihara, G., 2009. Early-warning signals for critical transitions. *Nature* 461, 53–59.
- Shnerb, N.M., Sarah, P., Lavee, H., Solomon, S., 2003. Reactive grass and vegetation patterns. *Phys. Rev. Lett.* 90 .
- Van Nes, E.H., Scheffer, M., 2007. Slow recovery from perturbations as a generic indicator of a nearby catastrophic shift. *Am. Nat.* 169, 738–747.
- Viebahn, J., Dijkstra, H.A., 2014. Critical transition analysis of the deterministic wind-driven ocean circulation – a flux-based network approach. *Int. J. Bifurc. Chaos* 24 .
- Wissel, C., 1984. A universal law of the characteristic return time near thresholds. *Oecologia* 65, 101–107.
- Yin, Z., Dekker, S.C., van den Hurk, B., Dijkstra, H.A., 2014. Bimodality of woody cover and biomass caused by vegetation structures in West Africa. *Earth Syst. Dynam. Discuss.* 5, 83–120.