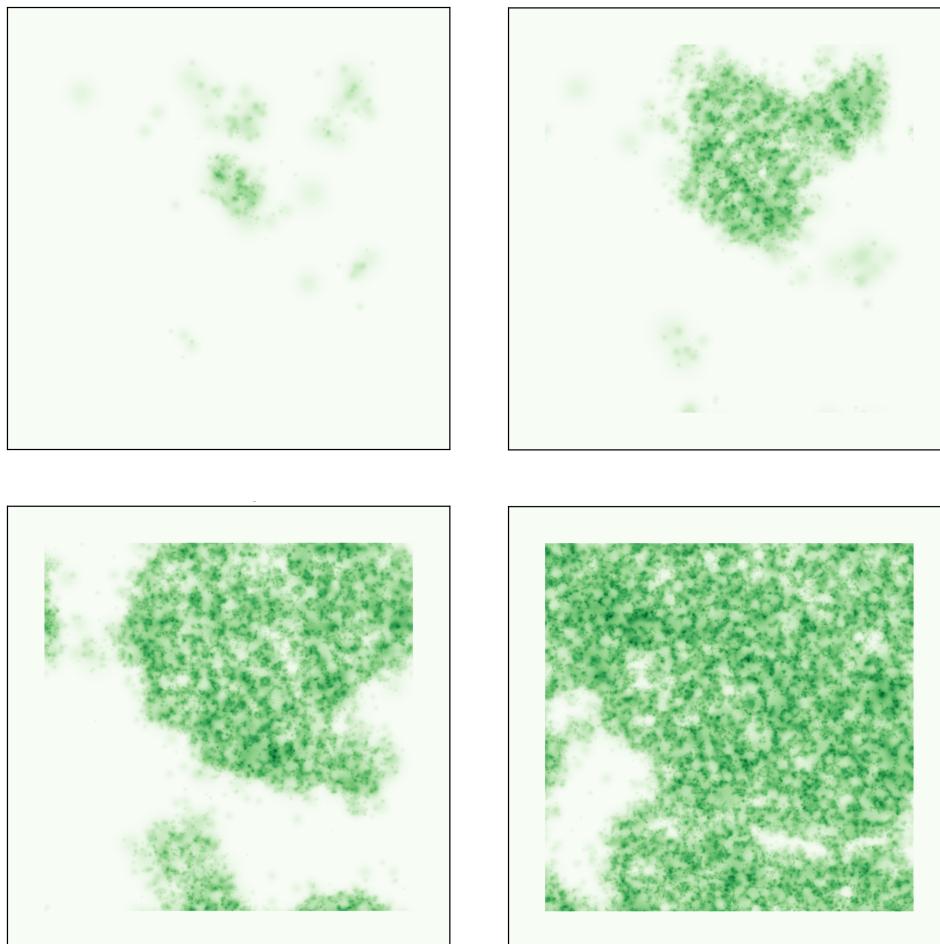


Tipping points and early warning signals in complex ecosystem models

Carl Ivarsen Askehave

Student ID: wfq585 — 60 ECTS MSc Thesis

Supervisor: Johannes Jakob Lohmann
University of Copenhagen
May 20th, 2025



Abstract

Several large ecosystems face irreversible, abrupt transitions to different states. Drylands in semiarid climates may experience desertification as water supplies decrease or wildfires intensify. Similarly, large rainforests could transition to savannahs. These shifts are driven by positive feedback loops that become dominant at tipping points (TPs), mathematically described as saddle-node bifurcations, leading to rapid, system-wide changes. “Critical slowing down” (CSD) can predict TPs by monitoring increased variability in observables, acting as early-warning signals (EWS). However, spatio-temporal effects, such as spatial heterogeneity and scale-dependent feedbacks are often not included in the modeling of TPs. This project aims to understand how ecosystem TPs change with increasing complexity of the positive feedbacks inherent in local-scale ecosystem models. To this end, different stable vegetation configurations in the models are found numerically, and the bifurcations of these patterns are determined as control parameters (e.g., mean annual rainfall) vary. It is investigated what happens at the TP of vegetation collapse, i.e., which variables or observables change most in terms of their mean and variability, and this may guide the choice of observables to use in EWS for similar systems. The model simulations performed in this project will be used to constrain under which circumstances abrupt, system-wide ecosystem collapse can be expected, and to assess whether a collapse may be predicted from real-world-observations.

Contents

1	Introduction	3				17
2	Theory	5				
2.1	Mathematical modelling of critical transitions	5	5.1	Collision detection using the KDTree	17	
2.1.1	Flow in a changing potential landscape	5	5.2	Density estimation	18	
2.1.2	Normal forms	6	5.3	Boundary conditions	19	
2.1.3	Bifurcation diagrams	6	5.4	Distributing non-overlapping plants	20	
2.1.4	Critical slowing down and early warning signals	7	5.5	Temporal EWS	20	
			5.6	Spatial EWS	21	
3	Conceptual models of tipping points	9	6	Results	22	
3.1	Daisy World and negative feedback	9	6.1	Multistability and hysteresis	22	
3.2	The Budyko energy balance model and positive feedback	10	6.2	Resilience and early warning signals	24	
3.3	The Stommel box model	10	7	Discussion and conclusion	29	
3.4	The Ising model and microscopic tipping points	11	8	Outlook	31	
4	Construction of forest ecosystem models	12	8.1	Future developments	31	
4.1	Dispersal and germination	12	8.2	Computational considerations	31	
4.2	Local competition	13	8.2.1	Structure of arrays	31	
4.3	Scale-dependent feedbacks	13	8.2.2	Interpreted vs. compiled languages	32	
4.4	Implementation of forest ecosystem models	14	A	Performance metrics of the density estimation	33	
			B	Overview of algorithms	34	
				References	35	

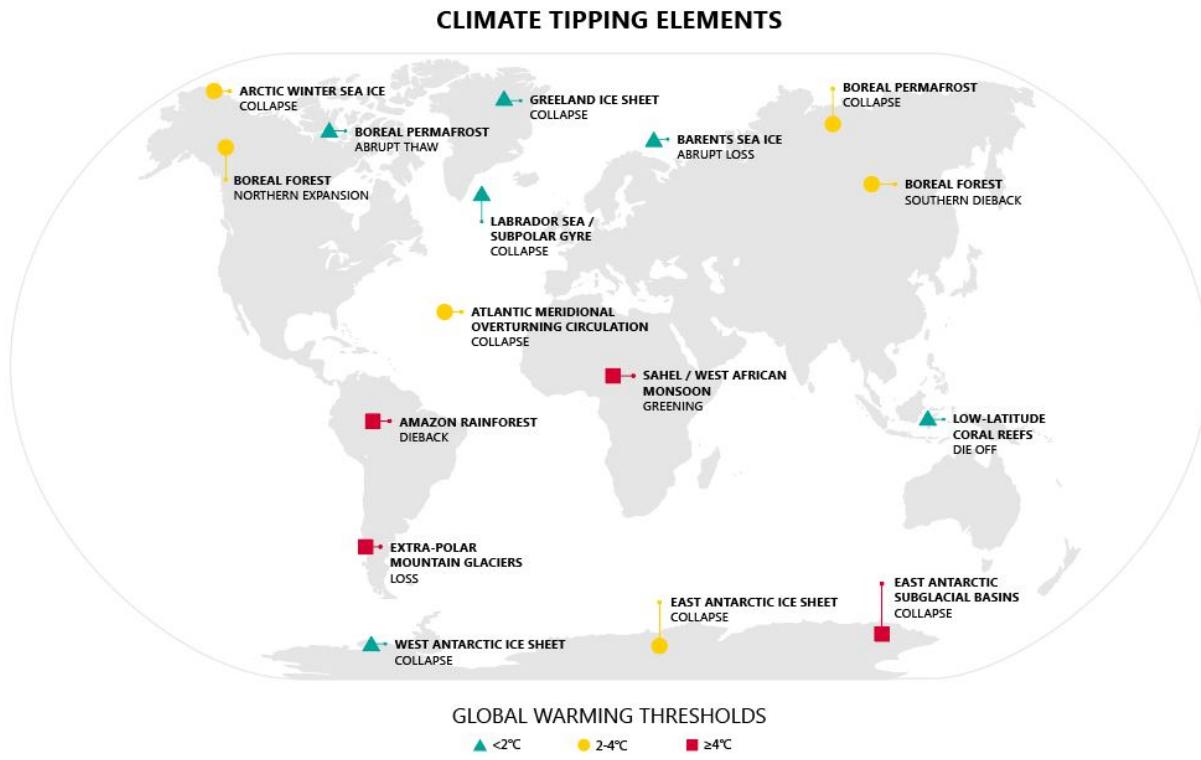


Figure 1: Tipping elements of the Earth's climate system along with the estimated increase in global average temperature that would induce their tipping. (Graphic taken from (Happold 2025)).

1 Introduction

Tipping points are critical thresholds beyond which a system can undergo an abrupt and often irreversible shift to a different state. These points mark the boundaries where small changes in external conditions can trigger large-scale transformations in the structure, function, or dynamics of the system. Examples of tipping points are financial crashes, epidemic outbreaks, the collapse of ecosystems etc.

In 2009 scientists identified specific parts of the climate system believed to be of special political interest because of their large scale and potential importance to the future state of the climate. These “climate tipping elements”, are large parts of the global climate system (at least subcontinental in scale) which are believed to feature multistability and thus the possibility for tipping or regime shifts (Lenton et al. 2008). The sudden shift in state of an ecosystem can have great ecological and

economic repercussions and restoring the system back to its old state might be hard and very expensive (Scheffer, Carpenter, et al. 2001).

Among the tipping elements that have been identified is included the possible collapse of the Atlantic Meridional Overturning Circulation (AMOC), dieback of the Amazon rainforest or collapse of the Greenland ice sheet among others (see Figure 1). It's important to note that even though the tipping elements are large systems, in general ecosystems on all scales can have tipping points, and evidence for tipping points has been found in several types of ecosystems where multistability has been observed such as lakes, coral reefs and oceans and also in ecosystems in which the existence of multistability is still unclear such as forests and deserts (Scheffer, Carpenter, et al. 2001).

Tipping points are inherent to systems governed by

non-linear dynamics, where positive feedback mechanisms are present. Negative feedback mechanisms counteract changes to the system and have a stabilizing effect, whereas positive feedback mechanisms amplify small changes making them more pronounced over time.

An example of a negative feedback is the Planck feedback, which occurs because of the Boltzmann-law, which implies that the amount of outgoing radiation from the Earth increases as its temperature increases, thus having a self cooling effect. This feedback is highly non-linear, increasing with the fourth power of the temperature. An example of a positive feedback mechanism is the ice-albedo effect, where a cooling climate forms more ice, which increases reflectivity and thus further decreasing temperatures and vice versa.

In systems where the amplifying effects of one or more positive feedback mechanisms gain prominence over the stabilizing effects of the negative feedback mechanisms present—due perhaps to a change in the external conditions of the system or just a small perturbation—the system might reach a tipping point, it undergoes a *critical transition*, abruptly transitioning to a qualitatively different state. In systems exhibiting *multistability*—that is, the existence of multiple stable states over a range of external conditions—the pre-tipping state might still be unattainable, even if the external conditions of the system return to their pre-tipping configuration (Scheffer, Carpenter, et al. 2001). This phenomenon is called *hysteresis*.

In general it is impossible to deduce whether and under which external conditions a system has positive feedbacks, and thus tipping points. This could in theory be achieved, but would require a complete model of the system, which doesn't exist for the climate. Therefore a mathematical theory of tipping points has been developed in the hope that it might provide methods to predict oncoming tipping points. Using this framework, specific statistical measures have been identified to sometimes indicate that a system is approaching a tipping point. These statistical measures are called *early warning signals* (EWS) and may be the only way of forecasting tipping, if no realistic model is available.

The most generic early warning signal is *critical slowing down* (CSD), where the rate of recovery of a system subject to perturbations decreases as the system nears a tipping point. This decrease in recovery rate is also sometimes referred to as the system losing *resilience*. CSD is a product of the understanding of tipping points as *saddle-node bifurcations*, which have a specific mathematical form, from which CSD is derived. CSD manifests itself as an increase in the variance and autocorrelation in the time series data of the system.

If the mathematical formulation of tipping points as

solitary saddle-node bifurcations is generally true, then CSD would also imply increases in the spatial variance and correlation in the system. However, computer modelling has suggested the existence of more complex tipping behavior in spatially extended systems through a series of smaller *Turing bifurcations* (Rietkerk, Boerlijst, et al. 2002), and studies of similar computer models have shown that the proposed spatial early warning signal sometimes fail to indicate tipping (Dakos et al. 2011), although the Turing bifurcations can generate spatial patterns, which might themselves constitute EWS.

The climate system is very complex and thus difficult to model, since there will always be more processes to include and since it is not clear when the model is realistic enough to capture the dynamics truthfully. Therefore it is unclear whether the understanding of tipping points as solitary saddle-node bifurcations with no spatial effects included fully encapsulates also dynamics of systems of higher complexity, i.e. systems which are extended over large spaces and have not only one or two but a plethora of known feedback mechanisms, the magnitudes and couplings of which might be distributed in a spatially heterogeneous way, and might change over time. Therefore uncertainty arises also with respect to the extent to which EWS via CSD can be expected in these systems. As noted in (Scheffer, Bascompte, et al. 2009): "The theoretical basis of the work on EWS in simple models is quite strong, and the first results from more elaborate models suggest that similar signals may arise in highly complex systems. Nonetheless, more work is needed to find out how robust these signals are in situations in which spatial complexity, chaos and stochastic perturbations govern the dynamics".

This project aims to contribute to this broader question by examining whether the simple paradigm of a single saddle-node bifurcation—and the associated EWS via CSD—hold in an individual-based, spatially explicit model, in the case where there is local interaction via a local positive feedback. To this end, first, two individual-based forest ecosystem models with two different types of positive feedbacks will be constructed, namely one local spatially explicit feedback and one non-local spatially averaged one. The individual-based approach is chosen to make certain that the saddle-node is not baked in to the model via e.g. governing differential equations. The effects of the local versus non-local feedbacks on the behavior of an otherwise unchanged model will then be analyzed for both static and changing external conditions to discern the existence of multistability and hysteresis and to examine to what extent EWS are present and can be detected.

2 Theory

2.1 Mathematical modelling of critical transitions

Suppose a dynamical system exists with state \mathbf{x} , the change of which is described by the system of equations

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{r}) \quad (1)$$

where $\dot{\mathbf{x}}$ refers to the change in the systems state \mathbf{x} over time and $\mathbf{f}(\mathbf{x}, \mathbf{r})$ is some function of the current state \mathbf{x} and some parameters \mathbf{r} . Written out this system of equations becomes

$$\begin{aligned} \dot{x}_1 &= f_1(x_1, \dots, x_n, r_1, \dots, r_m), \\ &\vdots \\ \dot{x}_n &= f_n(x_1, \dots, x_n, r_1, \dots, r_m). \end{aligned} \quad (1)$$

The variables $\mathbf{x} = (x_1, \dots, x_n)$ might describe the positions and velocities of particles in a force field, the prices and headings of the stocks on the stock market or even the sizes and reproductive rates of different species in an ecosystem, etc.

Example: Particle in the atmosphere

The equation governing a particle falling in a constant gravitational field with the inclusion of air drag is $\ddot{y} = -g - \kappa \dot{y}^2$, where $\kappa = k/m$ is the ratio between the coefficient of drag k and the mass of the particle m . Thus doing the simple reformulation $x_1 = y$ and $x_2 = \dot{y}$ the equation can be written as so

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -g - \kappa x_2^2 \end{aligned} \quad (2)$$

which is on the same form as (1), with state $\mathbf{x} = (y, \dot{y})$ and parameters $\mathbf{r} = (g, \kappa)$.

2.1.1 Flow in a changing potential landscape

Assuming that the function $\mathbf{f}(\mathbf{x})$ (we omit writing the parameters \mathbf{r} for clarity, but they are still there) can be written as the negative gradient of some potential function V , equation (1) becomes

$$\dot{\mathbf{x}} = -\nabla V(\mathbf{x}). \quad (3)$$

This is the general form of a *gradient system*, and although this assumption always holds in one dimension, for dimensions higher than 1 the existence of a function V for which equation (3) holds is not always guaranteed.

When a dynamical system can be described by an equation on the form of (3), then the changes in the state $\dot{\mathbf{x}}$ can be interpreted as the ball rolling along the contours of the potential surface, where the size of the change is directly proportional to the negative slope of the surface at the given state \mathbf{x} , meaning that the ball will roll towards the nearest local minimum (see Figure 2).

In this framework, the ball represents the current state of the system, and the local minima of the potential surface represent the local stable states, that the system will tend towards as time runs forward. When working with gradient systems it is very helpful to define the fixed points of the system, as the points \mathbf{x}^* where $-\nabla V(\mathbf{x}^*) = 0$. There are three distinct types of fixed points, namely stable, unstable and semi-stable ones. Stable fixed points are points at which if subject to a perturbation in a random direction, the system will return to the point after a short time. Unstable fixed points are points which the system will rapidly depart from when subject to a perturbation in a random direction. Semi-stable points are points that are stable when the system is perturbed in one direction, but unstable when perturbed in the other direction.

Now, the exact topology of the potential surface is dependent on the system at hand, which might change due to changes in the external conditions. This could cause the landscape to change in such a way that a previously stable state turns into a semi-stable one, an unstable one or might even disappear completely (see Figure 2). When this happens the system is said to pass a *bifurcation point*, and the state of the system will rapidly shift to the nearest still existing stable state or just diverge. This rapid shift in the state of the system when passing a bifurcation point is known as a *critical transition* (Scheffer, Bascompte, et al. 2009).

The specific type of bifurcation which is reached in our example with the double well potential is called a *saddle node bifurcation*, but is far from the only type of bifurcation.

It is well worth noting, that even if the parameters of the system return to the values they had before the critical transition, the system state might still be “trapped” in its new state, requiring a more dramatic change to return to its initial stable state. This is the phenomenon of *hysteresis*. We see that because the change to the system is implicitly dependent on its previous state, or its history, a critical transition can have a degree of irreversibility.

A growing concern in the scientific community is, whether or not the changes to the global climate—which is being brought about by the anthropogenic emissions of heat-trapping gasses—exhibit the same characteristic of irreversibility.

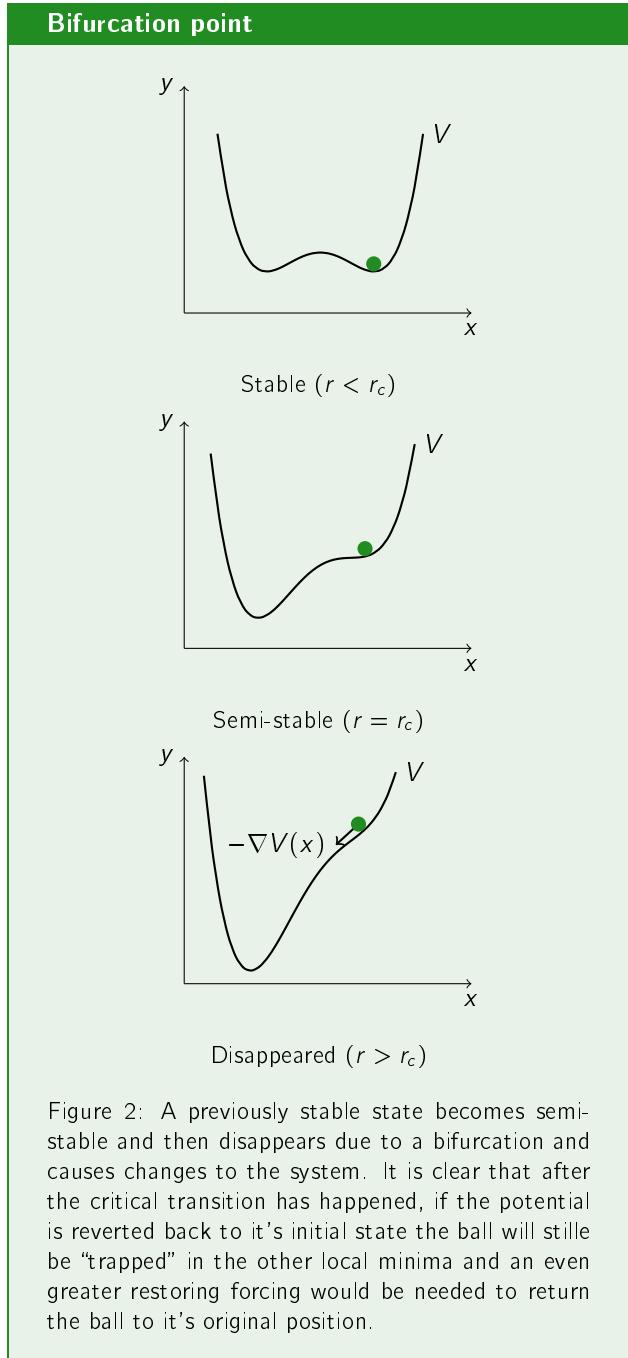


Figure 2: A previously stable state becomes semi-stable and then disappears due to a bifurcation and causes changes to the system. It is clear that after the critical transition has happened, if the potential is reverted back to its initial state the ball will still be “trapped” in the other local minima and an even greater restoring forcing would be needed to return the ball to its original position.

2.1.2 Normal forms

As already mentioned, the existence of a potential function that satisfies equation (3) is only guaranteed for

1-dimensional systems, but thankfully the *center manifold theorem* ensures that an equation on the form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, r), \quad (1)$$

where \mathbf{f} is sufficiently continuous and has a bifurcation point at $(\mathbf{x}, r) = (\mathbf{x}^*, r_c)$, can be reduced to its *normal form* around the bifurcation point by a series of appropriate coordinate transformations (for a thorough derivation see (Wiggins 2003, Section 3.3)). The crucial point is that the evolution of the state \mathbf{x} will generally converge towards a *center manifold*, which is generally a nonlinear function of the system coordinates $h = h(\mathbf{x})$. The evolution of the state along the center manifold can then be described by the normal form of the bifurcation $h = r + h^2$.

Example: Saddle-node bifurcation — The normal form of a saddle-node bifurcation is

$$f(x, r) = r + x^2, \quad x, r \in \mathbb{R}, \quad (4)$$

where r is the bifurcation parameter. It has two fixed points $x_{\pm}^* = \pm\sqrt{-r}$, and a bifurcation point at $(x^*, r_c) = (0, 0)$. The fixed points exist only for $r < r_c = 0$, and vanish otherwise giving us the bifurcation. The normal form also yields the potential function

$$V(x, r) = - \int f(x, r) dx = -rx - \frac{1}{3}x^3. \quad (5)$$

For any high dimensional system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, r)$ that exhibits a saddle node bifurcation the analysis of (4) is sufficient in describing fully the dynamics of the system near the saddle-node bifurcation point.

2.1.3 Bifurcation diagrams

Another helpful way to analyze the behavior of systems around bifurcation points is to look at the fixed points of the system when the control parameter r is varied. In the simplest case (of Figure 2), the potential might be modeled by

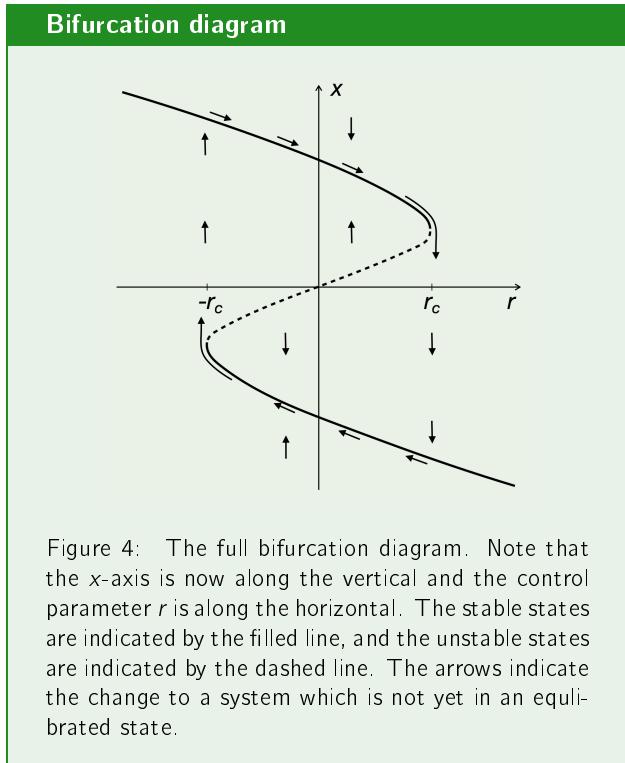
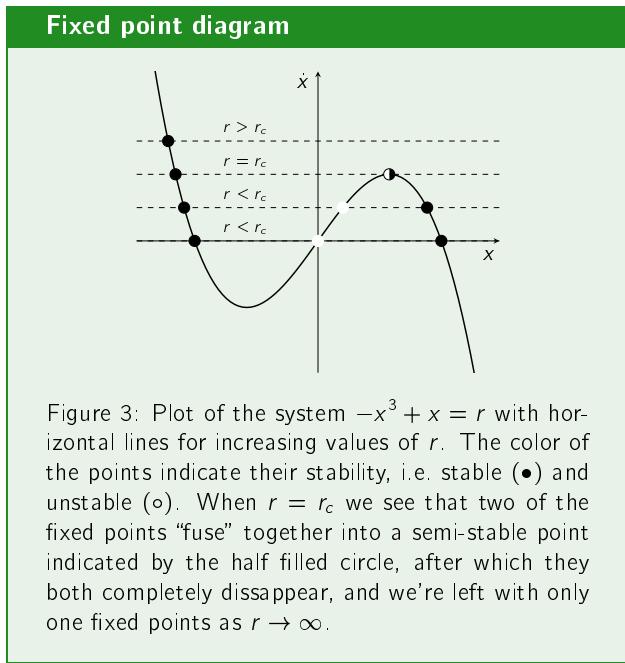
$$V(x, r) = \frac{1}{4}x^4 - \frac{1}{2}x^2 + rx, \quad (6)$$

which means that we’re interested in analyzing the equation

$$\dot{x} = -\nabla V(x, r) = -x^3 + x - r = 0. \quad (7)$$

For values of up to, but not including, some critical value $0 < r < r_c$, the equation has three roots, where two are stable and one is unstable. At the critical value $r = r_c$, the previously unstable point and one of the stable points seemingly fuse into a semi-stable point, which disappears completely for $r > r_c$, leaving only the last stable point (see Figure 3). The analysis is identical, but mirrored, for negative values of r .

Plotting the fixed points as a function of the control parameter, we get Figure 4.



2.1.4 Critical slowing down and early warning signals

In the scenario of the ball in the potential landscape (see Figure 2), one very important thing to note is that, as

the system approaches the bifurcation point, the curvature of the potential around the fixed point at which the ball is sitting is decreasing. This results in a increased recovery time of a state subject to a small perturbation from the stable state.

Because of the saddle-node bifurcation normal form

$$\dot{x} = r + x^2, \quad (8)$$

the vicinity of the ball can be described by a cubic potential (see Figure 5). As already discussed this has fixed points located at $x_{\pm}^* = \pm\sqrt{-r}$. Let’s examine what happens over time to a small perturbation from one of the fixed points $\tilde{x} = x - x^*$. Assuming that the perturbation is sufficiently small, Taylors theorem can be utilized to obtain

$$f(x^* + \tilde{x}) \approx f(x^*) + \frac{\partial f}{\partial x}\Big|_{x^*} \tilde{x} = \frac{\partial f}{\partial x}\Big|_{x^*} \tilde{x}. \quad (9)$$

Additionally, it is apparent, that

$$f(x^* + \tilde{x}) = \frac{d}{dt}(x^* + \tilde{x}) = \frac{dx^*}{dt} + \frac{d\tilde{x}}{dt} = \frac{d\tilde{x}}{dt}. \quad (10)$$

Combining these two results yeilds

$$\frac{d\tilde{x}}{dt} = \frac{\partial f}{\partial x}\Big|_{x^*} \tilde{x} = -\lambda \tilde{x}, \quad (11)$$

which has eigenvalues $\lambda_{\pm} = \pm 2\sqrt{-r}$ (the negative sign in front of λ stems from convention). From this can be seen that for a perturbation around the positive fixed point $x_+^* = \sqrt{-r}$ the corresponding eigenvalue $\lambda_+ = 2\sqrt{-r}$ is positive, and thus the perturbation will decrease over time making this fixed point stable. For a perturbation around the other fixed point the opposite is true.

Furthermore it is clear that at the bifurcation point $r = 0$ both of the recovery rates become zero and no perturbation will ever recover. Furthermore, we see that the eigenvalues, and thus the rate of recovery around the stable fixed point, is linearly proportional to the distance between the two fixed points $|x_+^* - x_-^*| = 2\sqrt{-r}$, which is sometimes called the *basin of attraction*. This means that when approaching the bifurcation point $r \rightarrow 0$, the basin of attraction shrinks $2\sqrt{-r} \rightarrow 0$ and the recovery rate slows, i.e. critical slowing down. The size of the basin of attraction is often also called the *resilience* of the system, and thus, as the systems nears the bifurcation, it loses resilience.

Increased variance and autocorrelation — So far, we have assumed that we can describe the changes in the state of the system entirely by the gradient of the potential function $-\nabla V(x)$, but this presumes that we know

everything about the system we're modelling, which is generally not the case. Therefore, systems are often modeled with an additional noise term driving small perturbations, as such

$$\dot{x} = -\nabla V + \delta\eta(t) \quad (12)$$

Where $\eta(t)$ is white noise and $\delta > 0$ is the noise intensity. Linearizing the saddle-node bifurcation normal form around the stable fixed point as in (11), yields

$$\dot{x} \approx -\lambda x + \delta\eta(t), \quad (13)$$

which is exactly the expression for an *Ornstein-Uhlenbeck process*, which has variance (Gardiner et al. 1985)

$$\text{Var}(t) = \frac{\delta^2}{2\lambda} [1 - e^{-2\lambda t}], \quad (14)$$

and lag1-autocorrelation of

$$\rho_1(t) = e^{-\lambda} - e^{-\lambda(2t-1)}. \quad (15)$$

we see that for asymptotic time $t \rightarrow \infty$ we get

$$\text{Var}(t) = \frac{\delta^2}{2\lambda} \quad (16)$$

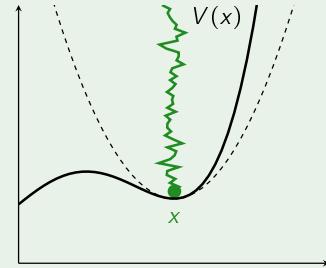
$$\rho_1(t) = e^{-\lambda} \quad (17)$$

This means that as the curvature of the potential decreases $\lambda \rightarrow 0$, the variance diverges as $\delta^2/2\lambda \rightarrow \infty$ and the autocorrelation converges as $e^{-\lambda} \rightarrow 1$ (see Figure 5). These are early warning signals, and are a general property of systems governed by the saddle-node normal form. In general the potential can be asymmetric and have varying higher order terms which might also influence other statistics of the signal over time such as for example the skewness. Therefore, these have also sometimes been proposed as candidates for early warning signals (Morales et al. 2015).

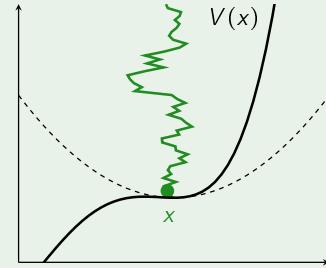
In systems governed by noise in this way, no states are really asymptotically stable. There is always a non-zero chance of transitioning to another state via *noise-induced tipping*, but if the noise is low, this chance of

noise-induced tipping is small and a system is called *metastable*.

Critical slowing down



The system trajectory is perturbed by the noise around the stable fixed point.



As the control parameter is increased, the curvature of the potential decreases, and the variance of the trajectory increases.

Figure 5: The state of the system (green ball) and its previous trajectory (green line) having been perturbed by noise. The potential (black filled line) and its 2nd order Taylor expansion (black dashed line).

3 Conceptual models of tipping points

The study of Earth's climate system, is the study of many complex interactions happening at a plethora of different scales from the macroscopic to the microscopic. Some of the processes can be studied in a laboratory, but the system as a whole can not and therefore, models have to be developed to carry out experiments "in-silico". Models, and the experiments carried out using them, can aid in understanding better the interactions between the climate systems constituting parts, and can be used to project future climate scenarios.

Global climate models refer to a class of models which are numerical representations of the whole of the climate system and the processes inherent to its different components: the atmosphere, the ocean, the cryosphere and the biosphere (Aurore 2020). The modelling of the global climate is in general based on two pillars. The first pillar is the modelling of the fluid mechanical equations governing atmospheric and oceanic circulation often called *global circulation models* (GCMs). The second pillar is the modelling of other physical, chemical and biogeochemical processes. Models that include both aspects is often referred to as *Earth system models* (ESMs) (Gerhard 2021).

In GCMs the domain of the atmosphere and ocean is sliced into discrete cubic meshes in order to discretize the governing equations which can then be solved numerically. For global studies on time scales relevant for long-term evolution the size of the meshes is on the order of hundreds of kilometers (~ 100 km), which means that finer-grained processes like turbulence, convection, cloud and precipitation formation, cloud interactions with radiation, etc. can not be modeled directly, but need to be included in the model via empirically derived *parameterizations* of the effects of these processes on the scale of the meshes.

The modelling of the biosphere is done by regarding the exchange of water and energy between the soil, the biosphere and the atmosphere based on energy and water conservation principles. The modelling of vegetation is represented in ESMs by dynamic global vegetation models (DGVMs), which simulate the distribution and structure of vegetation and its interactions with climate through physiological principles. DGVMs represent plant communities using single area-averaged concentrations of plant functional types, which interact with the rest of the elements of the ESM depending on the type of plant through its effect on the carbon and hydrological cycles and other biogeochemical processes (Fisher et al. 2018). The cryosphere is usually

included in models via the fluid mechanical modelling of glacier flow and mass balance (Aurore 2020). As with GCMs, finer-grained processes in ESMs have to be included through parameterization.

In contrast to ESMs, which are used to model the global climate and all its processes from first principles, conceptual models seek to model specific phenomena, mechanisms or processes inherent to the whole or just part of the climate system. Conceptual models are generally simpler than ESMs both in conceptualization and implementation, because a lot of the complexity which is not immediately relevant to the subject being modeled is often simplified or approximated away, maybe by disregarding the spatial extent of the system or aggregating discrete elements into continuous mathematical equations.

Here, different conceptual models from climate science and one from statistical physics are examined along with how they have been used to suggest mechanisms which might exist in real systems, such as the climate.

3.1 Daisy World and negative feedback

In 1983 Watson and Lovelock studied the close-coupling of the biota and the global environment in a highly idealized conceptual model called "Daisy World" (Watson and Lovelock 1983). In the model a cloudless planet is inhabited by two biological species, namely black and white daisies, which differ only in the amount of solar radiation they reflect back into space. The white daisies reflect more light than the bare ground, and therefore have a net cooling effect on the climate, and the opposite is true for black daisies which absorb more light than the bare ground, heating up the climate. Spatial effects are disregarded, meaning that only the proportions of the total area of bare ground which is covered by each type of daisy is considered. Additionally, the growth rate of the daisies is assumed to be a parabolic function of temperature being zero at 5°C and 40°C having a maximum in between at 22.5°C . Finally, the local temperatures of the daisies are assumed to vary with their respective albedos: black daisies, which absorb the most sunlight, are warmer than the ground, which in turn is warmer than the white daisies.

Watson and Lovelock showed that the interplay between the daisies' effect on the temperature, and the effect of the temperature on the population of the daisies, exhibited a negative feedback loop which had a stabilizing or *homeostatic* effect on the climate. And although

they heavily emphasize the theoretical nature of their work and the problems in trying to extrapolate the results to the Earth climate, they propose that the effect that the Earth's biota can have on decreasing the amount of atmospheric CO₂, and thus decreasing the greenhouse effect and lowering the mean global temperature, might be playing a similar stabilizing role in the climate.

3.2 The Budyko energy balance model and positive feedback

While studying the effects of changes in solar radiation on the climate Russian scientist Mikhail Budyko put forth an energy balance model (EBM) for the Earth in his paper 1969 paper "The effect of solar radiation variations on the climate of the Earth" (Budyko 1969).

He modeled the energy balance as the difference between the incoming solar radiation and the outgoing radiation

$$\text{Balance} = \text{Incoming} - \text{Outgoing}.$$

Budyko assumes a steady state of the balance, meaning that it's assumed to vanish when averaged globally, which is equivalent to the Earth's average temperature not changing over time. But locally, the energy balance is assumed to be linearly proportional to the difference between the local temperature and the global average, approximating the effect of energy diffusion through oceanic and atmospheric circulation.

Similarly to the total energy balance, both the incoming and outgoing radiations are also modeled as a balance between two factors. The incoming radiation at the surface is assumed to be some fraction of the incoming radiation at the top of the atmosphere since some will be deflected due to the albedo, while the outgoing radiation at the top of the atmosphere is a fraction of the value at the surface, since some will be contained due to the "cloudiness" or the greenhouse effect.

$$\text{Incoming at surface} = \text{Incoming} - \text{Reflected}$$

$$\text{Outgoing} = \text{Outgoing at surface} - \text{Contained}$$

Logically the outgoing radiation is a function of the temperature at the surface.

Budyko also includes the effect of a changing ice cover by introducing an albedo change which is linearly proportional to the fraction of total area covered in ice times the ratio between incoming radiation in the ice covered area and the mean planetary incoming radiation.

Using this model Budyko examines the change in the latitudinal temperature distribution due to relatively small changes in incoming solar radiation. Using his model, Budyko finds that a 1% reduction in radiation reduces the mean temperature of the Earth by 5°C, while a 1.5% reduction in radiation reduces the temperature by 9 °C. It is also noted that for the above temperature drop the latitude at which glaciation takes place moves (10 – 18) °C south.

In light of these results Budyko concludes that "the present thermal regime and glaciations of the Earth prove to be characterized by high instability. Comparatively small changes of radiation-only by 1.0 – 1.5%-are sufficient for the development of ice cover on the land and oceans that reaches temperate latitudes".

The reason for this instability is the ice-albedo feedback, which amplifies the effects of changes in incoming radiation due to for example changes in volcanic activity or shifts in the Earth's orbit, and provides a mechanism that can drive the system towards two separate states, namely the greenhouse and icehouse states as they are commonly called. The icehouse state is the state which the Earth is presently in and it is characterized by "the presence of polar ice and large thermal contrast between the pole and the equator", while the greenhouse state is characterized by "the absence of glaciation and small meridional mean gradient of temperature".

Thus Budyko's model suggests that subtle changes in the external conditions of the climate, (incoming solar radiation in this case), can cause a global scale tipping point driven by the positive ice-albedo feedback.

3.3 The Stommel box model

In 1961, while studying the effects of mixing in sea water due to changes in temperature and salinity, American oceanographer Henry Stommel introduced a few conceptual models in his paper "Thermohaline Convection with Two Stable Regimes of Flow" (Stommel 1961). The models were all characterized by their approximation of different ocean basins as boxes connected with pipes.

The third and most complex model from Stommel's paper consists of two boxes filled with water with some temperature and some salinity. The boxes are allowed to have differing values of temperature and salinity. Additionally each box is surrounded by porous walls which let the water outside the boxes—which has yet another set of values of temperature and salinity—diffuse into the boxes through the walls. The two boxes are connected with a pipe through which the rate of flow is determined by the density difference between the two boxes, which

in general is a function of the differences in temperature and salinity between the boxes. The boxes are also connected by an upper overflow which ensures that the volume of water in each box remains the same.

The model is meant to reflect the basic principles of thermohaline circulation, which the circulation of water around the Earth's oceans driven by global density gradients created by differences in local surface heat temperature and freshwater fluxes.

Stommel models the changes in temperature and salinity as the sum of the change due to diffusion through the porous walls and the change due to flow through the pipe. He then models the flow through the pipe as linearly proportional to the density gradient between the two boxes. Thus the direction and speed with which the water is flowing through the system can be determined by relating the density in each box to its respective temperature and salinity using an equation of state, inserting these densities in the equation for the flow in the capillary and solving the resulting equation.

In his work Stommel observes that for specific parameter values, his model allows for more than one stable solution, alluding to the fact that the global thermohaline circulation might have different stable states too, and thus the ability of tipping from one to the other. More broadly this suggests that multistability and tipping points are not limited to the global climate system and can happen even in climate subsystems. Subsequently evidence has been found for this, and it is now believed that sudden shifts in the state of the thermohaline circulation might have played a crucial role in rapid climate change in the past (Lenton et al. 2008).

3.4 The Ising model and microscopic tipping points

The Ising model, which is not a climate model, but rather a statistical physics model of the microscopic phenomenon of ferromagnetism, was first published in 1925 by German physicist Ernst Ising in his paper "Beitrag zur Theorie des Ferromagnetismus" (Ising 1925). The model is widely known for its importance in the study of phase transitions and criticality, concepts which are intimately related to that of tipping points. Both tipping points and phase transitions involve crossing a threshold beyond which abrupt change can occur and are both modeled using bifurcation theory. In addition critical slowing down is a general feature of phase transitions which therefore also have detectable early warning signals (Morales et al. 2015). The microscopic equivalent of a tipping point is thus in some sense a phase transition.

In the Ising model each one of many points on a

spatially extended lattice is associated with a spin state pointing either up or down. An interaction energy is defined by the sum of a global magnetic field, affecting the whole lattice, and a local interaction affecting only neighboring lattice points. This local interaction allows for spatial heterogeneity in the state of the system, a feature which makes it stand apart from the models discussed until now. A temperature of the system is defined, such that the higher the temperature, the more likely it is for a lattice point to randomly jump from a low energy state to a higher energy state, as determined by the interaction energy making the system inherently stochastic. The state of the system is fully described by the spin orientation of each of the lattice points, and can be summarized as the total magnetization defined as the sum of the spins of all the points divided by the number of points.

When the lattice has a dimensionality greater than 1, the model has a phase transition when the temperature is equal to a critical value $T = T_c$. For values below the critical temperature $T < T_c$, large, long lived patches of the lattice align with parallel spins in either the up or down state, and only small patches of opposing spin appear for short moments. For values above the critical temperature $T > T_c$, the magnetization of the lattice becomes 0 on average, and patches of aligned spin are small and short lived.

In 2015 the model was used to study the behavior of early warning signals specifically with interest not in the ferromagnetic properties of the model but in the temporal signatures produced by the local-scale interactions between model components and the changes in these signatures when approaching criticality. It was demonstrated that as the system is brought across the critical point, early warning signals are present in the mean, variability, skewness and kurtosis of the time series of the total magnetization (Morales et al. 2015). This is evidence for the simple paradigm of a single saddle-node bifurcation holding in spatially explicit systems.

4 Construction of forest ecosystem models

Many different approaches to the modelling of forest ecosystems can be found in the literature including but not limited to reaction-diffusion models, individual-based models, cellular automata, etc. It is important to note, that it is unclear whether forest ecosystems and grasslands are two different kinds of systems or just different states of the same systems (Scheffer, Carpenter, et al. 2001), but here these will be treated as the latter, in order to compare simulation methods.

In reaction-diffusion models, vegetation cover is treated as a quantity in and of itself while the individual organisms making it up are ignored. This is because the discrete nature of individuals is impossible to account for in a system of continuous differential equations. The continuous nature of the equations underlying reaction-diffusion models makes them well suited for modelling the interactions of vegetation with other physical phenomena commonly formulated in terms of continuous equations such as for example diffusion of rainfall across terrain contours and into the soil, as in (Rietkerk, Boerlijst, et al. 2002), but makes them inept at modelling stochastic individual-based dynamics, such as seed dispersal, plant-on-plant competition, species-on-species interactions etc., where the dynamics are directly dependent on the state of discrete individuals. Another characteristic of reaction diffusion models and other models governed by differential equations is that they have to be formulated on a discrete grid, in order for solvability by a computer. This means that the model has an inherent minimum spatial resolution at computation.

These shortcomings are addressed in individual-based models such as for example, *zone of influence* (ZOI) models (Gates and Westcott 1978; Ford and Diggle 1981; Weiner et al. 2001), and its derivative, the *field of neighborhood* (FON) models (Berger and Hildenbrandt 2000), in which the local interactions between individual organisms are modeled. These kinds of models generally lack the analytic nature of the coupled differential equations of for example reaction-diffusion models because the rules-based dynamics at the individual level are generally hard to formulate into a system of equations, but in turn they allow for the application of specialized knowledge about the biological and ecological properties of the specific organisms, making them ideal for addressing challenges where the diversity of species traits that have to be considered simultaneously are very high (DeAngelis and Grimm 2014). A tesselation based

approach using Voronoi cells also exists (Czaran 1998).

Like reaction-diffusion models, cellular automata are also defined on grids, but they share the rule-based approach of individual-based models, making them in some sense related to both. However, because of the rules based approach they lack the analytic nature of reaction-diffusion models, and they also inherit the minimal spatial resolution which are inherent to grid-based models of all kinds. Despite these flaws cellular automata are easy to implement and great for modelling spatial patterns, for which it has been used in various studies of grassland ecosystems (Thiéry, D'Herbès, and Valentin 1995; Rietkerk, C. Dekker, et al. 2004; Scanlon et al. 2007).

Additionally, some hybrid approaches which combine two or more of the aforementioned modelling techniques also exist as for example the individual-based model coupled with a reaction-diffusion model of (Vincenot et al. 2016) or the local-scale individual-based model coupled with the continuous soil and climate models of (Pałubicki et al. 2022).

As seen in the Daisy World model and Budyko's energy balance model (Watson and Lovelock 1983; Budyko 1969), feedback mechanisms, which change in strength as some control parameter r is changed, can facilitate tipping points. Two processes in forest ecosystems which constitute feedback mechanisms are the propagation of plants through dispersal and germination along with the local competition for crucial resources such as light, water and nutrients. The former is a positive feedback, since reproduction leads to a growing population, which in turn leads to even more reproduction and so on. Then it is clear that the process of local competition has a negative feedback. Here, some approaches to modelling these processes from the literature, will be examined.

4.1 Dispersal and germination

The mechanisms which facilitate the persistence of plant populations are well understood from a biological standpoint, and include most notably the formation of offspring in the form of diaspores, dispersal of offspring via different processes specific to the species of plant and the diaspores it produces and finally the establishment of offspring through germination after dispersal has occurred (Howe and Smallwood 1982).

In (Damgaard 2025) it is noted that plant fecundity, i.e. the reproductive rate, is highly species specific, but within species it is most often found to be approximately linearly proportional to plant size. Additionally, seed dispersal usually happens by the wind or by some

animal vector, but obtaining good seed dispersal data is difficult. Additionally it is noted that many approaches in the modelling of the spatial effects in play in plant populations have used fixed rates of fecundity, and a probability of seed germination being a function of the local vegetation density weighted by a competition kernel.

In (von Hardenberg et al. 2001) seed dispersal is modeled in a reaction-diffusion model as a diffusion term in the equation describing the temporal evolution of the vegetation density. In order to model the vegetation with differential equations, this approach approximates the actual process of dispersal as the diffusion in space of a continuous quantity representing the vegetation density. In contrast, in (Vincenot et al. 2016), seed dispersal is modeled explicitly using an individual-based approach, in which seed dispersal sites are drawn at random from an isotropic lognormal density kernel, which is more true to nature in the sense that individual plants and seeds are modeled, but nonetheless, it is noted, that “an improved accuracy in the representation of the spatial process of plant dispersal is sought.”

4.2 Local competition

Local competition in plant populations constitutes a negative feedback mechanism because the larger the population in a given area is, the more competition for resources takes place between nearby trees in that area.

In (Rietkerk, Boerlijst, et al. 2002) and similar reaction-diffusion models, competition is only implicitly modeled, because the negative change in available nutrients at any given point is modeled as being proportional to the vegetation density at that point. This again approximates the actual process in order to formulate a system of differential equations.

In (Ford and Diggle 1981) competition was modeled using a ZOI for each plant defined by its height and an angle with the ground, such that the ZOI grows wider as the plant grows taller. Competition was assumed to be multiplicative of the number of taller plants overlapping a given plant and size-assymmetric, in the sense that in competition between two plants only the smallest plant should be affected while the largest plant should be left undisturbed. Additionally the spatial distribution and growth rates were sampled stochastically. Interestingly, it was also discussed in (Ford and Diggle 1981), that the amount of water running down a plants trunk, i.e. the stemflow, was significantly correlated with crown size which, in turn, was directly correlated with plant height. This means that large plants not only capture more light than small plants but might also capture and conduct

to their roots a greater amount of rainfall, providing them with another competitive advantage, and another argument for size-assymmetry in local competition.

In (Weiner et al. 2001) a similar approach was used, where the ZOI was defined as the circle inscribing the area of resources potentially available to a plant. A growth rate was then calculated for each plant based on its total area minus the part of it that was overlapped by another ZOI weighted by a size-assymetry factor.

In (Berger and Hildenbrandt 2000) shortcomings of the ZOI approach to competition modelling was highlighted, specifically that the ZOI doesn't take into account the fact that the strength of competition of a single plant generally decreases with the distance from the stem. This isn't modeled in the ZOI, where the competition strength is constant over the whole ZOI. The FON approach, in which the competition strength was assumed to decrease exponentially with distance, reaching a minimal value at the ZOI edge, was developed to combat this inaccuracy in order to more accurately model competition between mangrove trees.

In conclusion the results of (Ford and Diggle 1981; Weiner et al. 2001; Berger and Hildenbrandt 2000) agree, that in order to reproduce experimental results from the field of plant population ecology, like the effect of self-thinning, bimodality in size distribution and regular spatial distribution of large (or tall) plants, at least two components were crucial in the modelling of local competition:

1. Competition has to be a size-assymmetric process, that is, competition should affect the smaller plants more dramatically than larger plants.
2. Competition has to be formulated as a spatial interaction between stochastically distributed individuals, in the sense that the size of any given plant's neighbors are chance events.

4.3 Scale-dependent feedbacks

As breifly mentioned in the introduction, arid ecosystems characterized by an extended dry period, where annual evaporation exceeds annual precipitation, vegetation has been found to self-organize into a number of different patterns, leading to an increased resilience of the ecosystem to dry conditions. A couple of mechanisms have been proposed, which all emphasize the importance of the interplay between scale-dependent positive feedbacks as the driving source for pattern formation in spatially extended systems (Rietkerk, C. Dekker, et al. 2004; Scanlon et al. 2007; Rietkerk, Bastiaansen, et al. 2021).

In (Rietkerk, Boerlijst, et al. 2002) it was hypothesized that one such feedback mechanism could be driven by the process by which water infiltrates faster into soil in the immediate vicinity of vegetated than into bare soil, leading to an aggregation of below-surface soil water around patches of vegetation. A spatially explicit reaction-diffusion model consisting of three coupled differential equations describing the evolution of plant density, surface water and soil water was introduced, which was shown to adequately recreate the different types of patterns found in arid ecosystems and thus provide a candidate for the mechanism by which they form.

In (Vincenot et al. 2016) this work was continued, by coupling a similar reaction-diffusion model to an individual-based vegetation model, where the individual-based approach was used to explicitly model plant metabolism and seed dispersal. The results of this paper supported that of (Rietkerk, Boerlijst, et al. 2002), but added, that the reaction-diffusion approach “does not allow the simulation of the effects of precipitation seasonality and droughts”. Mechanisms which were shown in (Vincenot et al. 2016) to be of some importance in the promotion of pattern formation.

The work was further built upon in (Pałubicki et al. 2022), where an individual-based model of plant metabolism and dispersal was coupled to a reaction-diffusion model derived from (Hille Ris Lambers et al. 2001) and local a scale atmospheric model allowing for the simulation cloud formation and the exchange of water between the air and the soil in the form of rainfall and evaporation. This paper recreated the patterns of the aforementioned papers, again stating the importance of correctly representing local positive feedbacks.

4.4 Implementation of forest ecosystem models

In order to answer whether the assumptions of tipping points hold for a local-scale spatially extended system and whether tipping points and EWS can still be detected, two individual-based models will be here constructed. The individual-based approach is chosen to ensure that a saddle-node bifurcation is not implicitly included in the model, which might have been the case if differential equations were used. The models will be constructed with two different types of feedback mechanisms, one spatially explicit and one spatially averaged, in order to investigate whether tipping happens as a system-wide bifurcation or whether it can happen locally.

Specifically both models will include the modelling of local competition between plant individuals, which con-

stitutes a negative feedback, as well as the modelling of the process of reproduction through dispersal and germination which constitutes a positive feedback. The only differing characteristic between the two models will be in the modelling of the positive feedback. In the first model the probability of germination of a newly dispersed seed will be estimated locally, allowing for a spatially dependent feedback, whereas in the second model this estimation will be spatially averaged, removing any spatial effects, but keeping the total strength the same as for the first model. For these reasons the two models will be referred to as the “local” and the “global” model.

Plants — A population of plants $N(t)$ inhabit a square domain with side lengths L . The plants are denoted p_i , and each plant has a position $\mathbf{x}_i = (x_i, y_i)$ and a radius r_i , and are thus assumed to be simple circles. For simplicity the biomass of each plant is modeled as equal to its area $b_i = \pi r_i^2$.

Each plant grows at a rate G_i which might in general be a function of both global and local environmental parameters, but for the purposes of this work will be regarded as constant

$$G_i = \text{constant}. \quad (18)$$

Additionally a mortality due to natural causes is set to happen when the plant reaches above a certain size

$$r_i > r_{\max} \implies \text{death}. \quad (19)$$

This process of growth and mortality is highly idealized but is sufficient for the purposes of this project.

Competition — When the plants grow, there is a chance that they will collide. Collision detection is done using a KDTree data structure as implemented in the `scipy` library for python (see Section 5.1).

When two plants collide, they engage in competition. The winner of the competition is the plant with the largest radius, and in the rare case that the two radii are equal, a survivor is chosen at random:

$$\begin{aligned} \text{if } r_1 > r_2 &\Rightarrow \text{plant 1 survives}, \\ \text{if } r_1 < r_2 &\Rightarrow \text{plant 2 survives}, \\ \text{if } r_1 = r_2 &\Rightarrow \text{random survivor}. \end{aligned} \quad (20)$$

Dispersal and germination — At every time step, each plant that has reached maturity attempts to reproduce a number of times given by that specific plants spawn rate attribute. The spawn rate is set to 1 for the purposes of this work. Each seed that a plant produces is dispersed to some position \mathbf{x}_j determined by an isotropic dispersal kernel centered at the plants position \mathbf{x}_i .

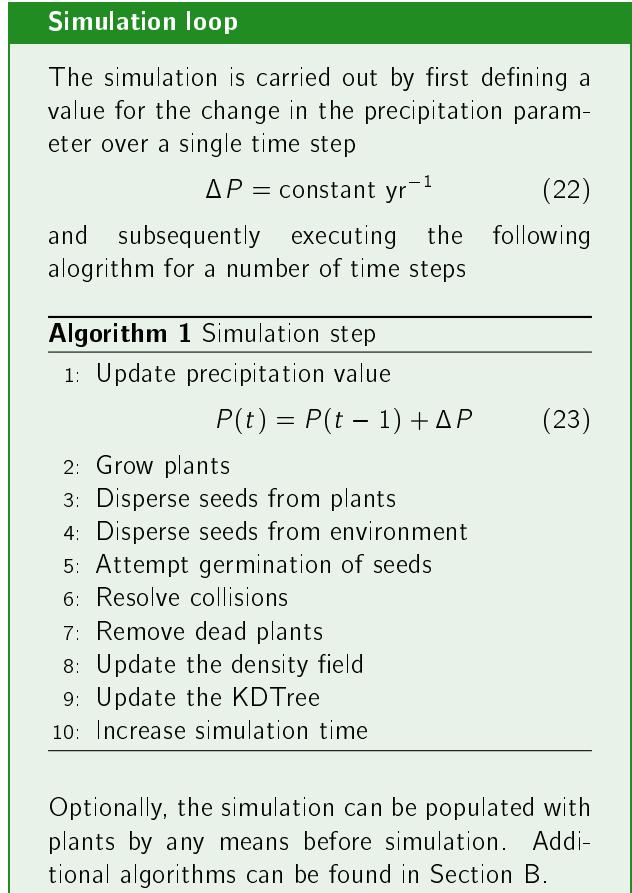
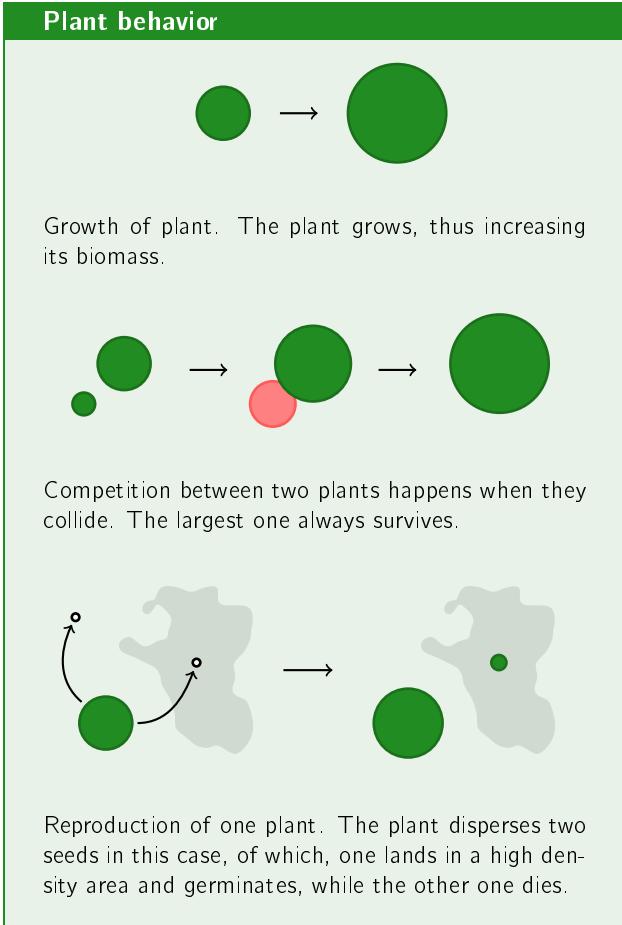
Along with the seeds dispersed by the plants inhabiting the domain, an extra effect is included of ambient spawning, by a constant *ambient spawn rate* which disperses a constant amount of seeds uniformly distributed across the whole domain each time step. If any seeds are dispersed to a position outside the domain boundary, it is handled according to the boundary conditions (See Section 5.3).

The probability of a seed germinating at its new position is given by the product between *soil quality* Q , the *precipitation* parameter P and the *plant density* at the position $D(\mathbf{x}_j)$, as such

$$P_{\text{germ}}(\mathbf{x}_j, t) = \max(Q, D(\mathbf{x}_j)) P(t). \quad (21)$$

It is crucial to note that the precipitation parameter is defined as a multiplicative probability factor scaling the germination probability globally, and thus it has no physical units (e.g. mm yr⁻¹). It is designed in this way in order to be analogous to the control parameter r as described in the mathematical theory of tipping points in Section 2. It is also sometimes referred to as the external condition of the models.

The crucial difference between the two models is how the density $D(\mathbf{x}_j)$ is calculated. In the local model, the density is estimated by taking into account only the plants within a certain distance from the seed position \mathbf{x}_j , while in the global model, the density is calculated globally, by taking into account all plants in the domain. A more detailed description of the density estimation can be found in Section 5.2



Simulation attributes and values*	
Plant position (x_i)	The position of a plant within the simulation domain.
Plant radius (r_i)	The current radius of a plant.
Plant biomass (b_i)	The biomass of a plant, equal to its area $b_i = \pi r_i^2$.
Minimum plant radius ($r_{\min} = 1 \text{ m}$)	The initial radius of a plant after germination.
Maximum plant radius ($r_{\max} = 30 \text{ m}$)	The maximum radius of a plant, above which it dies.
Plant growth rate ($G = 0.1 \text{ m yr}^{-1}$)	The rate at which a plant increases its radius.
Plant maturity size ($r_{\mathtt{mat}} = 1 \text{ m}$)	The size above which a plant can reproduce.
Plant dispersal range ($\sigma_{\text{disp}} = 90 \text{ m}$)	The standard deviation of the dispersal kernel of a plant.
Plant spawn rate ($S = 1 \text{ yr}^{-1}$)	The average amount of seeds a plant disperses during reproduction.

Domain side length ($L = 2000 \text{ m}$) The side lengths of the square domain which the plants inhabit.

Simulation time step ($\Delta t = 1 \text{ yr}$)

Ambient spawn rate ($S_A = 10^{-6} \text{ m}^{-2} \text{yr}^{-1}$) The average amount of seeds dispersed from the ambient environment outside the domain.

Soil quality ($Q = 0.01$) The ambient germination chance in the absence of plants and when precipitation is equal to 1.

Precipitation (P) The control parameter which scales the probability of germination for newly dispersed seeds. If $P = 0$, the germination of new seeds is impossible. Analogous to r in Section 2 and sometimes referred to as the external condition.

Total biomass density ($D = \sum_i b_i / L^2$) The total plant biomass currently in the simulation divided by the domain area.

*: The values given in this overview are used throughout unless otherwise explicitly stated in the text.

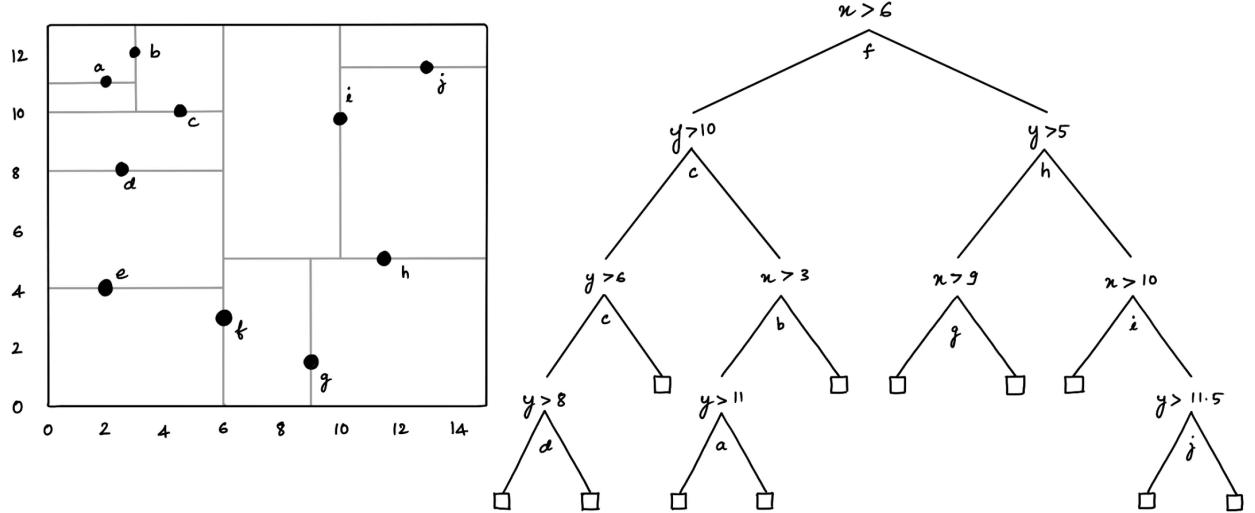


Figure 6: An example of a 2-dimensional KDTree data structure. (This graphic was taken from (Adimurthy 2022)).

5 Methods

5.1 Collision detection using the KDTree

The KDTree works by recursively dividing the domain of its elements into smaller and smaller pieces until each subpiece contains at most a preset number of data points. This results in a tree data structure which consists of a number of nodes n connected to its *children* by edges e . The *depth* of a node is the number of edges present in the path from the root node of the tree to that node (see Figure 6).

For a set of M data points

$$P = (p_1, \dots, p_M) \in \mathbb{R}^2 \quad (24)$$

we start by associating P with the first node n_1 in the tree. We now recursively do the following steps, until the size of the set of points associated with each node in the tree is less than or equal to the preset amount.

1. If the depth of the node is even, draw a vertical line through the median x -coordinate of the points. This splits the set of points associated with the node into two sets for which one contains all points to the left or on top of the line, and the other contains all points to the right of the line.
2. If the depth of the node is odd, draw a horizontal line through the median y -coordinate of the

points. This splits the set of points associated with the node into two sets for which the one contains all points below or on top of the line, and the other contains all points to above the line.

3. We create a new child node for each of the two new sets of points.
4. Repeat until the size of the set of points associated with each child node is less than or equal to the preset amount.

Note that for each point in the set there exists a node in the KDTree with two children representing the partition of the domain defined by the line running through that point.

The KDTree allows for fast calculation of distance matrix for its data points that are within a certain maximum distance of each other. In this case, the maximum distance is $2r_{\max}$ since this is the furthest two plants can be from each other and still have a chance of colliding. The distance matrix D becomes

$$D = \begin{bmatrix} 0 & d_{0,1} & \cdots & d_{0,N} \\ d_{1,0} & 0 & & d_{1,N} \\ \vdots & & \ddots & \vdots \\ d_{N,0} & d_{N,1} & \cdots & 0 \end{bmatrix}, \quad (25)$$

where $d_{i,j} = \|\mathbf{x}_j - \mathbf{x}_i\|$ is the distance between the two plants. These distances are then compared to the pairwise sum of radii $r_{i,j} = r_i + r_j$ to see whether the plants

are overlapping.

$$R = \begin{bmatrix} 0 & r_{0,1} & \cdots & r_{0,N} \\ r_{1,0} & 0 & & r_{1,N} \\ \vdots & \ddots & \ddots & \vdots \\ r_{N,0} & r_{N,1} & \cdots & 0 \end{bmatrix}, \quad (26)$$

$$d_{i,j} - r_{i,j} = \begin{cases} \leq 0 & \text{colliding} \\ > 0 & \text{not colliding,} \end{cases} \quad (27)$$

5.2 Density estimation

In the global model, the density is set to be the total biomass density of plants in the domain, such that no single place is more suited for germination than any other, but the overall probability rises with increasing total biomass and the maximum is 1.

This differs from the local model, where the probability of germination is linearly proportional to the biomass density at the position of the seed trying to germinate. The local density at a point is estimated using a *sum-of-kernels* approach, heavily inspired by the method presented in (Gingold and Monaghan 1977), which relies on the following analysis.

The value of some function A is approximated as the convolution with some smoothing function W

$$A(\mathbf{x}) = \int A(\mathbf{x}') W(\mathbf{x} - \mathbf{x}') d\mathbf{x}', \quad (28)$$

where W satisfies

$$\int W(\mathbf{x}) d\mathbf{x} = 1. \quad (29)$$

Of course, if we're trying to estimate the value of $A(\mathbf{x})$ we probably don't know the value of $A(\mathbf{x}')$. Therefore we can't evaluate (28), but if we have a set of N points $\mathbf{x}_i = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$, where we know the values of $A_i = A(\mathbf{x}_i)$, then the integral (28) can be approximated using a Monte Carlo method as such

$$A(\mathbf{x}) \approx \sum_i^N V_i A_i W(\mathbf{x} - \mathbf{x}_i), \quad (30)$$

where V_i is the volume of the i 'th particle, and A_i is the value of the quantity A for that particle.

Thus, when we wanted to estimate the density at some point in the domain $\rho(\mathbf{x})$, then from the above analysis, our approximation would yield

$$\rho(\mathbf{x}) \approx \sum_i^N m_i W(\mathbf{x} - \mathbf{x}_i), \quad (31)$$

where $m_i = \rho_i V_i$ is the mass of the i 'th particle.

Choosing a smoothing function — Many possibilities exist for the choice of a smoothing function. The only constraint is that the chosen function W integrates to 1. Some often mentioned possibilities are the normal, lognormal and cauchy distributions.

We choose the smoothing function W to be a 2-dimensional gaussian on the form

$$W(\mathbf{x}) = \frac{1}{2\pi h} \exp\left(-\frac{1}{2}\left(\frac{\mathbf{x} - \mathbf{x}_0}{h}\right)^2\right), \quad (32)$$

where \mathbf{x}_0 is the center of the smoothing function and h is the bandwidth of the function, a proxy for it's spatial extent.

One crucial way that our method differs from the one presented in (Gingold and Monaghan 1977), is that we take advantage of the fact that W integrates to 1 regardless of the size of the bandwidth h , meaning we are free to choose a specific bandwidth for each plant depending on the size of it. After some experimentation we've decided on the value $h = 2r_i$, where r_i is the radius of the plant. This has the effect that the biomass density contribution of larger plants reach further away from the plant position than that of smaller plants.

This all means that a given plant P_i with position \mathbf{x}_i , radius r_i and biomass b_i contributes to the density field with a gaussian on the form

$$\rho_i(\mathbf{x}) = b_i \exp\left(-\frac{1}{2}\left(\frac{\mathbf{x} - \mathbf{x}_i}{2r_i}\right)^2\right), \quad (33)$$

and the total density at a specific point can thus be approximated as the summation of the contribution of all of the individuals

$$\rho(\mathbf{x}) \approx \sum_i^N b_i \exp\left(-\frac{1}{2}\left(\frac{\mathbf{x} - \mathbf{x}_i}{2r_i}\right)^2\right). \quad (34)$$

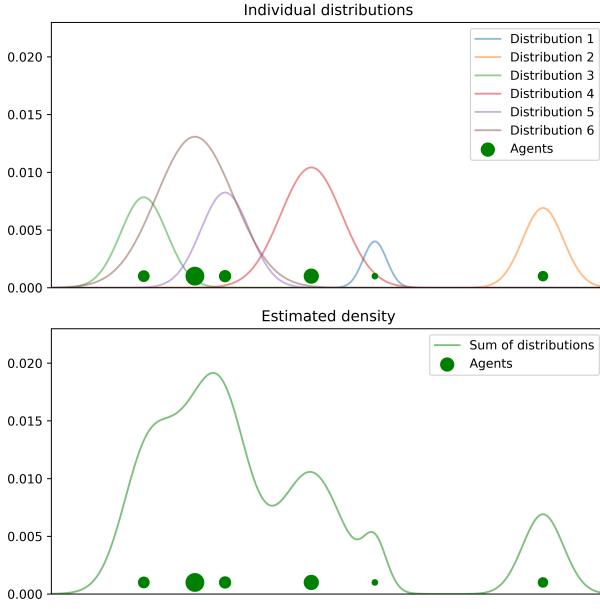


Figure 7: 1D demonstration of the density estimation approach.

Performance optimizations — To increase the speed of the simulation, the density estimation is calculated for each point in a grid with a predecided resolution each iteration and then when plants try to germinate they look-up the density value near them. This removes redundant density calculations for plants that are close to each other.

To further reduce the number of calculations, we can take advantage of the fact that a certain percentage of the density contribution of a plant will always fall inside a radius of width λh around the plant where $\lambda > 0$. Thus, the bigger λ is, the larger the percentage of the density contribution is guaranteed to be inside the radius. After some experimentation (see Section A), the value of $\lambda = 3$ was found to yield an average relative error around 1% for reasonable fields resolutions. Thus we define

$$r_{\text{query}} = 3h = 6r \quad (35)$$

as the “density query radius”, which allows for the construction of another KDTree structure to contain the grid points of the density field. This KDTree can then be searched to find all points x which are close enough to the plant for density calculation

$$|x - x_i| < r_{\text{query}}, \quad (36)$$

and skip the rest, speeding up performance.

5.3 Boundary conditions

The boundary conditions which were implemented for these models, assumes that the domain tiles the plain indefinitely. This means that plants close to one edge of the domain affects the collision detection and density estimation on the opposite side of the domain, extending to plants which are close to two edges, i.e. in a corner.

During reproduction, any seeds that are dispersed outside the domain are shifted to it's corresponding position inside the domain when tiled (see Figure 8)

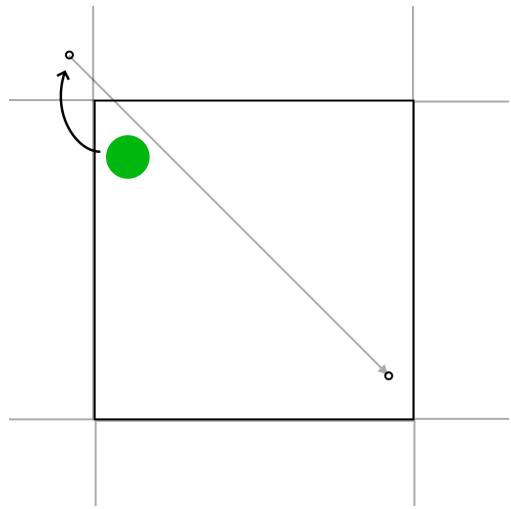


Figure 8: A seed dispersed to the bottom right corner of a tile outside the domain is shifted to the bottom right corner of the domain before attempting to germinate.

In the collision detection and density estimation we have to take into account that plants are not only points, but are spatially extended, both in the sense of their ZOI, but also now in the sense of their density query radius, wither one of which might reach beyond the domain boundaries, and affect the opposite side of the domain. Thus in the collision detection and density estimation, we create a number of duplicate “ghost” plants for all the plants that reach outside the domain depending on how many boundaries it extends beyond (see Figure 9). It should be noted, that the sense in which the plants can reach beyond the domain is different between whether the interest is in collision detection or density estimation, since the ZOI and query radius have different sizes, and thus different plants might get “ghost copies” depending on the specific calculation being made. In the collision detection, if a ghost plant either wins or loses in competition, the effect of this is carried over to the original plant. In the density estimation all plants including ghosts contribute to the density calculation (see Figure 10).

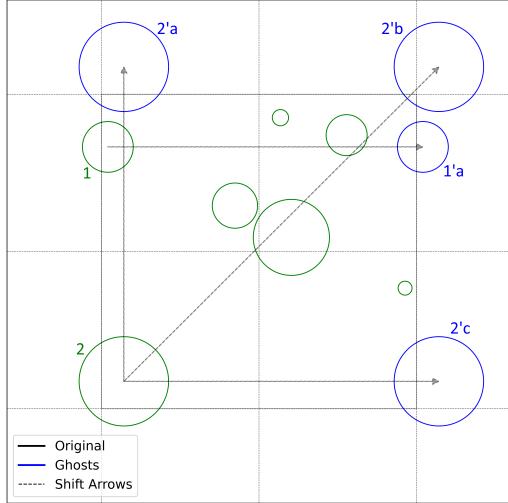


Figure 9: One plant (1) extending beyond the left edge is copied and shifted to its corresponding place on the tile to the right (1'a). Another plant (2), which reaches beyond both the left and bottom boundaries is copied and shifted three times according to which adjacent tiles it is overlapping (2'a) for the bottom tile, (2'b) for the diagonal tile to the bottom left and (2'c) for the left tile. After the shifting has happened some ghosts might collide with other plants or each other, and the consequences of these collisions are carried over to their original plant, but here no collision occurred. Note that the plants are large compared to the domain size here only for demonstration purposes.

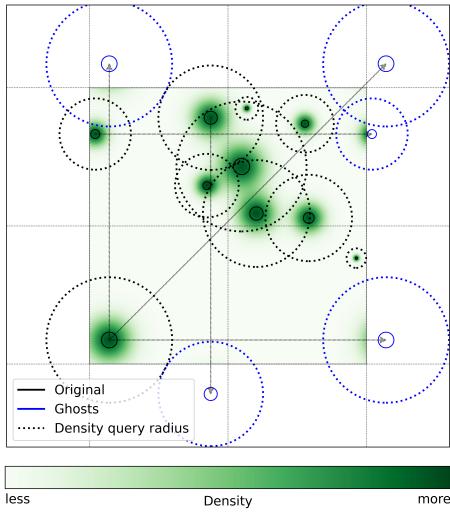


Figure 10: The plants (black circles) which have density query radii (dotted lines) that reach beyond the boundaries are duplicated and shifted (blue circles) accordingly, before the density calculation (green “clouds”). Note that the plants are large compared to the domain size here only for demonstration purposes.

5.4 Distributing non-overlapping plants

To have the ability to initialize simulation states at a specific target density with somewhat reasonable distributions of plant sizes and positions, an algorithm was created. The algorithm iteratively generates plants with a radius drawn uniformly from a predetermined range and a position drawn uniformly from the positions available in the domain. Each time a new plant is generated the algorithm makes a check of whether it collides with another, already initialized plant. If the answer is no the plant is planted, if not, it is discarded and another one is generated. An example of an initial distribution can be seen in Figure 11.

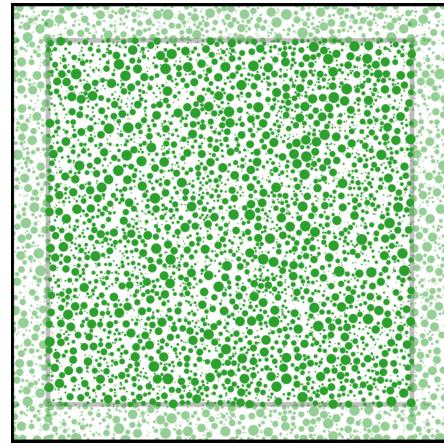


Figure 11: A distribution generated with the iterative spawning algorithm for a domain with side lengths of 2000 m.

5.5 Temporal EWS

Variance and Skewness — For a time series of N data points x_i , $i \in [1, N]$, the mean \bar{x} is given by

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i, \quad (37)$$

the variance σ^2 is given by

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \quad (38)$$

and the skewness γ_1 is then given by

$$\gamma_1 = \frac{1}{N} \sum_{i=1}^N \left(\frac{x_i - \bar{x}}{\sigma} \right)^3. \quad (39)$$

Autocorrelation — Given a set of N data points $x_i(t)$, defined by a rolling window of size N where t denotes

the central time of the rolling window, the data is detrended by fitting a 3rd degree polynomial $P_3(t)$ and subtracting it from the data

$$\hat{x}_i(t) = x_i(t) - P_3(t). \quad (40)$$

The autocorrelation at lag 1 of the detrended data is then given by

$$\rho_1 = \frac{1}{N} \sum_{i=1}^N \hat{x}_i(t-1) \cdot \hat{x}_i(t). \quad (41)$$

5.6 Spatial EWS

Spatial correlation — The spatial correlation is analogous to autocorellation, but where the lag is in distance and not in time. It thus quantifies how alike the system is at different timescales.

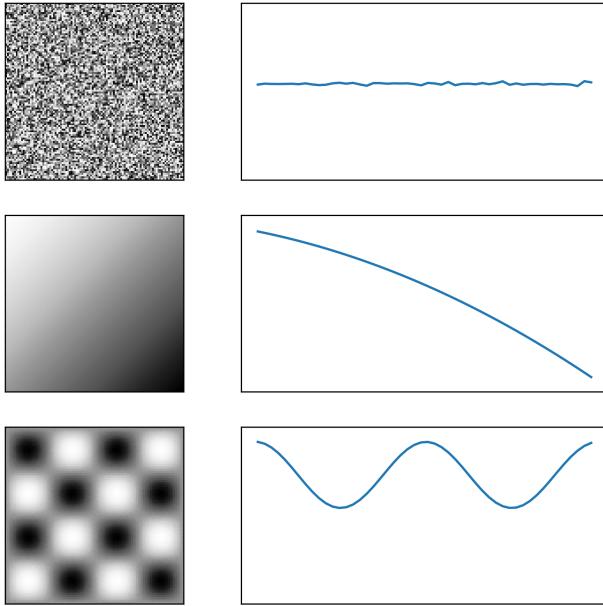


Figure 12: The spatial correlation calculated for three different example fields. The first case is for random white noise, and has very low correlation. The second example is a smooth gradient, which is reflected in the downward trend of the spatial correlation. The last example is one of a periodic sinusoidal signal, which shows up clearly in the correlation graph.

The spatial correlation as a function of distance r for a discrete real valued function in 2-dimensions $f(x, y) = f_{x,y}$ is (Kéfi et al. 2014)

$$C_2(r) = \frac{MN \sum_{i,m,j,n=1}^{M,M,N,N} w_{i,j;m,n} (f_{i,j} - \bar{f})(f_{m,n} - \bar{f})}{W \sum_{m,n=1}^{M,N} (f_{m,n} - \bar{f})^2}, \quad (42)$$

where the summation notation

$$\sum_{i,m,j,n=1}^{M,M,N,N} \text{ is equivalent to } \sum_{i=1}^M \sum_{m=1}^N \sum_{j=1}^M \sum_{n=1}^N,$$

and

$$\bar{f} = \frac{1}{MN} \sum_{m,n=1}^{M,N} f_{m,n} \quad (43)$$

is the mean of f and $w_{i,j;m,n}$ is 1 if the points (i, j) and (m, n) are separated by a distance of r and 0 otherwise and W is the total amount of points that are separated by distance r .

Spatial power spectrum — Another way to probe the spatial statistics of a system is by calculating the *Discrete Fourier Transform* (DFT) of the system data. The DFT $\hat{f}(x, y) = \hat{f}_{x,y}$ of a discrete real valued 2-dimensional function $f(x, y) = f_{x,y}$ is calculated as

$$\hat{f}_{x,y} = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} (f_{m,n} - \bar{f}) e^{-2\pi i (\frac{mx}{M} + \frac{ny}{N})}, \quad (44)$$

which by $e^{i\theta} = \cos(\theta) + i \sin(\theta)$, can be separated into real and imaginary parts with coefficients $a_{x,y}$ and $b_{x,y}$ respectively

$$\hat{f}_{x,y} = a_{x,y} + i b_{x,y} \quad (45)$$

The coefficients of the DFT $a_{x,y}$ and $b_{x,y}$ allow for the calculation of a power spectrum

$$I_{x,y} = \frac{1}{MN} (a_{x,y}^2 + b_{x,y}^2), \quad (46)$$

which describes the intensity of the spatial frequencies in the data.

The power spectrum can be used to calculate the r -spectrum, and the θ -spectrum, which describe the intensity of the spatial frequencies as a function of distance r or angle θ . The r -spectrum allows gives insight into the periodicity of the data, and the θ -spectrum gives insight into the *isotropy*, that is the circular symmetry, of the data. If the data is very periodic, the r -spectrum will show a peak at the distance of the periodicity, and if the data is anisotropic, then the θ -spectrum will show a peak at the angle of the anisotropy. The r - and θ -spectra are well suited for the characterization of spatial patterns in the data, and are calculated by

$$I(r) = \sum_{x,y} I_{x,y} \delta\left(r - \sqrt{x^2 + y^2}\right), \quad (47)$$

$$I(\theta) = \sum_{x,y} I_{x,y} \delta(\theta - \arctan(y/x)), \quad (48)$$

where δ is the Dirac delta function (Kéfi et al. 2014).

6 Results

6.1 Multistability and hysteresis

To examine the general stable and unstable regimes of the two models, a number of realizations of each model was generated for different fixed values of precipitation. For each unique precipitation value, plants were initialized at different densities using the non overlapping spawning algorithm (see Section 5.4), and runs were carried out for a maximum of 5000 iterations, or until sufficiently converged (see Figure 13). For each realization that converged, the metastable value of total biomass density calculated as the mean over the last 1000 steps, was recorded along with its precipitation value (see Figure 14).

It is apparent that for each model, when the precipitation values are above a certain value $P_c \approx 0.6$, the realizations reach stability in the sense that noisy perturbations are always recovered around a fixed stable value. For precipitation values close to but still above P_c , we see that stability is still achieved, but variability in biomass density is generally larger in magnitude here than for values well above the critical value. For precipitation values below P_c , the simulations all eventually die out in both models.

It is clear from Figure 14 that both models exhibit multistability for a range of precipitation values. In the local model, multistability was attained for precipitation values in the range $P \in (0.6, 0.18)$ whereas in the global model the range of values is $P \in (0.6, 3.8)$.

In light of these results, an experiment was carried out to examine to which degree, a uniformly distributed initial population, would traverse the window of multistability when subject to slow change of external conditions in the case of each model. 3 realizations of each model were calculated starting at a precipitation value of $P = 0.5$, which was decreased at a rate of $\Delta P = -4 \times 10^{-5} \text{ yr}^{-1}$ until collapse, and then subsequently increased again until the original biomass density was reached (see Figure 15).

It is seen that both models starting at 0.5 precipitation exhibit accelerating decay towards the tipping point and collapse during the decrease in precipitation. Subsequently as the precipitation is increased again, it is seen that both models transition back to the vegetated state, both following a different trajectory through the state space. Specifically, the local model transitions back for precipitation values of around $P > 0.16$, whereas the global model transitions back only for precipitation values of around $P > 0.36$.

It is worth noting that while this resembles hysteresis loops, in theory there is a high chance that the sizes of the windows of hysteresis for the two models are actually 0. This is because when the system enters the window of multistability, and is given infinite time, it will, at some point, transition to the alternative state due to noise-induced tipping.

This effect is seen in Figure 15, since the transitions from barren to vegetated do not happen for the same values of precipitation, as the precipitation is increased. This indeed happens due to the stochastic dynamics of the models.

Examining the dynamics of repopulation for one realization of each model yields Figure 16, in which it is seen that repopulation in the global model happens as a sudden, rapid filling of the domain, whereas repopulation in the local model happens as an aggregation around of nucleation site, where density is high and the local feedback allows for this site to grow in size.

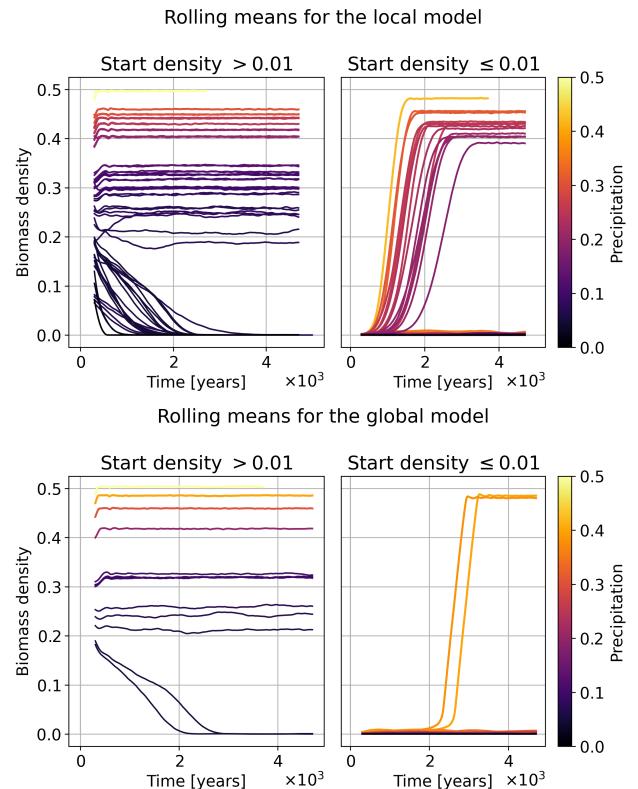


Figure 13: Rolling mean of the total biomass density for realizations with different fixed precipitation values initiated at different starting densities for both models.

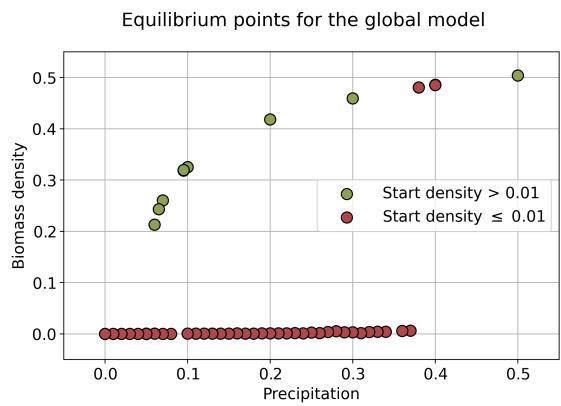
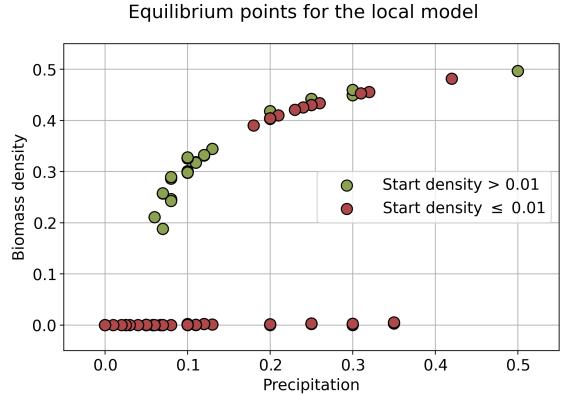


Figure 14: Stable points of realizations with different static, initial conditions for the local and the global model. Green points correspond to starting in a vegetated state, whereas red points correspond to starting in the bare state.

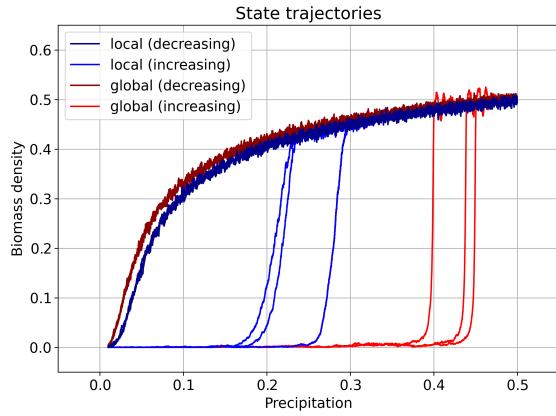


Figure 15: Plot of the state trajectories in terms of the total biomass density vs precipitation for 3 runs of each model with increasing and decreasing precipitation trajectories distinguished by color. It is clearly seen that repopulation happens for lower precipitation values in the local model than in the global one. The effect of the inherent stochasticity is also apparent in the fact that repopulation happens at different times even for realizations of the same model.

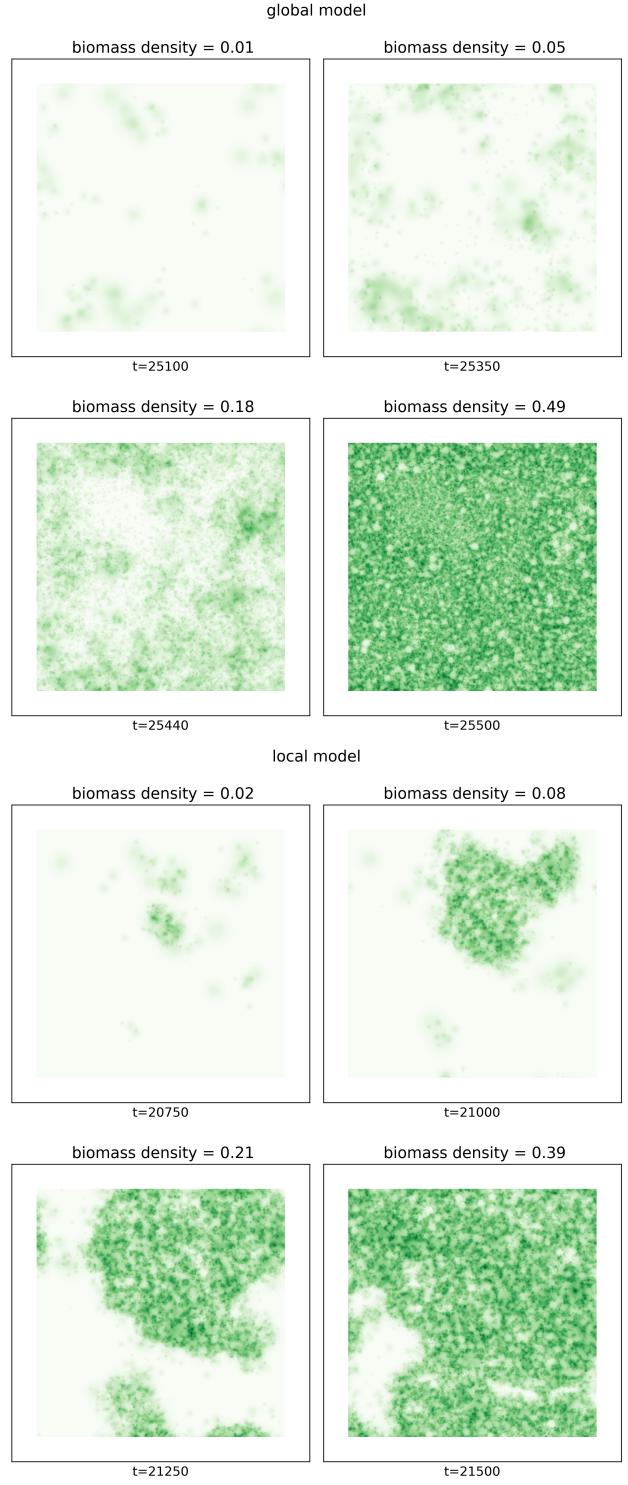


Figure 16: The course of repopulation plotted for one realization of each model. Differences is clearly seen in the manner in which repopulation happens. In the global model it happens suddenly and all over the domain, wheras in the local model it happens gradually stemming from an initial “nucleation site”. Note that both, the time and the rate at which repopulation happens is different between the two models.

6.2 Resilience and early warning signals

To analyze the effect of critical slowing down, an experiment was carried out to estimate the characteristic recovery time of the two models when subject to perturbations from their stationary state at different fixed values of precipitation. Each sample was taken by first letting a realization reach metastability for a fixed precipitation value, then subsequently perturbing it by removing a fraction equal to 25% of the total biomass from the simulation in one time step. The realization was then run until the realizations converged to metastability again. It should be mentioned that the size of the perturbations is not small as in theory required for linear response, which is the basis for critical slowing down. However a perturbation of this size was needed for detection through the noise. Finally, the time series data of the biomass was fitted to an exponential on the form

$$y = \alpha e^{-t/\tau} + \beta, \quad (49)$$

where τ is identified as the characteristic recovery time. In the end all samples reached stability following the perturbations.

A clear increase in the characteristic recovery time in response to perturbations values is seen for both models for decreasing precipitation.

In the climate system, there are rarely ways to perform controlled perturbations in the sense of the above experiment, but a measure of the resilience of the system is still sought. Therefore early warning signals are used to quantify the behavior of the systems inherent fluctuations, which are themselves in a way a kind of perturbations. For this purpose a 3 realizations of each model were analyzed.

Each realization had plants initiated uniformly on the domain using the non overlapping spawning algorithm and started at a precipitation value of $P = 0.5$, above the critical value $P_c \approx 0.06$. The precipitation was subsequently decreased slowly with time ($\Delta P = -4 \times 10^{-5} \text{ yr}^{-1}$) until the realizations collapsed. The point of tipping for each model was estimated by calculating the mean of the realization biomasses and finding the point of greatest slope. All subsequent analysis of EWS was done after a transient period of 1000 time steps, and before the estimated tipping point.

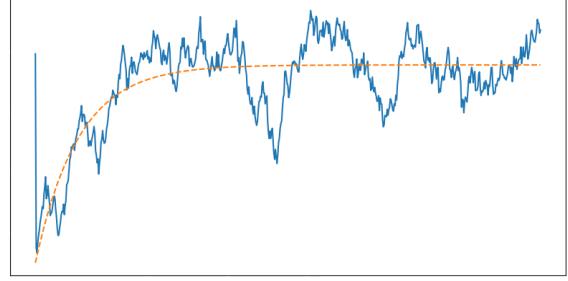


Figure 17: Plot of a sample run (blue) immediately following a perturbation along with its exponential fit (orange).

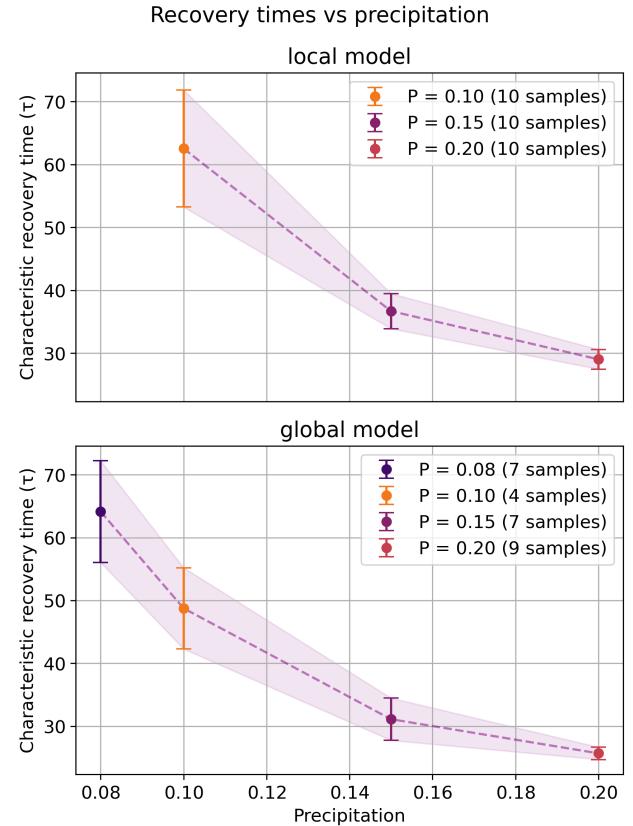


Figure 18: Plot of the mean and standard deviation of the characteristic recovery time of ensembles of perturbation recovery runs obtained for different precipitation values.

To calculate the temporal EWS of each realization of the two models, the time series of the realization was analyzed in the following way. Within a rolling window the signal was first detrended by subtracting a 3rd degree polynomial fit, and then subsequently the variance, skewness and lag1-autocorrelation was calculated of the detrended window. For each of the windows a mean and standard deviation was calculated for each of the four statistics across the different realizations of each model. Additionally, the mean biomass density of the realizations of each model is plotted (green) on top of the

the individual runs (grey), accompanied by the ensemble statistics in Figure 19.

In both models a clear increase in lag1-autocorrelation is observed as the precipitation is decreased with time, with a marginally larger increase in the global model. The variance increases slightly in both models, but since CSD predicts a divergence in variance these increases are deemed statistically insignificant. The errors on the skewness are large enough for the trends seen to be caused by fluctuations, and thus these too are deemed statistically insignificant.

On Figures 20 and 21, the time series data of the total biomass density, population size and precipitation of two sample realizations used for the analysis can be seen. It is seen that both models exhibit similar behavior. It is also seen, that there are obvious differences in the response of the biomass and population data to the changing precipitation. The biomass has a non-linear response, wheras the population has a linear response. It is also apparent that for both models, the variance in the biomass increases only very slightly if at all before tipping happens, which agrees with the analysis. However, it is seen from the figures that the variance of population data actually decreases with decreasing precipitation.

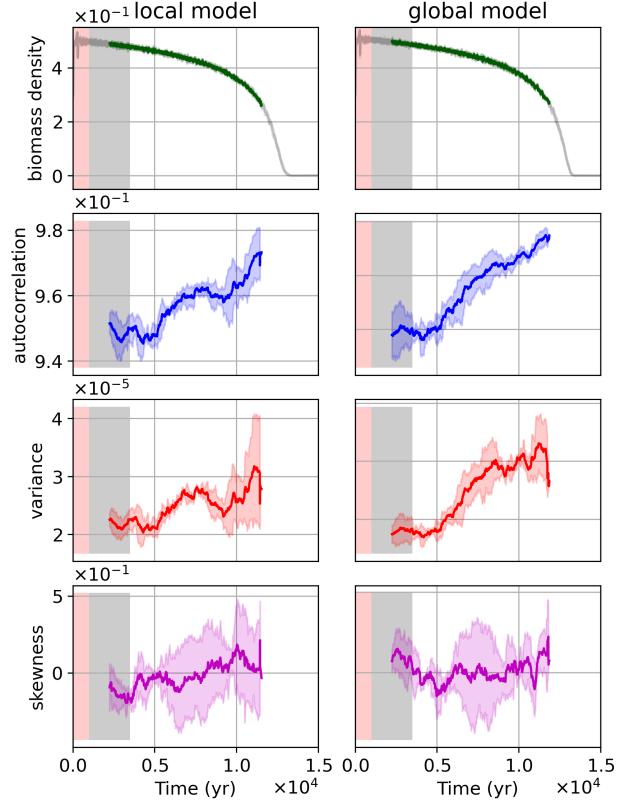


Figure 19: Temporal statistics of ensemble of runs for both models including uncertainties. The statistics were done in a rolling window of size 2500 (grey) starting after a transient period of size 1000 (red).

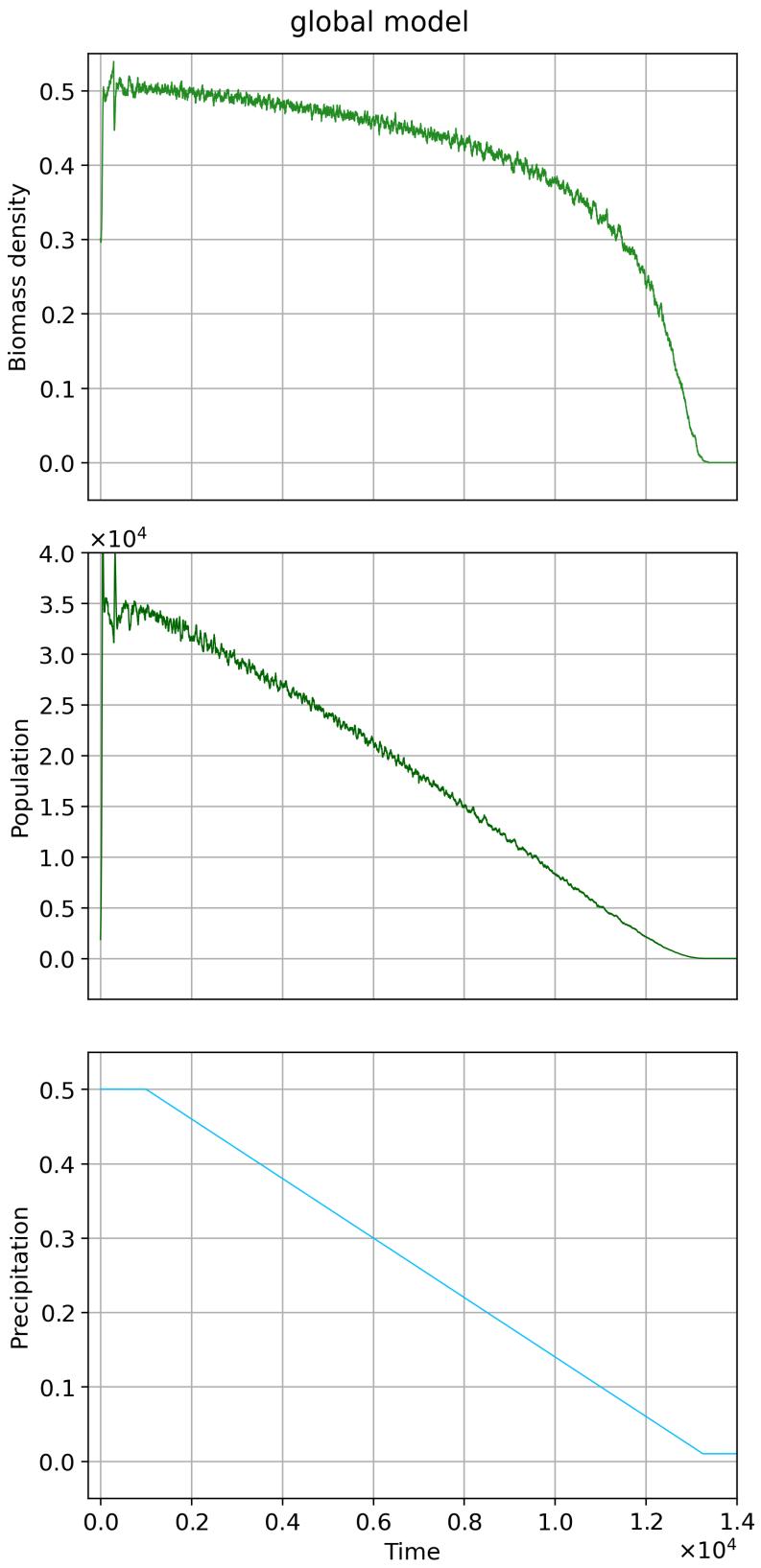


Figure 20: Plot of the time series data of one of the realizations of the global model which was used in the analysis of EWS. The data is plotted in this large format and with a thin line width to better make out the fluctuations in the data. Precipitation was decreased by $\Delta P = -4 \times 10^{-5} \text{ yr}^{-1}$. The biomass density is seen to have a non-linear response to the decrease in precipitation, whereas the population is seen to have a linear response.

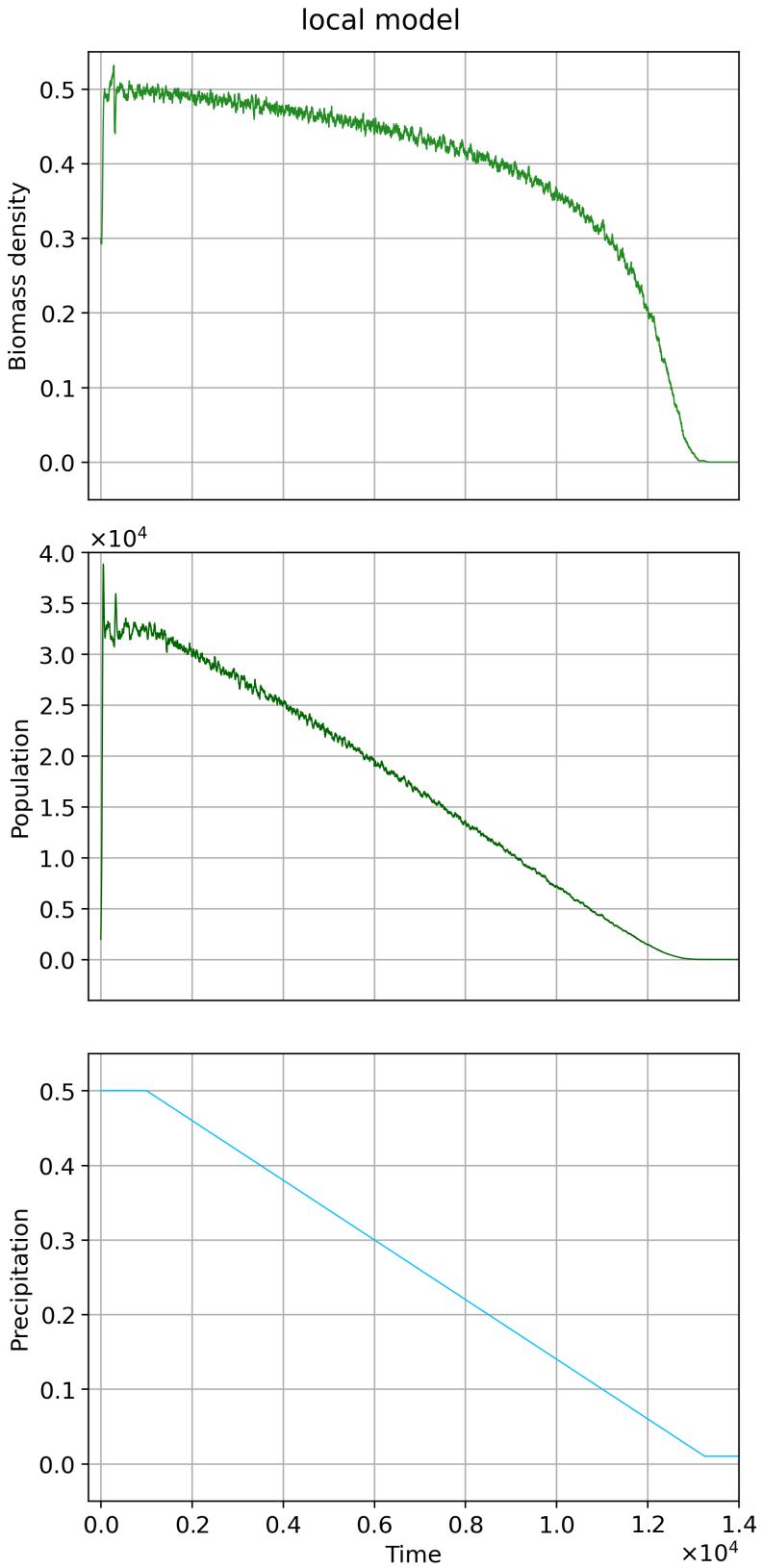


Figure 21: Plot of the time series data of one of the realizations of the local model which was used in the analysis of EWS. The data is plotted in this large format and with a thin line width to better make out the fluctuations in the data. Precipitation was decreased by $\Delta P = -4 \times 10^{-5} \text{ yr}^{-1}$. As was the case of the global realization of Figure 20, the biomass density is seen to have a non-linear response to the decrease in precipitation, whereas the population response is seen to be linear.

When a system is governed by spatially explicit rules, where the strengths of interactions between nearby elements vary with distance, the effect of CSD can manifest itself as a spatial phenomenon, being observable in the spatial correlation or as changes in the power spectrum of spatial patterns present in the system (Kéfi et al. 2014). Therefore, these additional spatial statistics were calculated for the above realizations in the following way. Four equidistant target times between the end of the transient period at 1000 time steps and the estimated tipping point were chosen as basis for the spatial analysis (see Figure 22). For each of the target times in each of the model realizations, a density field was estimated, from which the spatial correlation as a function of distance was calculated, along with a 2-dimensional power spectrum image, which was integrated over angles and over distances to get the r - and θ -spectra respectively (see Figures 23 and 24).

It is seen that as the precipitation is decreased spatial correlation increases in both models, increasing more at low distances. Both models exhibit the largest change in spatial correlation in the span of the last time step, which is also reflected in the plots of the density fields. Additionally, the r -spectrum for both models show an increase in power at low distances as the precipitation is reduced, however this signal is not very significant. The generally uniform distribution of power in the θ -spectrum remains throughout in both models.

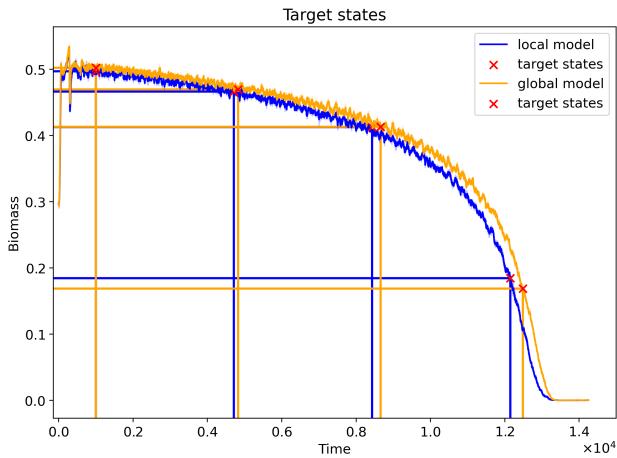


Figure 22: Plot of the times at which the spatial statistics were measured.

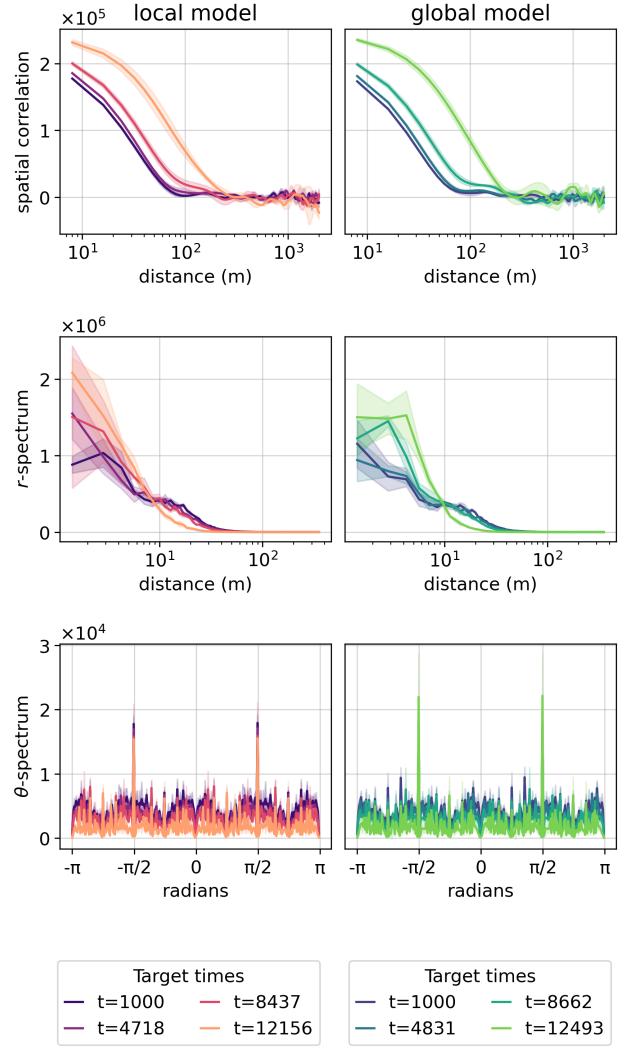


Figure 23: The spatial correlation as a function of distance along with the r - and θ -spectra for both models calculated at 4 equidistant times between the end of the transient period and the estimated tipping points. A clear increase in the spatial correlation is seen as the tipping point is approached. The specific times at which the calculations were done are written in the legend on the bottom. Note that the x-axis for the spatial correlation and r -spectrum plots scales logarithmically.

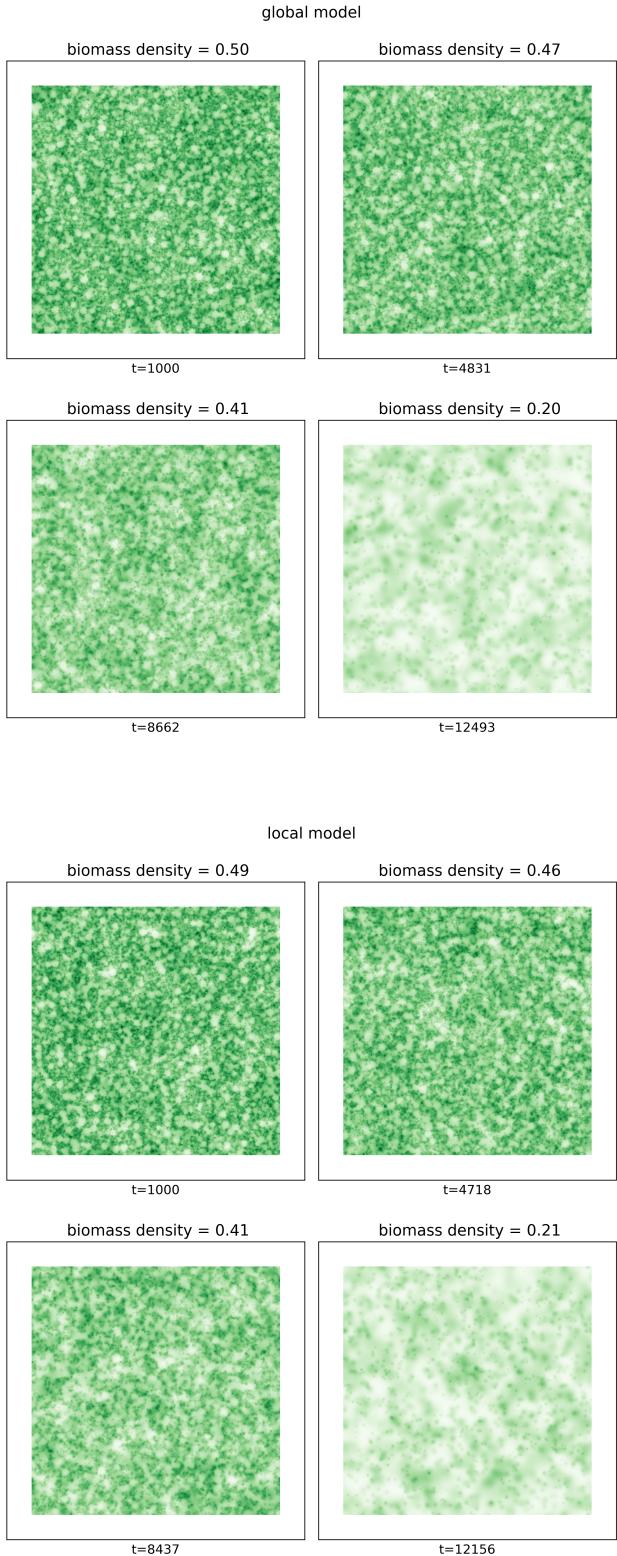


Figure 24: Plot of the density fields of one realization of each model at the four target times used as the basis for the spatial EWS analysis.

7 Discussion and conclusion

In the investigation of multistability, evidence was found in both models for the existence of a barren state for low precipitation values, and a vegetated state for high values, both being metastable states because of the inherent stochasticity of the system dynamics. No evidence was found for high multistability, i.e. more co-existing states, but this can't be ruled out as other states might be attainable from different initial conditions. This could be tested in the future with larger ensemble simulations.

Differences were found between the two models. The window of precipitation values in which multistability occurs was effectively wider for the global model. The fact that repopulation of the domain could be achieved for much lower precipitation values in the local model, suggests that the local density feedback is a greater promoter of plant repopulation than the global feedback. This might be caused by the fact that, for the global model to transition from a barren state to a vegetated state requires the reproduction of ambiently dispersed plants to sustain a global density large enough to promote population growth. In contrast, the local model only requires for local populations of plants to sustain themselves and grow, which is possible regardless of the population size outside their immediate vicinity, making the transition from a barren state to a vegetated state more attainable in this case. This is seen in Figure 16. In the future, experiments could feature changes in the size of the domain in order to address whether this effect persists for larger ecosystems.

Interestingly, the position of the tipping point was almost identical between the two models. For any given value of precipitation, the population, and thus the density of a realization fluctuates around a value fixed value determined by the relative strengths between the inhibitory pressure of local competition and the facilitative pressure of reproduction. This is true for both models. The reason for the tipping point being virtually identical in both models might then be because this is the precipitation threshold beyond which, the largest possible probability for reproduction (whether attained locally or globally) is not enough to sustain a population when subject to competition, rendering survival of the vegetated state impossible. Surprisingly though, the local feedback does not lead to an earlier collapse as may be expected since local feedbacks usually can destabilize the system in both directions, the idea being that spontaneously occurring local barren patches could in theory grow and cause a collapse "from within", but while this was not observed in the experiments done here, it would be an interesting subject of future experiments.

In the investigation of the resilience, clear evidence was found in both models for loss of resilience for decreasing precipitation values in both models, indicating the presence of critical slowing down regardless of the feedback mechanism at play. Additionally, the rate at which resilience decreased was found to be larger at lower precipitation values, with models exhibiting roughly the same rate of decrease. This is an interesting result since the models are individual-based and thus no general theory for tipping and early warning signals exist in this case. Although work is being done in developing such a theory for other kinds of individual-based models (Zagli, Lucarini, and Pavlou 2024). Nonetheless, it should be noted, that it is physically quite intuitive for recovery times to get longer at lower precipitation, as the probability for germination is lower and thus repopulation will on average take longer. However, the generality of this result is lacking since the removal of plants uniformly across the domain is a very specific kind of perturbation. Additional types of perturbations could be performed by removing plants locally to simulate for instance response to wildfires or adding or redistributing plants. Different types of perturbations might also affect the rate of increase of recovery times differently depending on the model. In future experiments these perturbations could be included in order to see whether the results found here hold more generally.

The decrease in resilience suggests that CSD is happening, which was indeed partly reflected in the results of the temporal EWS analysis, which showed an increase in autocorrelation for decreasing precipitation. The variance showed only a slight, roughly linear increase characterized by larger uncertainties, and not a divergence as expected from the theory of CSD. Skewness results showed no evidence of oncoming collapse.

Comparing these results with those of the resilience experiment, it is apparent, that the rising increase in the rate of recovery seen in the resilience experiment was not seen in the variance results. This might be caused by the fact that the population decreased as precipitation was decreased and thus the total contribution of plants through growth and mortality on the fluctuations in the biomass should in theory also decrease. Another reason might be that the observable used in the analysis, which was biomass in this case, might not be the right observable of the system with regards to detecting EWS. It could be that the variance of other observables might have different responses and thus EWS could maybe be observed in the analysis of these. This is backed by Figures 20 and 21, which show that the variance of the population observable has a different response, decreasing with decreasing precipitation and is also subject of on-

going research as in the case of (Lohmann et al. 2025). Future experiments could be carried out with the purpose of identifying for particular observables of the system, which do exhibit a significant increase in variance. Additionally, it should be noted, that the perturbations in the resilience experiment are 25% of the total biomass and are therefore not small, which is in theory needed for the linear response which CSD is based on. This might also be playing a role with respect to the rising increase in the rate of recovery for smaller precipitation values.

The spatial EWS analysis showed a clear increase in spatial correlation indicating the oncoming tipping point. Moreover a slight aggregation of power at low distances was seen in the r -spectrum was observed in both models, but a closer look at the density fields of the experiment (Figure 24) showed that very little if any significant pattern formation took place, and thus it is hard to conclude anything from this result. Finally a uniform distribution in the θ -spectrum was seen for both models. This reflects the isotropic nature of the density contributions of each plant and was expected, however spontaneous symmetry breaking could happen, leading to anisotropic patterns, but this was not observed here.

Future experiments involving different values of the parameters governing the spatial feedback, such as the density kernel and bandwidths, along with dispersal ranges and plants minimum and maximum sizes could be carried out in order to examine whether spatial pattern formation is attainable and if so whether this could avoid tipping in the local model as has been shown to happen in various studies (Rietkerk, Boerlijst, et al. 2002; Vincent et al. 2016; Pałubicki et al. 2022).

It is worth noting that the sample size underlying the EWS results is rather small (3 sample realizations for each model). This is due to the runtime performance of the models being quite low restricting the amount of data which could be calculated. Ideally, as many realizations as possible should be gathered to get more accurate statistics and lower uncertainties on all of the results featured in the EWS analysis.

In conclusion the local-scale individual-based models constructed in this project were found to exhibit tipping dynamics containing both multistability and a form of bifurcation. Localized feedback was found to decrease the window of multistability thus raising the question of whether further increases in complexity with regards to the modelling of feedback mechanisms could lead to a further decrease in the range of multistability. The existence of a relatively large multistability is a major factor in whether a tipping point is catastrophic or not, since it governs to what degree the tipping is irreversible. Lastly,

the main idea underlying the possibility of predicting tipping points, i.e. loss of resilience because of critical slowing down, seems to be preserved, despite no saddle-node bifurcation being explicitly modeled. Measuring this loss of resilience via a decrease in variance was found to be ineffective however, whereas examining autocorrelation and spatial correlation proved more insightful, drawing parallels to the origins of critical slowing down as observed in phase transitions, where transitions are preceded by an increase in correlation length.

8 Outlook

This project is meant as a contribution to the study of tipping points and early warning signals in the global climate system. It seeks to contribute to the theoretical understanding of tipping points and early warning signals in the context of systems with spatially heterogeneous feedbacks, of which many exist in the global climate system.

8.1 Future developments

The models created in this project have a large set of possible parameters which for the purposes of this project were set to constant values and not changed. Future experiments could probe this large unexplored parameter space for other interactions between the inherent feedbacks in the model. For example increasing the maturity size of plants could regulate the interplay between the feedbacks of reproduction and competition as only large plants, i.e. plants that survive for a while, get to reproduce. This regulation of the feedback mechanisms could also be explored by changing kernels governing density estimation or seed dispersal, which might also stimulate pattern formation in the plant populations allowing for the investigation of the viability of this mechanism as an early warning signal.

The global model featured here, is still technically spatially explicit, since the mechanisms of competition and dispersal are modeled as stemming from the plant individuals. Thus alternative implementations of these mechanisms are possible. For example dispersal could be modeled as a spatially uniform mechanism by scaling the already existing ambient dispersal mechanism by the number of plants present in the simulation and disabling dispersal by plant individuals. Similarly, competition could be implemented as a stochastic spatially averaged process, where overlapping plants stop growing instead of competing, and then a given number of plants which have stopped growing could be chosen to die ran-

domly at any given time step, removing the spatially explicit part of this mechanism. This might present further observable differences between systems with and without spatially explicit feedback mechanisms.

Another possibility for the future development of the methods in this paper is the implementation of additional mechanisms of forest ecosystems such as photosynthesis, nutrient uptake, the complex spatial extent of vegetation (roots, branches, leaves, etc.) along with interactions between vegetation and climate through evapotranspiration for example. The modelling of plants leaving behind biomass when dead is also disregarded, but could potentially serve as another local positive feedback mechanism.

Also it could be interesting to examine the effects of seasonal variations in precipitation as this might be a mechanism which facilitates spatial pattern formation (Vincenot et al. 2016).

The base processes modelled here present a plethora of possible future adaptations, such as introducing different plant species, the introduction of depopulation through grazing or forest fires, adding the effect of terrain as has been done in some studies of alpine trees along with (Pałubicki et al. 2022) and (Rietkerk, Boerlijst, et al. 2002), examining the effects of evolution, by having species evolve over time, by slightly mutating their attributes and letting natural selection find the fittest trees. This might uncover scenarios in which two or more species form symbiotic relations where each one contributes to a positive feedback, which wouldn't be strong enough to support either species alone.

8.2 Computational considerations

An obvious next step in all projects involving computational models is the optimization of the implementation. Some different ways in which the models featured could be optimized in the future are presented here.

8.2.1 Structure of arrays

One method for the optimization of the models featured in this project is a reconsideration of the data structures and the manner in which the bulk of the simulation data is stored while the simulation is running.

The models used in this project, store data by assigning a position and radius to each plant individual, and subsequently storing all plants individuals in an array. This approach is called the *array-of-structures* approach (AOS), the structures being the plants in this case. One benefit of the AOS approach is that the methods by which the plants behave can be modeled on a per in-

dividual basis making for very intuitive implementation with little abstraction from the real-world processes that inspired the behavior, but this can come at a cost. As is well known in scientific computing, when having to calculate the same thing a lot of times, vectorization of data can speed up performance dramatically.

But using the AOS approach, the vectorization of quantities has to be done by first iterating through all the plants in our list, and collecting the relevant data, and then subsequently gathering it in a vectorized fashion. This initial iteration through the plants is associated with a significant overhead in computation, making it a problematic option for optimized computation times. Another one of the problems with the AOS approach in our case, is that the amount of plants in the simulation is always varying, as some might die and need to be removed and some might be spawned due to reproduction. This means that under the hood, the computer has to allocate, disallocate and reallocate memory to the storing of plant data, which is also associated with significant computational overhead.

An alternative to the AOS is the *structure-of-arrays* approach (SOA), where relevant parameters of the plants (the structures) such as the positions and radii, are stored in one array, as opposed to the *array-of-structures* (AOS) approach, where all plants are stored in a list thus achieving vectorization.

Using instead the structure-of-arrays (SOA) approach, the data of all plants are kept in a single structure, in which data frequently used in calculation, such as the positions, radii, etc. are stored in a vectorized fashion, speeding up calculations. And even though the SOA approach doesn't completely remove the problem of the amount of plants varying over time, it is very easy to define a preset size for the data structure (using maybe an estimate of how many plants will be in a single run) and thus, the memory that the data structure needs can be assigned at the inception of the simulation, and only has to be resized if the amount of plants exceeds the initial estimate, reducing the amount of memory allocation and reallocation.

8.2.2 Interpreted vs. compiled languages

The computer models featured in this project were implemented in the programming language python, which is well known as a ubiquitous tool for scientific computing, statistics and data analysis. Python is an *interpreted* programming language as opposed to a *compiled* programming language, and thus it has the ability to run code as soon as it has been written, making development and debugging of features fast and flexible. This

is achieved by compiling the python code into intermediate bytecode which is then interpreted by a bytecode interpreter as the code runs. But this is associated with a significant overhead. Examples of interpreted languages are Python, MATLAB, JavaScript and R.

In contrast compiled programming languages, are languages where the code has to be compiled to machine code, before being executable. This means that every time the code has been edited a new version of the software has to be compiled, which can make the development of the program less flexible and more time consuming to debug etc. But one major upside to using compiled languages is that all of the overhead associated with compilation happens only when the program is built, and therefore not when it is run. Compiled languages can therefore generally be used to create much more optimized code, because when compiled to machine code, the computer is much faster at executing the commands. Examples of compiled languages are C, C++, C#, Java, Julia and Rust.

Some modules for python such as Cython and Numba do circumvent the interpreted nature of the language by allowing the compilation into machine code or the parallelization of specific compatible parts of the code before runtime.

For these reasons, rewriting the code in a compiled programming language or at least using one of the python libraries mentioned could be possible angles of attack in order to optimize the models used in this project.

A Performance metrics of the density estimation

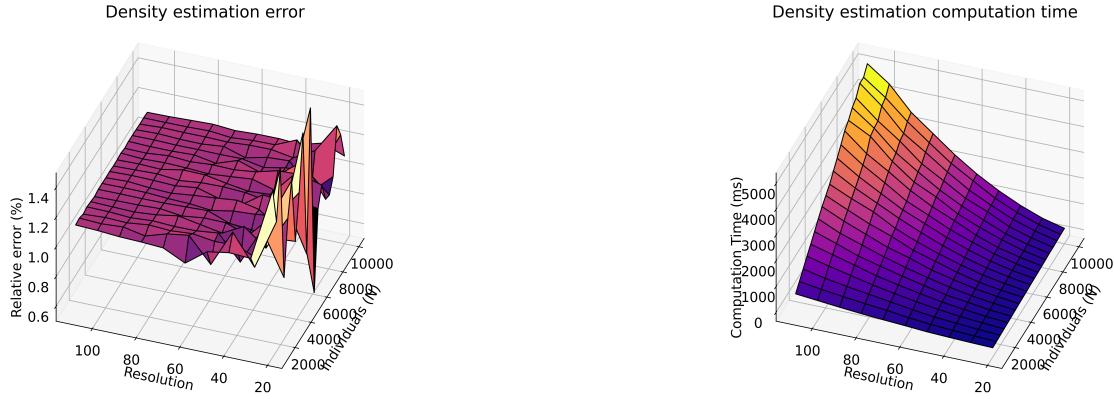


Figure 25: The relative error of the density estimation as a function of the number of plants contributing to the density and the resolution of the density field.

Figure 26: The computational time of the density estimation as a function of the number of plants contributing to the density and the resolution of the density field.

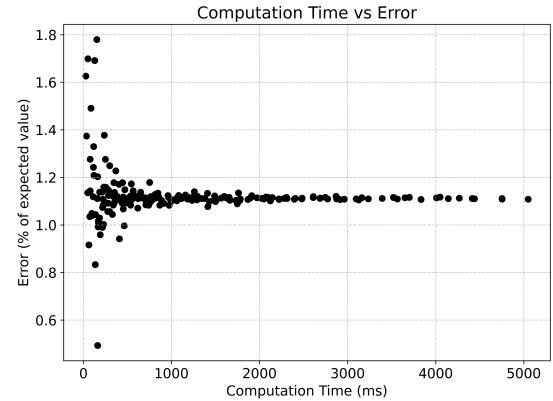


Figure 27: The error of the simulation as a function of computational time.

B Overview of algorithms

Algorithm 1 Simulation step

- 1: Update precipitation value (if specified)
- 2: Grow plants
- 3: Disperse seeds from plants
- 4: Disperse seeds from environment
- 5: Attempt germination of seeds
- 6: Resolve collisions
- 7: Remove dead plants
- 8: Update the density field
- 9: Update the KDTree
- 10: Increase time step

Algorithm 2 Plant growth and mortality

- 1: Input: Plant with growth rate
- 2: Output: Updated plant
- 3: Increase plant radius by growth rate
- 4: **if** plant radius exceeds maximum radius **then**
- 5: Mark plant as dead
- 6: **end if**
- 7: Return updated plant

Algorithm 3 Plant dispersal

- 1: Input: Plant with spawn rate
- 2: Output: List of seed positions
- 3: Determine number of seeds:
- 4: **if** random(0,1) is greater than decimal part of spawn rate **then**
- 5: Number of seeds = floor of spawn rate
- 6: **else**
- 7: Number of seeds = floor of spawn rate + 1
- 8: **end if**
- 9: Initialize empty list for seed positions
- 10: **for** each seed to be dispersed **do**
- 11: Generate random position based on plant position and dispersal kernel
- 12: Add position to seed list
- 13: **end for**
- 14: Return list of seed positions

Algorithm 4 Seed germination

- 1: **Input:** Seed position, germination chance
- 2: **Output:** New plant at seed position (if germination succeeds)
- 3: Compute germination probability using soil quality, precipitation, germination chance, and density (local or global depending on the model being used)
- 4: **if** random(0, 1) is less than germination probability **then**
- 5: Create a new plant with minimum possible radius at the seeds position
- 6: Return the new plant
- 7: **else**
- 8: Germination fails, return nothing
- 9: **end if**

Algorithm 5 Collision resolution

- 1: **Input:** Positions and radii of plants, boundary condition
- 2: **Output:** Indices of dead plants
- 3: **if** positions is empty **then**
- 4: **return** empty list of indices
- 5: **end if**
- 6: Initialize empty list to hold indices of dead plant
- 7: Generate additional “ghost” positions and radii according to the periodic boundary condition (See Section 5.3)
- 8: Determine collision losers based on the relative size of the plants
- 9: Find indices of dead plants (that are not “ghosts”)
- 10: Mark collision losers as dead
- 11: Mark plants as dead in the simulation
- 12: **return** indices of dead plants

References

- Happold, Buro (Jan. 2025). *Climate Referendum: How Can Berlin Become Climate Neutral by 2030?* Accessed: 2025-01-27. URL: <https://www.burohappold.com/news/climate-referendum-how-can-berlin-become-climate-neutral-by-2030/>.
- Lenton, Timothy M. et al. (2008). "Tipping elements in the Earth's climate system". In: *Proceedings of the National Academy of Sciences* 105.6, pp. 1786–1793. DOI: 10.1073/pnas.0705414105.
- Scheffer, Marten, Steve Carpenter, et al. (Oct. 2001). "Catastrophic shifts in ecosystems". In: *Nature* 413.6856, pp. 591–596. ISSN: 1476-4687. DOI: 10.1038/35098000. URL: <https://doi.org/10.1038/35098000>.
- Rietkerk, Max, Maarten C. Boerlijst, et al. (2002). "Self-Organization of Vegetation in Arid Ecosystems". In: *The American Naturalist* 160.4. PMID: 18707527, pp. 524–530. DOI: 10.1086/342078.
- Dakos, Vasilis et al. (June 2011). "Slowing Down in Spatially Patterned Ecosystems at the Brink of Collapse." In: *The American Naturalist* 177.6, E153–E166. ISSN: 0003-0147. DOI: 10.1086/659945. (Visited on 04/23/2025).
- Scheffer, Marten, Jordi Bascompte, et al. (2009). "Early-warning signals for critical transitions". In: *Nature* 461.7260, pp. 53–59. ISSN: 1476-4687. DOI: 10.1038/nature08227.
- Wiggins, Stephen (2003). *Introduction to Applied Non-linear Dynamical Systems and Chaos*. Vol. 2. Texts in Applied Mathematics. New York: Springer-Verlag. ISBN: 978-0-387-00177-7. DOI: 10.1007/b97481. (Visited on 04/25/2025).
- Gardiner, Crispin W et al. (1985). *Handbook of stochastic methods*. Vol. 3. Springer Berlin.
- Morales, Irving O. et al. (June 2015). "Behavior of Early Warnings near the Critical Temperature in the Two-Dimensional Ising Model". In: *PLOS ONE* 10.6, e0130751. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0130751. (Visited on 03/10/2025).
- Aurore, Voltaire (July 2020). *Climate Models*. URL: <https://www.encyclopedie-environnement.org/en/climate/climate-models-2/> (visited on 04/21/2025).
- Gerhard, Krinner (Nov. 2021). *On the Contributions to Climate Physics of Klaus Hasselmann and Syukuro Manabe, Nobel Prize 2021*. URL: <https://www.encyclopedie-environnement.org/en/climate/klaus-hasselmann-syukuro-manabe-nobel-prize-2021/> (visited on 04/21/2025).
- Fisher, Rosie A. et al. (2018). "Vegetation Demographics in Earth System Models: A Review of Progress and Priorities". In: *Global Change Biology* 24.1, pp. 35–54. ISSN: 1365-2486. DOI: 10.1111/gcb.13910. (Visited on 05/16/2025).
- Watson, Andrew J. and James E. Lovelock (1983). "Biological homeostasis of the global environment: the parable of Daisyworld". In: *Tellus B* 35B.4, pp. 284–289. DOI: <https://doi.org/10.1111/j.1600-0889.1983.tb00031.x>.
- Budyko, M. I. (1969). "The effect of solar radiation variations on the climate of the Earth". In: *Tellus* 21.5, pp. 611–619. DOI: <https://doi.org/10.1111/j.2153-3490.1969.tb00466.x>.
- Stommel, Henry (1961). "Thermohaline Convection with Two Stable Regimes of Flow". In: *Tellus* 13.2, pp. 224–230. DOI: <https://doi.org/10.1111/j.2153-3490.1961.tb00079.x>.
- Ising, Ernst (1925). "Beitrag zur Theorie des Ferromagnetismus". ger. In: *Zeitschrift für Physik* 31.1, pp. 253–258. ISSN: 0044-3328.
- Gates, D. J. and M. Westcott (1978). "Zone of influence models for competition in plantations". In: *Advances in Applied Probability* 10.3, pp. 499–537. DOI: 10.2307/1426632.
- Ford, E. D. and P. J. Diggle (Oct. 1981). "Competition for Light in a Plant Monoculture modeled as a Spatial Stochastic Process". In: *Annals of Botany* 48.4, pp. 481–500. ISSN: 0305-7364. DOI: 10.1093/oxfordjournals.aob.a086152. (Visited on 04/24/2025).
- Weiner, J. et al. (Oct. 2001). "The Effects of Density, Spatial Pattern, and Competitive Symmetry on Size Variation in Simulated Plant Populations". In: *The American Naturalist* 158.4, pp. 438–450. ISSN: 0003-0147, 1537-5323. DOI: 10.1086/321988. (Visited on 04/24/2025).
- Berger, Uta and Hanno Hildenbrandt (2000). "A New Approach to Spatially Explicit Modelling of Forest Dynamics: Spacing, Ageing and Neighbourhood Competition of Mangrove Trees". In: *Ecological Modelling* 132.3, pp. 287–302. ISSN: 0304-3800. DOI: 10.1016/S0304-3800(00)00298-2.
- DeAngelis, Donald L. and Volker Grimm (June 2014). "Individual-Based Models in Ecology after Four Decades". In: *F1000Prime Reports* 6, p. 39. ISSN: 2051-7599. DOI: 10.12703/P6-39. (Visited on 04/24/2025).
- Czaran, T. (1998). *Spatiotemporal Models of Population and Community Dynamics*. Population and Community Biology Series. Springer. ISBN: 978-0-412-57550-1.

- Thierry, J. M., J. M. D'Herbès, and C. Valentin (1995). "A Model Simulating the Genesis of Banded Vegetation Patterns in Niger". In: *Journal of Ecology* 83.3, pp. 497–507. ISSN: 0022-0477. DOI: 10.2307/2261602. JSTOR: 2261602. (Visited on 04/24/2025).
- Rietkerk, Max, Stefan C. Dekker, et al. (Oct. 2004). "(PDF) Self-Organized Patchiness and Catastrophic Shifts in Ecosystems". In: ResearchGate. DOI: 10.1126/science.1101867. (Visited on 11/19/2024).
- Scanlon, Todd M. et al. (Sept. 2007). "Positive Feedbacks Promote Power-Law Clustering of Kalahari Vegetation". In: *Nature* 449.7159, pp. 209–212. ISSN: 1476-4687. DOI: 10.1038/nature06060. (Visited on 04/24/2025).
- Vincenot, Christian E. et al. (May 2016). "Spatial Self-Organization of Vegetation Subject to Climatic Stress—Insights from a System Dynamics—Individual-Based Hybrid Model". In: *Frontiers in Plant Science* 7. ISSN: 1664-462X. DOI: 10.3389/fpls.2016.00636. (Visited on 09/11/2024).
- Pałubicki, Wojtek et al. (July 2022). "Ecoclimates: Climate-Response Modeling of Vegetation". In: *ACM Trans. Graph.* 41.4, 155:1–155:19. ISSN: 0730-0301. DOI: 10.1145/3528223.3530146. (Visited on 09/12/2024).
- Howe, H. F. and J. Smallwood (Nov. 1982). "Ecology of Seed Dispersal". In: *Annual Review of Ecology, Evolution, and Systematics* 13. Volume 13, pp. 201–228. ISSN: 1543-592X, 1545-2069. DOI: 10.1146/annurev.es.13.110182.001221. (Visited on 12/04/2024).
- Damgaard, Christian (2025). *Evolutionary Ecology of Plant-Plant Interactions*. Aarhus University Press. (Visited on 04/24/2025).
- von Hardenberg, J. et al. (Oct. 2001). "Diversity of Vegetation Patterns and Desertification". In: *Physical Review Letters* 87.19, p. 198101. DOI: 10.1103/PhysRevLett.87.198101. (Visited on 10/10/2024).
- Rietkerk, Max, Robbin Bastiaansen, et al. (Oct. 2021). "Evasion of Tipping in Complex Systems through Spatial Pattern Formation". In: *Science* 374.6564, eabj0359. DOI: 10.1126/science.abj0359. (Visited on 10/10/2024).
- HilleRisLambers, R. et al. (2001). "Vegetation Pattern Formation in Semi-Arid Grazing Systems". In: *Ecology* 82, pp. 50–61. ISSN: 0012-9658. DOI: 10.1890/0012-9658(2001)082[0050:vpfisa]2.0.co;2. (Visited on 04/24/2025).
- Adimurthy, Adesh Nalpet (Mar. 2022). *Hybrid Spatial Data Structure based on KD-tree and Quadtree*. Accessed: 2025-05-14. URL: <https://blog.devgenius.io/hybrid-spatial-data-structure-based-on-kd-tree-and-quadtrees-8c0c5eebdbbf>.
- Gingold, R. A. and J. J. Monaghan (1977). "Smoothed particle hydrodynamics: theory and application to non-spherical stars." In: *Monthly Notices of the Royal Astronomical Society* 181, pp. 375–389. DOI: 10.1093/mnras/181.3.375.
- Kéfi, Sonia et al. (Mar. 2014). "Early Warning Signals of Ecological Transitions: Methods for Spatial Patterns". In: *PLOS ONE* 9.3, e92097. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0092097. (Visited on 05/01/2025).
- Zagli, N, V Lucarini, and G A Pavliotis (July 2024). "Response Theory Identifies Reaction Coordinates and Explains Critical Phenomena in Noisy Interacting Systems". In: *Journal of Physics A: Mathematical and Theoretical* 57.32, p. 325004. ISSN: 1751-8121. DOI: 10.1088/1751-8121/ad6068. (Visited on 05/16/2025).
- Lohmann, Johannes et al. (Apr. 2025). "The Role of Edge States for Early Warning of Tipping Points". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 481.2312, p. 20240753. DOI: 10.1098/rspa.2024.0753. (Visited on 05/17/2025).