

Identifying vegetation patterns for a qualitative assessment of land degradation using a cellular automata model and satellite imagery

Hediye Yarahmadi ^{1,*} Yves Desille ^{1,2,†} John Goold ^{1,‡} and Francesca Pietracaprina ^{1,§}

¹*School of Physics, Trinity College Dublin, Dublin 2, Ireland*

²*Université Paris-Saclay, 91405 Orsay, France*

 (Received 28 September 2023; revised 6 April 2024; accepted 29 July 2024; published 27 August 2024)

We aim to identify the spatial distribution of vegetation and its growth dynamics with the purpose of obtaining a qualitative assessment of vegetation characteristics tied to its condition, productivity and health, and to land degradation. To do so, we compare a statistical model of vegetation growth and land surface imagery derived vegetation indices. Specifically, we analyze a stochastic cellular automata model and data obtained from satellite images, namely using the normalized difference vegetation index and the leaf area index. In the experimental data, we look for areas where vegetation is broken into small patches and qualitatively compare it to the percolating, fragmented, and degraded states that appear in the cellular automata model. We model the periodic effect of seasons, finding numerical evidence of a periodic fragmentation and recovery phenomenology if the model parameters are sufficiently close to the model's percolation transition. We qualitatively recognize these effects in real-world vegetation images and consider them a signal of increased environmental stress and vulnerability. Finally, we show an estimation of the environmental stress in land images by considering both the vegetation density and its clusterization.

DOI: [10.1103/PhysRevE.110.024136](https://doi.org/10.1103/PhysRevE.110.024136)

I. INTRODUCTION

Land degradation is a complex environmental process involving the loss of biological activity and economic productivity. It is caused both by peculiarities of the land and climate, and by unadapted human activity. Specifically, desertification is a transition that involves land degradation of drylands, resulting in the transformation of productive land into arid areas [1–4]. Drylands [5], covering approximately 41% of the Earth's land surface, are characterized by limited and fluctuating rainfall, exerting an environmental pressure, or stress, on soil and vegetation growth [6]. Several areas in the world, including Europe—which is the geographical focus of this work—have seen an increase in dryland extent [7].

Monitoring large land areas is fundamental to understand the progression of these environmental changes. To this aim, numerous studies [8–10] have explored the relationships among drylands, vegetation patterns, and external stresses, using a number of models and data sources. These include remote sensing [11], which uses satellite or airborne sensors to gather data on vegetation cover and health; geographic information systems [12] for analyzing spatial patterns and changes in vegetation; ecohydrological modeling [13] to examine the interaction between vegetation and water resources; agent-based modeling [14] to simulate individual-level behavior and its impact on ecosystems; and

species distribution modeling [15] to predict plant species distribution under different environmental conditions. A statistical physics approach, alongside these methods, also contributes to understanding vegetation distribution and dynamics, as well as the role of disorder in the vegetation processes. Particularly, vegetation patchiness is a valuable tool for assessing degradation risk and detecting early warning signals associated with it [16–18]. Disorder enters the vegetation dynamics through environmental fluctuations, stochastic processes such as weather patterns, disturbances (e.g., wildfires and human activities), and spatial heterogeneity in soil composition or topography. The presence of disorder affects both the existence and the way the transitions occurs, making them more gradual or continuous [19–21]. This, in turn, may affect the environmental resilience.

Our objective in this work is to describe the structure of vegetation cover through a numerical model of vegetation growth dynamics, and to conduct an analysis of satellite image data to recognize patterns of vegetation patchiness that are signs of land degradation in progress. We use a stochastic cellular automaton (SCA) model [3,4,22,23], which incorporates events such as vegetation mortality, survival and propagation. This SCA model is born out of individual plant lifecycle considerations (e.g., propagation of seeds, new plant colonization, plant competition, etc). These effects are coarse-grained into stochastic state transitions of macroscopic cells which are arranged in a grid, covering large land areas. The state changes in the cells are able to generate patterns observed in arid ecosystems, such as gaps, stripes, and labyrinth-like structures [3,4]. It also presents two phase transitions as the environmental stress (or vegetation mortality) changes: a desertification transition where vegetation disappears, and a

*Contact author: hediye.yarahmadi@tcd.ie

†Contact author: yves.desille@universite-paris-saclay.fr

‡Contact author: gooldj@tcd.ie

§Contact author: francesca.pietracaprina@tcd.ie

percolation transition where the vegetation breaks down into clusters and which acts as a precursor to the desertification transition [24–26].

A number of data sources are available to survey and identify changes in vegetation and land ecosystems. These include satellite images from several Landsat missions, Sentinel, MODIS, SPOT, and RapidEye [27], as well as databases of specific-purpose processed data (e.g., several vegetation intensity indices in the Copernicus Land [28] database). The availability of such data has enabled the investigation of phase transitions in drylands [29]. Here, we aim to verify whether the processes in the SCA—namely the clusterization processes acting under increased environmental stress—can be qualitatively detected in satellite image data at a large scale, focusing on several areas across Europe.

The work is structured as follows: Sec. II provides an overview of the theoretical and numerical methods employed, including the model used for simulations and the incorporation of seasonal effects. In Sec. III, the analysis is applied to recent satellite images of a selection of European countries. Section IV connects the SCA simulated process and the observed land data. Finally, our conclusions are outlined in Sec. V.

II. VEGETATION DYNAMICS AND CLUSTERIZATION

In order to capture the vegetation patchiness dynamics, we consider the already mentioned SCA model introduced in Ref. [4] and used to study the dynamics of vegetation patterns in ecological systems. In the following, we will provide a brief overview of the model, while we refer to the original papers for a comprehensive description [4,22].

We represent the vegetation by means of a three-state SCA model. In a $L \times L$ square lattice, each cell can exist in one of three states: (+) is a vegetation-covered state (living cell); (0) is an empty state, available for colonization (dead cell); and (−) is a degraded state (degraded cell). Cells undergo state transitions with transition rates W_{ij} . For instance, W_{0+} represents the rate of transition from an empty to a vegetated state. Not all transitions have a nonzero probability. For instance, a degraded cell must undergo recovery before it can be colonized. At the same time, only an empty cell is vulnerable to degradation. Consequently, transitions between a degraded state (−) and a vegetated state (+) are prohibited. The rates at which the allowed transitions take place are as follows:

$$W_{+0} = m, \quad (1)$$

$$W_{0+} = [\delta\rho_+ + (1 - \delta)q_{+|0}](b - c\rho_+), \quad (2)$$

$$W_{0-} = d, \quad (3)$$

$$W_{-0} = r + fq_{+|-}. \quad (4)$$

The system's dynamic evolution is governed by a Markov chain, with Eqs. (1)–(4) representing mortality, colonization, degradation, and recovery processes, respectively. Transitions from and to the dead (0) state are influenced by q_{ij} : Given a cell in state j , this represents the fraction of its nearest neighbors in state i . Besides the contribution of the neighboring cells, Eqs. (1)–(4) involve additional parameters related

to the lifecycle of plants, with the following interpretations: δ is the proportion of seeds dispersed by wind, animals, etc.; b is the colonization parameter, which accounts for various intrinsic properties of a vegetated cell, such as seed production rate, seed survival, germination, and survival probabilities (not including global competition effects); c is the strength of global competition effects; d represents the rate of soil degradation, incorporating intrinsic soil characteristics, climatic factors, and anthropogenic influences; f is the local facilitation parameter, describing cooperative interactions among plants and positive feedback between soil and plants; and finally, r is the spontaneous regenerative rate of a degraded cell in the absence of vegetation covering the neighboring cells. The parameter values used in the following analysis are as follows: $b = 0.6$, $c = 0.3$, $d = 0.2$, $\delta = 0.1$, $f = 0.9$, and $r = 0.0004$. These values reflect typical processes in semi-arid ecosystems (particularly: intermediate colonization, low competition, intermediate soil degradation, high facilitation, and low spontaneous regeneration) and align with those used in previous studies in order to reflect real field data simulations [4,23,24]. A wide range of parameter values has been analyzed in the literature with numerical simulations in order to qualitatively reproduce different climatic regimes [3]. In our analysis, the control parameter for the model is the mortality rate m , which represents the intensity of external stress.

The process to simulate vegetation dynamics is as follows: We start by initializing a randomized lattice configuration consisting of alive, barren, and dead cells. The system evolves through the transition probabilities given by Eqs. (1)–(4). We note that, while Eq. (1) and (4) can be applied independently, Eq. (2) and (3) have to satisfy the condition $W_{00} + W_{0+} + W_{0-} = 1$, on whether the zero cells will transition into the alive or barren states or preserve their current situation. After a transient dynamics due to the initialization of the system with a random configuration, the system converges to an equilibrium which depends on the chosen parameter set. This initial transient dynamic is discarded (typically 5000 iterations); once the system reaches equilibrium, we collect the vegetation fraction and vegetation cluster sizes for each iteration step. Throughout this work, unless explicitly specified, the simulations are conducted on square $L \times L$ lattices with a linear size of $L = 100$ with periodic boundary conditions. This system size is comparable with the sample size chosen for the satellite images in Sec. III and has weak finite size effects [24]. The time-series data typically consists of at least 10^4 records, but near the desertification threshold, the number of steps is increased up to 10^6 due to critical slowing down effects.

The properties of vegetation clusters are investigated through the alive vegetation density $\rho_+ = \frac{N_+}{L^2}$, where N_+ is the number of alive cells, and the size of the largest cluster of alive cells C_+ , which is the largest vegetation cluster size divided by the total number of lattice cells. We analyze a percolation transition in response to the external stress parameter m .

A. Numerical simulations of vegetation dynamics

In this section, we investigate the model for different values of the mortality rates m . The inset in Fig. 1(a) shows the

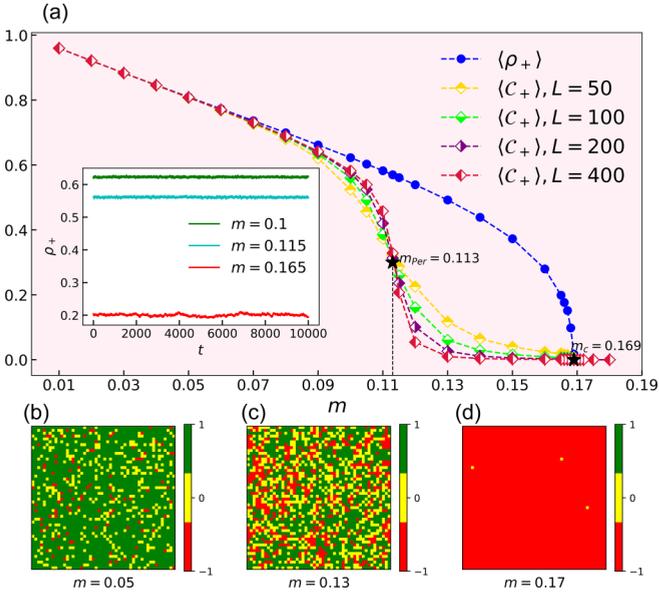


FIG. 1. Evolution of vegetation dynamics. Panel (a): The average of the vegetation density $\langle \rho_+ \rangle$ for $L = 100$ and the relative size of the largest cluster $\langle C_+ \rangle$ for $L = 50, 100, 200, 400$ as a function of the mortality rate m . The markers represent the percolation threshold $m_{per} = 0.113$, which is determined from the crossing point for the various system sizes, and the degradation transition point $m_c = 0.169$. Inset of (a): The typical evolution of the living cell density $\rho_+(t)$ for several values of the mortality rate m . The linear lattice size is $L = 100$, and the time unit corresponds to one iteration step. Panels (b)–(d): Vegetation distribution in a lattice of size $L = 50$ at (b) low ($m = 0.05$), (c) intermediate ($m = 0.13$), and (d) high ($m = 0.17$) mortality. Green cells are vegetated, yellow cells are empty, and red cells are degraded areas.

evolution of the living cell density $\rho_+(t)$, for several values of m . After a short transient (not shown), the system fluctuates around an equilibrium value $\rho_+(m)$. With the increase of external stress m , the average density of vegetated cells decreases, until a continuous phase transition occurs at $m_c = 0.169$ [see the main panel of Fig. 1(a)] corresponding to a desertification transition. Additionally, we determine the size of the largest cluster of vegetated cells $C_+(m)$. As m increases, at finite sizes C_+ has a sharp crossover that corresponds to a percolation transition which acts as a precursor to desertification [4,23,24]. Considering different system sizes allows to find the percolation transition point at $m_{per} = 0.113$ as the crossing point of the curves at different L . See Sec. I of Supplemental Material (SM) Ref. [30] for information about the transitions.

Thus, two distinct transitions are involved in land degradation processes [24,31–33]. The first transition, a percolation transition in the living vegetation clusters, serves as an early warning sign for the second transition, which is associated with desertification as the fraction of living vegetation goes to zero. Both of these transitions have been observed to occur in ecosystems [26]. The bottom panels of Fig. 1 shows snapshots of the model at different stages: healthy and percolating [Fig. 1(b)], vulnerable and broken into small clusters [Fig. 1(c)], and degraded [Fig. 1(d)].

B. Modeling the seasonal effects

So far, we have considered a SCA model with a constant mortality rate. In order to include the effect of seasonal cycles in real ecosystems, we introduce a time-dependent, periodic mortality rate $m(t)$. We aim to model a simplified time dependence in the environmental stress. To do so, we consider an asymmetric periodic function representing a fast increase in environmental stress, which is reminiscent of the phenomenological temperature and precipitation time series in semiarid areas [34].

Specifically, we are interested in investigating whether an ecosystem that seasonally is put in a situation of high environmental stress is able to recover, as well as what conditions need to be satisfied to avoid a permanent transition to a desertified state.

In our analysis, starting from a randomly generated configuration, we first let the system equilibrate with a constant mortality parameter $m_{in} = m_0$. At time t_{in} we apply a time-dependent $m(t)$ given by the following asymmetric functional form, which we choose as it is both simple in terms of number of parameters, and resembles phenomenologically the experimental time series of air temperature in the geographic areas considered:

$$m(t) = m_0 + A \sin[2\pi/T t + k \sin(2\pi/T t)] \quad t > t_{in}, \quad (5)$$

where $m(t)$ consists of a constant component, m_0 , representing the average environmental stress across the seasons. The additional parameters in the oscillatory term are: the amplitude A , representing the intensity of the seasonal effect, the period T representing the duration of a year in the simulation's arbitrary time units, and an asymmetry parameter k representing fast season changes particularly from winter to summer. Here we consider $T = 5000$ and $k = 0.2$. Note that T is sufficiently large to be greater than the timescale of the internal fluctuations of ρ_+ . The asymmetry k is chosen to resemble fits of experimental temperature data. The initial value m_0 is set to several values near the percolation and degradation transition points, $m_{per} = 0.113$ and $m_c = 0.169$, respectively. Values just below m_c are specifically interesting in order to determine the effects of periodic increases in environmental stress and whether they can drive a transition to a nonrecoverable dead or barren state. The numerical computation is performed for $m_0 = 0.08$ to 0.17 , varying A , and 10^6 time steps.

We investigate the impact of increasing the amplitude A on the behavior of ρ_+ and C_+ . Fig. 2(a) shows the effects of different amplitude values for $m_0 = 0.08$. ρ_+ shows a periodical response to $m(t)$, with a very small response delay in the system size considered. C_+ closely follows and overlaps with ρ_+ for the majority of the time, showing that most alive cells belong to the same contiguous cluster. For sufficiently high amplitudes, C_+ temporarily dips to lower values only near the minimums of $m(t)$, with a fast recovery to values overlapping ρ_+ .

In Fig. 2(b) we show results for a higher mortality which is close to the percolation transition value and periodically sent above it. In this case as well, C_+ and ρ_+ both show a periodical response to $m(t)$. For low amplitudes, C_+ is at all times significantly lower than ρ_+ while still being nonzero. For sufficiently high amplitude, however, it quickly goes to

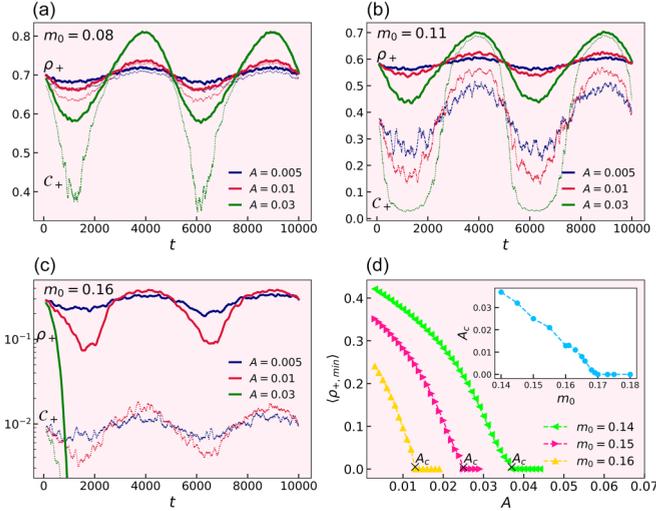


FIG. 2. Evolution of vegetation dynamics for time-dependent mortality rate $m(t)$. Panels (a), (b), and (c): Vegetation density ρ_+ (solid lines) and largest cluster size C_+ (dotted lines) as a function of time, for initial mortality $m_0 = 0.08, 0.11,$ and 0.16 respectively. Data are smoothed by means of a moving average with a window size of 100. Panel (d): Mean value of the minima of the vegetation density $\langle \rho_{+,min} \rangle$ as a function of the amplitude A . We mark the crossover amplitudes $A_c(m_0, T)$ where the curves drop to 0. Inset of panel (d): Crossover amplitude A_c as a function of m_0 .

very small values during the lower half-period of the oscillation, suggesting that the system has been temporarily driven to the highly fragmented and nonpercolating regime; although C_+ stays near 0 for a substantial part of the lower half-period, a similar rapid recover happens in the upper half-period, up to reaching values close to ρ_+ only near the maximum of $m(t)$.

Finally, in Fig. 2(c) we show the results for a system oscillating around $m_0 = 0.16$, that is a point close to the transition value m_c . Interestingly, we show that even when the time-dependent mortality rate goes above the critical value, the system can recover its alive vegetation cover ρ_+ and continue to respond to the oscillating $m(t)$, as long as the amplitude of the oscillation, and thus the time spent in a condition where $m(t) > m_0$, is sufficiently small. We find that there is a crossover amplitude A_c such that for $A > A_c$ the system enters a state of almost all dead or degraded cells and is no longer able to recover, even though the environmental stress $m(t)$ strongly decreases afterwards. The drop to a dead or degraded configuration happens as the system fluctuates near the crossover point with increasingly long metastable oscillating states. At this value of m_0 , the system is already firmly in the highly fragmented regime and C_+ , while responding to the oscillations alongside ρ_+ (as long as the system is not fully degraded and ρ_+ is not identically 0), is an order of magnitude smaller.

Next, we extract the crossover amplitude A_c for several values of average mortality m_0 . To do so, we look for the time series which identically reach 0. For each value of A , we determine the value of $\rho_{+,min}$, the minimum vegetation density for each period, which happens at the time of highest seasonal stress for each period. We use a cubic spline interpolation to determine the minima for each period. $\rho_{+,min}(t)$ reaching 0

quantitatively marks the drop to a nonrecoverable degraded configuration. Figure 2(d) shows $\langle \rho_{+,min} \rangle$, the average minimum across all periods, as a function of the amplitude A for $m_0 = 0.14, 0.15,$ and 0.16 . The crossover to full degradation is observed as $\langle \rho_{+,min}(A) \rangle$ goes to zero at $A = A_c$. For $A > A_c$, ρ_+ will be asymptotically zero after an unstable transient. Finally, in the inset of Fig. 2(d), we plot these crossover amplitude points $A_c(m_0)$. A_c is approximately linear in m_0 until $m_0 = 0.17 \approx m_c$, at which point it reaches zero. Beyond this point, for $m_0 \geq 0.17$, A_c is identically zero.

The system’s behavior and crossover to desertification are influenced by the period T , with shorter oscillations (compared to the system’s own intrinsic timescales) resulting in a more robust alive phase.

III. FINDING VEGETATION CLUSTERS THROUGH SATELLITE DATA

Data obtained from Earth observation satellites can be analyzed to recognize the patterns of vegetation cover in real ecosystems. In this section, we recognize areas of fragmented vegetation qualitatively matching the phases exhibited by the cellular automaton model of Sec. II.

We use two vegetation indices commonly used to detect vegetation from satellite imagery: the normalized difference vegetation index (NDVI) [35] and the leaf area index (LAI) [36]. NDVI is widely used in ecosystem monitoring due to its simple formulation. It measures the vegetation greenness by calculating the ratio of spectral reflectances in near-infrared and red light. Green, living plants have a high reflection in the near-infrared and high absorption in red and visible light frequencies, which is a markedly different response than bare soil, water, snow, or urbanized areas. For this reason, it is commonly used in remote sensing to qualitatively assess vegetation health and density in a specific area. We also use the LAI data, which is defined as half the total area of living vegetation elements in the canopy per unit of ground area. This quantity is obtained either through direct measurement (performed locally on a sampling basis) or through indirect methods such as image analysis. All of the data for our analysis are extracted from the European database *Copernicus Global Land Service*, which provides global vegetation data products [28].

The NDVI and LAI images were acquired from the Copernicus Land dataset for three days per month: For NDVI on the 1st, 11th, and 21st day of each month spanning the time period from 2014 to 2021, while for LAI on the 10th, 20th, and last day of each month, from 2014 up until August 2020 (there are gaps in the LAI data during the winter and autumn seasons, due to insufficient illumination in satellite observations). Both datasets have a resolution of 300 m and the data used are composed of a mosaic corrected for atmospheric differences, including the removal of cloud coverage. Our analysis focuses on specific regions, namely in France, Germany, Ireland, Spain, and Greece. We note that drylands are particularly present in Spain and Greece. We preprocessed the images in the dataset in order to normalize the individual image pixel data, a grayscale image with values from 0 to 255, to $[0,1]$ for land areas. Additionally, we exclude pixels corresponding to sea, lakes, rivers and any water bodies that have

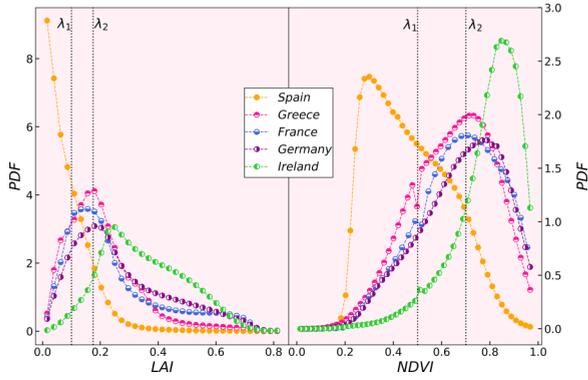


FIG. 3. Probability distribution of normalized NDVI and LAI values across examined land regions. The data are aggregated over samples taken at different dates and times. As a reference, the threshold values λ_1 and λ_2 for Greece, France, and Germany are marked as vertical dotted lines.

been premarked by the Copernicus dataset preprocessing, as well as those otherwise detectable through NDVI and LAI intensity values: indeed, some smaller water areas need to be manually excluded, which we do through a local thresholding method (we use the Threshold-Sauvola approach [37], see the SM Sec. II [30] for more detail). We note that the values of \mathcal{C}_+ and ρ_+ which will be computed from the experimental data and used in the following analysis have been normalized to the land area only, excluding water bodies.

A. Spatial distribution of vegetation and breakdown into clusters

We now analyze the satellite images of vegetation in order to identify a breakdown into disconnected clusters. Specifically, here we look for the breakdown in connectivity which in the vegetation model used in Sec. II translates to a percolation transition. To do so, we match the pixel intensity in the satellite image to three possible states, reminiscent of the SCA model. We thus introduce a simplification to discretize the data, which for a single pixel corresponds to an intensity between 0 and 1, based on two chosen threshold parameters. This effectively maps low intensity pixels to a degraded cell, medium intensity to an empty cell, and high intensity to an alive cell. We are then able to consider the fraction of alive vegetation ρ_+ and the largest vegetation cluster \mathcal{C}_+ analogously to the analysis of the SCA model in Sec. II.

Determining the threshold values is tied to the typical values of the NDVI and LAI indices, which itself are qualitative in nature, and their interpretations in terms of plant health. As a starting point, we consider the probability distribution function of NDVI and LAI pixels intensity across the analyzed regions, shown in Fig. 3. It is worth noting that LAI is a curated index, deriving from a mix of indirect and direct measurements, and it is not directly and quantitatively comparable to NDVI. Both indices clearly signal the presence of vegetation, with different distributions due to the climate and morphological characteristics of the country groups: especially intense values in Ireland, and less vegetation intensity in Spain.

We introduce the two cut-off parameters λ_1 and λ_2 , which classify the pixels as vegetated or alive for values $> \lambda_2$, empty or dead for values between λ_1 and λ_2 and degraded for values $< \lambda_1$. The specific values of λ_1 and λ_2 are chosen so that they are respectively below and around the typical values of the indices. Note that for different regions we may choose different λ_1 and λ_2 values due to differences in vegetation type and thus NDVI or LAI intensity; this is particularly relevant for the selected areas in Spain and Ireland. In the following analysis, we use the values $\lambda_1 = 0.5$, $\lambda_2 = 0.7$ (NDVI) and $\lambda_1 = 0.1$, $\lambda_2 = 0.175$ (LAI) for Greece, France, and Germany, all of which have similar NDVI and LAI intensities. We note that the qualitative results of the following analysis will not depend on the precise values chosen (see SM Sec. V [30] for more details).

The SCA model considered in Sec. II describes lifecycle processes at the scale of a single plant; it is however also suitable to coarse-graining, as the mortality, reproduction or colonization, degradation and recovery processes can be applied collectively and averaged over land areas. Our aim is to verify that the overall behaviors of in the simplified SCA model can be qualitatively detected in vegetation even at the rather coarse resolution of 300m of the considered satellite data.

The NDVI and LAI images include, as well as water bodies which are identified and excluded from each sample as described above, areas which are blocked from vegetation growth, and therefore locked in a “degraded” state, due to anthropic causes (e.g., cities). As well, additional human activities such as farming and grazing will introduce systematic effects in the classification of the image pixels into the three states. In the following analysis, we neglect these effects on the analysis of vegetation clusterization. Moreover, we neglect the internal heterogeneity of the data within each region, namely the diversity in soil type, local climate, prevalent plant type and any other small-scale properties. This choice is due to both the low resolution of the source image data and of the selected samples (which, as outlined in Sec. III B, will have a linear size of 30 km), and the simplified coarse-grained nature of the analysis of vegetation dynamics that will be performed.

B. Analysis of NDVI time series

Here we focus on NDVI data specifically for areas in France (first two samples) and Greece (last sample), which respectively have a higher or lower prevalence of semiarid conditions. The areas that we consider in each different country is split in samples of 100×100 pixels (that is, 30×30 km). For each sample, we analyze the fraction of alive vegetation ρ_+ and the relative size of the largest vegetation cluster \mathcal{C}_+ over time. Among the samples considered, we see a range of phenomenology, which we analyze qualitatively and for which we show some examples in Fig. 4.

In Fig. 4(a) we consider a sample with very high maximum values of $\rho_+(t)$ and $\mathcal{C}_+(t)$ and where, for the majority of the time, there is an overlap between $\rho_+(t)$ and $\mathcal{C}_+(t)$. This indicates that nearly all the vegetated pixels belong to the largest cluster of vegetation. This situation is akin to the phase at a low mortality rate in the simulated vegetation dynamics in Sec. II, which we associate to healthy vegetation and low

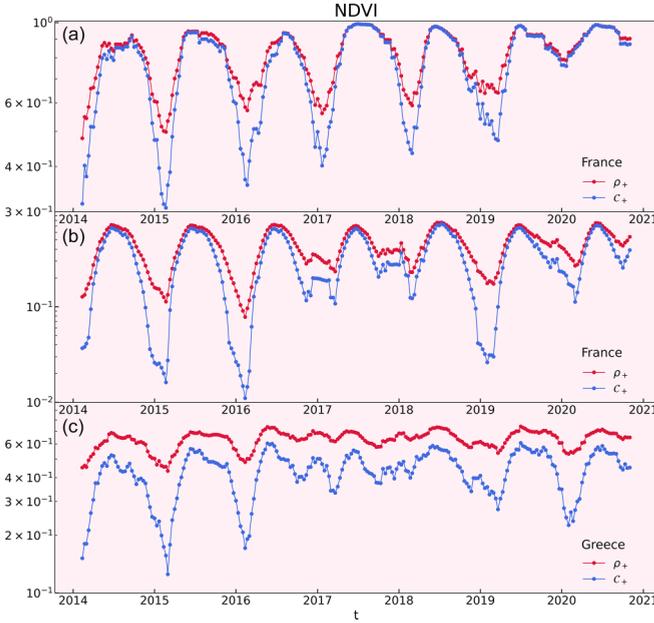


FIG. 4. Comparison of ρ_+ and \mathcal{C}_+ as a function of time for three samples [(a) and (b)] in France and (c) in Greece. We use NDVI data for the years 2014 to 2021. The data are smoothed by moving average over a window of size 10.

environmental stress. We note a seasonal periodic behavior in both $\rho_+(t)$ and $\mathcal{C}_+(t)$ with a temporary dip in \mathcal{C}_+ during the autumn or winter season. This is qualitatively similar to the behavior in Fig. 2(a) for low average mortality rate in the presence of strong seasonal oscillations.

In Fig. 4(b) we show a sample exhibiting a more marked difference between $\rho_+(t)$ and $\mathcal{C}_+(t)$, which only overlap at their maximum points. This situation is similar to Fig. 2(b), where a mortality rate close to the percolation transition value is periodically pushed above it.

Finally, in Fig 4(c) we show a sample with consistently different $\rho_+(t)$ and $\mathcal{C}_+(t)$, which themselves exhibit low values. This is qualitatively comparable with Fig. 2(c), where the average mortality rate is between the percolation and the degradation point and, if seasonal oscillation have a sufficiently small amplitude, $\mathcal{C}_+(t)$ oscillates as $\rho_+(t)$ but around a much smaller average value without pushing the system to full degradation.

We have repeated this analysis for the additional areas that have been considered in Spain, Germany, and Ireland. As for Greece and France, respectively, we are able to identify a higher prevalence of situations of stress associated with the semiarid climate for Spain, and generally its absence in Germany and Ireland.

IV. MAPPING VEGETATION STRESS

In this section, we qualitatively compare the vegetation fraction and clusterization data obtained from the SCA model simulations with the experimental one obtained from satellite images. We aim here to identify the areas that show vulnerability: These areas will have a relatively dense vegetation (high ρ_+), but are noncontiguous and broken down into many

small clusters (low \mathcal{C}_+). As shown in Sec. II, this is a percolation transition that precedes the situation where vegetation is not surviving. Increased vulnerability is especially detectable in many areas during the autumn and winter months, as observed by the analysis of the time series of ρ_+ and \mathcal{C}_+ in Sec. III B.

In order to better visually identify the situation in which each image sample lies, we consider a simultaneous scatter plot of ρ_+ and \mathcal{C}_+ . We first simulate the SCA model for several (non-time-dependent) values of the mortality rate m and we construct the curve $(m, \overline{\rho_+}, \overline{\mathcal{C}_+})$. The averages $\overline{\rho_+}$ and $\overline{\mathcal{C}_+}$ are computed by taking the centroid of the scatter cloud obtained from a large ($10^4 - 3 \times 10^5$) number of snapshots of the equilibrated system. Interestingly, the projection of the curve in the ρ_+, \mathcal{C}_+ plane is independent of the value of the parameters of the SCA model for a wide range of realistic parameters (see Sec. III of the SM [30]). Given their universality, we use this simulated curve as a reference for the ρ_+ and \mathcal{C}_+ computed from the experimental image data.

We then again consider the subimage samples of size 30×30 km and consider the values of ρ_+ and \mathcal{C}_+ for each sample. We summarize again here the three main scenarios. First, when we encounter areas with high ρ_+ and \mathcal{C}_+ , it is an indication of regions with low mortality rate: this corresponds to a large portion of the image covered by unbroken, contiguous percolating vegetation. The second case corresponds to regions characterized by low ρ_+ and \mathcal{C}_+ values. These areas thus have low vegetation coverage, which include genuinely degraded land, rocky terrains, and urban areas. The last case involves instances of high ρ_+ values but low \mathcal{C}_+ values. This combination suggests that despite the presence of relatively dense vegetation, the environmental stress and thus the mortality rate is high, causing the vegetation to break down into small clusters, as a precursor to full degradation. The data points obtained from satellite images can be compared with two guidelines. The first is the already mentioned curve numerically simulated with the SCA model at various values of the mortality rate; the other is the line $\rho_+ = \mathcal{C}_+$, which corresponds to a fully connected vegetation. This condition represents a low mortality rate situation, and all data points will lie in the $\rho_+ > \mathcal{C}_+$ sector.

We show in Fig. 5 a scatter plot of ρ_+ against \mathcal{C}_+ for all NDVI image samples obtained for France [Fig. 5(a)] and Greece [Fig. 5(b)] from 2014 to 2021; we additionally show the results for the LAI dataset in the inset. In this figure we display all available sample points, aggregated for all dates, in order to highlight their distribution in the (ρ_+, \mathcal{C}_+) plane. For additional geographical areas, we refer to the SM Sec. IV [30]. The data points in Fig. 5 approximately lie between the numerical curve obtained from the numerical simulations of the SCA model and the diagonal $\mathcal{C}_+ = \rho_+$. In order to understand the significance in terms of environmental stress, we should first highlight the systematic effect of permanently unvegetated pixels in each sample. These pixels might correspond to locations not covered by soil, e.g., urban or rocky areas. These reduce the effective available size of each sample, in a way that is not accounted by the model (as these pixels are not recoverable even in the most favorable conditions). This explains the presence of data points close to the diagonal for even mid and low values of \mathcal{C}_+ . As well, points spread

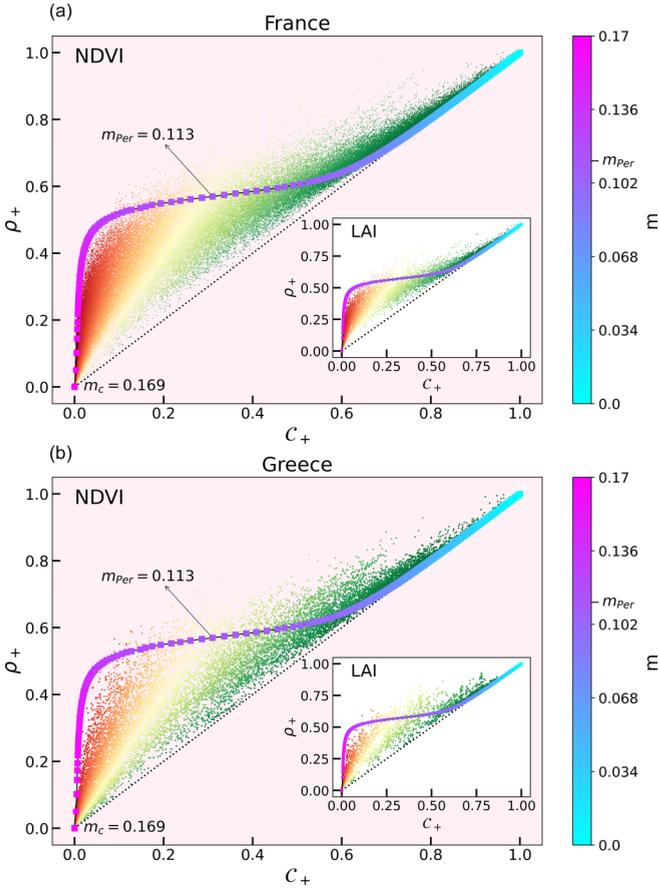


FIG. 5. Vegetation density ρ_+ and the relative size of the largest cluster C_+ for 30×30 km samples. The main panels (a) and (b) present the cumulative NDVI data for all available dates spanning the years 2014–2021, showing the results for France and Greece, respectively. The points are color-coded with C_+/ρ_+ representing a qualitative classification of the degradation in each sample. The diagonal $\rho_+ = C_+$ is shown, as well as the (m, ρ_+, C_+) data points obtained from the numerical simulation of the SCA model with the parameters' values outlined in Sec. II; the values of m are represented through a color scale. Insets: Results from the LAI data for all available dates spanning the years 2014 to 2021, for France (a) and Greece (b).

between the diagonal and the numerical curve correspond to increasingly degraded samples with reduced effective size. One could attempt to quantify this systematic effect by introducing stochastically arranged blocked pixels in the numerical model; we leave this detailed analysis to future work. Here, we give a simple qualitative assessment of vulnerability by evaluating the ratio C_+/ρ_+ . Given its value, we classify the data points in three levels of vegetation health. The healthiest vegetation samples are located close to the diagonal and in the top right area (green points), where the vegetation is above the percolation transition. The points in the bottom left area (red points), close to the numerical curve, are samples whose vegetation is broken down into small nonpercolating clusters, which therefore have the lowest vegetation health. The points in the central area have a reduced effective soil availability; among them, the points close to the diagonal (green points) still present contiguous vegetation, while, farther away from

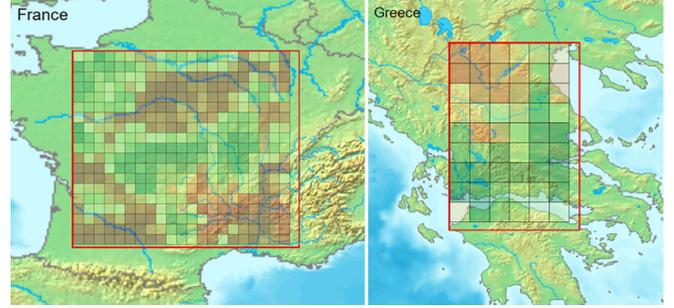


FIG. 6. Qualitative assessment of land degradation, obtained by comparing each sample's vegetation density and largest vegetation cluster size. The color coding refers to the value of C_+/ρ_+ for each sample, representing a qualitative classification of its degradation. We show here France (left) and Greece (right) on 11 March 2020.

the diagonal and closer to $C_+ = 0$, the samples show intermediate (yellow points) to low (red points) vegetation health. In Fig. 5 we mark the different classifications for all points with green-yellow-red colors, for which we consider these approximate threshold values for C_+/ρ_+ : 0.3 for red to yellow and 0.7 for yellow to green. Finally, we show this qualitative classification by color coding each sample with its C_+/ρ_+ value in Fig. 6 for areas in France and Greece, superimposed to a map of the land that has been analyzed, where each square correspond to one 30km^2 sample, as defined in Sec. III B.

V. CONCLUSIONS

In this work, we have analyzed the spatial distribution of vegetation through a numerical stochastic model and through the analysis of land data from satellite images. The model is known to exhibit three phases, namely showing contiguous, fragmented, and degraded vegetation, and the corresponding two transitions, a percolation and a degradation transition which happen as the environmental stress increases. Considering samples of 30 km in size, we looked at the vegetation fraction and the fraction of contiguous vegetation in each sample with the aim of identifying these processes in real land data.

Although the model has simplifications and the experimental data have relatively low resolution, we have shown that the model's description of the vegetation dynamics can be qualitatively compared to the land data. Particularly, we extended the model with a time-dependent environmental stress resembling the periodic seasonal changes. Some key aspects of its phenomenology can be recognized in the experimental time series for a variety of samples. Finally, the values of the vegetation fraction and the largest cluster size allowed us to offer a qualitative categorization of each sample between the three scenarios of contiguous, fragmented and degraded vegetation. As a future perspective, we highlight the potential usage of this classification as a tracker of land health (see an example of this use in the SM Sec. VI [30]). Namely, after accounting for seasonal variations, the time series of vegetation fraction and cluster size for a land area can be used to detect worsening or recover of vegetation over time based on the historical data presented in this work.

ACKNOWLEDGMENTS

F.P. has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant No. 838773. J.G. is supported by a SFI-Royal Society University Research Fellowship and acknowledges funding from European Research Council Starting Grant ODYSSEY (Grant No. 758403). This work has received support from ERC PoC "Emerald," Grant No.

101069222. Funded by the European Union. Views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Union or ERCEA. Neither the European Union nor the granting authority can be held responsible for them. The Irish Centre for High End Computing (ICHEC) has supported the computational aspects of this work through Grants No. tcpy205c and No. tcpy223c.

- [1] J. von Hardenberg, E. Meron, M. Shachak, and Y. Zarmi, *Phys. Rev. Lett.* **87**, 198101 (2001).
- [2] N. Shnerb, P. Sarah, H. Lavee, and S. Solomon, *Phys. Rev. Lett.* **90**, 038101 (2003).
- [3] S. Kéfi, M. Rietkerk, M. V. Baalen, and M. Loreau, *Theor. Popul. Biol.* **71**, 367 (2007).
- [4] S. Kéfi, M. Rietkerk, C. L. Alados, Y. Pueyo, V. P. Papanastasis, A. ElAich, and P. C. De Ruiter, *Nature (London)* **449**, 213 (2007).
- [5] The categorization of land regions into arid, semiarid, and dry subhumid climatic zones is based on aridity indices (e.g., Thornthwaite and Palmer Drought Severity Index) and Holdridge life zones, considering factors such as precipitation and potential evapotranspiration [38].
- [6] R. J. Hobbs and L. F. Huenneke, *Conserv. Biol.* **6**, 324 (1992).
- [7] J. Spinoni, J. Vogt, G. Naumann, H. Carrao, and P. Barbosa, *Int. J. Climatol.* **35**, 2210 (2015).
- [8] J. F. Reynolds, D. M. S. Smith, E. F. Lambin, B. Turner, II, M. Mortimore, S. P. Batterbury, T. E. Downing, H. Dowlatabadi, R. J. Fernández, J. E. Herrick *et al.*, *Science* **316**, 847 (2007).
- [9] O. E. Sala, F. Stuart Chapin, J. J. Armesto, E. Berlow, J. Bloomfield, R. Dirzo, E. Huber-Sanwald, L. F. Huenneke, R. B. Jackson, A. Kinzig *et al.*, *Science* **287**, 1770 (2000).
- [10] S. Archer, T. W. Boutton, and K. A. Hibbard, *Global Biogeochemical Cycles in the Climate System* (Elsevier, Amsterdam, 2001), pp. 115–137.
- [11] J. R. Jensen, *Remote Sensing of the Environment: An Earth Resource Perspective* (Pearson Education India, Chennai, 2009).
- [12] M. F. Goodchild, *Prog. Hum. Geogr.* **15**, 194 (1991).
- [13] C. L. Tague and L. E. Band, *Earth Interact.* **8**, 1 (2004).
- [14] V. Grimm, U. Berger, F. Bastiansen, S. Eliassen, V. Ginot, J. Giske, J. Goss-Custard, T. Grand, S. K. Heinz, G. Huse *et al.*, *Ecol. Model.* **198**, 115 (2006).
- [15] J. Elith, C. H. Graham, R. P. Anderson, M. Dudík, S. Ferrier, A. Guisan, R. J. Hijmans, F. Huettmann, J. R. Leathwick, A. Lehmann *et al.*, *Ecography* **29**, 129 (2006).
- [16] S. R. Levick and G. P. Asner, *Biol. Conserv.* **157**, 121 (2013).
- [17] K. M. Dahlin, G. P. Asner, and C. B. Field, *Proc. Natl. Acad. Sci. USA* **110**, 6895 (2013).
- [18] G. P. Asner, J. R. Kellner, T. Kennedy-Bowdoin, D. E. Knapp, C. Anderson, and R. E. Martin, *PloS ONE* **8**, e60875 (2013).
- [19] P. Villa Martín, J. A. Bonachela, S. A. Levin, and M. A. Muñoz, *PNAS* **112**, E1828 (2015).
- [20] P. Villa Martín, J. A. Bonachela, and M. A. Muñoz, *Phys. Rev. E* **89**, 012145 (2014).
- [21] F. Vazquez, J. A. Bonachela, C. López, and M. A. Muñoz, *Phys. Rev. Lett.* **106**, 235702 (2011).
- [22] S. Kéfi, M. B. Eppinga, P. C. de Ruiter, and M. Rietkerk, *Theor. Ecol.* **3**, 257 (2010).
- [23] S. Kéfi, V. Guttal, W. A. Brock, S. R. Carpenter, A. M. Ellison, V. N. Livina, D. A. Seekell, M. Scheffer, E. H. Van Nes, and V. Dakos, *PloS ONE* **9**, e92097 (2014).
- [24] R. Corrado, A. M. Cherubini, and C. Pennetta, *Phys. Rev. E* **90**, 062705 (2014).
- [25] A. Manor and N. M. Shnerb, *J. Theor. Biol.* **253**, 838 (2008).
- [26] J. von Hardenberg, A. Y. Kletter, H. Yizhaq, J. Nathan, and E. Meron, *Proc. R. Soc. B: Biol. Sci.* **277**, 1771 (2010).
- [27] K. Ose, T. Corpetti, and L. Demagistri, *Optical Remote Sensing of Land Surface: Techniques and Methods* (Elsevier, Amsterdam, 2016), pp. 57–124.
- [28] European Commission JRC, Copernicus Global Land Service, <https://land.copernicus.eu/global/products>.
- [29] F. Meloni, G. M. Nakamura, C. R. Granzotti, and A. S. Martinez, *Physica A* **534**, 122048 (2019).
- [30] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevE.110.024136> for information about the methods and comparison with alternative datasets for vegetation intensity.
- [31] R. Corrado, A. M. Cherubini, and C. Pennetta, *ISCS 2014: Interdisciplinary Symposium on Complex Systems* (Springer, Berlin, 2015), pp. 99–107.
- [32] R. Corrado, A. M. Cherubini, and C. Pennetta, *Commun. Nonlin. Sci. Numer. Simul.* **22**, 3 (2015).
- [33] D. Bengochea Paz, K. Henderson, and M. Loreau, *Ecol. Lett.* **25**, 163 (2022).
- [34] Global temperature and precipitation datasets are available through several meteorological databases (e.g., European data portal, NOAA). As a simplification, we consider higher temperatures as an indicator of general environmental stress.
- [35] N. Pettorelli, *The Normalized Difference Vegetation Index* (Oxford University Press, Oxford, 2013).
- [36] H. Fang, F. Baret, S. Plummer, and G. Schaepman-Strub, *Rev. Geophys.* **57**, 739 (2019).
- [37] J. Sauvola and M. Pietikäinen, *Pattern Recogn.* **33**, 225 (2000).
- [38] T. A. McMahon, M. C. Peel, L. Lowe, R. Srikanthan, and T. R. McVicar, *Hydrol. Earth Syst. Sci.* **17**, 1331 (2013).