NORTHWESTERN UNIVERSITY

Investigations of Pattern Formation in Dryland Vegetation

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Engineering Sciences and Applied Mathematics

By

Karna Gowda

EVANSTON, ILLINOIS

September 2017

ABSTRACT

Investigations of Pattern Formation in Dryland Vegetation

Karna Gowda

Vegetation in dryland environments is often patchy in response to water limitation. This patchiness can take the form of periodic patterning at length scales much larger than that of an individual plant. Instances of patterns resembling leopard spots and tiger stripes are widespread in dry regions of Africa, Australia, and North America. Mathematical modeling efforts over the past two decades have sought to account for dryland vegetation patterning via a self-organizing interaction between vegetation and water resources. These efforts have led to predictions for the response of the vegetation to environmental variation, and have generated speculation that the morphology of the patterns encodes information about the fragility of the environment.

For a class of patterns that occurs on flat terrain, we use a bifurcation-theoretic framework to analyze the sequence of morphologically distinct patterns that are predicted by models to occur in a scenario of increasing aridity. Applying insights from

the analysis of an idealized problem, we find a strong indication that a specific sequence of patterns occurs in the ecologically-relevant parameter space of a widely-studied dryland vegetation model. This finding is a step towards establishing credibility for a specific pattern sequence as a signal of dryland ecosystem fragility.

For a class of patterns that occurs on gradually sloping terrain, we use aerial imagery over the Horn of Africa to assess the nature of change in the vegetation over six decades. We observe that the most substantial vegetation change is associated with human impact rather than climatic pressure, and that a signature of this impact appears in a readily measurable property of the pattern. These findings illuminate time scales of patterned vegetation dynamics, and suggest a focus on human impacts for future observations and theoretical explorations.

Acknowledgements

My time at Northwestern was made rich by my mentors, colleagues, friends, and family. It is these people I credit for who I am today (or at least the better parts of me).

Above all, I am indebted to my advisor Mary Silber for taking a chance on me. I arrived to Northwestern as a master's student with no research experience to speak of, but perhaps some enthusiasm. Our work together began with a project in her special topics course. This thesis is essentially still that project, continued farther than I could have envisioned. Mary instilled in me her deep sense of skepticism, and the integrity and rigor that follows. She is at once patient and demanding; I am grateful that she pushed me to be better, and always with a sincerity and warmth that made me believe myself capable. Mary's humor and openness to new opportunity conveys an outlook on life that I strive to emulate. I am lucky to learn from her.

I am also grateful to others who have helped me find my way as a researcher. To Sarah Iams, for the acuity and vitality she brings to everything she does. She inspires me to be smarter, and to try harder, though I cannot help but fall short of her example. To Madhav Mani, whose zeal for science and learning is unparalleled. He has helped me come to a new appreciation and understanding of nature.

The faculty of the ESAM department at Northwestern have put together a useful curriculum that spans many aspects of applied mathematics. Thanks to Hermann Riecke, a collaborator on a project in this thesis and a member of my qualifying exam

committee, who was an entertaining instructor and an energetic colleague. To Bill Kath, another member of my qualifying exam committee, whose broad expertise, from Green's functions to high-throughput sequencing algorithms, was inspiring. To Danny Abrams, for serving on my thesis committee, and for sharing his enthusiasm for soil science and hydrology. To David Chopp, for serving on my thesis committee, and for putting together an outstanding numerical methods curriculum. To Ed Olmstead, for his excellent instruction and sage advice throughout the years.

My friends, both within and outside of the department, supported me and grew with me during my time at Northwestern. Thanks to Avinash Karamchandani for being such a patient roommate for so many years. I am inspired by his continual willingness to sit down and think aloud with others. To Danny Wells, whose enthusiasm for biology was certainly contagious. To Ashley Meulemans, Mark Meulemans, and Matthew Wilkinson, for sticking with me all these years. I see no end to our friendship.

Time has brought my family closer together, and has made it larger. Thanks to my parents, Mangala and Mohan Gowda, for bringing me this far, and for accepting me for who I am. Thanks to my brother, Yadav Gowda, whose friendship has brought me contentment beyond almost anything else in recent years. Thanks also to Cheryl, Bob, and Karen Clifton, for welcoming me so warmly into your family.

Thanks finally to Sara Clifton, who understands me better than I understand myself. She astounds me with the depth of her empathy and compassion. I am lucky to be walking through life by her side.

Table of Contents

List of Tables

List of Figures

CHAPTER 1

Introduction

Dryland environments make up over 40% of Earth's terrestrial area and are home to more than a third of the global human population [1] (Figure 1.1). Drylands are commonly classified by the degree of imbalance between the influx of rainfall and the potential[1] outflux of water due to evaporation and plant transpiration. The vegetation phenomena discussed in this dissertation primarily occur in arid and semiarid drylands, where potential water losses due to evaporation and transpiration can exceed rainfall by as much as factor of 20. Such an imbalance means that scarce water dictates the pace and extent of life in these environments.

But not all dryland is wasteland. Though rainfall is meager, many drylands are productive within their means. Hydrological factors often permit spatially-patchy plant growth. Dryland rainfall is typically seasonal and concentrated in a small number of large events [3]. Such events can produce surface runoff that redistributes water from high ground to low ground, resulting in favorable growth conditions in topographic channels and local depressions. In cases where existing plants enhance the water infiltration properties of the soil, runoff can redistribute water from bare areas to patches of vegetation [4]. Tightly clustered vegetation can also reduce local soil water evaporation

---

[1]"Potential" refers to the amount of water that would be lost due to the processes of evaporation and plant transpiration if water were in excess. Potential evaporation and transpiration are typically estimated using surface and air temperature, relative humidity, incoming solar radiation, and wind information via the Penman-Monteith [2] equation.
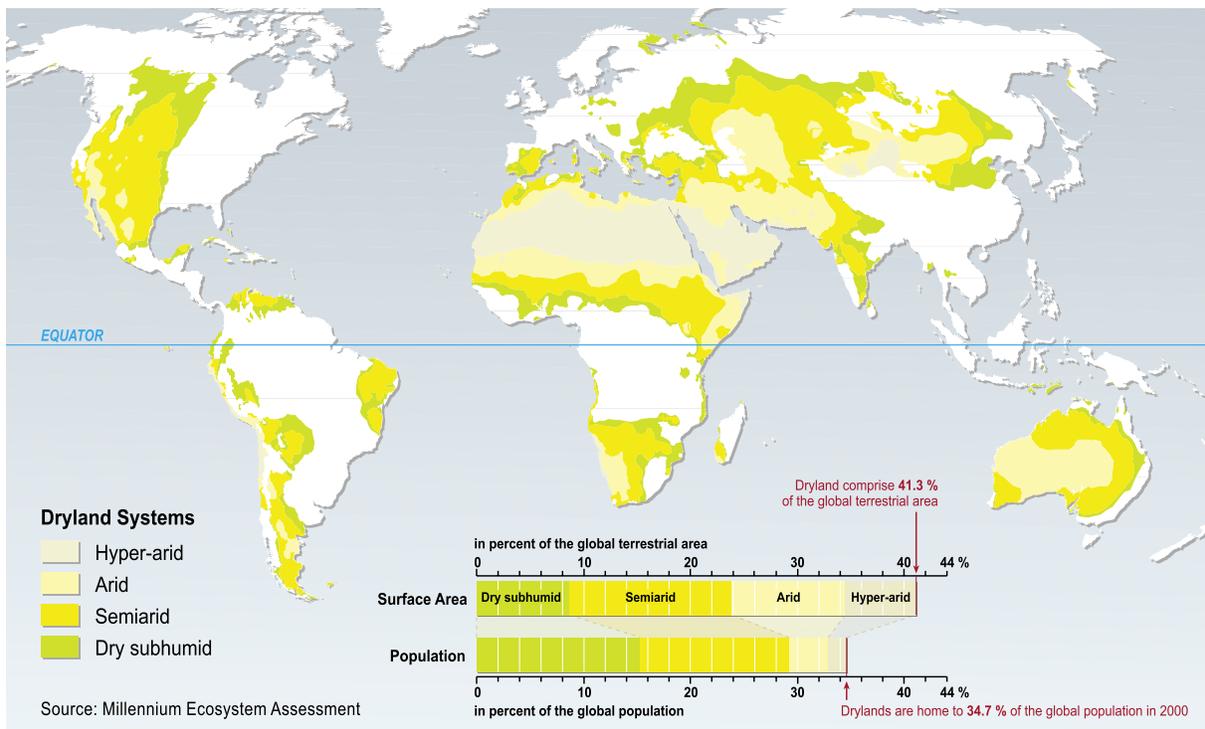
Figure 1.1. Global extent of drylands. Taken from [1].

through shading, and fertilize the surrounding area with leaf litter, thus facilitating the growth of additional vegetation [5].

In particularly striking cases, dryland vegetation patchiness occurs as periodic patterning at length scales much larger than that of an individual plant. This patterning most frequently manifests as bands of vegetation separated by stretches of bare ground aligned along gradual slopes, but can also occur as patches that resemble leopard spots when viewed from the air. Such vegetation patterning is widespread in Africa, Australia, and North America, and isolated examples have also been reported in South America and Asia [6] (Figure 1.2). These patterns typically comprise some combination of trees, shrubs, and grasses, with the particular species and composition differing between regions [6].
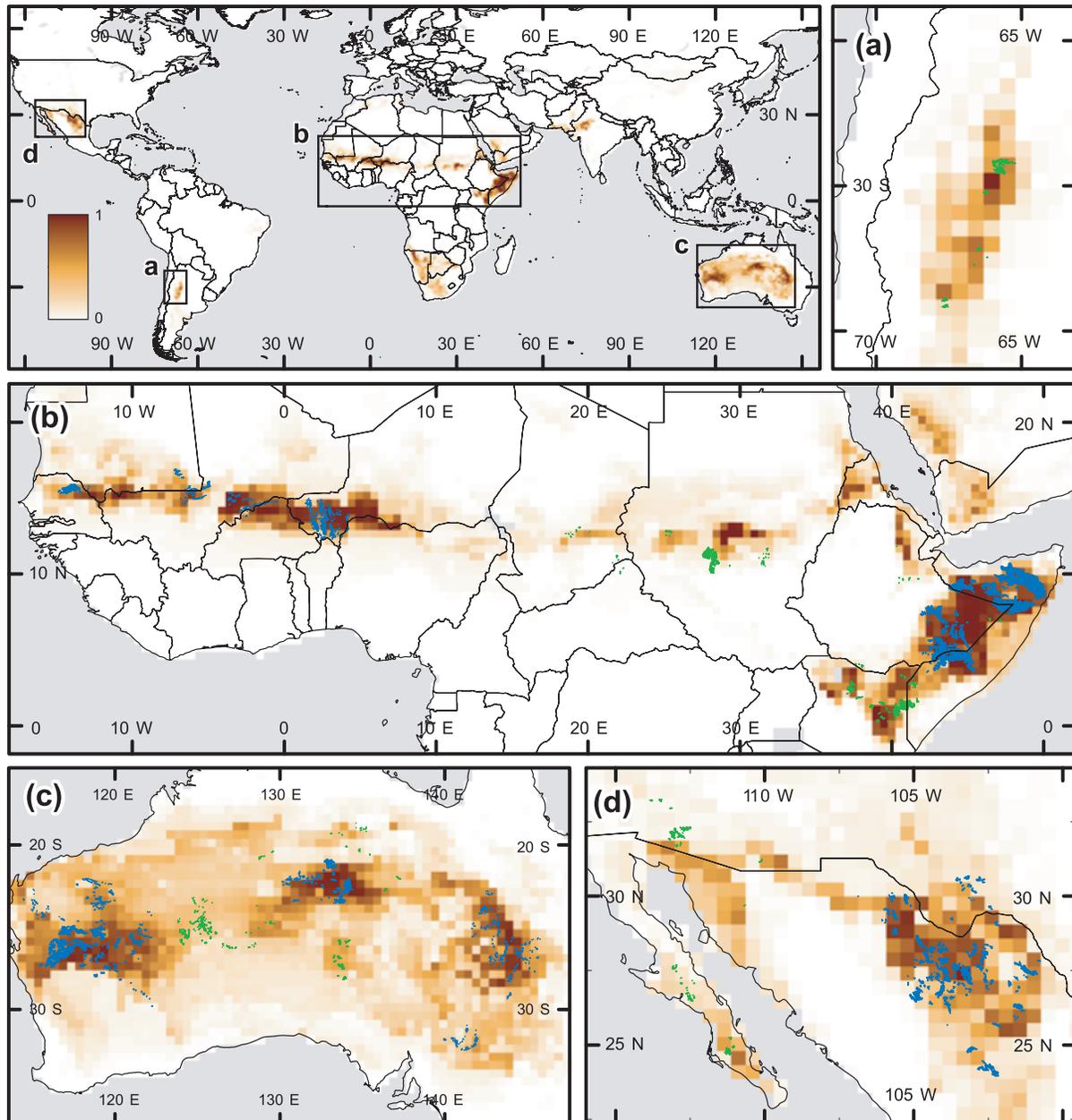
Figure 1.2. Biogeography of dryland vegetation patterns. Confirmed patterned areas are indicated by blue and green shading. Taken from [6].

The remarkable appearance of these patterns led their discoverer in the Horn of Africa to speculate in 1950 that "the causes . . . must be investigated by physics and

mathematics" [7]. Indeed much work in the last two decades has gone into mathematizing the water redistribution and plant facilitation mechanisms into idealized models that are able to qualitatively reproduce the patterns (see Ref. [8] for a review). One family of models captures the spatial and temporal dynamics of coarsely-defined water and vegetation fields through reaction-advection-diffusion partial differential equations (PDEs). A schematic representation of an early such model by Klausmeier [9] is as follows:

$$
\begin{aligned}
\frac{\partial \text{ soil water}}{\partial t} &= \text{rainfall} - \text{evapotranspiration} + \text{transport}, \\
\frac{\partial \text{ plant biomass}}{\partial t} &= \text{growth} - \text{mortality} + \text{dispersal}.
\end{aligned}
\tag{1.1}
$$

In these equations, the rates of evapotranspiration and plant growth are nonlinear functions of both soil water and plant biomass variables, while water transport and dispersal are idealized using first and second spatial derivatives in the form of advection and diffusion terms. The form and behavior of this model is similar to models for pattern formation in chemical reactions, morphogenesis, and patterns in animal fur [10, 11].

Drylands and the human populations they support are imperiled by global climate change. Projected temperature increases of over 3°C in many dryland environments will drive up rates of evaporation, reducing the soil moisture available for plant growth [12]. Climate models also project that drylands will expand to cover nearly 60% of Earth's terrestrial area by the end of the century under probable emissions scenarios [13]. A majority of these changes will play out in developing countries [13], whose

environments are already under strain due to warming and increasing drought severity coupled with population growth and land-use shifts [1]. Dryland populations are often strongly dependent on their local environment for sustenance, which can lead to destructive feedback between climate change and land-use shifts. In an area we study in this dissertation, the Sool Plateau of Somalia, it is documented that a severe drought spanning 2000–2004 killed off nearly all livestock within the region, forcing the region's inhabitants to transition from pastoral farming to woodcutting for charcoal production [14]. This transition likely played a large role in the degradation we observed in the region.

Given this bleak picture of increasing aridity and environmental strain in global drylands, many have asked where theoretical work might have some impact. Models of vegetation pattern formation offer predictions about the response of vegetation to environmental variation [8], suggest strategies for optimal reforestation [15], and have contributed to the discussion about tipping points in global climate [16, 17]. In ecology more broadly, there has been much discussion in recent years about whether catastrophic transitions (e.g., bifurcations that lead to undesirable configurations of ecological phase space) can be anticipated using limited data [18, 19]. These ideas have been applied to infer proximity to desertification from vegetation patch size distribution scaling behavior, for the case where the dryland vegetation patches lack a distinct scale [20]. For periodic dryland vegetation, the PDE models have suggested that morphologically distinct states of the vegetation [16] (Figure 1.3) or changes in length scale of the patterning [21] might signal desertification. These suggestions have generated
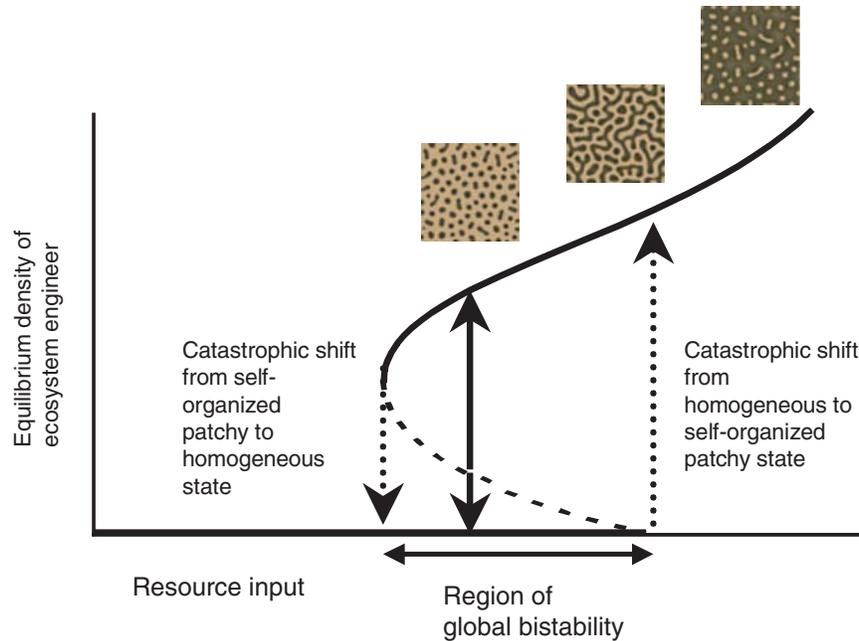
Figure 1.3. Early-warning sign of desertification in the sequence of spatial patterns that occur as an environment becomes more arid. Taken from [16].

subsequent model analysis and empirical work, which includes the investigations presented in this dissertation.

———————————

In this dissertation, I present two self-contained investigations of pattern formation in dryland vegetation. In a general sense, the goal of these investigations to better understand the response of self-organized dryland vegetation to climatic change and human impact.

Chapter 2 presents a theoretical investigation of vegetation patterns on flat terrain, and is supported by Appendices A and B. The content of this chapter and its appendices is culled from material published in Refs. [22] and [23]. The results in Section 2.3

come from joint work with Hermann Riecke and my advisor Mary Silber. The results in Section 2.4 and Appendices A and B come from joint work with Yuxin Chen, Sarah Iams, and Mary Silber.

Chapter 3 presents a primarily empirical investigation of vegetation patterns on sloped terrain via analysis of aerial imagery spanning six decades, and is supported by Appendix C. The content of this chapter and its appendix is drawn from Ref. [24], which has been submitted for review. This project was conducted jointly with Sarah Iams and Mary Silber.

CHAPTER 2

A theoretical investigation of flat-terrain patterns in a scenario of increasing aridity

Many studies of spatially periodic patterns in models of dryland vegetation focus on patterns as potential indicators of ecosystem fragility [25–36]. In particular, patterned states that occur in a scenario of flat terrain (i.e., isotropic topography) in several models have been shown in simulations to evolve through a sequence of morphologies, "gaps $\rightarrow$ labyrinth $\rightarrow$ spots" (Figure 2.1), as ecosystem aridity increases [25, 28, 37, 38]. This sequence precedes the collapse of vegetation in the models, which has led to the suggestion that real ecosystems may evolve through this predictable sequence of patterns en route to desertification [16, 18]. In this way, vegetation patterns may serve as early-warning signs of catastrophic ecosystem transitions.

The "gaps $\rightarrow$ labyrinth $\rightarrow$ spots" pattern sequence prediction emerges from a modeling framework comprising a number of different ecological hypotheses, functional formulations, and restrictions on plausible parameter sets. It is therefore important to investigate whether the prediction is robust within this framework. In this chapter, we ask whether easily calculable quantities, derived from bifurcation theoretic analysis of pattern-forming instabilities, can be used to predict where the sequence occurs in a model's parameter space. In Section 2.1, we give background information about flat-terrain patterns and model predictions for how they vary with aridity. In Section 2.2, we introduce a framework in which pattern transitions between two Turing
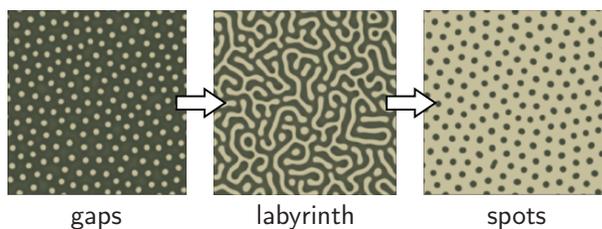
Figure 2.1. Example of the "gaps → labyrinth → spots" sequence in numerical simulations of the vegetation model by Rietkerk *et al*. (2.13). Qualitatively different patterns occur at successively smaller values of a precipitation parameter. Darker shading denotes higher levels of vegetation biomass.

bifurcations can be analyzed at small amplitude. We compute the coefficients of equations which describe the amplitudes of Fourier modes on a 2D hexagonal lattice near a pattern-forming instability. In Section 2.3, for limits of the model parameter values where transitions between patterns all occur in a regime of weak nonlinearity, we show that the coefficient of the quadratic-order term in the amplitude equations affects the sequence that occurs. In Section 2.4, we show that this quadratic coefficient appears to serve as a proxy for the sequence of nonlinear patterns that will manifest for any parameter set of a specific model, even when the system is far from degeneracy. Additionally, in the ecologically-relevant limit of the model parameters, we find that quadratic coefficient takes values consistent with the "gaps → labyrinth → spots" sequence, providing strong evidence that the sequence is a model prediction that is robust to parameter variation.

2.1. Flat-terrain patterns

Periodic vegetation patterns occurring on approximately flat terrain have been observed in African and Australian drylands [8]. These patterns can be roughly classified
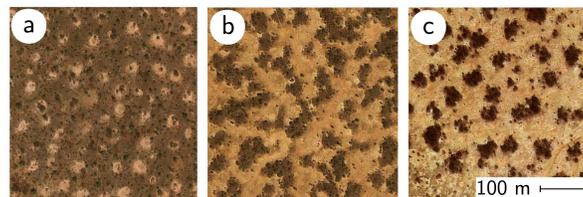
Figure 2.2. Aerial images of flat terrain vegetation patterns in Sudan. (a) Gap (10.7549, 28.5955), (b) labyrinth (11.1024, 27.8228), and (c) spot patterns (11.6280, 27.9177). Images © Google, DigitalGlobe.

on the basis of their morphology as either gaps, labyrinths, or spots (see Figure 2.2 for examples from Sudan). Gapped patterns appear as mostly uniform cover with a pattern of bare patches occurring at a distinct length scale. Labyrinth patterns appear as elongated but disordered patches of vegetation. Spotted patterns resemble an inversion of gapped patterning, with vegetation concentrated in small patches separated by bare areas.

These flat-terrain patterns are thought to arise due to self-organization of vegetation through water redistribution mediated by plant roots [8, 39]. Roots may increase the permeability of the surrounding soils, resulting in accumulation of pooled surface water in vegetated areas following large rainfall events [28, 37]. Roots can also induce a flow of soil water towards vegetated areas due to suction pressure [25]. Additionally, plants can facilitate the growth of other plants via shading, which reduces the local soil evaporation rate [28, 40]. These mechanisms have been incorporated into reaction-diffusion PDE models for water-vegetation interactions, and where they can generate patterns via a Turing-type instability.

Numerical simulations of PDE dryland vegetation models show gap, labyrinth, and spot patterns occurring at successively lower values of a parameter representing mean

annual precipitation [25, 28, 37] (Figure 2.1). Additionally, a study by LeJeune *et al.* [38] used bifurcation analysis to demonstrate analytically that this sequence occurs in a tractable 1-field model for a particular parameter set. The apparent agreement between these observations, which come from different model formulations, provides some support for the standard sequence as a robust prediction of this suite of models.

Empirical support for the "gaps $\rightarrow$ labyrinth $\rightarrow$ spots" sequence comes chiefly from two studies of remotely-sensed imagery. A 2006 study by Barbier *et al.* [41] used imagery over southwest Niger to demonstrate that gap patterns emerged from uniform vegetation cover during a period of prolonged drought. This result is consistent with model observations of gap patterns occurring near the onset of pattern formation. A 2011 study by Deblauwe *et al.* [42] classified pattern morphologies in imagery over Sudan (see Figure 2.2 for examples of such patterns) and found that different morphologies vary over spatial precipitation gradients in accordance with the standard sequence prediction. Gaps tended to occur in areas with relatively high mean annual precipitation, spots occurred in areas with relatively low precipitation, and labyrinths occurred in between. Pattern dynamics were also assessed using three sets of images taken over a 35-year span. Gaps in some areas were shown to transition to labyrinths over a period of time again coinciding with a sustained regional drought. Labyrinths transitioned to spots in different areas over the same period of time. Though neither of these studies show the standard sequence preceding the collapse of vegetation, they demonstrate consistency between some model predictions and empirical observations.

2.2. Bifurcation-theoretic framework for studying pattern sequences on flat-terrain

2.2.1. Motivating Example

To motivate the framework for our subsequent analysis, we briefly review the bifurcation structure of the PDE vegetation model by von Hardenberg *et al.* [25] (VH01). In nondimensional form, the model is:

$$
\frac{\partial w}{\partial t} = \underbrace{p}_{\text{precip.}} - \underbrace{(1 - \rho n)w}_{\text{evap.}} - \underbrace{w^2 n}_{\text{transpir.}} + \underbrace{\delta \nabla^2 (w - \beta n)}_{\text{transport}},
$$

(2.1)

$$
\frac{\partial n}{\partial t} = \underbrace{\frac{\gamma w}{1 + \sigma w} n}_{\text{growth}} - \underbrace{\nu n}_{\text{mort.}} - \underbrace{n^2}_{\text{satur.}} + \underbrace{\nabla^2 n}_{\text{dispersal}} .
$$

The variable n represents vegetation density and w represents ground water density. The nonlinear functions in (2.1) capture the effects of plant facilitation and competition. Spatial terms model the diffusive spread of vegetation and the transport of water, with cross-diffusion accounting for suction by plant roots in the latter. A parameter representing mean annual precipitation, p, imposes resource scarcity. Additionally, γ, σ, ν, ρ and β are $\mathcal{O}(1)$ positive parameters and δ, which characterizes the diffusivity of water relative to that of vegetation, is usually taken to be a large, positive parameter.

Spatially uniform equilibria of VH01 satisfy

(2.2)
$$0 = p - (1 - \rho n)w - w^2 n \equiv F(n, w),$$

$$0 = \frac{\gamma w}{1 + \sigma w} n - \nu n - n^2 \equiv G(n, w).$$

One such equilibrium, corresponding to a non-vegetated "desert" state, is given by $(n, w) = (0, p)$. The desert state is linearly stable for $p < p_0 = \nu/(\gamma - \nu\sigma)$, and undergoes a transcritical bifurcation at $p = p_0$ to a uniform "vegetated" equilibrium, (n^*, w^*), for which $n^* > 0$. The bifurcation diagram depicting these spatially uniform states and their stability as a function of p is plotted in Figure 2.3. Turing bifurcations on the uniform vegetated equilibrium produce patterned states at the lower and upper Turing points $p = p_\ell, p_u$. One can determine these points by first linearizing (2.1) about the uniform vegetated equilibrium and assuming solutions take the form of Fourier mode perturbations $(n - n^*), (w - w^*) \sim e^{iqx}$ to this equilibrium, for arbitrary perturbing wave number q. This linearization results in the Jacobian matrix

(2.3)
$$J(p, q) \equiv \left. \begin{pmatrix} \frac{\partial F}{\partial n} + \beta\delta q^2 & \frac{\partial F}{\partial w} - \delta q^2 \\ \frac{\partial G}{\partial n} - q^2 & \frac{\partial G}{\partial w} \end{pmatrix} \right|_{(n,w)=(n^*,w^*)},$$

where $F(n, w)$ and $G(n, w)$ are given in (2.2). $\text{Det}(J(p, q)) = 0$ defines a stability boundary for the spatially uniform state in the p-q plane, which in this case forms a closed bubble that is plotted in Figure 2.4. The Turing points p_ℓ and p_u are the left and right endpoints of this stability bubble, and they are each associated with a preferred critical wave number $q_\ell, q_u > 0$. These points are obtained by solving $\text{Det}(J(p, q)) = \partial\text{Det}(J(p, q))/\partial q = 0$.
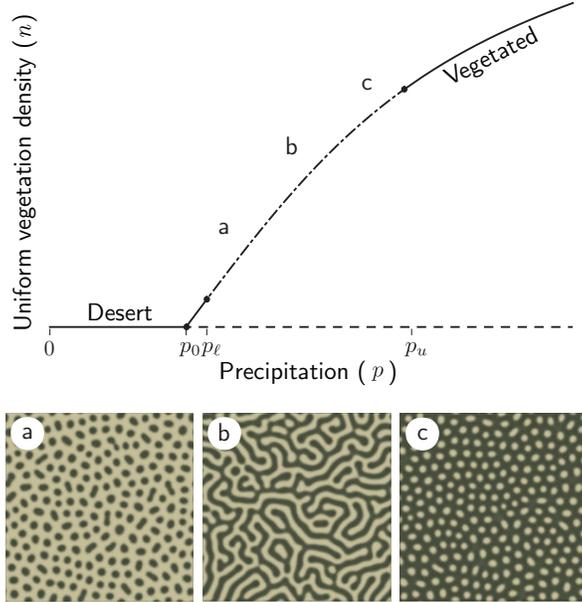
Figure 2.3. Bifurcation diagram depicting the spatially uniform states of VH01 with numerical solutions, using the default parameter set $\gamma = \sigma = 1.6$, $\nu = 0.2$, $\rho = 1.5$, $\beta = 3$ and $\delta = 100$. The desert state loses stability to a vegetated state at $p = p_0 \approx 0.157$. The vegetated equilibrium is unstable to perturbations $n, w \sim e^{iqx}$ in the dashed region $p_\ell < p < p_u$, where $p_\ell \approx 0.169$ and $p_u \approx 0.413$ are the Turing points. Numerical simulations use precipitation values (a) $p = 0.20$, (b) $p = 0.30$, and (c) $p = 0.40$, with higher vegetation density plotted in darker shading.

Numerical simulations of VH01 using the default parameter set yield asymptotic states that follow the standard "gaps $\rightarrow$ labyrinth $\rightarrow$ spots" sequence for decreasing values of p (see Figure 2.3). It is natural to idealize these solutions as regular patterns on a 2D hexagonal lattice [38, 43]. A uniform vegetated state that develops dry gaps resembles a "down-hexagons" pattern (H^-), a labyrinthine intermediate state resembles a distorted stripes pattern (S), and a state of isolated vegetation spots resembles "up-hexagons" (H^+).
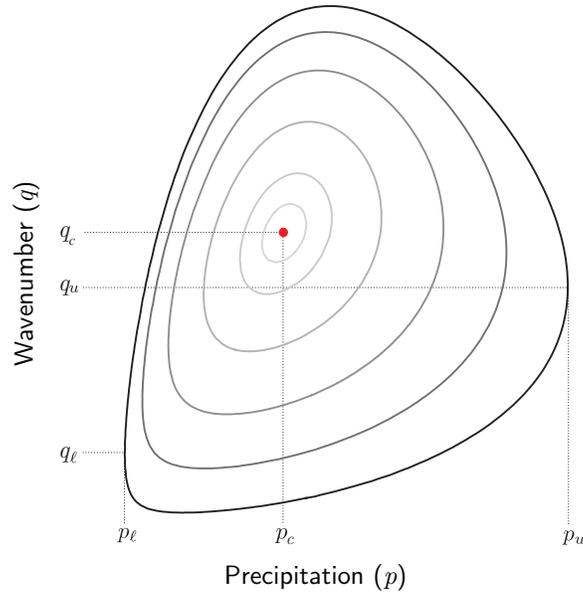
Figure 2.4. Linear stability boundary of uniform vegetated states of VH01 plotted in the p-q plane, where q is the perturbing wave number. The boundary forms a closed bubble whose left and right endpoints are the Turing points p_ℓ and p_u, which have corresponding critical wave numbers q_ℓ and q_u. As the diffusion coefficient δ is decreased, the stability bubble collapses to a degenerate Turing point, (p_c, q_c).

2.2.2. Formulating the bifurcation problem

Near a Turing point (p_{crit}, q_{crit}), the behavior of a solution to a pattern-forming system is characterized by the modes of a Fourier expansion corresponding to wave number q_{crit}. Aspects of pattern formation can be analyzed through the time-varying amplitudes of these critical modes. The form of the amplitude equations for Fourier modes on a 2D hexagonal lattice can be derived through a standard calculation, described for instance in [11, 44, 45]. Here, we summarize key points and results of this derivation, and then formulate the degeneracies that allow us to analyze transitions between patterned states at small amplitude.

We consider the evolution of critical Fourier modes on a hexagonal lattice perturbing the uniform state,

(2.4)
$$z_1 e^{i\mathbf{q_1}\cdot\mathbf{x}} + z_2 e^{i\mathbf{q_2}\cdot\mathbf{x}} + z_3 e^{i\mathbf{q_3}\cdot\mathbf{x}} + c.c. + \cdots .$$

The wave vectors $\mathbf{q_1}$, $\mathbf{q_2}$, $\mathbf{q_3}$ are chosen such that

(2.5)
$$\mathbf{q_1} = q_{crit}(1, 0),$$
$$\mathbf{q_2} = q_{crit}\left(-1/2, \sqrt{3}/2\right),$$
$$\mathbf{q_3} = -(\mathbf{q_1} + \mathbf{q_2}).$$

These vectors lie on the critical circle $|\mathbf{q}| = q_{crit}$ and z_1, z_2, z_3 (as well as their complex conjugates $\bar{z}_1$, $\bar{z}_2$, $\bar{z}_3$) are the complex amplitudes of the corresponding critical modes. Near p_{crit}, all other complex Fourier modes associated with the hexagonal lattice are linearly damped.

The form of the equations describing the evolution of the critical mode amplitudes near a Turing point can be determined using an equivariant bifurcation theory approach [44]. To cubic order, these equations are

(2.6)
$$\frac{dz_1}{dt} = \mu z_1 + a\bar{z}_2\bar{z}_3 - \left(b|z_1|^2 + c(|z_2|^2 + |z_3|^2)\right) z_1,$$
$$\frac{dz_2}{dt} = \mu z_2 + a\bar{z}_1\bar{z}_3 - \left(b|z_2|^2 + c(|z_1|^2 + |z_3|^2)\right) z_2,$$
$$\frac{dz_3}{dt} = \mu z_3 + a\bar{z}_1\bar{z}_2 - \left(b|z_3|^2 + c(|z_1|^2 + |z_2|^2)\right) z_3.$$

The coefficients μ, a, b, and c are real-valued and are determined by the specific terms and parameter values of a system in a neighborhood of a Turing point. These equations

Solution and branching equation	Eigenvalues
Stripes (S) $\mathbf{z} = (x_s, 0, 0)$ $0 = \mu x_s - b x_s^3$	$-2b x_s^2, 0,$ $(b - c)x_s^2 + a x_s \ (\times 2),$ $(b - c)x_s^2 - a x_s \ (\times 2)$
Up-hexagons (H^+) $\mathbf{z} = (x_h, x_h, x_h)$ $0 = \mu x_h + a x_h^2 - (b + 2c)x_h^3$	$-3a x_h, 0 \ (\times 2),$ $a x_h - 2(b + 2c)x_h^2,$ $-2a x_h - 2(b - c)x_h^2 (\times 2)$
Down-hexagons (H^-) $\mathbf{z} = -(x_h, x_h, x_h)$ $0 = \mu x_h - a x_h^2 - (b + 2c)x_h^3$	$3a x_h, 0 \ (\times 2),$ $-a x_h - 2(b + 2c)x_h^2,$ $2a x_h - 2(b - c)x_h^2 \ (\times 2)$
Mixed-modes (MM) $\mathbf{z} = (x_1, x_2, x_2)$ $0 = \mu x_1 + a x_2^2 - (b x_1^2 + 2c x_2^2)x_1$ $0 = \mu x_2 + a x_1 x_2 - (b + c)x_2^3 - c x_1^3$	Always unstable

Table 2.1. Branching equations for solutions on a hexagonal lattice, and eigenvalues of linearizations of (2.6), together with their multiplicities. $\mathbf{z} = (z_1, z_2, z_3)$, and $x_s, x_h, x_1, x_2 > 0$.

describe the branching and relative linear stability of stripes and hexagons solutions to perturbations on the hexagonal lattice. Solutions and their eigenvalues are listed in Table 2.1. From these eigenvalues, one may determine conditions for the stability of a solution. For example, a necessary condition for the stability of stripes is $b > 0$, which ensures that the eigenvalue $-2b x_s^2$ is negative. Eigenvalues are often repeated (indicated by $\times 2$ in Table 2.1) due to the spatial symmetries of a solution, and zero eigenvalues reflect the neutral stability of solutions to translation.

2.3. Pattern transitions in a degenerate scenario

2.3.1. Formulating the Turing degeneracy

We introduce a distinguished control parameter λ. The coefficients in equations (2.6) are generally functions of this parameter, i.e. $\mu = \mu(\lambda)$, etc. We constrain these parameters to vary with λ in the following way. First, we force the bifurcation problem to occur in the vicinity of a degenerate Turing point at $\lambda = 0$ (without loss of generality), which corresponds to μ varying quadratically in λ to leading order (i.e. $\mu(\lambda) = \mu''(0)\lambda^2/2+\cdots$, where $\mu''(0) < 0$). Unfolding the degenerate point results in the expansion

$$(2.7) \qquad \qquad \mu(\lambda) = \mu_0 + \mu_1\lambda + \mu_2\lambda^2 + \cdots ,$$

where μ_0 and μ_1 are small parameters and $\mu_2 < 0$. This unfolding corresponds to a small parameter perturbation to a Turing degeneracy, and two Turing points may emerge as real solutions to $\mu(\lambda) = 0$. An illustration of the Turing instability growth rate, which is proportional to μ, is shown near a Turing degeneracy in Figure 2.5. Second, if the quadratic coefficient $a \neq 0$, it can be shown that all solutions to (2.6) bifurcate unstably [44]. Hence, to capture stable solutions of (2.6), it is standard to consider (2.6) near $a = 0$. We therefore unfold a as

$$(2.8) \qquad \qquad a(\lambda) = a_0 + a_1\lambda + \cdots ,$$
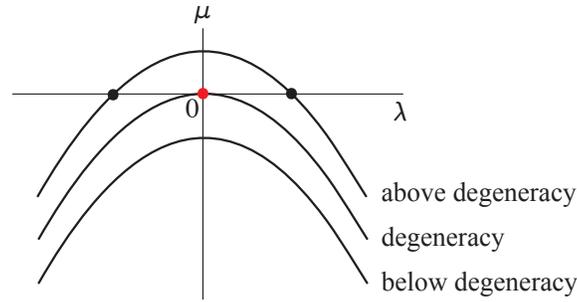
Figure 2.5. Growth rate of the critical Fourier mode perturbing a spatially uniform equilibrium, μ, plotted as a function of the control parameter λ. The points denote Turing bifurcations.

where a_0 is another small parameter. Last, we do not consider the variation of b or c

with λ, which is reasonable if we avoid degeneracies involving those coefficients.

Through a rescaling of time and amplitudes and ignoring higher order terms, we

can write unfoldings of the degenerate linear and quadratic coefficients of (2.6) as

(2.9) $$\mu(\lambda) = \mu_0 + \mu_1\lambda - \lambda^2, \; a(\lambda) = a_0 + \text{sgn}(a_1)\lambda,$$

where μ_0, μ_1 and a_0 are small parameters. The equations (2.6) with coefficients (2.9)

now show transitions between pattern solutions that can occur entirely at small ampli-

tude.

2.3.2. Pattern transitions near the Turing degeneracy

For fixed values of the coefficients b and c, the eigenvalues in Table 2.1 specify regions

in the a-μ plane where solutions to (2.6) are stable, and the coefficients (2.9) describe

paths through these regions that are parameterized by the control parameter λ. A path

through the stability region of a solution corresponds to the existence of a stable solution for an interval of the control parameter. Exiting a region through a stability boundary represents a solution losing stability and points to a transition between patterned states. In the following, we outline our procedure for identifying small-amplitude pattern transition scenarios.

By requiring the non-identically zero eigenvalues in Table 2.1 to be negative, stability regions in the a-μ plane for solutions to (2.6) divide consideration into three distinct cases: (i) $0 < b < c$, (ii) $-c < b < 0$, and (iii) $-b/2 < c < b$. Together, these make up all cases in which stable solutions to (2.6) exist. Notably, case (i) is the only case in which small-amplitude stripes are stable.

Table 2.2 lists stability regions for each case, which are bounded by the curves

$$(2.10) \qquad \mu = -a^2/4(b + 2c),$$

$$(2.11) \qquad \mu = a^2(2b + c)/(b - c)^2,$$

$$(2.12) \qquad \mu = a^2 b/(b - c)^2.$$

(2.10) is derived from the existence condition for hexagons, and (2.11) and (2.12) come from necessary conditions for stability of hexagons and stripes, respectively. These stability regions are plotted in Figure 2.6, and the bounding curves are labeled I-VI and distinguish between regions of stability on either side of the line $a = 0$.

Varying the control parameter λ in the coefficients (2.9) leads to parabolic paths in the a-μ plane. Depending on the sign of a_1 and the small parameters μ_0, μ_1, a_0, these parabolic paths cut through the regions of pattern existence/stability in Table
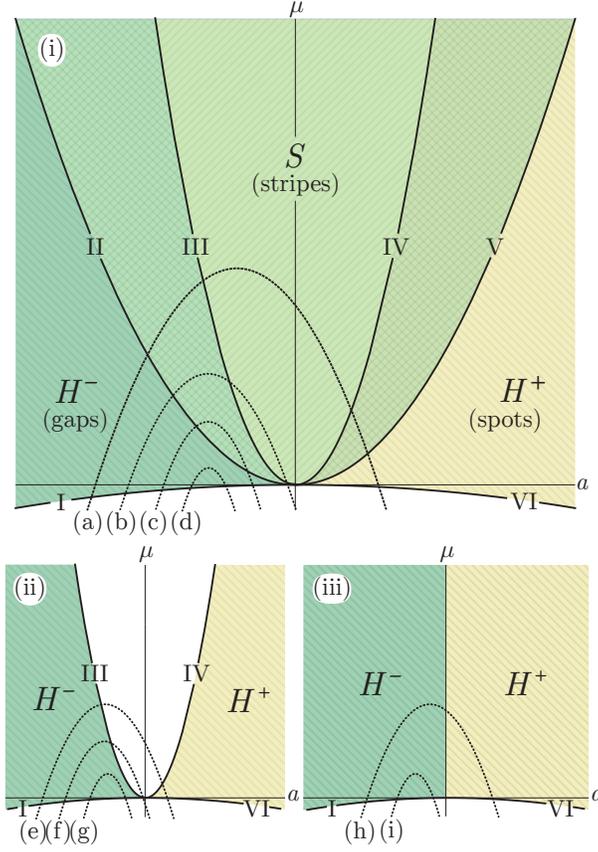
Figure 2.6. Stability regions in the a-μ plane for S and $H^{+/-}$ solutions to (2.6) for cases (i) $0 < b < c$, (ii) $-c < b < 0$ and (iii) $-b/2 < c < b$, with sample paths (a-i) formed by varying λ in (2.9). The boundaries given by (2.10)-(2.12) are labeled I-VI and distinguish between boundaries on either side of $a = 0$. Cross hatching denotes regions of bistability between S and $H^{+/-}$ solutions.

Case	Stripes (S)	Hexagons ($H^{+/-}$)
(i)	$\mu > a^2 b/(b-c)^2$	$-a^2/4(b+2c) < \mu < a^2(2b+c)/(b-c)^2$
(ii)	—	$-a^2/4(b+2c) < \mu < a^2(2b+c)/(b-c)^2$
(iii)	—	$\mu > -a^2/4(b+2c)$

Table 2.2. Regions of stability for stripes and hexagon solutions to (2.6) in the a-μ plane for cases (i) $0 < b < c$, (ii) $-c < b < 0$, and (iii) $-b/2 < c < b$. H^+ is stable when $a > 0$, and H^- is stable when $a < 0$.
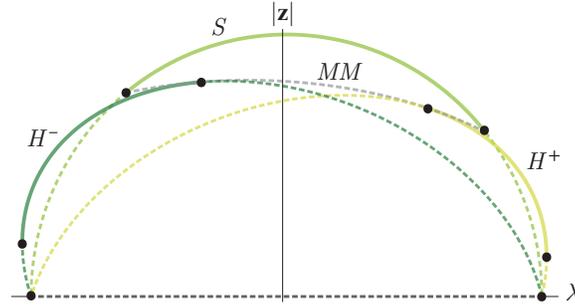
Figure 2.7. Schematic bifurcation diagram corresponding to path (a) in Figure 2.6, taking $\text{sgn}(a_1) = 1$. The Euclidean norm of $\mathbf{z} = (z_1, z_2, z_3)$ is plotted, with stable solutions solid and unstable solutions dashed. Bifurcation points are indicated by dots.

2.2. Some example paths through these regions are sketched in Figure 2.6. We observe that example path (a) crosses through boundaries I-VI. This means varying λ along this path leads (2.6) sequentially through the regions of stability for H^-, S, and H^+ solutions. A bifurcation diagram corresponding to this path is sketched in Figure 2.7, from which we infer the transition sequence "$H^- \to S \to H^+$" (i.e. "gaps $\to$ labyrinth $\to$ spots") occurs as λ varies.

It follows from this example that for each case, each distinct boundary crossing sequence is linked to a pattern transition sequence. Distinct boundary crossings and inferred transitions for all three cases are listed in Table 2.3. "Symmetric" sequences occur where boundaries are crossed on one side of the line $a = 0$ only (e.g. case (i) sequence "I,I"), while "asymmetric" sequences occur where boundaries are crossed on both sides of the μ-axis (e.g. case (i) sequence "I,II,III,IV,V,VI"). Path (a) in Figure 2.6 results in an asymmetric sequence, while paths (b-d) result in symmetric sequences. Only one asymmetric sequence, corresponding to a transition involving both H^+ and H^- solutions, exists for each case.

Case	Ex.	Sequence	Transition scenario
(i)	(a)	I-VI	$H^- \to S \to H^+$,
	(b)	I,II,III,III,II,I	$H^- \to S \to H^-$
	(c)	I,II,II,I	H^- (w/ S bistability)
	(d)	I,I	H^-
(ii)	(e)	I,III,IV,VI	$H^- \to - \to H^+$,
	(f)	I,III,III,I	$H^- \to - \to H^-$
	(g)	I,I	H^-
(iii)	(h)	I,VI	$H^- \to H^+$
	(i)	I,I	H^-

Table 2.3. Distinct boundary crossing sequences and inferred transition scenarios for cases (i) $0 < b < c$, (ii) $-c < b < 0$, and (iii) $-b/2 < c < b$, with example paths from Figure 2.6 indicated. Dashes in case (ii) indicate that no intermediate state is stable at small amplitude. "Symmetric" sequences occurring for $a > 0$ (e.g. VI,VI) are omitted.

One type of scenario not explicitly described above occurs when a parabolic path crosses through a region of hexagon stability without crossing the line $\mu = 0$. Since Turing bifurcations occur when $\mu = 0$, such a scenario can be interpreted as a small-amplitude hexagon state that occurs without Turing bifurcations. These states coexist with spatially uniform states that are stable to Fourier mode perturbations, and thus may only be accessed through a finite amplitude perturbation. Transition scenarios for such patterned states must necessarily involve only one type of hexagon solution.

This analysis shows that a number of transition scenarios between patterned states, beyond the standard "gaps $\to$ labyrinth $\to$ spots," are possible. Many of these scenarios correspond to "symmetric" sequences and therefore only involve one type of hexagons solution (e.g. case (i) sequence "I,II,III,III,II,I"). This demonstrates that it is possible for pattern transitions to occur without the appearance of spots, for instance,

as in paths (b-d) in Figure 2.6. The occurrence of spots as precursors to pattern collapse may therefore be sensitive to the specific parameters of a system.

2.3.3. Pattern transitions near degeneracy in the model by von Hardenberg *et al.*

In this section, we analyze patterned states near a Turing degeneracy in the model (2.1) by von Hardenberg *et al.* [25] (VH01). Note that we only consider the dependence of pattern transition scenarios on the diffusion parameter δ as an illustration, but that one could just as well consider the dependence on other parameters. We obtain the coefficients of the amplitude equations (2.6) following the results of Judd & Silber [46], who derive these coefficients perturbatively for general two-component reaction-diffusion systems with diagonal diffusion. (2.1) is transformed into this problem by diagonalizing the diffusion matrix.

For the default parameter set $\gamma = \sigma = 1.6$, $\nu = 0.2$, $\rho = 1.5$, $\beta = 3$ and $\delta = 100$, Turing instabilities occur along the spatially uniform vegetated equilibrium at $(p_\ell, q_\ell) \approx (0.169, 0.106)$ and $(p_u, q_u) \approx (0.414, 0.206)$. The coefficients b and c of (2.6) are negative at both Turing points, which renders unstable small-amplitude stripes and hexagons. Weakly nonlinear theory is therefore unable to describe stable patterned solutions to VH01 near these points. Numerical simulations, which are shown in Figure 2.3, reveal that the "gaps $\rightarrow$ labyrinth $\rightarrow$ spots" sequence occurs at large amplitude as precipitation decreases.

By modifying the value of δ in the default parameter set, we find that VH01 is amenable to weakly nonlinear analysis in a neighborhood of $\delta = \delta_c \approx 24.7$ (the two
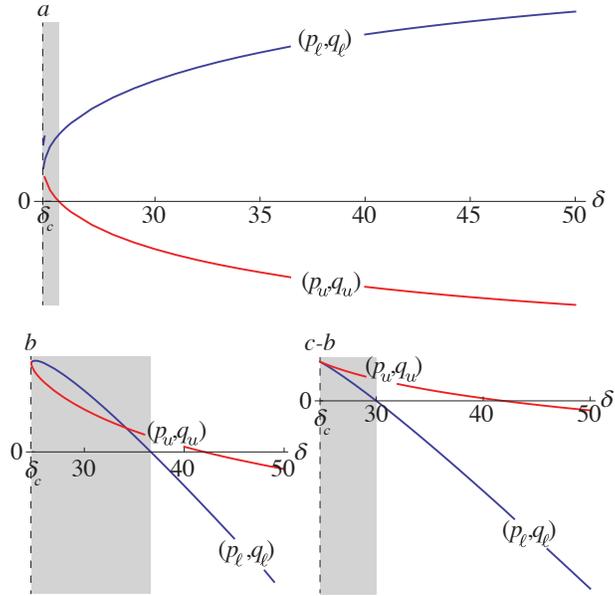
Figure 2.8. The coefficients a, b, and $c - b$, plotted at Turing points (p_ℓ, q_ℓ) and (p_u, q_u) as functions of the parameter δ (default parameter values used otherwise). Grey-shaded intervals indicate where $a > 0$, $b > 0$, and $c - b > 0$ at both Turing points.

Turing points collapse to a single degenerate point at $\delta = \delta_c$). This is observed through the coefficients b and c of (2.6). We saw in case (i) of Section 2.3.2 that both small-amplitude stripes and hexagons can be stable when $0 < b < c$. Plotting b and $c - b$ evaluated at Turing points (p_ℓ, q_ℓ) and (p_u, q_u) as functions of δ in Figure 2.8, we observe that $0 < b < c$ when $\delta < 30.0$. Plotting a in a similar way shows that a changes sign between the Turing points when $\delta > 25.5$, which permits the assymetric sequences described in Section 2.3.2, i.e. sequences that involve both gap and spot patterns. Together, these coefficients indicate that the sequence "$H^- \to S \to H^+$" appears at small amplitude in the interval $\delta \in [25.5, 30.0]$, which parallels what is observed at large amplitude for larger values of δ. A bifurcation diagram depicting this scenario is shown in Figure 2.9.
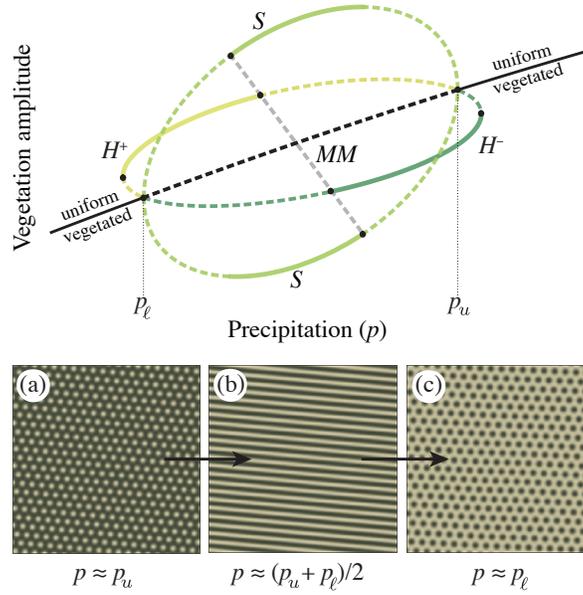
Figure 2.9. Schematic bifurcation diagram for solutions of VH01 with $\delta \in [25.5, 30.0]$ and numerical solutions with $\delta = 27.0$ (default parameter values used otherwise). The numerical simulation (a) is evolved to a steady state resembling an H^- solution from spatially noisy initial conditions at $p = 0.292 \approx p_u$. p is then decreased to $0.259 \approx (p_\ell + p_u)/2$ and the solution is allowed to settle to the S steady state shown in (b). This procedure is repeated once more at $p = 0.225 \approx p_\ell$, and the S solution initial condition settles to an H^+ steady state, shown in (c).

To simulate this sequence numerically, we set $\delta = 27.0$ so that two Turing points emerge at $(p_\ell, q_\ell) \approx (0.225, 0.214)$ and $(p_u, q_u) \approx (0.292, 0.247)$. A numerical solution of VH01 at $p \approx p_u$ is evolved using a pseudospectral RK4 scheme from spatially random initial conditions for n and w drawn uniformly from the interval $[0.2, 0.4]$. After this solution reaches a steady state, the parameter p is decreased in two steps, and the solution is allowed to reach a steady state after each step. These steady states, shown in Figure 2.9, closely resemble H^-, S, and H^+ solutions at $p \approx p_u$, $(p_\ell + p_u)/2$, and $p \approx p_\ell$ respectively, which is consistent with the picture suggested by analysis.
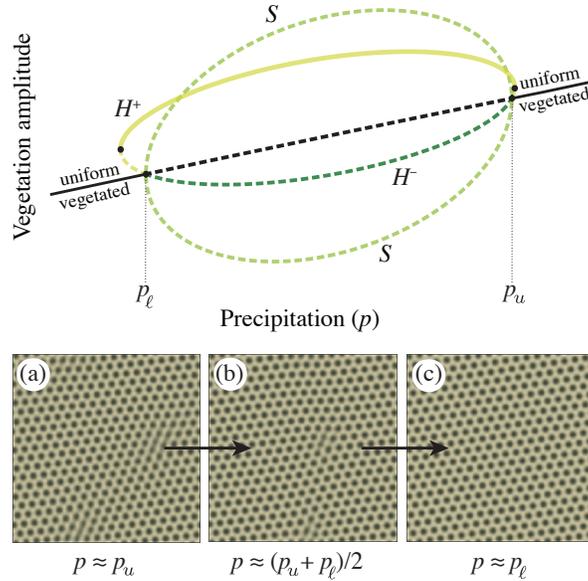
Figure 2.10. Bifurcation diagram for solutions of VH01 and numerical solutions with $\delta = 25.0$ (default parameter values used otherwise). The numerical simulation (a) is evolved to a steady state resembling an H^+ solution from spatially noisy initial conditions at $p = 0.269 \approx p_u$. p is then decreased to $0.257 \approx (p_\ell + p_u)/2$ and the H^+ state remains stable, as shown in (b). This procedure is repeated once more at $p = 0.245 \approx p_\ell$, and the H^+ state continues to remain stable, as shown in (c).

Figure 2.8 also shows that $a > 0$ at both Turing points when $\delta \in [24.7, 25.5]$, resulting in "symmetric" pattern transition sequences that exclude stable H^- (gaps) patterns in this interval. In Figure 2.10, we plot a bifurcation diagram of solutions to VH01 as functions of p with $\delta = 25.0$, which shows that only H^+ (spot) solutions are stable at small amplitude between the two Turing points. For this value of δ, the two Turing bifurcations occur at $(p_\ell, q_\ell) \approx (0.245, 0.231)$ at $(p_u, q_u) \approx (0.269, 0.243)$. Numerical simulations at $p = p_\ell, (p_\ell + p_u)/2$, and p_u all yield small amplitude up-hexagon solutions (see Figure 2.10). We remark that although this analysis is evidence of an alternative

to the standard sequence in the model by von Hardenberg *et al.*, this alternative occurs only in a small interval of δ.

2.3.4. Summary and discussion

In order to analyze aspects of transition between patterned states in PDE vegetation models, we have formulated a bifurcation problem on a 2D hexagonal lattice. Amplitude equations capture dynamics that are dominated by the evolution of critical Fourier mode perturbations to a uniform equilibrium. We enforced degeneracies of the amplitude equation coefficients that enable transitions between patterned states to occur at small amplitude, where they can be investigated analytically. We found that a number of scenarios beyond the standard "gaps $\rightarrow$ labyrinth $\rightarrow$ spots" transition sequence are possible in this generic setting, and that the appearance of these scenarios is distinguished by the coefficients of the amplitude equations near the degenerate point. Since the bifurcation problem exploits symmetries present in a number of vegetation pattern models near Turing bifurcation points, one can place these models within the framework we have described. We did this for the model by von Hardenberg *et al.* [25] and observed the standard sequence as well as a "spots-only" scenario for two near-degenerate parameter sets.

Based on our analysis and observations of patterned states in VH01, we speculate that specific assumptions make the standard sequence relatively robust in some PDE models for vegetation pattern formation. This is despite our observation of the "spots-only" scenario in VH01, which we found to occur only for a small interval of the parameter we vary. In Section 2.3.2, we found that an analog of the standard sequence

occurs at small amplitude due in part to the quadratic coefficient of (2.6) changing sign between the Turing bifurcation points. This weakly nonlinear assessment forms an organizing center for more strongly nonlinear model behavior away from a degeneracy, and we conjecture that the quadratic coefficient changing sign between Turing points serves as a potential signpost for the standard sequence. Our expectation is that if a small-amplitude analog of this sequence is observed near a degeneracy, then the quadratic coefficient will often give a qualitative description of behavior even when no small-amplitude patterned states are stable away from the degeneracy. In this way, our approach suggests a framework for assessing the robustness of the standard sequence under parameter variations in VH01, as well as in other models.

2.4. Pattern transitions in the model by Rietkerk *et al.*

In the previous section we demonstrated that, in certain limits of model parameter values, pattern sequences can be studied analytically using bifurcation theory. In limits where transitions between patterns all occur in a regime of weak nonlinearity, we showed that the coefficient of the quadratic-order term in these equations affects the sequence that occurs. If this coefficient changes its sign from negative to positive as a precipitation parameter is decreased in value, an analog of the standard sequence occurs in certain cases. Otherwise, alternative sequences, such as one consisting only of spot patterns, can occur. Based on a preliminary numerical investigation of the model by von Hardenberg *et al.* [25], we speculated that this coefficient also encodes information about pattern sequences that occur in more fully nonlinear cases.

In this section, we ask whether the standard sequence can be identified in the model by Rietkerk *et al*. [37] (R02) using the sign of the quadratic coefficient of the hexagonal lattice amplitude equations. Specifically, we investigate whether the standard sequence occurs at parameter values where the quadratic coefficient changes its sign from negative to positive as the precipitation parameter decreases. We calculate the amplitude equation coefficients across broad parameter spaces, and evaluate these coefficients at two values of the precipitation parameter which correspond to pattern-forming bifurcation points. We conduct numerical simulations of the model over the same parameter spaces in order to identify the pattern sequences that occur. We compare the results of the analysis with those of the simulations in order to address whether the standard sequence in this model is signaled by the quadratic coefficient changing signs as the precipitation parameter decreases. Of particular interest is whether this is true in regimes where the weakly nonlinear analysis provides no direct information about pattern stability.

If coefficients from bifurcation analysis can serve as proxies for the standard sequence, then this would allow for more efficient exploration of a model's parameter space than by direct numerical simulation. This would be useful for assessing the robustness of the standard sequence in 3-field models, such as the models by Rietkerk *et al*. [37] and Gilad *et al*. [34, 47], which depend on a large number of non-dimensional parameters.

2.4.1. Model by Rietkerk *et al.*

We study the PDE vegetation model by Rietkerk *et al.* [37] (R02), which consists of three fields: surface water h, soil water w, and plant biomass n. Using the non-dimensional form given by Zelnik *et al.* [34], the model is written as

(2.13)

$$\frac{\partial h}{\partial t} = \underbrace{p}_{\text{precip.}} - \underbrace{I(n)h}_{\text{infil.}} + \underbrace{D_h \nabla^2 h}_{\text{diffusion}},$$

$$\frac{\partial w}{\partial t} = - \underbrace{\nu w}_{\text{evap.}} + \underbrace{I(n)h}_{\text{infil.}} - \underbrace{\gamma G(w)n}_{\text{transpir.}} + \underbrace{D_w \nabla^2 w}_{\text{diffusion}},$$

$$\frac{\partial n}{\partial t} = - \underbrace{\mu n}_{\text{mort.}} + \underbrace{G(w)n}_{\text{growth}} + \underbrace{\nabla^2 n}_{\text{dispersal}},$$

where

$$I(n) = \alpha \frac{n+f}{n+1} \text{ and } G(w) = \frac{w}{w+1}.$$

In this model, precipitation is a constant input to the surface water field. Surface water infiltrates and becomes soil water. The infiltration rate (i.e. the conversion rate of surface water to soil water) increases in the presence of biomass via the function $I(n)$ to model the increased permeability of the soil due to plant roots. $I(n)$ saturates to α as $n \to \infty$. Water leaves the soil via evaporation, and is also transpired by plants. The growth rate of biomass is directly proportional to the transpiration rate, and increases with the availability of soil water via the saturating function $G(w)$. Together, these terms make a positive feedback between infiltration and biomass growth: biomass growth increases with soil water, soil water increases with infiltration, and the infiltration rate increases with biomass.

This model includes surface and soil water diffusion terms. Plant dispersal, which encompasses seed dispersal and clonal growth, is also modeled using a diffusion term. The diffusion terms are in two spatial dimensions (2D), i.e. $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$. Surface water diffusion is typically assumed to occur much more rapidly than soil water diffusion, so $D_h \gg D_w$. Among three-field PDE vegetation models, soil water diffusion and plant dispersal have been modeled as occurring on either similar [37] or different [47] scales with $D_w \geq 1$. An advection term present in the original form of R02 is neglected here, because the focus of our investigation is on flat-terrain patterns. The dynamics of water on a slope modeled via advection break the symmetry that causes 2D patterns such as gaps or spots at pattern onset.

In general, the form of the growth term varies between models [9, 47], and it determines the number of uniform steady state solutions that occur for a given system. For R02, the rate of biomass growth depends linearly on the amount of biomass. The growth rate is also a saturating function of soil water, so that it is linear in the amount of soil water for small values of this variable, and constant for large values. This growth rate permits two spatially uniform steady state solutions, which satisfy the equations

$$0 = p - I(n)h = p - \alpha\frac{n+f}{n+1}h,$$

$$0 = -\nu w + I(n)h - \gamma G(w)n = -\nu w + \alpha\frac{n+f}{n+1}h - \gamma\frac{w}{w+1}n,$$

$$0 = (-\mu + G(w))n = \left(-\mu + \frac{w}{w+1}\right).$$
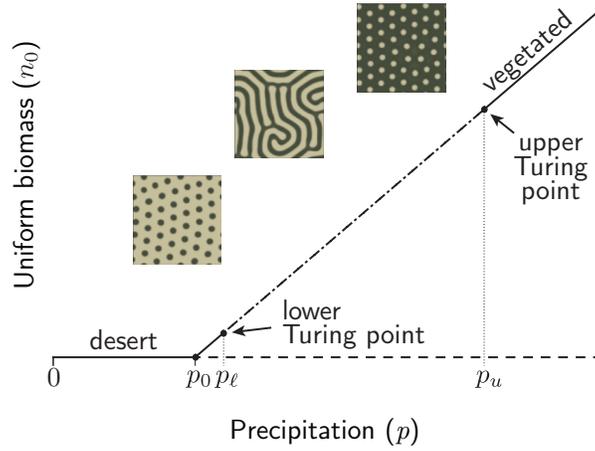
Figure 2.11. Schematic diagram depicting the uniform steady state solutions of R02, with insets showing examples of patterned states occurring at different values of p. The uniform desert state is stable on the interval $p \in [0, p_0)$, and the uniform vegetated state is stable to spatially uniform perturbations for $p > p_0$. The vegetated state is unstable to spatially periodic perturbations at a range of wavelengths on the interval $p \in (p_\ell, p_u)$. We refer to the endpoints p_ℓ and p_u as the lower and upper Turing points, respectively.

One solution, $(h, w, n) = (p/f\alpha, p/\nu, 0)$, represents a zero-biomass desert state. The other solution $(h, w, n) = (h_0, w_0, n_0)$ represents a vegetated state with nonzero biomass,

$$(2.14) \qquad h_0 = \frac{p}{I(n_0)}, \ w_0 = \frac{\mu}{1 - \mu}, \ n_0 = \frac{1}{\gamma\mu}\left(p - \frac{\nu\mu}{1 - \mu}\right),$$

for which $n_0 > 0$ when $p > \mu\nu/(1 - \mu) \equiv p_0$. We note a peculiarity of this model is that the soil water w_0 in the vegetated state depends only on the mortality parameter μ, and not on the precipitation level p. A diagram of uniform steady state biomass as a function of the precipitation parameter is shown in Figure 2.11. The desert state is stable to spatially uniform perturbations at low values of precipitation ($0 \leq p < p_0$), while the vegetated state is stable to such perturbations at higher values of precipitation ($p > p_0$). These steady states exchange stability in a transcritical bifurcation at $p = p_0$.

For a range of precipitation values within the interval $p \in (p_\ell, p_u)$, the vegetated state may be unstable to spatially periodic perturbations at a range of wavelengths. When this is the case, pattern-forming instabilities occur at p_ℓ and p_u via the Turing mechanism [48], and we refer to these points as the lower and upper Turing points respectively.

2.4.2. Model parameter variation

We studied variations of the non-dimensional R02 parameters given by Zelnik *et al.* [34], which are based on the dimensioned parameter values estimated by Rietkerk *et al.* [37]. These parameter values are summarized in Table 2.4. The f parameter in R02 controls the strength of the infiltration feedback, and is bounded between 0 and 1. The default value given in [34] is $f = 0.2$. We used $f \in [0.1, 0.9]$ in numerical simulations. The α parameter controls the rate of infiltration, and can plausibly take on a large range of values depending on the soil type that is modelled. The default value given in [34] is $\alpha = 0.4$, and we considered $\alpha \in [10^{-1}, 10^3]$. The D_h parameter is the ratio of the surface water diffusion rate and the biomass dispersal rate. Rietkerk *et al.* [37] use $D_h = 10^3$ for R02 and Zelnik *et al.* [34] use $D_h = 10^4$ for the corresponding parameter value in a simplified version of the model by Gilad *et al.* [47]. We varied $D_h \in [10^{0.5}, 10^4]$. To consider the dependence of results on co-variation of parameters, we studied the α–f and D_h–f parameter spaces. In additional analysis described in Appendix A and summarized in Section 2.4.5, we considered an ecologically relevant limit where $D_h \gg \alpha$. Our parameter exploration was conducted primarily in this limit, since α was held fixed at 0.4 while D_h was varied, and D_h was held fixed at 10^3 while α was varied.

	Interpretation	Value in [34]	Constraints	Variation studied
μ	mortality rate	0.5	$0 < \mu < 1$	—
α	infiltration rate	0.4	$\alpha > 0$	$10^{-1} - 10^3$
f	infiltration feedback strength	0.2	$0 < f < 1$	$0.1 - 0.9$
ν	evaporation rate	0.4	$\nu > 0$	—
γ	transpiration rate	0.1	$\gamma > 0$	—
D_w	soil water:biomass diffusion rate	1	$1 \leq D_w < D_h$	—
D_h	surface water:biomass diffusion rate	10^3	$D_h > D_w$	$10^{0.5} - 10^4$

Table 2.4. Parameters given by Zelnik *et al*. [34] for the R02 model, and parameters varied in this study.

2.4.3. Amplitude equation calculations

We computed coefficients of the amplitude equations (2.6) for R02 using the procedure outlined in Judd & Silber [46], which takes a perturbative approach to obtaining these coefficients for a 2-field reaction diffusion system. These coefficients are written as expressions of the reaction functions and diffusion parameters. We adapted the Judd & Silber [46] procedure for a three-field reaction-diffusion system to obtain expressions for the amplitude equation coefficients; the aspect of this procedure specific to the calculation of the quadratic coefficient a is described in Appendix A.1. The values of the amplitude equation coefficients a, b, and c are computed at the lower and upper Turing points via *Mathematica*. These calculations are performed on a grid of points in the α–f and D_h–f parameter spaces. The results of these calculations were verified at a few non-degenerate points using a center manifold reduction approach (the general approach is described in [49]). A Nelder-Mead minimization library in *Mathematica* was used to find roots of the quantities a, b, $c - b$, and $b + 2c$, which are relevant to assessing the stability of solutions to (2.6).

We find that the α–f and D_h–f parameter spaces of the model by Rietkerk *et al.* [37] (R02) can be divided into regions where the amplitude equations (2.6) give different qualitative predictions. The results of the calculations at the upper Turing point are summarized in Figure 2.12, and the lower Turing point calculations are summarized in Figure 2.13. In the unlabelled white regions, no Turing points occur because the uniform vegetated steady state is stable to spatially periodic perturbations, and no calculations are performed. The black curves separating the white and shaded regions denote a degeneracy of the Turing points, where the upper and lower Turing points come together at a single precipitation value. In the shaded regions below this curve, two Turing points occur on the vegetated state, and analysis is performed at each point.

Each shaded region in Figures 2.12 and 2.13 is associated with a qualitatively distinct bifurcation diagram applicable to a neighborhood of the Turing point. The qualitative aspects (e.g. stability, branching direction) of the bifurcation diagrams are determined by the signs of the quantities a, b, $c - b$, and $b + 2c$. These quantities arise from the amplitude equation steady state eigenvalues and branching equations listed in Table 2.1. Notably, the sign of the a coefficient serves as a necessary condition for the stability of either small-amplitude gaps or spots solutions. A necessary condition for the stability of gaps is given by $a < 0$, and a necessary condition for the stability of spots is given by $a > 0$.

For the analyses summarized in Figures 2.12 and 2.13, the regions are arrayed similarly in both α–f and D_h–f parameter spaces. For example, regions A-E in Figure 2.12 corresponding to the upper Turing point occur in the same order when varying parameters away from the Turing degeneracy curve (e.g. when increasing D_h compared
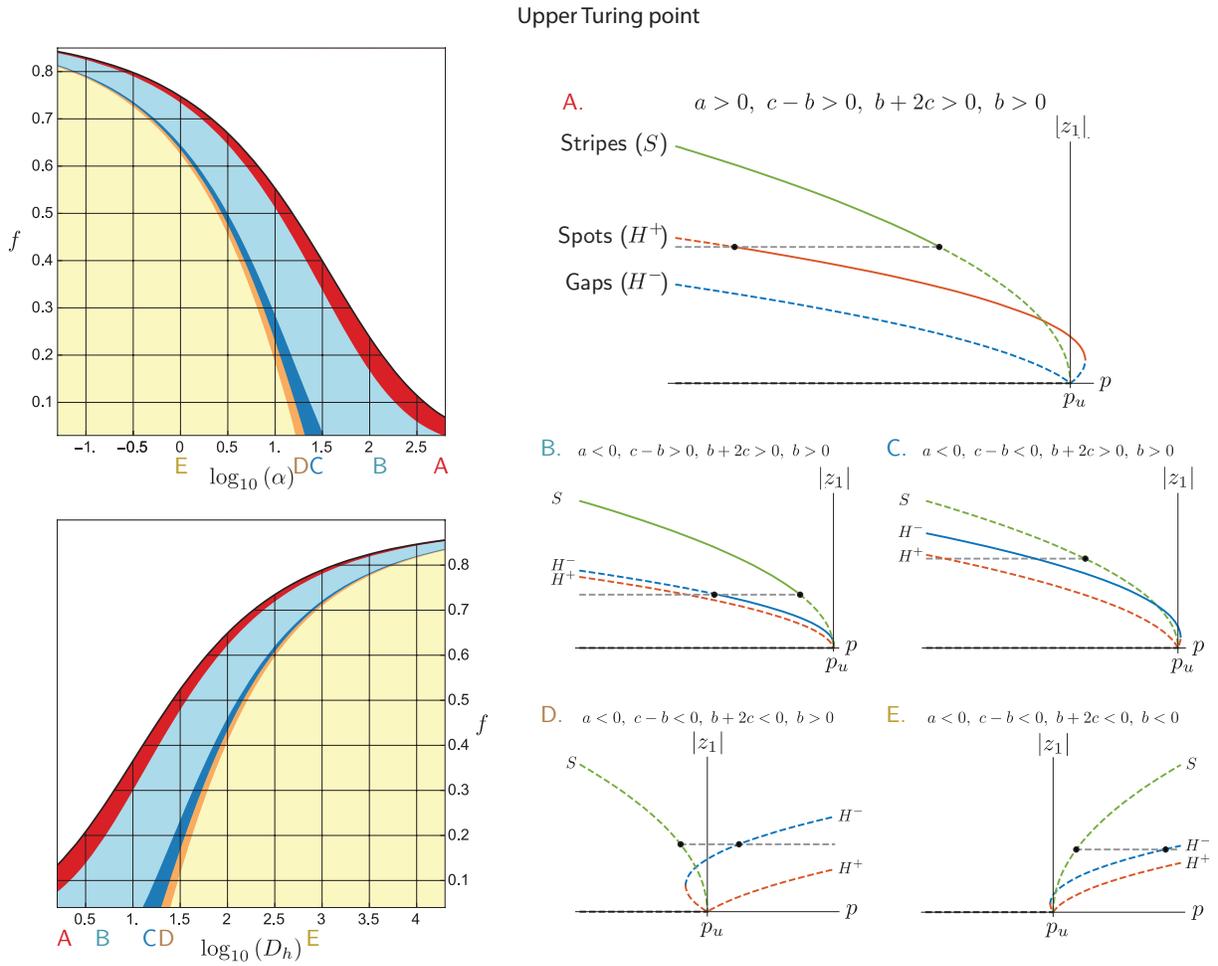
Figure 2.12. Summary of upper Turing point amplitude equation calculations over α–f and D_h–f parameter spaces, along with schematic bifurcation diagrams. The coefficients of the amplitude equations (2.6) are computed, and the curves $a = 0$, $b - c = 0$, $c + 2b = 0$, and $b = 0$ separate the parameter spaces into regions labelled A-E. Qualitatively similar bifurcation structures occur within each region. In the white region, no Turing points occur on the uniform vegetated steady state of R02 and no calculations are performed.

to decreasing α, with f fixed). We find that this occurs because the Turing point calculation and the coefficients a, b, and c depend on the quantity α/D_h, and not α and D_h independently. In the appendices, we show how the Turing point calculation and

the quadratic coefficient a depend on α/D_h via the quantity $\delta \propto \alpha/D_h$. Though the amplitude equation coefficients do not depend on α and D_h independently, this does not translate to the invariance of full solutions to R02 for fixed ratios of α/D_h. This can be seen in weakly nonlinear solutions, which have linear eigenfunctions that depend on α and D_h independently.

We first interpret the results of bifurcation analysis at the upper Turing point, which are summarized in Figure 2.12. We consider a scenario in which precipitation decreases slowly over time, so that the upper Turing point threshold is crossed from above. The sequence of pattern morphologies observed in such a scenario begins with patterns born near the upper Turing point. The regions in Figure 2.12 specify whether the amplitude equations predict a stable patterned state in some neighborhood of the upper Turing point, and also the morphology of that state. A gap (H^-) patterned state stable near the upper Turing point accords with the standard "gaps $\rightarrow$ labyrinth $\rightarrow$ spots" sequence prediction.

In region A of Figure 2.12, the quantities a, $c - b$, $b + 2c$, and b are all positive, which allows a stable spot solution (H^+) to the amplitude equations in a neighborhood of the upper Turing point. This analysis predicts that pattern sequences begin with spot patterns in region A of the parameter space, which is inconsistent with the standard sequence. The stripes solution (S) can also be stable in region A. It stabilizes away from the Turing point, so that spots may transition to stripes as precipitation decreases. However, since the predictions of the amplitude equations break down outside a small neighborhood of the Turing point, it is uncertain whether this stable stripes solution will manifest in the full system as a successor to spot patterns as precipitation

decreases. We expect that as we approach the $a = 0$ boundary of region A, the interval of stability for the spots branch will diminish in size, allowing stable stripes to appear as precipitation decreases.

In regions B and C of Figure 2.12, a is negative and $b + 2c$ and b are positive, which allows a stable gaps solution to the amplitude equations in a neighborhood the upper Turing point. This predicts pattern sequences that begin with gap patterns, which is consistent with the standard sequence. The stripes solution to the amplitude equations can also be stable in region B, as in region A, since $c - b > 0$. Here, gap patterns may transition to stripes as precipitation decreases. In region C, $c - b < 0$ prevents the stability of stripe steady states. The analysis therefore provides no information in this region about the patterns that may follow gaps as precipitation decreases.

In regions D and E of Figure 2.12, a, $c-b$, and $b+2c$ are negative, which means small-amplitude steady state solutions cannot be stable near the upper Turing point. Regions D and E differ only by the branching direction of the always-unstable stripes solution. The stripes solution branches towards the Turing instability interval (i.e. stripes bifurcate supercritically) for region D, since $b > 0$. The stripes solution branches away from the Turing instability interval (i.e. stripes bifurcate subcritically) for region E, since $b < 0$. In both D and E, the gaps solution branches away from the Turing instability region. Since there are no small-amplitude steady state solutions stable in regions D and E, we cannot directly infer from this analysis what patterned states occur near the Turing point. Here, patterned states of the full system likely arise from more strongly nonlinear behavior than the states in regions A-C.
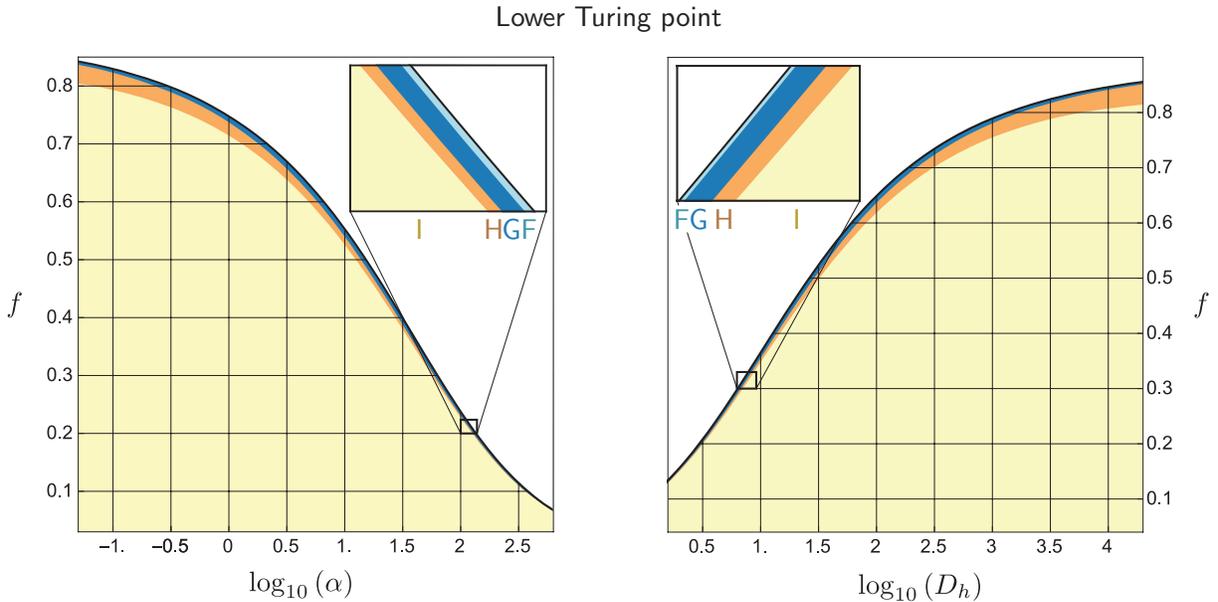
Figure 2.13. Summary of lower Turing point amplitude equation calculations over α–f and D_h–f parameter spaces. The coefficients of the amplitude equations (2.6) are computed, and the curves $c - b = 0$, $b + 2c = 0$, and $b = 0$ separate the parameter space into regions labelled F-I, each of which exhibits a qualitatively distinct bifurcation structure. Bifurcation diagrams applicable to regions F-I resemble diagrams B-E respectively in Figure 2.12, with the roles of gaps and spots exchanged and the solutions reflected so that supercritical branches bifurcate in the direction of increasing precipitation.

Figure 2.13 summarizes the results of bifurcation analysis at the lower Turing point. Over the entire α–f and D_h–f parameter spaces, a is positive, which is a necessary condition for the stability of spot solutions to the amplitude equations. Regions F-H all occur in close proximity to the degenerate Turing point curve, while region I fills the majority of the space. Stable solutions to the amplitude equations occur only in regions F and G. In region F, the spots solution to the amplitude equations is stable near the lower Turing point. The stripes solution is also stable away from the Turing point in this region, so that spots may transition to stripes as precipitation increases.

In region G, the spots solution is stable near the lower Turing point, but the small-amplitude stripes solution can never be stable. In regions H and I, solutions to the amplitude equations are never stable near the lower Turing point, and differ only in the branching direction of the stripes solution. The stripes solution branches towards the Turing instability region for region H, and away for region I.

At both the upper and lower Turing points, our bifurcation analysis cannot provide direct information about stable patterned states for a large region of the parameter space, where small-amplitude solutions are unstable. The central investigation of this chapter is whether the a coefficient, obtained via local analysis, contains information about patterned states near the Turing points in these other regions. In regions B and C of Figure 2.12, it is expected from the analysis that gap patterns are stable near the upper Turing point. We conjecture that the same is true in regions D and E, where no solutions to the amplitude equations are stable, but $a < 0$. Similarly, our analysis only shows that spot patterns are stable near the lower Turing point in regions F and G of Figure 2.13. We conjecture that spots patterns will be stable near the lower Turing point in regions H and I as well, since $a > 0$ there. Taken together, this would result in pattern sequences that begin with gaps and end with spots in a scenario of decreasing precipitation over time (i.e. analogs of the standard sequence) in the region of parameter space where $a < 0$ at the upper Turing point and $a > 0$ at the lower Turing point. We find that these conditions are satisfied over nearly all of the studied α–f and D_h–f parameter spaces, excluding only region A of Figure 2.12. Region A lies adjacent to the Turing degeneracy curve, where the two Turing points approach one another and thus the quadratic coefficients at the Turing points approach the same non-zero value. In

this region, $a > 0$ at the both Turing points, and we anticipate pattern sequences that begin and end with spots.

2.4.4. Numerical simulations

We conducted numerical simulations to identify pattern sequences that occur as precipitation decreases in the R02 model. We employed a numerical procedure which simulates the environment undergoing a slow monotonic change in precipitation over time. These simulations were run using grids of parameter values covering regions of the α–f and D_h–f parameter spaces. The procedure is outlined schematically in Figure 2.14, and is described in detail in Appendix B.1. For each set of parameter values, precipitation is incremented in small discrete steps, and the solution is allowed to reach a steady state between these steps in precipitation. The final state at the previous precipitation value is used as the initial condition for the new precipitation value. The precipitation increment step size was chosen based on the distance between the upper and lower Turing points, p_u–p_ℓ, so that approximately 30-100 end states were saved per simulation. Discrete steps were chosen instead of continuously varying precipitation to avoid transient effects, i.e. simulation results that are sensitive to the rate at which precipitation changes. Simulations were run using the exponential time differencing Runge-Kutta 4 (ETDRK4) scheme [50, 51] modified for 2D systems [52].

The procedure is constructed to run simulations over the interval of p where patterns are stable and to identify any possible hysteresis in the pattern sequences. This is accomplished by first incrementing p upward until patterns die out to yield a uniform
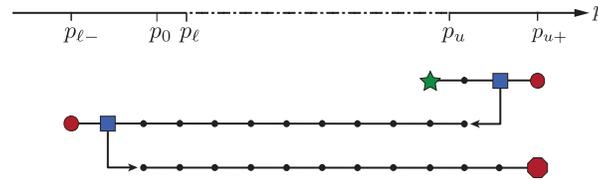
Figure 2.14. Diagram of numerical simulation procedure. Numerical simulations are run at discrete values of p marked by dots. The procedure is initialized with p just below the upper Turing point p_u (star) and run forward in time until a steady state is reached. Then p is stepped upward by a small increment and the simulation is once again allowed to reach steady state. This is repeated until patterns lose stability to a uniform vegetated state at $p = p_{u+}$ (right circle). Using the previous patterned steady state (right square) as an initial condition, p is then stepped downward in the same way until patterns lose stability to a uniform state at $p = p_{\ell-}$ (left circle). Then p is incremented upward a final time, and the procedure terminates when patterns once again lose stability (octagon). Note that p_{u+} and $p_{\ell-}$ do not necessarily coincide with the Turing points p_u and p_ℓ, because patterns may persist outside of the Turing instability interval.

state. We denote this point of pattern die-off as $p = p_{u+}$. Precipitation is then decremented, simulating a scenario in which an ecosystem slowly becomes more arid. This continues until patterns die out again, which yields another uniform state, which we denote $p = p_{\ell-}$. Precipitation is incremented upward a final time to assess hysteresis in the pattern sequence (i.e. whether the sequence occurs differently when p is slowly increasing versus decreasing). The procedure terminates when patterns die out once more. An approximate interval for the stability of patterns is given by $p \in (p_{\ell-}, p_{u+})$, which contains the Turing instability interval $p \in (p_\ell, p_u)$. The Turing instability interval is determined via a linear stability calculation and does not capture the nonlinear stabilization of patterns. In cases where the amplitude equations (2.6) predict stable hexagons solutions near a Turing point, these solutions bifurcate in such a way as to be stable outside the Turing interval. When amplitude equation solutions branch away

from the Turing interval (e.g. when stripes bifurcate subcritically), these solutions may also stabilize at large amplitude outside the Turing interval.

In numerical simulations, we find that the quadratic coefficient a of the amplitude equations (2.6) signals where the standard sequence occurs in the studied parameter spaces of R02. A summary of pattern sequences observed in these numerical simulations is shown in Figure 2.15. These simulations were conducted at sets of parameter values marked by letters, and pattern morphologies were identified by visual inspection. For comparison, the region of the parameter space where $a > 0$ at both the upper and lower Turing points (i.e. region A of Figure 2.12) is shaded. Elsewhere, $a < 0$ at the upper Turing point and $a > 0$ at the lower Turing point. From the results of weakly nonlinear analysis described in Section 2.4.3, we expect pattern sequences beginning with spots (in a scenario of decreasing precipitation) to appear in the shaded region. Outside the thin shaded region, we speculated that analogs of the standard sequence would occur. Indeed, only spot patterned states are observed in numerical simulations in the shaded region; in addition, analogs of the standard sequence are primarily observed in simulations elsewhere.

Examples of the numerical patterned states using $f = 0.2$ and different values of $\log_{10}(D_h)$ are also shown in Figure 2.15. The simulation output is accompanied by lines which plot the locations of the upper and lower Turing points, p_u and p_ℓ respectively, the transcritical point p_0, and upper and lower pattern stability boundaries, p_{u+} and $p_{\ell-}$ respectively. We observed only spot-patterned states in the thin shaded region of Figure 2.15. Examples of such states are shown in simulation output from $f = 0.2$ and $\log_{10}(D_h) = 0.6$. Near the upper Turing point, solutions approximately resemble the
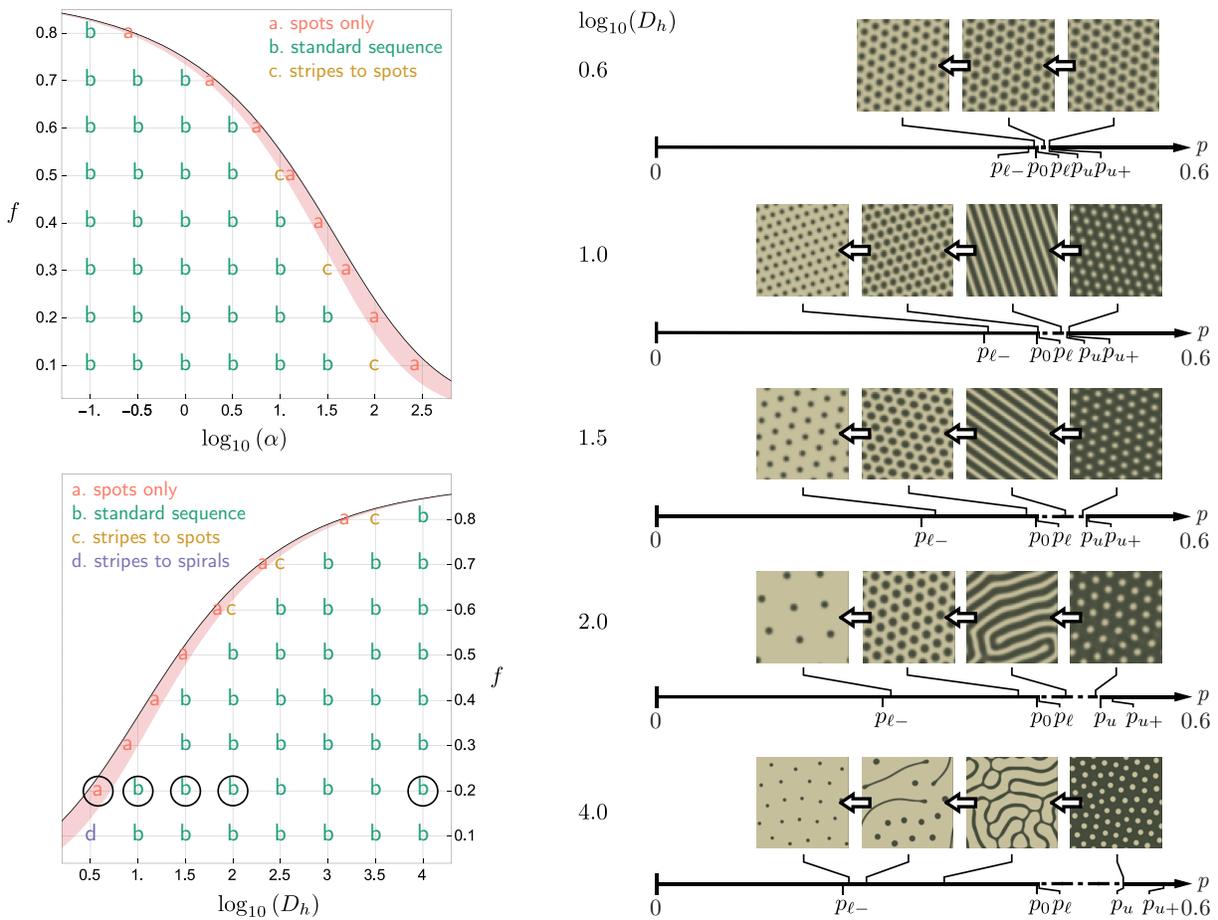
Figure 2.15. Summary of pattern transitions observed numerical simulations over α–f and D_h–f parameter spaces in R02, along with representative examples of transitions from numerical simulations at $f = 0.2$ and $\log_{10}(D_h) = 0.6 - 4.0$. Number lines plot the relative locations of the upper and lower Turing points (p_u and p_ℓ respectively), the transcritical point (p_0), and upper and lower pattern stability boundaries (p_{u+} and $p_{\ell-}$) for the example simulations shown. The parameter values corresponding to the example simulations are circled. Though p_ℓ and p_0 are nearly coincident, the distance between these points is exaggerated to illustrate that $p_\ell > p_0$.

spots (H^+) solution to the amplitude equations (2.6). The profiles of these spot patterns are roughly sinusoidal about the uniform vegetated steady state. An example profile is shown in Figure 2.16. At lower values of p, spot patterns remain stable. The spacing

between spots increases, and the individual spots of vegetation become more sharply peaked, quickly decaying to zero away from the centre of a spot. An example of a sharply peaked profile is also shown in Figure 2.16. Patterns other than spots are not observed in simulations conducted in the shaded region. No notable difference in the qualitative appearance of the spot patterns was observed as precipitation increased in discrete steps.

We primarily observed analogs of the standard "gaps $\rightarrow$ labyrinth $\rightarrow$ spots" sequence in the unshaded region of Figure 2.15, which agrees with our conjecture from analysis. Examples of this sequence in simulation output for $f = 0.2$ and $\log_{10}(D_h)$ ranging from $1.0 - 4.0$ are shown in Figure 2.15. The sets of simulations at $\log_{10}(D_h) = 1.0$ and $\log_{10}(D_h) = 1.5$ use parameter sets from regions B and C of Figure 2.12 respectively, where gaps solutions to the amplitude equations are expected to be stable near the upper Turing point. As precipitation decreases in the simulations, patterns resembling the gaps (H^-) solution to the amplitude equations are first observed near the upper Turing point. Gaps then transition to well-ordered stripe patterns in both sets of simulations. As precipitation decreases further, stripes become disordered before transitioning to spot patterns. The sets of simulations at $\log_{10}(D_h) = 2.0$ and $\log_{10}(D_h) = 4.0$ both use parameter sets from region E of Figure 2.12, where no small-amplitude patterns are stable near the upper Turing point. Still, gaps are observed in numerical simulations near the upper Turing point in both sets of simulations. As precipitation decreases, gaps transition directly to disordered labyrinthine stripes. These stripes eventually transition to spots, which take on non-sinusoidal profiles as shown
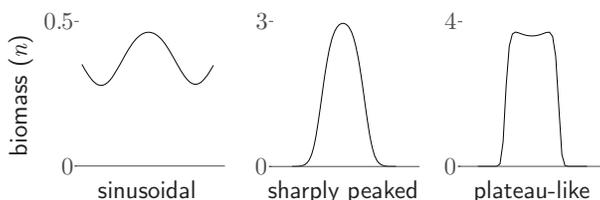
Figure 2.16. Example profiles of individual spot patterns taken from numerical simulations at $f = 0.2$ and $\log_{10}(D_h) = 0.6$ (sinusoidal), $\log_{10}(D_h) = 2.0$ (sharply peaked), and $\log_{10}(D_h) = 4.0$ (plateau-like). The example sinusoidal profile comes from a spot-patterned state near the upper Turing point. The sharply peaked and plateau-like profiles come from spot-patterned states well below the lower Turing points in their respective simulations.

in Figure 2.16. Hysteresis occurs in the points of transition between pattern morphologies, and this hysteresis is larger in the simulation with the larger value of D_h. The transitions between gaps and labyrinths occur at a lower value of p when decreasing precipitation compared to increasing precipitation. The same applies to the transition between labyrinths and spots.

The simulation examples in Figure 2.15 demonstrate a trend of increasingly nonlinear behavior as D_h increases. This trend is generally representative of what we observe in the other simulations when parameters are varied away from the Turing degeneracy curve, e.g. when α decreases. One aspect of the increasing nonlinearity can be accounted for via our bifurcation analysis. The regions in Figure 2.12 order the parameter spaces by nonlinearity at the upper Turing point. Moving away from the Turing degeneracy curve, the amplitude equations first predict stable weakly nonlinear patterns in regions A-C, and then imply strongly nonlinear patterns in regions D and E since small-amplitude patterns are unstable. This manifests in simulations as

small-amplitude sinusoidal patterns occurring near the upper Turing point when parameters are near to the Turing degeneracy curve, sharply peaked patterns occurring beyond this, and plateau-like patterns occurring when very far away from the curve. Other aspects of increasing nonlinearity are apparent in certain qualitative behaviors observed in the simulations. Patterned states increase in their disorder and begin to exhibit coexistence as distance from the degeneracy curve increases. The interval of pattern stability, $(p_{\ell-}, p_{u+})$, increases in length as well. In particular, $p_{\ell-}$ decreases to extend the length of the interval $(p_{\ell-}, p_{\ell})$. As this interval increases with increasing distance from the Turing degeneracy curve, the transition point between stripe and spot states decreases to lower values of precipitation. This causes spot patterns to remain stable at values of precipitation well below the lower Turing point. This implies the stabilization of a strongly nonlinear patterned state far from the Turing instability interval.

In addition to the standard sequence, we observed a few instances of "stripes $\rightarrow$ spots" sequences in the unshaded region of Figure 2.15. We determined that these are actually instances of the standard sequence, where gaps do not appear in simulations. Our bifurcation analysis indicates that gap solutions to the amplitude equations are stable only very near to the Turing point in these parameter sets. Because our numerical procedure increments precipitation in discrete steps of fixed size, the gaps branch may be bypassed. To test whether gaps can exist stably for parameter sets where "stripes $\rightarrow$ spots" transitions are observed, we conducted additional numerical simulations, which are described in the electronic supplementary material. In these simulations, gap patterns were assessed to be stable.

We also observed time-varying spiral wave patterns in one instance of the numerical simulations, at $f = 0.1$ and $\log_{10}(D_h) = 0.5$. We ran additional simulations, described in the electronic supplementary material, at nearby parameter values and found that spiral patterns are confined to values of D_h that are smaller than typically considered ecologically applicable. To our knowledge, there are no previous reports of spiral wave patterns occurring in R02 or any other vegetation model. However, we remark that the waves observed in R02 resemble spiral wave patterns observed in other reaction-diffusion contexts such as chemical reaction systems [53] and models of phytoplankton dynamics [54].

2.4.5. Quadratic coefficient analysis

In Appendix A.1, we derive a closed-form expression for the quadratic coefficient a. This expression involves derivatives of nonlinear terms in R02, the infiltration term $I(n)h$ and the growth term $G(w)n$, where $I(n) = \alpha(n + f)/(n + 1)$ and $G(w) = w/(w + 1)$. This expression also involves the null vector $(H_1, W_1, N_1)^T$ and left null vector $(\widetilde{H}_1, \widetilde{W}_1, \widetilde{N}_1)$ of the system linearized about the uniform vegetated steady state (2.14), both of which are defined in Appendix A.1. Explicitly,

$$(2.15) \quad a = \left(I'(n_0)H_1 N_1 + \frac{1}{2}I''(n_0)h_0 N_1^2 \right)\left(\widetilde{W}_1 - \widetilde{H}_1 \right)$$
$$+ \left(G'(w_0)W_1 N_1 + \frac{1}{2}G''(w_0)n_0 W_1^2 \right)\left(\widetilde{N}_1 - \gamma\widetilde{W}_1 \right).$$

The infiltration function $I(n)$ is an increasing concave-down function of n, and so $I'(n) > 0$ and $I''(n) < 0$. The growth function $G(w)$ is similarly increasing and concave-down with respect to w.

In a natural limit of the parameters which corresponds to $D_h \gg \alpha$, we find in Appendix A.3 that a negative term dominates quadratic coefficient a (2.15) at the upper Turing point and a positive term dominates at the lower Turing point. These arise through distinct scalings of the critical wavenumber and the onset parameter values with the quantity $\delta \equiv I(n_0)/D_h \propto \alpha/D_h$ at the different Turing points, and are calculated in Appendix A.2. These scalings at onset in turn determine distinct scalings for the left and right null vector components. At the lower Turing point, for δ sufficiently small, we find that $(H_1, W_1, N_1)^T = (\mathcal{O}(1), \mathcal{O}(1), \mathcal{O}(1))^T$ and $(\widetilde{H}_1, \widetilde{W}_1, \widetilde{N}_1) = (\mathcal{O}(\delta), \mathcal{O}(\delta), \mathcal{O}(1))$. Additionally we find that $n_0 = \mathcal{O}(\delta)$. Then at the lower Turing point,

$$a = \frac{2Q_\ell^2 G'(w_0)}{N_\ell} N_1^\ell + \mathcal{O}(\delta),$$

where Q_ℓ^2 and N_ℓ are the positive scaling constants given by (A.14). Thus a is positive at leading order at the lower Turing point. This corresponds to spot patterns. Sufficiently far from the degeneracy at which the two Turing points merge, the upper Turing point has the opposite sign. Specifically, at the upper Turing point, we find that $(H_1, W_1, N_1)^T = (\mathcal{O}(\delta^{1/2}), \mathcal{O}(\delta^{1/2}), \mathcal{O}(1))^T$ and $(\widetilde{H}_1, \widetilde{W}_1, \widetilde{N}_1) = (\mathcal{O}(\delta^{1/2}), \mathcal{O}(1), \mathcal{O}(1))$. Then at the upper Turing point

$$a = \frac{G'(w_0)I''(n_0)h_0 n_0}{\nu + \gamma G'(w_0)n_0} + \mathcal{O}(\delta^{1/2}).$$

This expression is negative at leading order since $I''(n) < 0$, and thus corresponds to gap patterns.

In order for this result to hold, we must also be sufficiently far from a Turing degeneracy. The quadratic coefficient at the two Turing points takes on the same sign in a neighborhood of the degeneracy, marked by the shaded region in Figure 2.15 (see also the discussion of region A in Section 2.4.3). The conditions that $D_h \gg \alpha$ and that we are sufficiently far from a Turing degeneracy are not independent. As f increases, the Turing degeneracy occurs at smaller values of δ. This can be observed in Figures 2.12-2.15, where α decreasing and D_h increasing correspond to smaller δ. At smaller δ, the regions where a takes the same sign at both Turing points diminish, and less distance from the Turing degeneracy is necessary for the scaling results given above to hold. This can be seen in the narrowing of the shaded regions in Figure 2.15 with increasing f.

2.4.6. Summary and discussion

We find that the quadratic coefficient a from a bifurcation analysis divides the studied parameter spaces of the Rietkerk *et al.* [37] (R02) model into two regions. In a thin region of the parameter space adjacent to the degenerate Turing point curve, where the Turing points are very close to each other, a is positive at both points. Correspondingly, we observe only spot patterns in numerical simulations. Elsewhere a takes opposite signs at the two Turing points. When this happens, we primarily observe the standard sequence. This strongly suggests that the a coefficient resulting from weakly nonlinear analysis holds predictive value for assessing the nonlinear behavior of the system.

Specifically, it appears to serve as a proxy for the sequence of nonlinear patterns that will manifest for any parameter set of R02.

Since a is computed analytically, it is possible to trace the influence of model terms and parameter values on the sign of a, and thus on the sequence of patterns that are predicted. This presents an approach for comprehensively exploring the full 7-dimensional parameter space of the R02 model. Our analysis of the quadratic coefficient in Appendix A shows that in any parameter regime where the surface water diffusion rate D_h is sufficiently large compared to the infiltration rate α, the quadratic coefficient takes values consistent with standard sequence.

We believe that $D_h \gg \alpha$ is an appropriate limit for the R02 model. We have treated the parameters $D_h > 1$, $\alpha > 0$ and infiltration feedback strength $0 < f < 1$ as essentially unconstrained in this study. Our analysis shows, however, that the Turing point calculation is invariant to fixed values of the ratio α/D_h (see Appendix A.2), and that Turing bifurcations only occur for α/D_h sufficiently small. This can be seen in Figures 2.12-2.15, where Turing bifurcations only occur for α sufficiently small with $D_h = 10^3$, and for D_h sufficiently large with $\alpha = 0.4$. The amount of separation between α and D_h that is required for Turing points to occur depends on f, which can also be seen in Figures 2.12-2.15. For values of f greater than 0.6, holding other parameters fixed at default values, α and D_h must be separated by at least two orders of magnitude for Turing points to occur. For any value of $f \in (0, 1)$, not much additional separation between α and D_h is required for the quadratic coefficient a to take opposite signs at the Turing points and for the standard sequence to occur.

Moreover, the limit where $D_h \gg \alpha$ is ecologically relevant. For all soil types, we expect D_h to be large, because it represents the ratio of surface water diffusion to biomass dispersal, which occur at quite different scales. Additionally, a global study of the factors associated with the existence of vegetation patterns by Deblauwe *et al.* [6] finds that patterns favor environments with non-sandy soils, where relatively small rates of infiltration allow for substantial water redistribution via surface runoff. For such non-sandy soils, Rietkerk *et al.* [37] estimate an infiltration rate and surface diffusion rate that results in $\alpha/D_h \approx 5 \times 10^{-4}$. Holding the other model parameters fixed at their default values, this corresponds to the standard sequence for 99% of the range of f over which patterns are present.

Although vegetation patterns tend to occur in non-sandy soils, a notable exception is the fairy circle phenomenon in Namibia. These patterns occur in a sandy soil environment marked by a high infiltration rate [55]. In empirical comparisons between clayey soils and sandy soils, infiltration rates differ only by a factor of approximately 10-20 [56]. In the Namibian system specifically, estimates for the infiltration rate parameter used in the model by Gilad *et al.* [28, 34, 47] are at most one order of magnitude larger than those in the non-sandy R02 system[1] [57]. Given similar surface water diffusion rates to R02, this results in the standard sequence over about 95% of the range of f for which patterns are present in the R02 model ($\alpha/D_h \approx 5 \times 10^{-3}$).

---

[1]Getzin *et al.* [57] estimate a nondimensional infiltration rate for the model by Gilad *et al.* [28, 34, 47] as $\alpha_{Getzin} \equiv A/M = 6$, where A is a dimensioned infiltration rate and M is a mortality rate. The nondimensional infiltration and mortality rates for R02 are given by Zelnik *et al.* [34] as $\alpha_{R02} \equiv A/(cg_{max})$, $\mu_{R02} \equiv M/(cg_{max})$, where $1/(cg_{max})$ sets the time scale of the nondimensional system. Therefore, $\alpha_{R02} = \alpha_{Getzin}\mu_{R02} < \alpha_{Getzin}$.

We conclude that the standard sequence prediction appears robust to parameter variation in the R02 model. This conclusion is based on evidence that the quadratic coefficient serves as a proxy for the standard sequence, along with our finding that the coefficient takes on values consistent with the standard sequence in the ecologically relevant region of the parameter space.

CHAPTER 3

An empirical investigation of vegetation band dynamics in the Horn of Africa

Bands of vegetation separated by stretches of bare ground arrayed along gradually-sloping terrain have been reported at multiple sites in Africa, North America, and Australia [58]. Vegetation bands often occur on land used for pastoral farming [42, 59, 60] and can serve as crucial buffers against land degradation [61, 62]. Many model investigations have focused on the dynamics and resilience of vegetation in response to changes in environmental conditions [16]. Evidence for these predictions has been limited by an irregular record of remotely-sensed observations from the pre-Landsat-era.

In this chapter, we investigate the dynamics and resilience of vegetation bands in two distinct regions of Somalia. We collected and precisely aligned digital scans of British Royal Air Force (R.A.F.) aerial survey photography taken in 1952, reconnaissance photography taken in 1967, and multispectral satellite imagery taken in the last decade. Among our study areas are locations that remain relatively pristine, and also locations that experienced a dramatic increase in human pressure. Both regions have experienced multi-year fluctuations in rainfall and a warming trend over the last half-century. The aim of this study is to assess the nature of measurable change in vegetation banding on a multi-decadal timescale via a systematic visual comparison of imagery. We also identify band properties for which changes can be readily observed

in remotely-sensed imagery. Our approach relies on transect-based measurements and comparative Fourier analysis of the bands.

3.1. Vegetation bands on gradual slopes

Spatially periodic bands of vegetation are widespread in the drylands of Africa, North America, and Australia [58] (see Figure 3.1 for examples). Due to large scale of the phenomenon, vegetation banding was not recognized until the systematic use of aerial survey photography in the 1940s [7, 63]. In general, vegetation bands typically occur in areas of low topographic relief, with slopes of less than one meter of vertical change per 200 meters of horizontal change [61]. They reliably orient perpendicularly to the local elevation gradient, to the extent that they can be used to infer the flow of surface water and to identify water catchment basins [7].

Vegetation bands are thought to arise due to redistribution of water via surface flow. Plant roots may increase the permeability of soils, reducing infiltration rates of surface water runoff following significant rainfall events, thus reinforcing the growth of vegetation. This infiltration feedback is relevant in many areas investigated by field-work [58, 64–67]. Compacted clay soils and microbial soil crusts strongly inhibit infiltration in unvegetated areas, making the effect of plant roots on local water availability significant. Additionally, it is proposed that plant growth facilitation via local shading and laterally-extended plant root systems which gather water over large areas can lead to vegetation banding [28, 47].
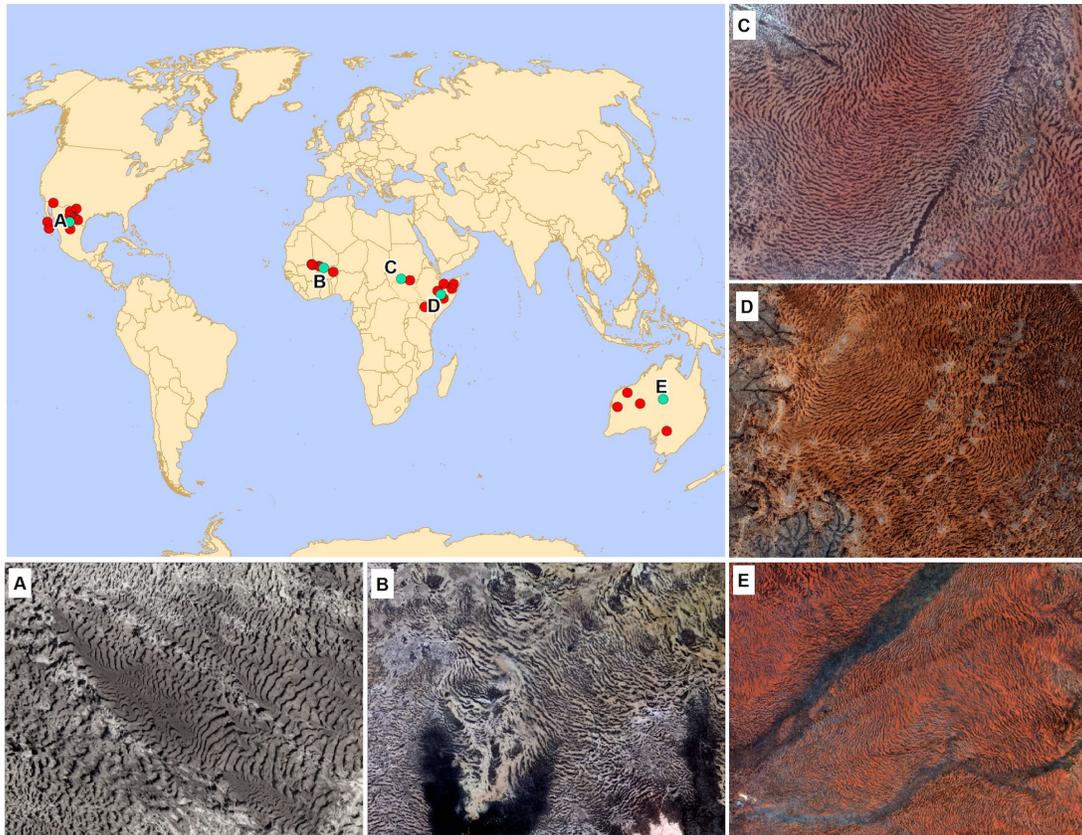
Figure 3.1. Global distribution of vegetation bands with example imagery. Points represent banded locations identified by Deblauwe *et al.* [6]. (A) Sierra del Diablo, Mexico (27.93° N, 103.35° W, 5/11/2016). (B) Mopti Region, Mali (14.64° N, 1.56° W, 3/6/2016). (C) Muglad, Sudan (11.27° N, 28.23° E, 3/20/2016). (D) Gode, Ethiopia (5.95° N, 43.96° E, 3/19/2016). (E) MacDonnell Range, Australia (23.39° S, 133.86° E, 3/20/2016). Figure courtesy of Jake Ramthun.

Vegetation bands are typically composed of a mixture of grasses, shrubs, and trees [58], and display a characteristic plant composition structure that varies along the direction of slope [66, 68] (Figure 3.2). Since most of the water each band receives after a rainfall event arrives due to runoff from upslope, the upslope edge of the band is
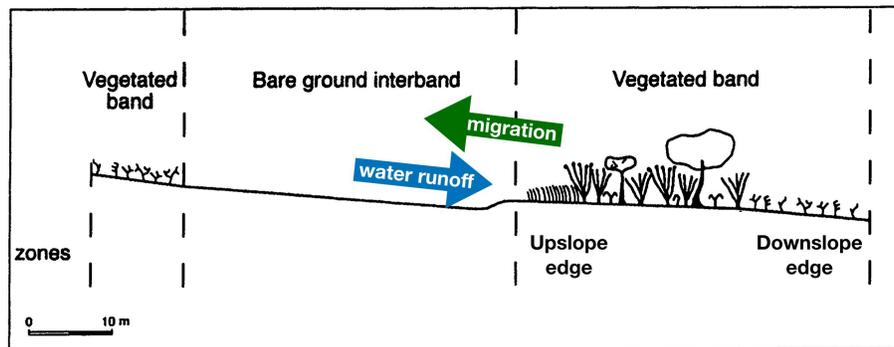
Figure 3.2. Schematic cross-section of a vegetation band, adapted from [68].

nourished preferentially, and can support fast-growing grasses which colonize the sur-rounding bare area. These grasses create conditions beneficial for further plant growth, supporting the growth of woody, long-lived vegetation in the heart of the band. The downslope edges of bands receive relatively less water, and experience a higher rate of plant mortality than other parts of the band. Thus, the spatial structure of the band results from a dynamic ecological process, which over long time periods can cause the bands to migrate uphill. This movement was first hypothesized and estimated in field work conducted in the 1950s [64], and only recently confirmed via analysis of satellite imagery [69].

3.2. Model predictions from previous studies

Modeling efforts since the late 1990s have focused on mathematizing these mecha-nisms through partial differential equation approaches [9, 28, 37, 70]. Such models are able to qualitatively reproduce the appearance of the bands, as well as the slow up-hill migration of bands estimated in early fieldwork [64]. One such model [28, 34, 47] that incorporates some of the mechanisms described in Section 3.1 can be written in

nondimensional form as

$$\frac{\partial h}{\partial t} = \underbrace{p}_{\text{precip.}} - \underbrace{\alpha \frac{n+qf}{n+q} h}_{\text{infil.}} + \underbrace{D_h \nabla^2 (h^2) + 2D_h \nabla h \cdot \nabla \zeta + 2D_h h \nabla^2 \zeta}_{\text{surface water transport}},$$

(3.1)
$$\frac{\partial w}{\partial t} = \underbrace{\alpha \frac{n+qf}{n+q} h}_{\text{infil.}} - \underbrace{\nu w}_{\text{evap.}} + \underbrace{\nu \rho n w}_{\text{shading}} - \underbrace{\gamma (1+\eta n)^2 wn}_{\text{transp.}} + \underbrace{D_w \nabla^2 w}_{\text{diffusion}},$$

$$\frac{\partial n}{\partial t} = \underbrace{\nu w n (1-n)(1+\eta n)^2}_{\text{growth}} - \underbrace{n}_{\text{mort.}} + \underbrace{D_n \nabla^2 n}_{\text{dispersal}}.$$

The variable h represents a surface water height field, w a soil water field, and n a plant biomass field. Surface water transport terms derived from a shallow-water approximation describe the flow of thin sheet of water on a surface $\zeta(x, y)$. These terms can be written in conservation form as $2D_h \nabla \cdot \mathbf{J}$, where $\mathbf{J} = h \nabla (h + \zeta)$, so that the speed of flow is proportional to $h + \zeta$. In the case of an idealized planar elevation surface, e.g., $\zeta = vx$, this model can generate periodic vegetation bands that migrate in the positive x direction over time. Often in this topographic scenario, the dynamics of models such as (3.1) are simulated and analyzed in one spatial dimension only.

Many model investigations have focused on the dynamics and resilience of vegetation in response to changes in aridity or rainfall conditions [16]. As aridity increases, banded vegetation is predicted to increase in wavelength [21, 35, 36, 71, 72], to undergo abrupt shifts in migration speed [73, 74], and to become susceptible to break-up that results in a spotted state [75]. Rainfall fluctuations are predicted to affect migration speed and the widths of bands [72, 73]. Grazing pressure, commonly viewed interchangeably with human impact, has been modeled as an increase to the plant mortality rate [21, 76], and by disturbances to the state of vegetation. Such disturbances

can cause an increase in band wavelength [34] or band break-up [47]. In many of these investigations, the simulated dynamics of environmental change and the resulting ecological shifts play out over the equivalent of decades to centuries.

Evidence for predictions related to the dynamics and resilience of vegetation bands has been limited by an irregular record of remotely-sensed observations from the pre-Landsat-era. Wu *et al.* used aerial survey photography of tiger bush banding in Niger over three decades to observe vegetation fragmentation [77], a phenomenon that appears related to decreases in band width and break-up. Valentin & d'Herbès used similar imagery over Niger to observe that the widths of bands fluctuated in response to rainfall history over four decades [61]. Deblauwe *et al.* used declassified reconnaissance imagery to measure appreciable band migration over four decades in Texas, USA, and the Haud region of Somalia. They also observed band width fluctuation in response to rainfall variation in Texas [69]. Additional studies of dryland vegetation patterning on flat terrain found that the vegetation transitioned between qualitatively distinct states during a period of significant drought [42], and that human pressure hastened this type of transition [41].

3.3. Band dynamics in the Horn of Africa

3.3.1. Regional information

We studied areas within two distinct regions of the Horn of Africa (Figure 3.3), the Sool Plateau and Haud pastoral regions of Somalia. Both regions are generally characterized by an arid climate (aridity index = 0.04-0.1) [78]. Due to a lack of continuous rainfall station monitoring in and around our regions of study, we assessed the historical regional climate using 20th Century Reanalysis [79] and the CPC/Famine Early Warning System Dekadal Estimates datasets. Mean annual rainfall in both regions ranges between 100-300 mm. We found no evidence that rainfall conditions have improved in either region in recent decades, and we identified a warming trend in average yearly temperature of 1-2 °C over the last half-century. Additional details of our analysis of climate data are given in Appendix C.1.

Regional soils are claylike and prone to crust formation, resulting in low permeability and surface water runoff following high-intensity rainfall [14, 66]. Hemming found that soils are wetter beneath bands in the Haud, indicating greater soil permeability in vegetated areas [66]. Vegetation bands in both regions are dominated by *Andropogon kelleri* grasses [63, 66]. Bands also contain a mix of trees and shrubs, most notably *Acacia bussei*. In recent decades, *Acacia bussei* has diminished in abundance in the Sool Plateau due to cutting for charcoal production [14]. Disruption of traditional grazing patterns has resulted in overgrazing in many areas of the Sool Plateau, including Dhahar (SP4) [14].
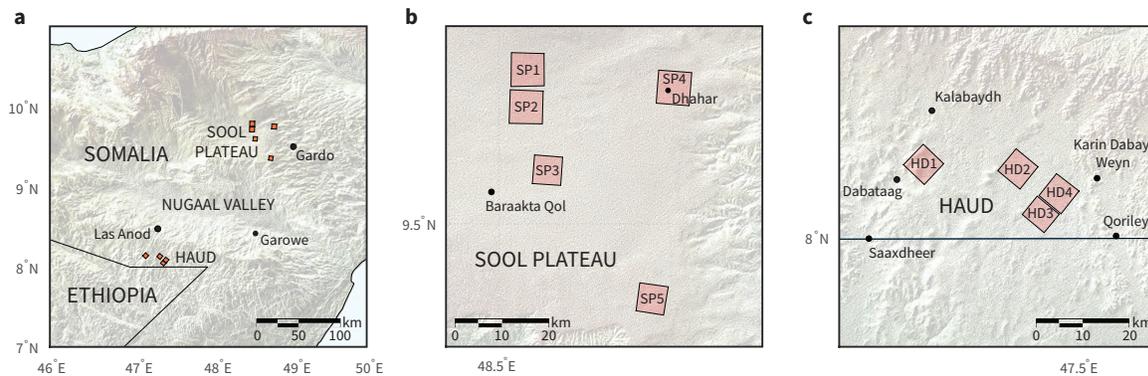
Figure 3.3. Areas studied in this investigation are defined by nine distinct aerial survey photographs taken in 1952. (a) Study areas are clustered in two regions of Somalia separated by the Nugaal Valley. (b) SP1–SP5 are located in the Sool Plateau pastoral region of Somalia. (c) HD1–HD4 are located in the Haud pastoral region of Somalia. Each photograph covers an approximate area of 50 km^2. Elevation is shown as shaded relief, and nearby settlements are labeled.

3.3.2. Imagery and elevation data

We studied approximately 260 km^2 of imagery within the Sool Plateau and 200 km^2 of imagery within the Haud. Study areas were chosen based on a combination of factors; in particular, we wished to include areas with different development and degradation outcomes, areas with recorded soil and floristic information based on field studies, areas in geographically distinct regions, and areas featuring well-defined banding. Study area boundaries are defined by our choice of British Royal Air Force (R.A.F.) aerial survey photography, which comprise our earliest image datasets. Aerial survey photographs were taken in 1951–52 over broad areas of British Somaliland, and specific photographs were scanned on request by the Bodleian Library at the University of Oxford. We also studied declassified reconnaissance satellite imagery taken in 1967, and DigitalGlobe imagery for dates spanning 2004–2016. Resolution of imagery used

in this study ranges between 1.4–2.4 m/pixel. Satellite images taken between 2004–2016 containing red and near-infrared channels were used to compute a Soil-adjusted Vegetation Index (SAVI) [80]. We manually georeferenced R.A.F. scans and the 1967 reconnaissance image using visually identified control points. We estimated alignment error to be approximately 1–2 pixels. Additional information about data sources is given in Table C.1, and georeferencing analysis and information about calculation of SAVI are given in Appendix C.2.1.

We estimated local gradient within our study areas using the Shuttle Radar Topography Mission Global 1 arc second elevation dataset [81]. Because of the noise characteristics of the dataset and the low relief of our study areas, we used a second-order finite difference operator with noise-suppressing properties to estimate gradient and slope [82]. Additional information about calculation of slope and a methodological analysis are given in Appendix C.2.2.

3.3.3. Human pressure and band degradation

We assessed changes in human activity and vegetation banding over time through a systematic visual comparison of R.A.F. photography and recent satellite imagery. We developed a graphical user interface in MATLAB for comparing images. For each area, we split both R.A.F. scans and recent imagery into 1 km × 1 km boxes, and evaluated qualitative features within these boxes. Roads and dirt tracks can be visually identified in both the aerial photographs and the satellite imagery, and their presence and qualitative appearance served as our primary proxy for inferring the extent of human pressure. Vegetation in both the aerial photos and satellite imagery contrasts sharply
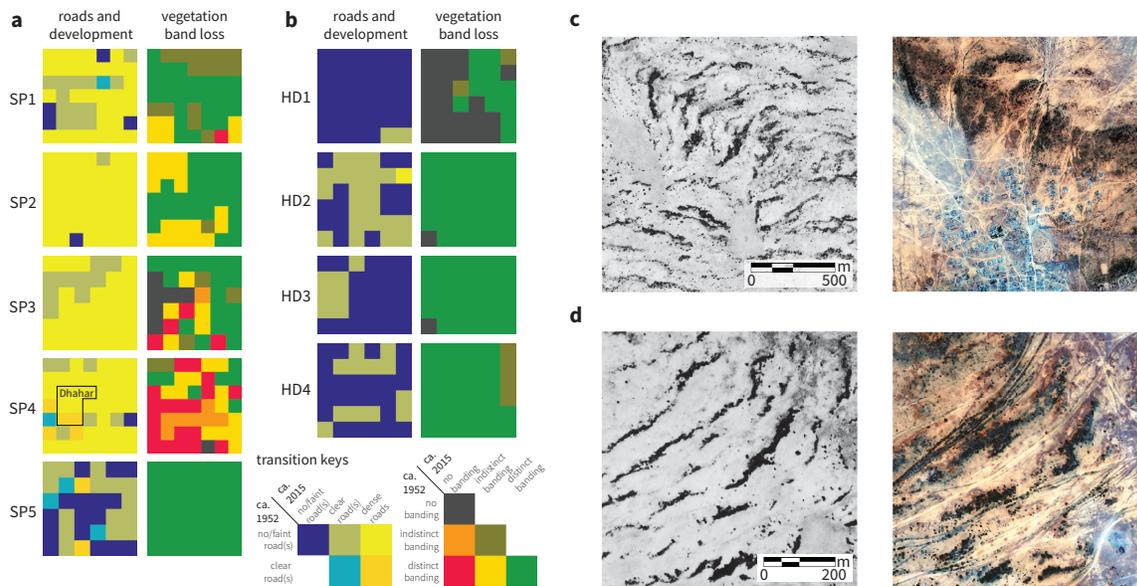
Figure 3.4. Band loss occurs in areas with substantial increases in human activity. Transitions between the qualitative state of road and track cover observed in aerial photography and recent satellite imagery are shown here for all study areas. (a) SP1–SP4 in the Sool Plateau show a high degree of road and track development, and a moderate to high degree of band loss. A large settlement (Dhahar) developed within SP4, and is indicated with a black border. SP5 saw little increase in road cover, and no band loss or degradation was observed. (b) Areas in the Haud (HD1–HD4) show only a small increase of road and track development, and no substantial band loss or degradation is observed. (c) An example of band loss and degradation due to land development in SP4 ($9.76°$ N, $48.82°$ E, 02/22/1952, 08/16/2016). (d) An example of band degradation amid dense track cover in SP3 ($9.58°$ N, $48.57°$ E, 11/29/1952, 12/03/2011). Satellite images courtesy of the DigitalGlobe Foundation.

with the light background of bare soil, and bands are clearly identifiable. We defined degradation in this context as either the breakdown in regularity or the disappearance of banding. A detailed description of this procedure is given in Appendix C.3.1.

Substantial road and track development occurred in much of the Sool Plateau, with most areas (SP1–SP4) transitioning from having either no roads or faint roads in 1952

to having roads or tracks that densely cover the landscape in the modern images (Figure 3.4a). The settlement Dhahar was founded within SP4 after the 1952 photograph, and now supports a population of approximately 13,000 (Figure 3.4c). We observed much less road and track development in SP5 and in the Haud (HD1–HD4) (Figure 3.4b). At many sites within the Haud, human-made structures visible in the 1952 images appear to persist into the current decade, suggesting no major change in land use over the intervening time (Appendix C.3.2, Figure C.5a,b).

Band degradation is prevalent in the human-impacted areas SP1–SP4 (Figure 3.4a). Bands have disappeared entirely from the landscape in large parts of SP3 and SP4. Only part of the band loss in these areas appears directly related to clearing for land development, since loss also occurs in areas without human-made structures. Where faint remnants of bands are visible in SP4, degradation appears to have occurred without a visible change in band wavelength (Appendix C.3.2, Figure C.5c). In SP2–SP4, dense tracks often appear between bands (Figure 3.4d). Frequently we observed vegetation growing within roads and tracks, which suggests that these structures likely disrupt the flow of water on the landscape. In SP5 and HD1–HD4, individual bands often remain identifiable after six decades based on visible details of their morphology, and we observed no substantial band degradation.

3.3.4. Band wavelength

Models predict that band wavelength should increase in response to sufficient increases in environmental pressure [21, 35, 36, 71]. We measured changes in band wavelength using the Fourier window method by Penny *et al.* [83]. The method measures

band wavelength and orientation in a sliding window using a 2D FFT, and computes a uniqueness metric based on the unimodality of the power spectrum. We discard data points which correspond to sites without banding using a manually-drawn mask. We additionally discard data points with uniqueness values below a threshold.

In areas where bands have not degraded from the landscape, we found that changes in wavelength are imperceptible. We defined wavelength change as $W_2/W_1 - 1$, where W_1 and W_2 are wavelengths measured in the 1952 R.A.F. photographs and recent satellite images, respectively. Typical change ranges between 0–10% for all study areas except SP3 and SP4, where change ranges between 0–20%. We visually inspected sites with $> 25\%$ measured wavelength change. In some cases, it appears that these detected changes occur due to the loss of an individual band, often near evidence of human activity (Appendix C.4.2, Figure C.7). In most cases, however, we saw no clear reason for a detected wavelength change, and attributed these false detections to wavelength measurement error.

Band wavelength is predicted by models to vary with local slope, though the nature of this relationship can be parameter and history-dependent [33, 84]. We found significant a negative correlation between wavelength and slope at SP3 only (Table 3.1). The sign of this correlation agrees with empirical findings in other studies [42, 69, 83].

3.3.5. Band widening in human-impacted areas

During the course of visual comparison, we noted that bands in some areas appear to widen over time in the direction of local slope, and that they appear to migrate uphill. We quantified band width change and band migration in all study areas using

Area	Slope (%) S	Wavelength (m) W_1	W_2	corr(S, W_1) (p, t, df)
SP1	0.3–0.4	130–170	130–190	-0.15 (0.29, 1.3, 51)
SP2	0.1–0.3	130–170	140–170	0.02 (0.89, 0.0, 83)
SP3	0.1–0.3	120–150	140–180	-0.34 (0.04, 4.5, 35)
SP4	0.1–0.2	130–160	150–180	0.08 (0.53, 0.4, 67)
SP5	0.2–0.4	120–140	120–140	-0.23 (0.06, 3.7, 63)
HD1	0.4–0.6	80–100	80–120	-0.25 (0.33, 1.0, 16)
HD2	0.3–0.5	90–110	90–120	-0.19 (0.20, 1.7, 45)
HD3	0.3–0.5	80–100	80–110	-0.14 (0.17, 1.9, 94)
HD4	0.4–0.5	100–120	100–120	0.08 (0.58, 0.3, 47)

Table 3.1. Band wavelength measured using the Fourier window method by Penny *et al.* [83]. Ranges shown are the 25th and 75th percentiles. Wavelengths W_1 were measured in the 1952 aerial photography datasets, and W_2 were measured in recent satellite imagery datasets. Significance of correlations was assessed using a t-test corrected for spatial autocorrelation [85], and p values, t values, and degrees of freedom are given. Analysis details are given in Appendix C.4.

automated transect measurements of individual bands. We gathered grayscale image intensity profiles along transects through the bands in the direction of slope. We used the same transects for multiple images at the same study area. We then fit a simple plateau function to each intensity profile to extract band width, as well as the positions of band edges which were used to measure band migration. Intensity profiles were measured along multiple parallel transects, and data points with high variance in measured widths were discarded. A detailed description of this methodology is given Appendix C.5.1.

We found that bands have widened appreciably since 1952 in the human-impacted areas SP1–SP4, while band widths remained approximately constant at the other study areas (Figure 3.5). We computed the ratio of band widths measured in recent imagery

Area	Slope (%) S	Width ratio R	Migration (m/yr) M_c	M_r	corr(S, R) (p)	corr(S,M_c) (p)	corr(S,M_r) (p)
SP1	0.3–0.4	0.95–1.5	0.26–0.68	0.12–0.48	-0.21 ($<$ 0.005)	-0.19 (0.01)	0.01 (0.91)
SP2	0.1–0.3	1.3–2.2	0.57–1.2	0.20–0.55	-0.12 ($<$ 0.005)	-0.1 (0.01)	0.03 (0.43)
SP3	0.1–0.3	1.0–1.7	0.23–0.79	0.08–0.51	-0.1 (0.13)	-0.09 (0.16)	-0.05 (0.47)
SP4	0.1–0.2	1.2–1.9	0.33–0.93	0.08–0.41	-0.12 (0.14)	-0.12 (0.17)	0.00 (0.97)
SP5	0.2–0.4	0.74–1.1	0.17–0.38	0.20–0.43	-0.13 (0.01)	-0.14 ($<$ 0.005)	-0.06 (0.21)
HD1	0.4–0.6	0.93–1.4	0.22–0.55	0.07–0.39	-0.13 (0.16)	-0.32 ($<$ 0.005)	-0.24 (0.01)
HD2	0.3–0.5	0.85–1.2	0.25–0.47	0.21–0.46	0.12 (0.06)	-0.06 (0.31)	-0.16 ($<$ 0.005)
HD3	0.3–0.5	0.87–1.2	0.26–0.48	0.22–0.46	0.03 (0.65)	-0.14 (0.01)	-0.18 ($<$ 0.005)
HD4	0.4–0.5	0.85–1.2	0.19–0.41	0.15–0.39	-0.07 (0.18)	-0.14 (0.01)	-0.16 (0.01)

Table 3.2. Typical slopes, width ratios and band migration rates in each study area. Ranges shown are the 25th and 75th percentiles. The width ratios (R) and upslope colonization and retreat rates, M_c and M_r respectively, were measured between 1952 and c. 2010. Significance of correlations was assessed using a one-tailed t-test corrected for spatial autocorrelation [85], and p values are given.

to the widths in 1952. The median ratio exceeded 1.2 at the human-impacted areas SP1–SP4 (Figure 3.5a). The most substantial widening occurs at SP2, where the median ratio is 1.8. We measured band widths at SP2 using additional images taken in 1967, 2004, 2006, 2011, and 2013. We found that widths did not change between 1952 and 1967, and then nearly doubled between 1967 and 2004 (Figure 3.5b,c). From 2004 onward, band width held approximately constant. Similar analyses over multiple time points at SP1, SP3, and SP4 showed a similar pattern (Appendix C.5.2, Figure C.9). Since recent images were taken in a variety of seasonal and rainfall history conditions, we conclude that the widening observed in SP1–SP4 is not an artifact of seasonality.

Vegetation bands in Africa and North America are reported to migrate uphill over time due to vegetation colonization at the upslope edge of the band and mortality-driven retreat at the downslope edge [58, 69]. In all our study areas, we similarly observed that bands gradually migrate uphill over six decades (Table 3.2). In addition,
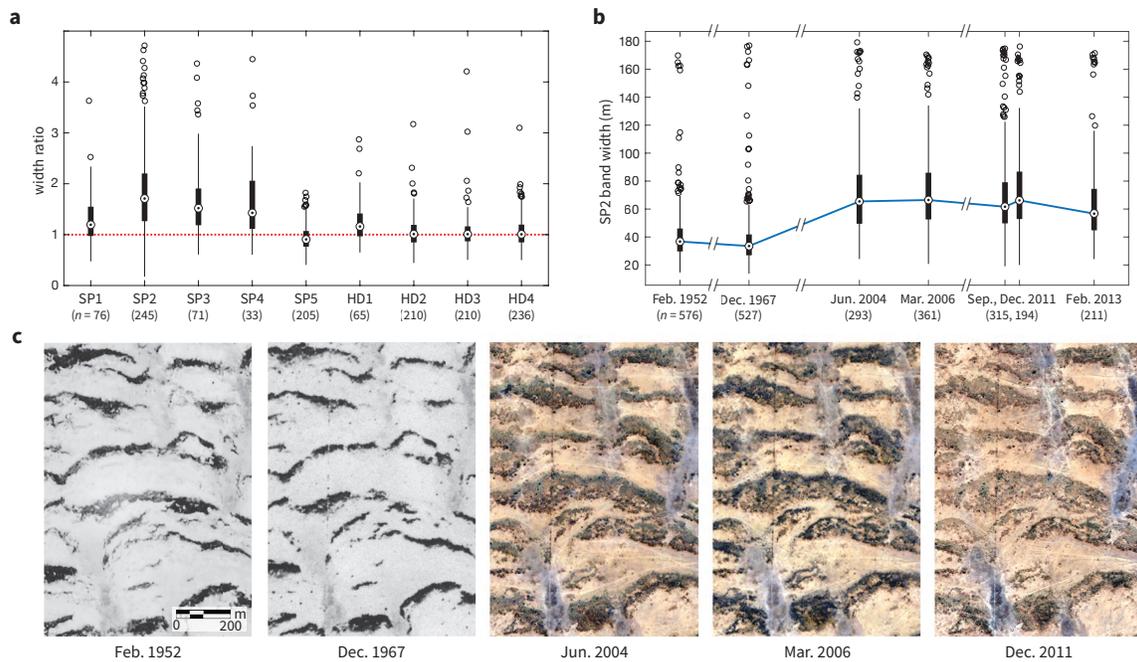
Figure 3.5. Bands widen appreciably in the direction of slope in the most heavily human-impacted areas, SP1–SP4. (a) The ratio of widths measured in a recent image to widths measured in a 1952 photograph are shown for all study areas. (b) The band widths measured in SP2 are shown at six points in time. Widths change little between 1952 and 1967, and nearly double between 1967 and 2004. (c) An example of band widening in SP2 ($9.73°$ N, $48.55°$ E). Satellite images courtesy of the DigitalGlobe Foundation.

we found that bands widen at SP1–SP4 due to increased uphill migration rates at the upslope edges of the bands (colonization rate) and to decreased rates at the downslope edges (retreat rate) (Figure 3.6). Between 1952 and 1967, both colonization and retreat rates are comparable in all Sool Plateau study areas, resulting in band widths that remain unchanged over this period. Between 1967 and c. 2010, colonization rates increase and retreat rates decrease in SP1–SP4, resulting in band widening. In contrast, colonization and retreat rates both decrease in SP5 by the same factor during this period, resulting in slower migration and no widening in this study area.

We found evidence that migration rates and band width ratios vary inversely with local slopes, which are typically shallow and range between 0.1–0.6% (Table 3.2). Colonization rates are negatively correlated with slope in a majority of study areas ($p \leq 0.01$), including the areas in the Sool Plateau where width ratios are also negatively correlated with slope ($p \leq 0.01$). Retreat rates are only negatively correlated with slope in the Haud study areas ($p \leq 0.01$). Since widening in SP1–SP4 occurred due to increased colonization rates, this suggests that bands tend to widen to a greater degree on shallower slopes due to faster rates of colonization. An inverse relationship between slope and migration rate seems to contradict previous investigations of a mathematical model for vegetation banding, which indicate either a negligible [84] or an increasing relationship [86]. Moreover, a positive relationship between migration and slope is expected in the limit of vanishing slope, since the migration rate must approach zero as the anisotropy induced by the slope vanishes. The bands in regions of study may rely critically on the slopes being above some threshold, yet not too steep, and this could explain why we do not observe positive relationships between slope and migration.

3.3.6. Model simulations of band widening

In the previous section, we describe a widespread and persistent increase in band width at the human-impacted areas SP1–SP4. In earlier studies of Niger and Texas, USA, researchers observed sizable fluctuations in band width on sub-decadal time scales in response to rainfall variation, with favorable rainfall conditions resulting in wider bands [61, 69]. In the Sool Plateau, we did not observe appreciable band
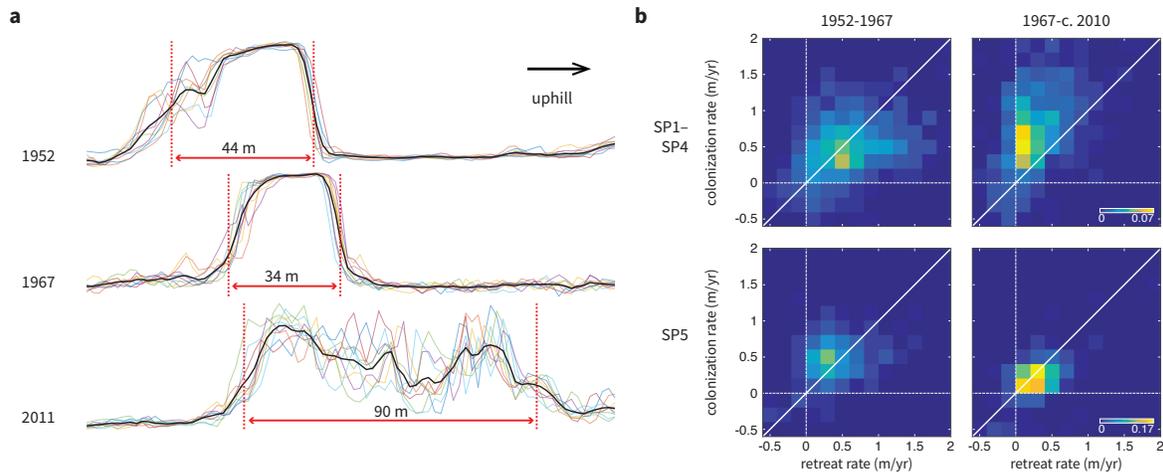
Figure 3.6. Bands widen in areas SP1–SP4 due to increased upslope edge migration (colonization) and decreased downslope edge migration (retreat). (a) Example image intensity profiles along a transect in SP2. Profiles along multiple parallel transects are shown in color, and the mean profile is shown in black. The estimated band edges are indicated in red. (b) Bivariate distribution of front and back migration rates shown for the periods 1952-1967 (first column) and 1967-c. 2010 (second column), with study areas SP1-SP4 (first row) compared in aggregate with SP5 alone (second row). Band widening in SP1–SP4 results from front migration rates increasing and back migration rates decreasing during the period 1967-c. 2010.

width fluctuations in response to multi-year rainfall variation, nor did we find evidence that rainfall conditions have become more favorable in the region over time (Appendix C.1). The fact that widening is localized to Sool Plateau areas SP1–SP4, that it occurs to different degrees in each of these areas, and that it is not observed at nearby area SP5 strongly suggests that non-climatic factors have driven the apparent changes.

We performed a sensitivity analysis of the PDE vegetation model by Klausmeier [9] (K99) to examine how changes in band width can be achieved through factors other

than rainfall. In dimensional form, the model is written as

(3.2)
$$\frac{\partial W}{\partial T} = \underbrace{A}_{\text{precip.}} - \underbrace{LW}_{\text{evap.}} - \underbrace{RWN^2}_{\text{transp.}} + \underbrace{V\frac{\partial W}{\partial X}}_{\text{advec.}},$$

$$\frac{\partial N}{\partial T} = \underbrace{-MN}_{\text{mort.}} + \underbrace{JRWN^2}_{\text{growth}} + \underbrace{D_N\frac{\partial^2 N}{\partial X^2}}_{\text{dispersal}}.$$

The variable N represents vegetation density and W represents ground water density. The spatial terms are in one spatial dimension, and $\partial W/\partial X$ models the downhill movement of water due to runoff and subsurface flow. Descriptions, units, and values of the parameters used are given in Table 3.3. The baseline parameter set we use for K99 is based on the values given in [9]. Since bands in our regions of study are composed of a mix of grasses and shrubs, parameters which are stated in [9] to differ between grasses and trees (M, J, and R) are set at intermediate values so that the spatial scale of banding resembles the scales in our regions of study. Water flow rate V was also approximately tuned so that a comparable time scale of migration is reproduced in the simulations to that of our regions of study. We set the mean annual rainfall parameter to 150 mm, which is within the range of typical rainfall levels in our regions of study. We simulated K99 using the exponential time differencing Runge-Kutta 4 pseudospectral scheme [50, 51], with 2048 grid points, a 1000 m domain, and a time step of 0.1 years.

Acacia cutting for charcoal production has been prevalent in the Sool Plateau since at least the 1980s [14], and has likely caused a decrease in woody biomass within the vegetation bands in many areas. A shift in composition from woody to grass biomass

Parameter	Units	Description	Value
A	mm H_2O yr^{-1}	mean annual rainfall	150
L	yr^{-1}	evaporation rate	4
J	kg m^{-2} (mm H_2O)$^{-1}$	biomass yield per unit H_2O	0.0025
M	yr^{-1}	mortality rate	0.75
R	mm H_2O yr^{-1} (kg dry mass)$^{-2}$	transpiration rate	50
D_N	m^2 yr^{-1}	plant dispersal rate	1
V	yr^{-1}	water flow speed	35

Table 3.3. Baseline parameter set used for sensitivity analysis of K99.

plausibly increases the effective transpiration rate, biomass yield per unit water, dispersal rate, and mortality rate of the vegetation. We found that individually increasing transpiration, yield, and dispersal rates resulted in band widening, while increasing mortality rate had the opposite effect (Figure 3.7). We found that band widening can be achieved through a simultaneous increase in transpiration rate, biomass yield per unit water, dispersal rate, and mortality rate parameters (Appendix C.6, Figure C.10), and conclude that a shift in species composition is a viable explanation for the changes we have observed in the Sool Plateau.

3.3.7. Summary and discussion

In many areas, remarkably little about the vegetation bands has changed. We found no evidence of systematic changes in wavelength despite apparent increases in environmental pressure in the region over the last 65 years (Appendix C.1). Moreover, individual bands remained largely identifiable between images over time, which allowed us to observe modest uphill migration. Our findings suggest that large-wavelength ($\sim$150
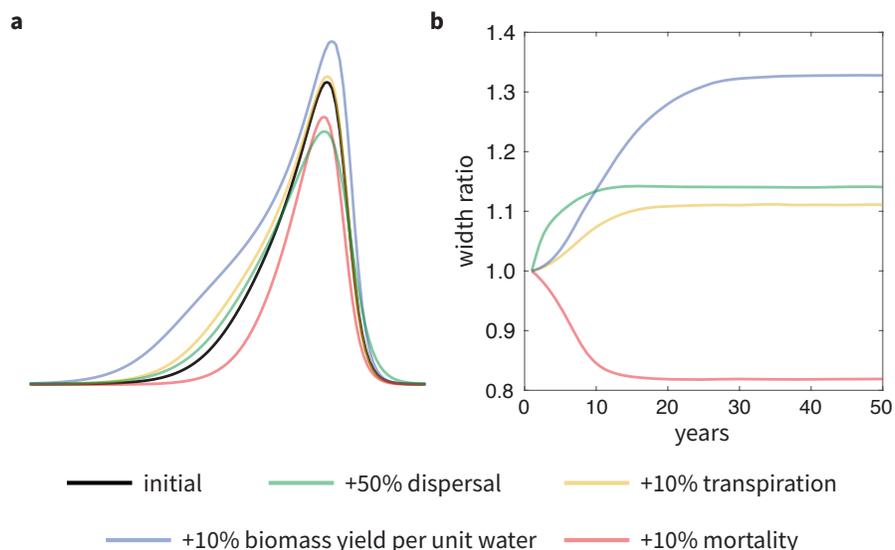
Figure 3.7. Simulations of the model by Klausmeier [9, 87] show that increases in band width can be achieved through parameter increases in plant dispersal rate, transpiration rate, and biomass yield per unit water. Widening occurs continuously on sub-decadal time scale. (a) A comparison of equilibrium band profiles, simulated first using an initial parameter set, and then simulated after applying perturbations to individual parameters. (b) Width ratio computed between perturbed and initial band profiles, plotted as a function of time.

m) bands should migrate appreciably (~5 m) over the span of a decade. Given sufficiently accurate image alignment, migration of this magnitude should be detectable in current high-resolution satellite imagery, and may in principle be brought to bear on relevant model predictions [21, 74]. Additionally our findings suggest a counter-intuitive relationship between migration and slope, one that may be stronger than the relationship between wavelength and slope. Further theoretical and empirical investigation of this relationship may shed light on the connection between topography, hydrology, and the emergent scales of vegetation banding.

In other areas, the most profound changes in vegetation banding are associated with increases in human pressure. This mirrors the findings of previous empirical studies [41, 42, 77], and underscores the importance of human impacts to the resilience of dryland vegetation. Apart from land clearing for development, we observed widespread degradation in areas with few human-made structures, suggesting more subtle but important forms of impact. Roads and dirt tracks now densely cover the landscapes of many of the study areas, and have likely affected the flow and availability of water to the vegetation bands [88]. Biomass harvesting, here in the form of *Acacia* cutting for charcoal production [14], has no doubt also played a role in the degradation we observed. Both changes in road cover and vegetation harvesting can potentially be monitored remotely, and such monitoring efforts should be well-informed by theoretical investigations that account for these and other forms of human impact.

Band wavelength has been a prime focus of theoretical and empirical investigations of vegetation pattern resilience. The dependence of wavelength on model parameters can often be studied analytically, and it is conceived that wavelengths are easily measured in remotely-sensed imagery. However, as we and others [83] have found, the spacing between bands is often quite irregular, making imprecise both the notion of wavelength and its measurement. Moreover, because a range of band wavelengths may be stable over a range of environmental conditions, wavelength changes in model scenarios of environmental change are history-dependent and discontinuous [21, 35, 36, 72]. These shortcomings make wavelength changes a poor signal of ecosystem regime shift, and explain the dearth of evidence for such changes in this and other studies.

We have provided new evidence that vegetation band widths change on observable time scales, and we argue that they represent an underutilized window into the response of dryland vegetation to climatic and human pressure. Our model investigation suggests that band widths change continuously in response to parameter variation, and that these changes should be observable on sub-decadal time scales (Figure 3.7b). Band width is straightforward to measure in remotely-sensed imagery, and comparisons over time need not depend sensitively on the quality of image alignment. Future theoretical investigation of this pattern property will be important to establishing its utility to dryland vegetation monitoring.

CHAPTER 4

Discussion & outlook

Much remains unknown about periodic dryland vegetation patterns. Although mathematical models can qualitatively reproduce the appearance of the patterns, the empirical study in this thesis has brought forward some apparent contradictions with model predictions. In particular, in Chapter 3, we found that the local wavelength of vegetation bands in Somalia is seldom correlated with local slope, while models can predict increasing [33, 84] or decreasing [33] relationships between these quantities. Also we found that the migration rates of bands in Somalia are often inversely correlated with slope, while models indicate either a negligible [84] or an increasing relationship [86]. These findings underscore the importance of quantitative characterization of dryland vegetation patterning. It is possible that our observations could be reconciled with the models through alternative approaches for determining theoretical relationships between slope and wavelength or migration rate. For instance, Sherratt [33] found slope-wavelength relationship to be negative in the model by Klausmeier [9] in the scenario where vegetation bands arise due to colonization of bare soil, and positive when bands emerge from uniform vegetation. It is also possible that important, relevant processes are not represented in the models (e.g., the downhill flow

of plant litter), or are not represented with the necessary realism (e.g., microtopography). Refinement and iteration of the models through comparison with quantitative observations will likely lead to a fuller understanding of patterned dryland vegetation.

The models discussed in this thesis are not ideal tools for many types of exploration. In Chapter 3, we considered a shift in plant species composition of vegetation bands from woody to grassy vegetation through changes in aggregate biomass parameters. In reality, plant types are distributed along the bands inhomogeneously, with more grasses at the colonizing upslope edges of the bands, and more woody vegetation in the mid-band and downslope edges [66, 68]. Models that explicitly account for different plant functional types may provide more nuanced (or even qualitatively different) predictions that can be compared with our observations (see Gilad *et al*. [89] for a modified model in this direction). Hybrid modeling approaches, which represent water fields as continuums and biomass through discrete agents (see Vincenot *et al*. [90]), could also provide a platform for testing scenarios with multiple plant functional types. Hybrid models may also facilitate direct comparisons between field observations and simulations, since micro-scale details of field sites such as the spatial positions of individual plants can be incorporated.

Periodic vegetation patterning has been observed on five continents. Cloud-based resources such as Google Earth Engine will make global-scale studies of vegetation patterning more feasible. Already this resource has been used to map global forest change [91], and to create more accurate inventories of trees in drylands [92]. Google Earth Engine contains a vast store of satellite imagery as well as global elevation data,
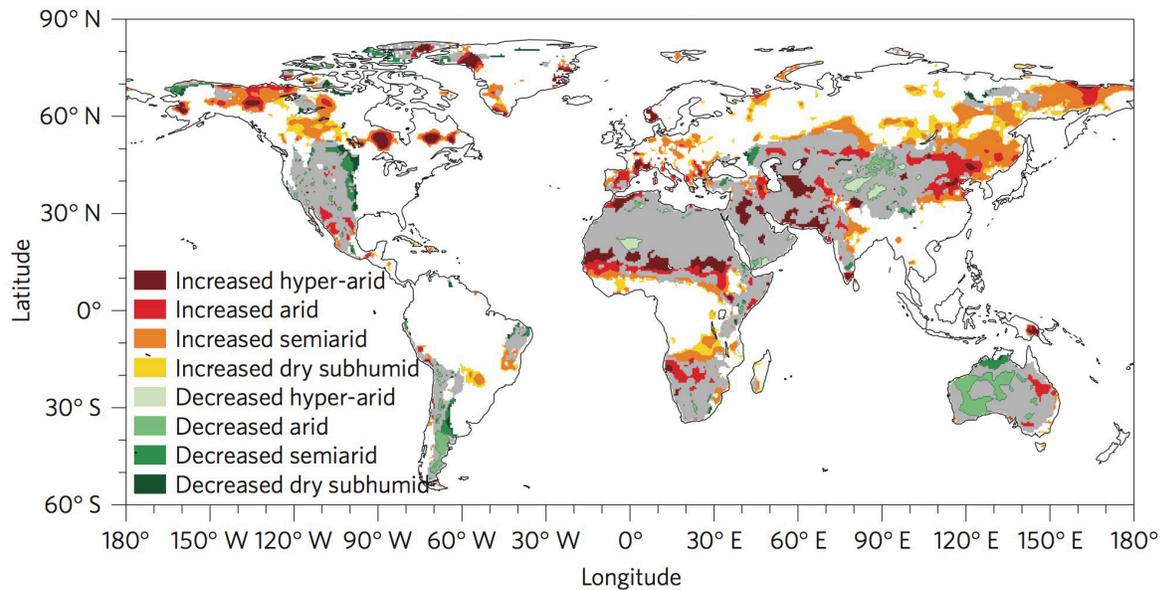
Figure 4.1. Projected changes in the global extent and classification of dry-lands by the end of the century. Taken from [13].

making it well-suited to investigate the role of topography on the morphology and spatial scale of vegetation patterning. How flat are the terrains on which gap, labyrinth, and spot patterns occur? What range of slopes supports vegetation bands, and are there relationships between slope and wavelengths of bands when comparing different parts of the world? These questions may be answered by porting established tools (e.g., Fourier analysis of vegetation bands [83]) to the rich computational environment of Google Earth Engine.

Global drylands are projected to expand and intensify in many parts of the world by the end of the century [13] (Figure 4.1), and such changes are likely to disproportionately impact the poor, who depend strongly on their environments for subsistence [1]. In Chapter 3, we found evidence that human activity exacerbates the degradation of

vegetation in the Sool Plateau. This highlights a potential feedback between climate change and human activity: severe droughts devastated livestock populations in the region, causing human populations to take up destructive biomass harvesting for subsistence [14]. The vegetation loss induced by this harvesting may eventually contribute to increased regional aridity [17]. The integrity of dryland vegetation is of key importance to the capacity of drylands to support life. Efforts to monitor vegetation and anticipate when human activity is destructive may lead to conservation and relief that can mitigate the effects of a changing climate.

References

[1] Millennium Ecosystem Assessment. Ecosystems and human well-being: desertification synthesis. World Resources Institute (2005).

[2] Allen, R. G., Pereira, L. S., Raes, D., Smith, M. *et al.* Crop evapotranspiration-guidelines for computing crop water requirements-fao irrigation and drainage paper 56. *FAO, Rome* **300**, D05109 (1998).

[3] Noy-Meir, I. Desert Ecosystems: Environment and Producers. *Annual Review of Ecology and Systematics* **4**, 25–51 (1973).

[4] Aguiar, M. R. & Sala, O. E. Patch structure, dynamics and implications for the functioning of arid ecosystems. *Trends in Ecology & Evolution* **14**, 273–277 (1999).

[5] Belsky, A. J. Influences of Trees on Savanna Productivity: Tests of Shade, Nutrients, and Tree-Grass Competition. *Ecology* **75**, 922–932 (1994).

[6] Deblauwe, V., Barbier, N., Couteron, P., Lejeune, O. & Bogaert, J. The global biogeography of semi-arid periodic vegetation patterns. *Global Ecology and Biogeography* **17**, 715–723 (2008).

[7] Macfadyen, W. A. Vegetation Patterns in the Semi-Desert Plains of British Somaliland. *The Geographical Journal* **116**, 199 (1950).

[8] Borgogno, F., D'Odorico, P., Laio, F. & Ridolfi, L. Mathematical models of vegetation pattern formation in ecohydrology. *Reviews of Geophysics* **47**, RG1005 (2009).

[9] Klausmeier, C. A. Regular and Irregular Patterns in Semiarid Vegetation. *Science* **284**, 1826–1828 (1999).

[10] Murray, J. D. Mathematical Biology. Springer Berlin Heidelberg, Berlin, Heidelberg (1993).

[11] Cross, M. C. & P. C. Hohenberg. Pattern formation outside of equilibrium. *Reviews of Modern Physics* **65**, 851–1112 (1993).

[12] Maestre, F. T., Salguero-Gomez, R. & Quero, J. L. It is getting hotter in here: determining and projecting the impacts of global environmental change on drylands. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**, 3062–3075 (2012).

[13] Huang, J., Yu, H., Guan, X., Wang, G. & Guo, R. Accelerated dryland expansion under climate?change. *Nature Climate Change* **316**, 847–7 (2015).

[14] Oduori, S. M., Alim, M. S. & Gomes, N. Environmental study of degradation in the Sool Plateau and Gebi Valley: Sanaag Region of Northern Somalia. Tech. Rep., Horn Relief and Oxfam, Netherlands (2003).

[15] Mau, Y., Haim, L. & Meron, E. Reversing desertification as a spatial resonance problem. *Physical Review E* **91**, 012903 (2015).

[16] Rietkerk, M., Dekker, S. C., de Ruiter, P. C. & van de Koppel, J. Self-Organized Patchiness and Catastrophic Shifts in Ecosystems. *Science* **305**, 1926–1929 (2004).

[17] Rietkerk, M. *et al.* Local ecosystem feedbacks and critical transitions in the climate. *Ecological Complexity* **8**, 223–228 (2011).

[18] Scheffer, M. *et al.* Early-warning signals for critical transitions. *Nature* **461**, 53–59 (2009).

[19] Dakos, V. *et al.* Methods for Detecting Early Warnings of Critical Transitions in Time Series Illustrated Using Simulated Ecological Data. *PLOS ONE* **7**, e41010 (2012).

[20] Kéfi, S. *et al.* Spatial vegetation patterns and imminent desertification in Mediterranean arid ecosystems. *Nature* **449**, 213–217 (2007).

[21] Sherratt, J. A. & Lord, G. J. Nonlinear dynamics and pattern bifurcations in a model for vegetation stripes in semi-arid environments. *Theoretical Population Biology* **71**, 1–11 (2007).

[22] Gowda, K., Riecke, H. & Silber, M. Transitions between patterned states in vegetation models for semiarid ecosystems. *Physical Review E* **89**, 022701 (2014).

[23] Gowda, K., Chen, Y., Iams, S. & Silber, M. Assessing the robustness of spatial pattern sequences in a dryland vegetation model. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science* **472**, 20150893 (2016).

[24] Gowda, K., Iams, S. & Silber, M. Dynamics and resilience of vegetation bands in the Horn of Africa. *arXiv.org* (2017). `1705.05308v2`.

[25] von Hardenberg, J., Meron, E., Shachak, M. & Zarmi, Y. Diversity of Vegetation Patterns and Desertification. *Physical Review Letters* **87**, 198101 (2001).

[26] Bel, G., Hagberg, A. & Meron, E. Gradual regime shifts in spatially extended ecosystems. *Theoretical Ecology* **5**, 591–604 (2012).

[27] Dakos, V., Kéfi, S., Rietkerk, M., van Nes, E. H. & Scheffer, M. Slowing Down in Spatially Patterned Ecosystems at the Brink of Collapse. *The American Naturalist* **177**, E153–E166 (2011).

[28] Gilad, E., von Hardenberg, J., Provenzale, A., Shachak, M. & Meron, E. Ecosystem Engineers: From Pattern Formation to Habitat Creation. *Physical Review Letters* **93**, 098105 (2004).

[29] Guttal, V. & Jayaprakash, C. Self-organization and productivity in semi-arid ecosystems: Implications of seasonality in rainfall. *Journal of Theoretical Biology* **248**, 490–500 (2007).

[30] Kéfi, S., Eppinga, M. B., de Ruiter, P. C. & Rietkerk, M. Bistability and regular spatial patterns in arid ecosystems. *Theoretical Ecology* **3**, 257–269 (2010).

[31] Lejeune, O., Tlidi, M. & Couteron, P. Localized vegetation patches: A self-organized response to resource scarcity. *Physical Review E* **66**, 010901 (2002).

[32] Meron, E., Gilad, E., von Hardenberg, J., Shachak, M. & Zarmi, Y. Vegetation patterns along a rainfall gradient. *Chaos, Solitons & Fractals* **19**, 367–376 (2004).

[33] Sherratt, J. A. Using wavelength and slope to infer the historical origin of semiarid vegetation bands. *Proceedings of the National Academy of Sciences* **112**, 4202–4207 (2015).

[34] Zelnik, Y. R., Kinast, S., Yizhaq, H., Bel, G. & Meron, E. Regime shifts in models of dryland vegetation. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **371**, 20120358–20120358 (2013).

[35] van der Stelt, S., Doelman, A., Hek, G. & Rademacher, J. D. M. Rise and Fall of Periodic Patterns for a Generalized Klausmeier–Gray–Scott Model. *Journal of Nonlinear Science* **23**, 39–95 (2012).

[36] Siteur, K. *et al.* Beyond Turing: The response of patterned ecosystems to environmental change. *Ecological Complexity* **20**, 81–96 (2014).

[37] Rietkerk, M. *et al.* Self-Organization of Vegetation in Arid Ecosystems. *The American Naturalist* **160**, 524–530 (2002).

[38] Lejeune, O., Tlidi, M. & Lefever, R. Vegetation spots and stripes: Dissipative structures in arid landscapes. *International Journal of Quantum Chemistry* **98**, 261–271 (2004).

[39] Getzin, S. *et al.* Discovery of fairy circles in Australia supports self-organization theory. *Proceedings of the National Academy of Sciences* 201522130 (2016).

[40] Barbier, N., Couteron, P., Lefever, R., Deblauwe, V. & Lejeune, O. Spatial decoupling of facilitation and competition at hte origin of gapped vegetation patterns. *Ecology* **89**, 1521–1531 (2008).

[41] Barbier, N., Couteron, P., Lejoly, J., Deblauwe, V. & Lejeune, O. Self-organized vegetation patterning as a fingerprint of climate and human impact on semi-arid ecosystems. *Journal of Ecology* **94**, 537–547 (2006).

[42] Deblauwe, V., Couteron, P., Lejeune, O., Bogaert, J. & Barbier, N. Environmental modulation of self-organized periodic vegetation patterns in Sudan. *Ecography* **34**, 990–1001 (2011).

[43] Kealy, B. J. & Wollkind, D. J. A Nonlinear Stability Analysis of Vegetative Turing Pattern Formation for an Interaction–Diffusion Plant-Surface Water Model System in an Arid Flat Environment. *Bulletin of Mathematical Biology* **74**, 803–833 (2011).

[44] Golubitsky, M., Stewart, I. & Schaeffer, D. G. *Singularities and Groups in Bifurcation Theory* (Springer Verlag, 2012).

[45] Hoyle, R. *Pattern Formation: An Introduction to Methods* (Cambridge University Press, 2006), 1 edn.

[46] Judd, S. L. & Silber, M. Simple and superlattice Turing patterns in reaction–diffusion systems: bifurcation, bistability, and parameter collapse. *Physica D: Nonlinear Phenomena* **136**, 45–65 (2000).

[47] Gilad, E., von Hardenberg, J., Provenzale, A., Shachak, M. & Meron, E. A mathematical model of plants as ecosystem engineers. *Journal of Theoretical Biology* **244**, 680–691 (2007).

[48] Turing, A. M. The Chemical Basis of Morphogenesis. *Philosophical Transactions of the Royal Society B: Biological Sciences* **237**, 37–72 (1952).

[49] Wiggins, S. *Introduction to Applied Nonlinear Dynamical Systems and Chaos* (Springer, 2003).

[50] Cox, S. M. & Matthews, P. C. Exponential Time Differencing for Stiff Systems. *Journal of Computational Physics* **176**, 430–455 (2002).

[51] Kassam, A.-K. & Trefethen, L. N. Fourth-Order Time-Stepping for Stiff PDEs. *SIAM Journal on Scientific Computing* **26**, 1214–1233 (2005).

[52] Kassam, A. K. Solving reaction-diffusion equations 10 times faster (2003).

[53] Vanag, V. K. Waves and patterns in reaction–diffusion systems. Belousov–Zhabotinsky reaction in water-in-oil microemulsions. *Phys. Usp.* **47**, 923–941 (2007).

[54] Malchow, H., Hilker, F. M., Petrovskii, S. V. & Brauer, K. Oscillations and waves in a virally infected plankton system. *Ecological Complexity* **1**, 211–223 (2004).

[55] Zelnik, Y. R., Meron, E. & Bel, G. Gradual regime shifts in fairy circles. *Proceedings of the National Academy of Sciences* **112**, 12327–12331 (2015).

[56] Hillel, D. *Introduction to Soil Physics* (Academic Press, 2013).

[57] Getzin, S. *et al.* Adopting a spatially explicit perspective to study the mysterious fairy circles of Namibia. *Ecography* **38**, 1–11 (2014).

[58] Valentin, C., d'Herbes, J. M. & Poesen, J. Soil and water components of banded vegetation patterns. *Catena* **37**, 1–24 (1999).

[59] Ludwig, J. A., Tongway, D. J. & Marsden, S. G. Stripes, strands or stipples: modelling the influence of three landscape banding patterns on resource capture and productivity in semi-arid woodlands, Australia. *Catena* **37**, 257–273 (1999).

[60] Gomes, N. Access to water, pastoral resource management and pastoralists' livelihoods: Lessons learned from water development in selected areas of Eastern Africa (Kenya, Ethiopia, Somalia). Tech. Rep., Livelihood Support Programme (LSP) (2006).

[61] Valentin, C. & d'Herbes, J. M. Niger tiger bush as a natural water harvesting system. *Catena* **37**, 231–256 (1999).

[62] Ludwig, J. A., Wilcox, B. P., Breshears, D. D., Tongway, D. J. & Imeson, A. C. Vegetation patches and runoff-erosion as interacting ecohydrological processes in semiarid landscapes. *Ecology* **86**, 288–297 (2005).

[63] Macfadyen, W. A. Soil and Vegetation in British Somaliland. *Nature* **165**, 121–121 (1950).

[64] Greenwood, J. E. G. W. The Development of Vegetation Patterns in Somaliland Protectorate. *The Geographical Journal* **123**, 465 (1957).

[65] Boaler, S. B. & Hodge, C. A. H. Observations on Vegetation Arcs in the Northern Region, Somali Republic. *The Journal of Ecology* **52**, 511 (1964).

[66] Hemming, C. F. Vegetation Arcs in Somaliland. *The Journal of Ecology* **53**, 57 (1965).

[67] Janeau, J. L., Mauchamp, A. & Tarin, G. The soil surface characteristics of vegetation stripes in Northern Mexico and their influences on the system hydrodynamics. *Catena* **37**, 165–173 (1999).

[68] Thiery, J. M., d'Herbes, J. M. & Valentin, C. A Model Simulating the Genesis of Banded Vegetation Patterns in Niger. *The Journal of Ecology* **83**, 497 (1995).

[69] Deblauwe, V., Couteron, P., Bogaert, J. & Barbier, N. Determinants and dynamics of banded vegetation pattern migration in arid climates. *Ecological Monographs* **82**, 3–21 (2012).

[70] Lefever, R. & Lejeune, O. On the origin of tiger bush. *Bulletin of Mathematical Biology* **59**, 263–294 (1997).

[71] Yizhaq, H., Gilad, E. & Meron, E. Banded vegetation: biological productivity and resilience. *Physica A: Statistical Mechanics and its Applications* **356**, 139–144 (2005).

[72] Sherratt, J. A. History-dependent patterns of whole ecosystems. *Ecological Complexity* **14**, 8–20 (2013).

[73] Sherratt, J. A. Pattern solutions of the Klausmeier model for banded vegetation in semi-arid environments II: patterns with the largest possible propagation speeds. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science* **467**, 3272–3294 (2011).

[74] Dagbovie, A. S. & Sherratt, J. A. Pattern selection and hysteresis in the Rietkerk model for banded vegetation in semi-arid environments. *Journal of The Royal Society Interface* **11**, 20140465–20140465 (2014).

[75] Siero, E. *et al.* Striped pattern selection by advective reaction-diffusion systems: Resilience of banded vegetation on slopes. *Chaos: An Interdisciplinary Journal of*

Nonlinear Science **25**, 036411 (2015).

[76] HilleRisLambers, R., Rietkerk, M., van den Bosch, F., Prins, H. H. T. & de Kroon, H. Vegetation Pattern Formation in Semi-Arid Grazing Systems. *Ecology* **82**, 50 (2001).

[77] Wu, X. B., Thurow, T. L. & Whisenant, S. G. Fragmentation and changes in hydrologic function of tiger bush landscapes, south-west Niger. *Journal of Ecology* **88**, 790–800 (2000).

[78] Muchiri, P. W. Climate of Somalia. Tech. Rep. W-01, FAO-SWALIM, Nairobi, Kenya (2007).

[79] Compo, G. P. *et al.* The Twentieth Century Reanalysis Project. *Quarterly Journal of the Royal Meteorological Society* **137**, 1–28 (2011).

[80] Huete, A. R. A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment* **25**, 295–309 (1988).

[81] Farr, T. G. *et al.* The Shuttle Radar Topography Mission. *Reviews of Geophysics* **45**, 1485–33 (2007).

[82] Holoborodko, P. Noise robust gradient operators. `http://www.holoborodko.com/pavel/image-processing/edge-detection/` (2009).

[83] Penny, G. G., Daniels, K. E. & Thompson, S. E. Local properties of patterned vegetation: quantifying endogenous and exogenous effects. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **371**, 20120359–20120359 (2013).

[84] Sherratt, J. A. An Analysis of Vegetation Stripe Formation in Semi-Arid Landscapes. *Journal of Mathematical Biology* **51**, 183–197 (2005).

[85] Dutilleul, P., Clifford, P., Richardson, S. & Hemon, D. Modifying the t Test for Assessing the Correlation Between Two Spatial Processes. *Biometrics* **49**, 305 (1993).

[86] Sherratt, J. A. Pattern solutions of the Klausmeier Model for banded vegetation in semi-arid environments I. *Nonlinearity* **23**, 2657–2675 (2010).

[87] Ursino, N. The influence of soil properties on the formation of unstable vegetation patterns on hillsides of semiarid catchments. *Advances in Water Resources* **28**, 956–963 (2005).

[88] Hemming, C. F. The vegetation of the northern region of the Somali Republic. *Proceedings of the Linnean Society of London* **177**, 173–250 (1966).

[89] Gilad, E., Shachak, M. & Meron, E. Dynamics and spatial organization of plant communities in water-limited systems. *Theoretical Population Biology* **72**, 214–230 (2007).

[90] Vincenot, C. E., Carteni, F., Mazzoleni, S., Rietkerk, M. & Giannino, F. Spatial Self-Organization of Vegetation Subject to Climatic Stress—Insights from a System Dynamics—Individual-Based Hybrid Model. *Frontiers in Plant Science* **7**, 93–18 (2016).

[91] Hansen, M. C. *et al.* High-resolution global maps of 21st-century forest cover change. *Science* **342**, 850–853 (2013).

[92] Bastin, J.-F. *et al.* The extent of forest in dryland biomes. *Science* **356**, 635–638 (2017).

[93] Cross, M. & Greenside, H. *Pattern Formation and Dynamics in Nonequilibrium Systems* (Cambridge University Press, 2009), 1 edn.

[94] Dawes, J. H. P. & Williams, J. L. M. Localised pattern formation in a model for dryland vegetation. *Journal of Mathematical Biology* **73**, 63–90 (2015).

[95] Osorio, F., Vallejos, R. & Cuevas, F. SpatialPack: Computing the Association Between Two Spatial Processes. *arXiv.org* (2016). `1611.05289v1`.

APPENDIX A

Quadratic coefficient analysis of model by Rietkerk *et al.*

A.1. Quadratic coefficient calculation

In this appendix, we summarize the calculation of the quadratic coefficient, a, of the amplitude equations (2.6) for the Rietkerk *et al.* [37] model (R02). We do so in a manner that illustrates the role played by the nonlinear functions in R02.

We expand R02 (2.13) to quadratic order about the uniform vegetated steady state (h_0, w_0, n_0), which is a function of precipitation p given by (2.14). We take $H = h - h_0$, $W = w - w_0$, $N = n - n_0$:

$$
\text{(A.1)} \qquad \frac{\partial}{\partial t} \begin{pmatrix} H \\ W \\ N \end{pmatrix} = \mathcal{L} \begin{pmatrix} H \\ W \\ N \end{pmatrix} + \begin{pmatrix} -I_2(H,N) \\ I_2(H,N) - \gamma G_2(W,N) \\ G_2(W,N) \end{pmatrix} + \ldots.
$$

The linear operator $\mathcal{L}$ is

$$
\text{(A.2)} \qquad \mathcal{L} = \begin{pmatrix} -I(n_0) + D_h \nabla^2 & 0 & -I'(n_0)h_0 \\ I(n_0) & -\nu - \gamma G'(w_0)n_0 + D_w \nabla^2 & I'(n_0)h_0 - \gamma\mu \\ 0 & G'(w_0)n_0 & \nabla^2 \end{pmatrix},
$$

the quadratic order terms $I_2(H, N)$ and $G_2(W, N)$ are

(A.3) $$I_2(H, N) = I'(n_0)HN + \frac{1}{2}I''(n_0)h_0 N^2,$$

(A.4) $$G_2(W, N) = G'(w_0)WN + \frac{1}{2}G''(w_0)n_0 W^2,$$

and the ellipsis denotes terms of cubic order in (H, W, N). The linear stability of the uniform vegetated state to spatially periodic perturbations with wave number q is determined by substituting $(H, W, N)^T = \boldsymbol{\xi}e^{iqx}e^{\sigma t}$ into equation (A.1) linearized about $H = W = N = 0$. This gives the eigenvalue problem $\sigma\boldsymbol{\xi} = J(q^2, p)\boldsymbol{\xi}$, where $J(q^2, p)$ is the Jacobian matrix

(A.5) $$J(q^2, p) = \begin{pmatrix} -I(n_0) - D_h q^2 & 0 & -I'(n_0)h_0 \\ I(n_0) & -\nu - \gamma G'(w_0)n_0 - D_w q^2 & I'(n_0)h_0 - \gamma\mu \\ 0 & G'(w_0)n_0 & -q^2 \end{pmatrix}.$$

The Jacobian matrix depends explicitly on the wave number q, as well as on the precipitation parameter p through the uniform vegetated steady state (h_0, w_0, n_0). We consider p to be the bifurcation parameter in this analysis. A Turing point occurs at a parameter value $p = p_c$ and wave number $q^2 = q_c^2 > 0$ for which the maximum real part of an eigenvalue is zero, and all other modes are damped. Necessary conditions for a Turing point (q_c^2, p_c) are given by

(A.6) $$\mathrm{Det}(J(q_c^2, p_c)) = 0,$$

(A.7) $$\left.\frac{\partial\mathrm{Det}(J(q^2, p_c))}{\partial q^2}\right|_{q=q_c} = 0.$$

We follow a standard procedure [45, 46, 93] to obtain the quadratic coefficient a of the amplitude equations (2.6). We write a small amplitude hexagonal (spots/gaps) solution to (A.1) near (q_c^2, p_c) as

(A.8)
$$\begin{pmatrix} H \\ W \\ N \end{pmatrix} = \epsilon \left(z(t_1) f(\boldsymbol{x}) + c.c. \right) \begin{pmatrix} H_1 \\ W_1 \\ N_1 \end{pmatrix} + \epsilon^2 \boldsymbol{U_2} + \mathcal{O}(\epsilon^3),$$

where $\epsilon \ll 1$, $t_1 = \epsilon t$,

$$f(\boldsymbol{x}) = e^{i\boldsymbol{q_1} \cdot \boldsymbol{x}} + e^{i\boldsymbol{q_2} \cdot \boldsymbol{x}} + e^{i\boldsymbol{q_3} \cdot \boldsymbol{x}},$$

$(H_1, W_1, N_1)^T$ is to be determined, and $\boldsymbol{U_2}$ is a higher-order term which must be bounded. The wave vectors lie on a 2D hexagonal lattice, with $\boldsymbol{q_1} = q_c(1, 0)$, $\boldsymbol{q_2} = q_c(-1/2, \sqrt{3}/2)$, $\boldsymbol{q_3} = -(\boldsymbol{q_1} + \boldsymbol{q_2})$. Plugging (A.8) into (A.1) gives the $\mathcal{O}(\epsilon)$ equation

$$0 = J(q_c^2, p_c) \begin{pmatrix} H_1 \\ W_1 \\ N_1 \end{pmatrix}.$$

Since $J(q_c^2, p_c)$ has a zero eigenvalue, we take $(H_1, W_1, N_1)^T$ to be the associated right null vector to satisfy this equation. We choose the convention that $H_1^2 + W_1^2 + N_1^2 = 1$ and $N_1 > 0$.

At $\mathcal{O}(\epsilon^2)$, we have

$$
f(\boldsymbol{x})\frac{\partial z}{\partial t_1}\begin{pmatrix} H_1 \\ N_1 \\ W_1 \end{pmatrix} = \mathcal{L}\boldsymbol{U_2} + \begin{pmatrix} -I_2(H_1, N_1) \\ I_2(H_1, N_1) - \gamma G_2(W_1, N_1) \\ G_2(W_1, N_1) \end{pmatrix} (z(t_1)f(\boldsymbol{x}) + c.c.)^2 .
$$

The term $(z(t_1)f(\boldsymbol{x}) + c.c.)^2$ generates modes with wave vectors of magnitude q_c:

$$
(z(t_1)f(\boldsymbol{x}) + c.c.)^2 = 2\bar{z}^2 \left(e^{i\boldsymbol{q_1}\cdot\boldsymbol{x}} + e^{i\boldsymbol{q_2}\cdot\boldsymbol{x}} + e^{i\boldsymbol{q_3}\cdot\boldsymbol{x}} \right) + c.c. + ...
$$

These modes result in secular terms in the solution (A.8). To eliminate these terms, we apply the Fredholm alternative theorem to obtain the solvability condition $\partial z/\partial t_1 = a\bar{z}^2$, where

(A.9) $$\qquad a = I_2(N_1, H_1)\left(\widetilde{W}_1 - \widetilde{H}_1\right) + G_2(N_1, W_1)\left(\widetilde{N}_1 - \gamma\widetilde{W}_1\right),$$

and $(\widetilde{H}_1, \widetilde{W}_1, \widetilde{N}_1)$ is the left null vector of $J(q_c^2, p_c)$. We choose the convention $\widetilde{H}_1 H_1 + \widetilde{W}_1 W_1 + \widetilde{N}_1 N_1 = 2$, which eliminates an overall factor of 2 in (A.9).

A.2. Turing point calculation and scaling

In this appendix, we derive approximations for the critical wave number and the onset parameter value at the upper and lower Turing points, which give scaling relations with respect to the quantity $\delta \equiv I(n_0)/D_h$, where $I(n)$ is the R02 infiltration function, n_0 is the uniform vegetated equilibrium biomass, and D_h is the nondimensional surface water diffusion parameter.

We introduce the abbreviations $I_0 \equiv I(n_0)$, $\hat{h}_0 \equiv I'(n_0)h_0$, $\hat{n}_0 \equiv G'(w_0)n_0$, where (h_0, w_0, n_0) is the uniform vegetated steady state given in (2.14). In the following analysis, we will treat $\hat{n}_0$ as the bifurcation parameter, which is justified by the following observations regarding relationships between $\hat{n}_0$, $\hat{h}_0$, and the precipitation parameter p:

I. $\hat{n}_0$ is a linear increasing function of p. Specifically, $n_0 = (p - p_0)/\gamma\mu$ where $p_0 \equiv \nu\mu/(1 - \mu)$, and $\hat{n}_0$ is proportional to n_0 by the constant factor $G'(w_0) = (1 - \mu)^2$ (w_0 depends only on the parameter μ).

II. $\hat{h}_0$ can be expressed as a function of $\hat{n}_0$, since $\hat{h}_0 \equiv I'(n_0)h_0 = p(n_0)I'(n_0)/I(n_0)$.

III. We will show that Turing points are confined to regions of the parameter space where $\hat{h}_0 > \gamma\mu$ (see (A.12)).

IV. $\hat{h}_0$ is a monotonic decreasing function of $\hat{n}_0$ when $\hat{h}_0 > \gamma\mu$. This follows from

$$\frac{d\hat{h}_0}{d\hat{n}_0} = \frac{dn_0}{d\hat{n}_0}\frac{d\hat{h}_0}{dn_0} = \frac{1}{G'(w_0)}\frac{d\hat{h}_0}{dn_0} = \frac{1}{(1-\mu)^2}\left(I''(n_0)h_0 + I'(n_0)\frac{dh_0}{dn_0}\right),$$

and

$$\frac{dh_0}{dn_0} = \frac{d}{dn_0}\frac{p(n_0)}{I(n_0)} = p'(n_0)\frac{1}{I(n_0)} - I'(n_0)\frac{p}{I(n_0)^2}$$

$$= \frac{1}{I(n_0)}\left(\gamma\mu - I'(n_0)h_0\right) = -\frac{\hat{h}_0 - \gamma\mu}{I(n_0)} < 0 \text{ for } \hat{h}_0 > \gamma\mu.$$

Since $I(n)$ is a concave-down function (i.e. $I''(n_0) < 0$), it follows that $d\hat{h}_0/d\hat{n}_0 < 0$. Together, these relationships give a one-to-one correspondence between $\hat{n}_0$, $\hat{h}_0$, and precipitation p in the parameter regime relevant for this analysis. We note that if $\hat{h}_0 < \gamma\mu$ at the lower boundary of the domain, $\hat{n}_0 = 0$, it can never exceed $\gamma\mu$, and no Turing

points exist. For Turing points to exist, it must be the case that $\hat{h}_0 > \gamma\mu$ at $\hat{n}_0 = 0$. This gives a necessary condition for a Turing bifurcation in R02:

(A.10)
$$\hat{h}_0(0) = \frac{\nu\mu(1-f)}{f(1-\mu)} > \gamma\mu.$$

This condition explains the asymptotic approach of the degeneracy curves in Figures 2.12-2.15 to $f = \nu/(\nu + \gamma(1-\mu)) = 8/9$ as α decreases or D_h increases.

We write the Jacobian matrix (A.5) as

(A.11)
$$J(q^2, \hat{n}_0) = \begin{pmatrix} -I_0 - D_h q^2 & 0 & -\hat{h}_0 \\ I_0 & -\gamma\hat{n}_0 - \nu - D_w q^2 & \hat{h}_0 - \gamma\mu \\ 0 & \hat{n}_0 & -q^2 \end{pmatrix}.$$

The determinant of this matrix is

$$-D_h D_w q^6 - (I_0 D_w + \gamma D_h(\hat{n}_0 + \nu)) q^4 - \left(I_0(\gamma\hat{n}_0 + \nu) - D_h(\hat{h}_0 - \gamma\mu)\hat{n}_0\right) q^2 - I_0\gamma\mu\hat{n}_0.$$

The zero-eigenvalue condition (A.6) and the onset condition (A.7) are necessary for a Turing point $(q_c^2, \hat{n}_0^c)$, and result in the equations

(A.12) $$D_w q_c^6 + (\delta D_w + \gamma\hat{n}_0^c + \nu) q_c^4 + \left(\delta(\gamma\hat{n}_0^c + \nu) - (\hat{h}_0^c - \gamma\mu)\hat{n}_0^c\right) q_c^2 + \delta\gamma\mu\hat{n}_0^c = 0,$$

(A.13) $$3D_w q_c^4 + 2(\delta D_w + \gamma\hat{n}_0^c + \nu) q_c^2 + \delta(\gamma\hat{n}_0^c + \nu) - (\hat{h}_0^c - \gamma\mu)\hat{n}_0^c = 0,$$

where

$$\delta \equiv \frac{I_0}{D_h} \text{ and } \hat{h}_0^c \equiv \hat{h}_0(\hat{n}_0^c).$$

We treat δ as a small parameter by assuming $D_h \gg 1$, noting that I_0 is bounded away from zero ($I_0 \in (\alpha f, \alpha)$). The only possible negative term in (A.12) and (A.13) comes from the factor $-(\hat{h}_0^c - \gamma\mu)$, and so Turing points are confined to regions of the parameter space where $\hat{h}_0 > \gamma\mu$ (as anticipated above in observation III). We note that these equations depend on α and D_h only through the quantity $\delta \propto \alpha/D_h$. Therefore the Turing point calculation is invariant to fixed ratios of α/D_h.

We proceed by seeking approximate Turing point solutions $(q_c^2, \hat{n}_0^c)$ to (A.12) and (A.13) in the form of asymptotic expansions in δ. At $\mathcal{O}(1)$ in δ, (A.12) and (A.13) are

$$D_w q_c^6 + (\gamma\hat{n}_0^c + \nu)q_c^4 - (\hat{h}_0^c - \gamma\mu)\hat{n}_0^c q_c^2 = 0,$$

$$3D_w q_c^4 + 2(\gamma\hat{n}_0^c + \nu)q_c^2 - (\hat{h}_0^c - \gamma\mu)\hat{n}_0^c = 0.$$

The only real-valued solution in q_c^2 is $q_c^2 = 0$, which yields $(\hat{h}_0^c - \gamma\mu)\hat{n}_0^c = 0$. It can be shown that the equation $(\hat{h}_0^c - \gamma\mu)\hat{n}_0^c = 0$ has only two unique solutions, $\hat{n}_0^c = 0$ and $\hat{h}_0^c = \gamma\mu$. Thus there are two solutions at this order: (i) $q_\ell^2 \equiv q_c^2 = 0$ and $\hat{n}_0^\ell \equiv \hat{n}_0^c = 0$, which corresponds to the lower Turing point, and (ii) $q_u^2 \equiv q_c^2 = 0$ and $\hat{h}_0^u - \gamma\mu = 0$ where $\hat{h}_0^u \equiv \hat{h}_0^c = \gamma\mu$, which corresponds to the upper Turing point. We calculate corrections to these solutions, which yield distinct scalings for the wave numbers and onset parameter values with δ at the different Turing points.

For the lower Turing point solution, we assume the leading-order correction takes the form $(q_\ell^2, \hat{n}_0^\ell) = (Q_\ell^2 \delta^{\beta_1}, N_\ell \delta^{\beta_2})$, and we seek values of β_1 and β_2 which achieve a balance between terms in (A.12) and (A.13). For this correction, equations (A.12)

and (A.13) become

$$\delta^{3\beta_1} D_w Q_\ell^6 + \left(\delta^{2\beta_1+1} D_w + \delta^{2\beta_1+\beta_2} \gamma N_\ell + \delta^{2\beta_1} \nu \right) Q_\ell^4$$

$$+ \left(\delta^{\beta_1+\beta_2+1} \gamma N_\ell + \delta^{\beta_1+1} \nu - \delta^{\beta_1+\beta_2} (\hat{h}_0^\ell - \gamma\mu) N_\ell \right) Q_\ell^2 + \delta^{\beta_2+1} \gamma\mu N_\ell = 0,$$

$$3\delta^{3\beta_1} D_w Q_\ell^4 + 2 \left(\delta^{2\beta_1+1} D_w + \delta^{2\beta_1+\beta_2} \gamma N_\ell + \delta^{2\beta_1} \nu \right) Q_\ell^2$$

$$+ \delta^{\beta_1+\beta_2+1} \gamma N_\ell + \delta^{\beta_1+1} \nu - \delta^{\beta_1+\beta_2} (\hat{h}_0^\ell - \gamma\mu) N_\ell = 0.$$

We observe that

- $\delta^{3\beta_1} D_w Q_\ell^6$, $\delta^{2\beta_1+1} D_w Q_\ell^4$ and $\delta^{2\beta_1+\beta_2} \gamma N_\ell Q_\ell^4$ are always higher order terms than $\delta^{2\beta_1} \nu Q_\ell^4$.

- $-\delta^{\beta_1+\beta_2} (\hat{h}_0^\ell - \gamma\mu) N_\ell Q_\ell^2$ must appear at leading order for real solutions in Q_ℓ^2.

- $\delta^{\beta_1+\beta_2+1} \gamma N_\ell Q_\ell^2$ is always a higher order term than $-\delta^{\beta_1+\beta_2} (\hat{h}_0^\ell - \gamma\mu) N_\ell Q_\ell^2$.

- $\delta^{\beta_2+1} \gamma\mu N_\ell$ must appear at leading order for solutions with $Q_\ell^2, \ N_\ell > 0$.

A balance is achieved by $\beta_1 = \beta_2 = 1$. Thus $(q_\ell^2, \hat{n}_0^\ell) = (Q_\ell^2 \delta + \mathcal{O}(\delta^2), N_\ell \delta + \mathcal{O}(\delta^2))$, and Q_ℓ^2 and N_ℓ are solutions to the equations at $\mathcal{O}(\delta^2)$:

(A.14)

$$\nu Q_\ell^4 + \left(\nu - (\hat{h}_0^\ell - \gamma\mu) N_\ell \right) Q_\ell^2 + \gamma\mu N_\ell = 0,$$

$$2\nu Q_\ell^2 + \nu - (\hat{h}_0^\ell - \gamma\mu) N_\ell = 0,$$

where $\hat{h}_0^\ell = \hat{h}_0(N_\ell \delta) = \nu\mu(1-f)/(f(1-\mu)) + \mathcal{O}(\delta)$. It can be shown that a unique physical solution to this system (i.e. $Q_\ell^2, \ N_\ell > 0$) exists when the necessary condition (A.10) is satisfied. See also Dawes *et al.* [94] for a comparable scaling of the lower Turing point

critical wave number with the nondimensional water diffusion parameter in the model by von Hardenberg *et al.* [25].

For the upper Turing point solution, we assume the leading-order correction takes the form $q_u^2 = Q_u^2 \delta^{\beta_3}$, $\hat{h}_0^u - \gamma\mu = H_u \delta^{\beta_4}$. Invoking a similar argument as for the lower Turing point correction, we balance terms of order $\delta^{2\beta_3}$, $\delta^{\beta_3+\beta_4}$ and δ and find $\beta_3 = \beta_4 = 1/2$. Thus $(q_u^2, \hat{h}_0^u - \gamma\mu) = (Q_u^2 \delta^{1/2} + \mathcal{O}(\delta), H_u \delta^{1/2} + \mathcal{O}(\delta))$, and Q_u^2 and H_u are solutions to the leading order equations at $\mathcal{O}(\delta)$:

(A.15)
$$\left(\gamma \hat{n}_0^u + \nu\right) Q_u^4 - \hat{n}_0^u H_u Q_u^2 + \gamma\mu \hat{n}_0^u = 0,$$
$$2\left(\gamma \hat{n}_0^u + \nu\right) Q_u^2 - \hat{n}_0^u H_u = 0,$$

where $\hat{h}_0(\hat{n}_0^u) = \gamma\mu + H_u \delta^{1/2}$. It can be shown that a unique solution to this system, with $Q_u^2 > 0$, always exists.

A.3. Quadratic coefficient analysis

In Appendix A.2, we derived scaling relations between $\delta \equiv I(n_0)/D_h$, the critical wave numbers q_c and the Turing point parameter values in terms of $\hat{n}_0 \equiv G'(w_0)n_0$. In this appendix, we use these scaling relations in an analysis of the terms that are important in setting the sign of a when δ is small.

The leading order scaling behavior of the quadratic coefficient a is determined by the scaling of the right and left null vectors, $(H_1, W_1, N_1)^T$ and $(\widetilde{H}_1, \widetilde{W}_1, \widetilde{N}_1)$ respectively. We will use the relations above to derive the null vector scalings. We recall that a can

be written as

$$\text{(A.16)} \quad a = \left(I'(n_0)H_1N_1 + \frac{1}{2}I''(n_0)h_0N_1^2 \right)\left(\widetilde{W}_1 - \widetilde{H}_1 \right)$$
$$+ \left(G'(w_0)W_1N_1 + \frac{1}{2}G''(w_0)n_0W_1^2 \right)\left(\widetilde{N}_1 - \gamma\widetilde{W}_1 \right).$$

We obtain the following relations between the right null vector components from the first and third rows of the Jacobian matrix (A.5):

$$\left(D_h q_c^2 + I(n_0) \right)H_1 + I'(n_0)h_0N_1 = 0,$$

$$G'(w_0)n_0W_1 - q_c^2N_1 = 0.$$

In addition, we recall the right null vector convention $H_1^2 + W_1^2 + N_1^2 = 1$ with $N_1 > 0$. Together, these equations give

$$H_1 = -\frac{I'(n_0)h_0}{D_h q_c^2 + I(n_0)}N_1,$$

$$\text{(A.17)} \qquad W_1 = \frac{q_c^2}{G'(w_0)n_0}N_1,$$

$$N_1 = \left(1 + \frac{q_c^4}{G'(w_0)^2n_0^2} + \frac{I'(n_0)^2h_0^2}{(D_h q_c^2 + I(n_0))^2} \right)^{-1/2}.$$

Similarly we obtain the following relations between the left null vector components from the first and second columns of (A.5):

$$\left(D_h q_c^2 + I(n_0) \right)\widetilde{H}_1 - I(n_0)\widetilde{W}_1 = 0,$$

$$\left(\nu + D_w q_c^2 + \gamma G'(w_0)n_0 \right)\widetilde{N}_1 + G'(w_0)n_0\widetilde{W}_1 = 0.$$

Together with the left null vector convention $\widetilde{H}_1 H_1 + \widetilde{W}_1 W_1 + \widetilde{N}_1 N_1 = 2$, we find

(A.18)

$$\widetilde{H}_1 = \frac{I(n_0)}{D_h q_c^2 + I(n_0)} \widetilde{W}_1,$$

$$\widetilde{W}_1 = 2N_1^{-1} \left(\frac{-I(n_0)I'(n_0)h_0}{(D_h q_c^2 + I(n_0))^2} + \frac{\nu + (1 + D_w)q_c^2}{G'(w_0)n_0} + \gamma \right)^{-1},$$

$$\widetilde{N}_1 = \left(\frac{\nu + D_w q_c^2}{G'(w_0)n_0} + \gamma \right) \widetilde{W}_1.$$

At the lower Turing point, we found that $q_\ell^2 \equiv q_c^2 = Q_\ell^2 \delta + \mathcal{O}(\delta^2)$ and $G'(w_0)n_0 = N_\ell \delta + \mathcal{O}(\delta^2)$, where Q_ℓ^2 and N_ℓ are given by (A.14). The right null vector at the lower Turing point, (A.17), is thus given by

$$H_1^\ell = -\frac{I'(n_0)h_0}{(1 + Q_\ell^2)I(n_0)} N_1^\ell + \mathcal{O}(\delta),$$

$$W_1^\ell = \frac{Q_\ell^2}{N_\ell} N_1^\ell + \mathcal{O}(\delta),$$

$$N_1^\ell = \left(1 + \frac{Q_\ell^4}{N_\ell^2} + \frac{I'(n_0)^2 h_0^2}{(1 + Q_\ell^2)^2 I(n_0)^2} \right)^{-1/2} + \mathcal{O}(\delta).$$

Similarly the left null vector components are

$$\widetilde{H}_1^\ell = \delta \frac{2N_\ell}{(1 + Q_\ell^2)\nu} (N_1^\ell)^{-1} + \mathcal{O}(\delta^2),$$

$$\widetilde{W}_1^\ell = \delta \frac{2N_\ell}{(1 + Q_\ell^2)\nu} (N_1^\ell)^{-1} + \mathcal{O}(\delta^2),$$

$$\widetilde{N}_1^\ell = 2(N_1^\ell)^{-1} + \mathcal{O}(\delta).$$

Substituting these expressions in a (A.16) gives

$$a_\ell = \frac{2Q_\ell^2 G'(w_0)}{N_\ell} N_1^\ell + \mathcal{O}(\delta)$$

as the leading order behavior of the quadratic coefficient at the lower Turing point. Since $G'(w_0)$, N_ℓ, $N_1^\ell > 0$, a_ℓ is positive at leading order, corresponding to a prediction of spot patterns near the lower Turing point.

At the upper Turing point, given that we are sufficiently far from the degeneracy with the lower Turing point, we found that $q_u^2 \equiv q_c^2 = Q_u^2 \delta^{1/2} + \mathcal{O}(\delta)$, where Q_u^2 is determined by (A.15). Given this, the right null vector components, (A.17), evaluated at the upper Turing point are

$$H_1^u = \frac{-I'(n_0)h_0}{Q_u^2 I(n_0)} \delta^{1/2} + \mathcal{O}(\delta),$$

$$W_1^u = \frac{Q_u^2}{n_0 G'(w_0)} \delta^{1/2} + \mathcal{O}(\delta),$$

$$N_1^u = 1 + \mathcal{O}(\delta)$$

The left null vector components (A.18) are

$$\widetilde{H}_1^u = \frac{2G'(w_0)n_0}{Q_u^2 \nu + \gamma Q_u^2 G'(w_0)n_0} \delta^{1/2} + \mathcal{O}(\delta),$$

$$\widetilde{W}_1^u = \frac{2G'(w_0)n_0}{\nu + \gamma G'(w_0)n_0} + \mathcal{O}(\delta^{1/2}),$$

$$\widetilde{N}_1^u = 2 + \mathcal{O}(\delta^{1/2}).$$

Putting these expressions in a (A.16) gives

$$a_u = \frac{G'(w_0)I''(n_0)h_0 n_0}{\nu + \gamma G'(w_0)n_0} + \mathcal{O}(\delta^{1/2}).$$

Since $I(n_0)$ is a concave-down function, $I''(n_0) < 0$, and thus a_u is negative at leading order at the upper Turing point. This corresponds to a prediction of gap patterns near the upper Turing point.

APPENDIX B

Numerical simulations of model by Rietkerk *et al.*

B.1. Numerical simulation procedure

We numerically solved the Rietkerk *et al.* [37] (R02) model using the exponential time differencing Runge-Kutta 4 (ETDRK4) scheme [50, 51] modified for 2D systems [52]. ETDRK4 achieves pseudospectral accuracy in space and fourth-order accuracy in time. This scheme alleviates issues of stiffness often associated with reaction-diffusion systems [52], allowing R02 to be simulated efficiently. Simulations were run using a time step near the scheme's empirically derived stability limit, since our primary concern is with qualitative aspects of patterned solutions. This scheme was implemented in MATLAB and C.

For each set of parameter values marked by a letter in Figure 7, we simulated R02 over a range of precipitation values using a procedure described in detail here. Simulations are initialised with spatially random initial conditions with fields taking uniformly distributed values in $[1, 1.5]$. The simulation begins just below the upper Turing point, at $p^0 = p_\ell + 0.95(p_u - p_\ell)$. Then the following loop is run to identify the upper bound of pattern stability beginning with $k = 0$:

1. R02 is solved using ETDRK4 until either a steady state stop condition or a uniformity stop condition is reached. The end state of the simulation, $(h, w, n)^k$ at p^k, is saved.

2. If the steady state stop condition is reached, precipitation is incremented upward by a small value Δp, so that $p^{k+1} = p^k + \Delta p$. The procedure returns to step 1, using the saved end state $(h, w, n)^k$ as the initial condition for a new simulation. If the spatial uniformity stop condition is reached, the loop ends.

The steady state stop condition occurs when either (a) the root mean square difference between the current biomass state and the state 400 time units earlier drops below 10^{-4}, or (b) $t = 2 \times 10^5$. The uniformity stop condition occurs when the root mean square difference between the current biomass state and the mean value of that state drops below 10^{-4}. We infer the upper bound of pattern stability, $p_{u+} = p^N$, to be the value of precipitation where the uniformity condition is reached.

After the first loop terminates at p^N and the first uniformity condition is reached, precipitation is decremented by $2\Delta p$ so that $p^{N+1} = p^N - 2\Delta p$. The previous patterned (i.e. non-uniform) end state, $(h, w, n)^{N-1}$, is used as the initial condition for a new simulation. We run the following loop to identify pattern morphologies that occur as precipitation decreases, starting with $k = N + 1$:

1. R02 is solved using ETDRK4 until a stop condition is reached. The end state of the simulation, $(h, w, n)^k$, is saved.

2. If a steady state stop condition is reached, precipitation is decremented by Δp so that $p^{k+1} = p^k - \Delta p$. The procedure returns to step 1, using the saved end state $(h, w, n)^k$ as the initial condition for a new simulation. If the spatial uniformity stop condition is reached, the loop ends.

When this loop ends at p^{N+M}, we infer this point to be the lower bound of pattern stability, $p_{\ell-} = p^{N+M}$.

We increment precipitation upward in one final loop to identify any potential hysteresis effects. Precipitation is first incremented by $2\Delta p$ so that $p^{N+M+1} = p^{N+M} + 2\Delta p$. The previous patterned (non-uniform) end state $(h, w, n)^{N+M-1}$ is used as the initial condition for a new simulation. Then the following loop is run, starting with $k = N + M + 1$:

1. R02 is solved using ETDRK4 until a stop condition is reached. The end state of the simulation, $(h, w, n)^k$, is saved.

2. If a steady state stop condition is reached, precipitation is incremented by Δp so that $p^{k+1} = p^k + \Delta p$. The procedure returns to step 1, using the saved end state as the initial condition for a new simulation. If the spatial uniformity stop condition is reached, the loop ends.

The end result of this procedure is a series of saved end states at a range of precipitation values over the interval $p \in (p_{\ell-}, p_{u+})$.

The precipitation iteration step size Δp was chosen based on the distance between the upper and lower Turing points, $p_u - p_\ell$, so that 30-100 end states were saved per parameter set. The value of Δp used ranged between 0.0025 and 0.01. For most simulations, a time step size of $\Delta t = 0.4$ was used, which was near the empirical stability limit of the scheme for most parameter sets. On certain parameter sets, values of Δt as small as 0.01 were needed for numerical stability. We tested for time step size errors using a particular parameter set, with $\Delta t = 0.4, \ 0.2,$ and 0.1, and the end states were found to be qualitatively indistinguishable. A preliminary set of simulations was run on all parameter sets to identify a simulation domain size for each that permitted at least 7 wavelengths of patterns. Square $L \times L$ domains with $L = 400, \ 800,$ and 1600

were used. Corresponding $N \times N$ grid sizes were used, with $N = 64, \ 128,$ and 256 respectively.

B.2. Investigating the "stripes → spots" sequences

Our bifurcation analysis indicates that gap solutions to the amplitude equations should be stable near the upper Turing point (region B of Figure 5). When a is sufficiently small, as it may be for the "stripes → spots" observations, the gaps solution is stable for only a relatively small interval of the precipitation parameter compared to the stripes branch. This is illustrated schematically in bifurcation diagram B in Figure 7. Because our numerical procedure increments precipitation in discrete steps of fixed size, the gaps branch may be bypassed. Also, since gaps are stable only very near to the Turing point for these parameter sets, gaps may only be stable at the critical wavelength and may not appear in a domain of arbitrary size.

To test whether gaps can exist stably for parameter sets where "stripes → spots" transitions are observed, we seeded numerical simulations at $p = p_u$ with a hexagonal lattice gaps pattern at the critical wavelength. This initial condition was perturbed by spatially random noise drawn from a uniform distribution on the interval $[0, 0.1]$. An aspect ratio of $1 : \sqrt{3}$ and a domain size that permits an integer multiple of the critical wavelength were used. In all simulations, gap patterns persisted as steady states through $t = 1 \times 10^6$ and were assessed to be stable in these instances.
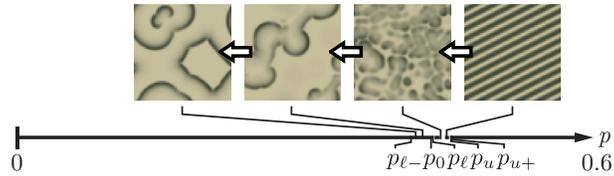
Figure B.1. Spiral wave patterns are observed in simulations at $f = 0.1$ and $\log_{10}(D_h) = 0.5$. An ordered stripes state transitions directly to spirals at $p = 0.414$.

B.3. Investigating the "stripes $\rightarrow$ spirals" sequence

We observed time-varying spiral wave patterns in one instance of the numerical simulations, at $f = 0.1$ and $\log_{10}(D_h) = 0.5$. States from this simulation are shown in Supplementary Figure B.1. We verified that the observation of spirals is robust to increased spatial and temporal resolution in the numerical scheme. We ran additional simulations at nearby parameter values to determine whether spirals are confined to a region of the parameter space. Holding $f = 0.1$ fixed, spirals appear in simulations at $\log_{10}(D_h) = 0.6$ but not $\log_{10}(D_h) = 0.7$. Holding $\log_{10}(D_h) = 0.5$ fixed, spirals appear at $f = 0.16$ but not at $f = 0.18$. We conclude that spirals are confined to smaller values of D_h than are typically considered ecologically applicable.

APPENDIX C

Vegetation band dynamics in the Horn of Africa

C.1. Regional climate

Rainfall in Somalia is bimodally distributed between the Gu season, spanning Apr.-May, and the Deyr season, spanning Oct.-Nov. Separating the rainy seasons are two dry seasons, Xagaa (Jun.-Sept.) and Jilaal (Dec.-Mar.). Deyr rainfall events are typically shorter and less significant than those of the Gu. The Jilaal season is typically the hottest and driest time of year.

Due to a lack of continuous rainfall station monitoring in and around our regions of study, we assessed the historical regional climate using climate reanalysis and remotely-sensed rainfall estimation datasets. The 20th Century Reanalysis (V2c) dataset assimilates surface pressure observations, sea-surface temperature, and sea ice extent into a global climate model to obtain a reconstruction of Earth's climate spanning 1871-2011 [79]. The V2c dataset is available at 6-hour temporal and $2°$ spatial resolution. The coarse spatial resolution of the data prevents us from distinguishing the Sool Plateau and the Haud regions. Uncertainty estimates can be derived from 56 replicate model simulations. The CPC/Famine Early Warning System Dekadal Estimates (RFEv2) dataset uses satellite microwave sensing and ground station observations to estimate total rainfall over the African continent for dates spanning 2000 to present. RFEv2 data is available at daily intervals and $0.25°$ spatial resolution.

To assess the rainfall conditions surrounding our imagery datasets, we obtained annual total rainfall estimates from the V2c and RFEv2 datasets (Figure C.1a-d). In the absence of ground confirmation, we exercise caution in interpreting the V2c estimates for the 1940s-60s, and conclude only that there is no evidence that rainfall conditions have improved in either region in recent decades. We speculate that rainfall conditions surrounding the 1952 and 1967 datasets were quite favorable. We also speculate that conditions have either declined or reverted to a regional mean in recent decades.

We assessed rainfall conditions for the recent imagery in greater detail by calculating seasonal rainfall totals from the RFEv2 dataset (Figure C.1c-d). In the Sool Plateau, the images used in this study were taken in a variety of seasons and rainfall history conditions. The 2004 image was taken shortly after the return of rains that followed a very severe multi-year drought. The 2006, 2011, and 2013 images were taken amid more typical rainfall conditions. The 2016 image was taken during a period of drought, which is ongoing at the time of writing. In the Haud, the images used in this study were taken in years (2012 and 2016) with robust rainfall during the wet seasons.

We examined regional temperature history using surface temperature estimates from the V2c dataset. We computed the average yearly temperature, defined as a yearly average over the daily midpoint between minimum and maximum temperatures (Figure C.1e). We identified a distinct linear warming trend between 1960 to the present of 1-2 °C.
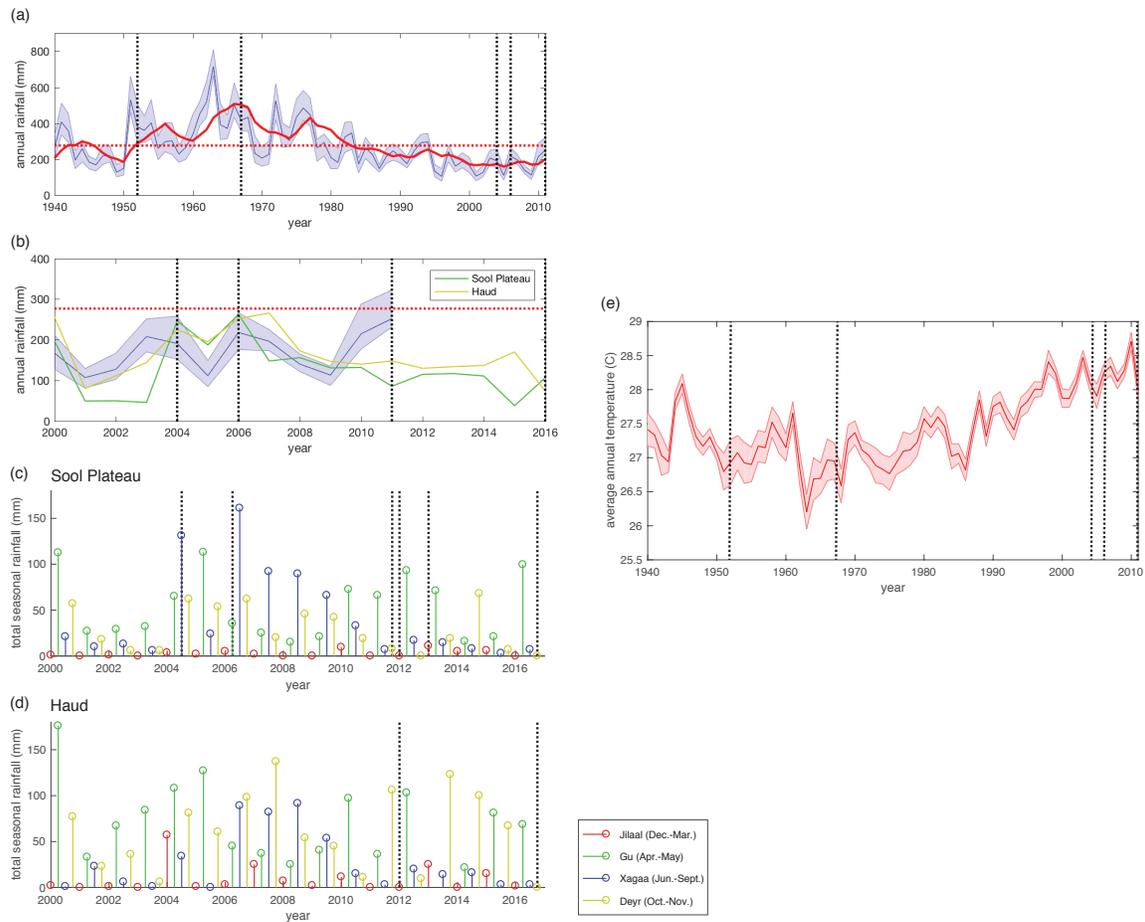
Figure C.1. Annual total rainfall estimates from V2c and RFEv2 datasets, and temperature estimates from V2c dataset. (a) shows the median V2c annual total rainfall estimates between 1940 and 2011 for a large region that includes both Sool Plateau and Haud sites. The area between 25th and 75th percentiles is shaded. The running average of median precipitation over the previous 5 years is plotted with a red solid line. The average rainfall over the entire interval is indicated with a red dashed line. (b) shows V2c (2000-2011) and RFEv2 (2000-2016) annual rainfall datasets. The average V2c rainfall estimate over 1940-2011 is indicated with a red dashed line. (c) shows seasonal rainfall totals in an area containing the Sool Plateau study areas, and (d) shows totals in an area containing Haud study areas. (e) shows average yearly temperature (°C) computed from V2c reanalysis dataset. Average yearly temperature is defined as a yearly average over the daily midpoint between minimum and maximum temperatures. One standard deviation about the mean based on 56 reanalysis simulations is indicated with shading. Dates of imagery datasets are indicated in black dashed lines.

C.2. Data

C.2.1. Imagery

We studied approximately 260 km^2 of imagery in areas of the Sool Plateau and 200 km^2 of imagery in areas of the Haud. Study area boundaries are defined by our choice of British Royal Air Force (R.A.F.) aerial survey photography, which comprise our earliest image datasets. Aerial survey photographs were taken in 1951-52 over broad areas of British Somaliland, and are archived at the Bodleian Library at the University of Oxford. The aerial photographs used in this study were scanned on request by the Bodleian Library using British Ordnance Survey maps to identify images.

The coordinates of study areas and additional information about imagery used in this study are given in Table C.1. R.A.F. images were scanned at a nominal resolution of 1.4-2.5 m/pixel. We obtained more recent imagery through the USGS and Digital-Globe Foundation. We purchased declassified reconnaissance satellite imagery[1] taken in 1967 from the USGS Earth Explorer site. We downloaded freely available Orbview-3 imagery taken in 2005 from the USGS Earth Explorer site. We were granted Quickbird-2, WorldView-1, and WorldView-2 imagery for dates spanning 2004-2016 by the Digi-talGlobe Foundation.

We manually georeferenced aerial survey photograph scans in ArcMap 10.3 against the ArcGIS World Imagery layer using the WGS84 Web Mercator coordinate system (EPSG:3857). Because vegetation bands migrate over time, we could not match scans with geospatial coordinates using the appearance of the bands themselves. Instead

---

[1]Corona program, Mission No. 1102-1

Area	(Lat, Lon)	Area (km^2)	Date	Res. (m)	Bands used	Source
SP1	(9.79°, 48.55°)	57	02/22/1952†	1.9*	Grayscale scan	Bodleian
			12/12/1967	2.0*	Grayscale scan	USGS
			06/10/2004	2.4	R,G,B,NIR	DigitalGlobe
			03/23/2006	2.4	R,G,B,NIR	DigitalGlobe
			09/29/2011†	2.0	R,G,B,NIR-2	DigitalGlobe
			12/03/2011	2.0	R,G,B,NIR-2	DigitalGlobe
			02/24/2013	2.0	Panchromatic	DigitalGlobe
SP2	(9.72°, 48.55°)	58	02/22/1952†	1.9*	Grayscale scan	Bodleian
			12/12/1967	2.0*	Grayscale scan	USGS
			06/10/2004	2.4	R,G,B,NIR	DigitalGlobe
			03/23/2006	2.4	R,G,B,NIR	DigitalGlobe
			09/29/2011†	2.0	R,G,B,NIR-2	DigitalGlobe
			12/03/2011	2.0	R,G,B,NIR-2	DigitalGlobe
			02/24/2013	2.0	Panchromatic	DigitalGlobe
SP3	(9.60°, 48.59°)	46	11/29/1952†	1.4*	Grayscale scan	Bodleian
			12/12/1967	2.0*	Grayscale scan	USGS
			06/10/2004	2.4	R,G,B,NIR	DigitalGlobe
			03/23/2006	2.4	R,G,B,NIR	DigitalGlobe
			12/03/2011†	2.0	R,G,B,NIR-2	DigitalGlobe
			02/24/2013	2.0	Panchromatic	DigitalGlobe
SP4	(9.75°, 48.83°)	58	02/22/1952†	1.6*	Grayscale scan	Bodleian
			12/12/1967	2.0*	Grayscale scan	USGS
			11/06/2005	1.0	Panchromatic	USGS
			08/16/2016†	2.0	R,G,B,NIR-2	DigitalGlobe
SP5	(9.36°, 48.79°)	44	02/14/1952†	1.5*	Grayscale scan	Bodleian
			12/12/1967	2.0*	Grayscale scan	Bodleian
			08/16/2016†	2.0	R,G,B,NIR-2	DigitalGlobe
HD1	(8.14°, 47.21°)	46	02/17/1952†	2.5*	Grayscale scan	Bodleian
			11/24/2016†	2.0	R,G,B,NIR-2	DigitalGlobe
HD2	(8.14°, 47.39°)	50	02/14/1952†	2.5*	Grayscale scan	Bodleian
			12/25/2011	2.0	R,G,B,NIR-2	DigitalGlobe
			01/21/2012†	2.0	R,G,B,NIR-2	DigitalGlobe
HD3	(8.06°, 47.44°)	50	01/24/1952†	2.5*	Grayscale scan	Bodleian
			12/25/2011	2.0	R,G,B,NIR-2	DigitalGlobe
HD4	(8.09°, 47.47°)	50	01/24/1952†	2.5*	Grayscale scan	Bodleian
			12/25/2011†	2.0	R,G,B,NIR-2	DigitalGlobe

Table C.1. Study area locations and imagery datasets used in this investigation. Datasets used in visual comparison and Fourier analysis indicated with †. Nominal resolutions of photograph scans are indicated with *. NIR and NIR-2 denotes near-infrared band data.

we relied upon apparent geological features, such as limestone outcrops, and geometrically distinct clusters of individual trees or shrubs that persisted over time. Aerial survey photographs were matched using no fewer than 10 control points per image, and were aligned by fitting a projective transformation. Control points were stored in a tab-delimited file. A projective transformation is overdetermined for greater than 4 control points, and the root mean squared error (RMSE) of the transformed control points served as our estimate of alignment error.

To estimate the effect on RMSE of adding additional control points, we used a re-sampling procedure that calculates the alignment RMSE for different subsets of the control points. For an image that was aligned using n control points, we computed the RMSE for permutations of $5 \leq k \leq n$ control points (Figure C.2). The average RMSE values over the permutations were then computed for each value of k. In this procedure, if the total number of such permutations $\binom{n}{k}$ exceeded 10^3, a random sampling of 10^3 distinct permutations were used. Otherwise, all permutations were used. The resulting curves were well fit by the saturating function $a\tilde{k}/(1+b\tilde{k})$, where $\tilde{k} = k - 5$, to extrapolate the saturating value of the average RMSE curve. In all cases the saturating RMSE value was comparable to the resolution of the imagery, suggesting an alignment error on the order of 1-2 pixels.

A reconnaissance satellite image was also manually georeferenced in ArcMap 10.3 using a third-order polynomial transformation with 18 control points. The image covers a much broader area than the aerial photographs, and due to distortions arising from the imaging methodology a projective transformation did not produce a suitable
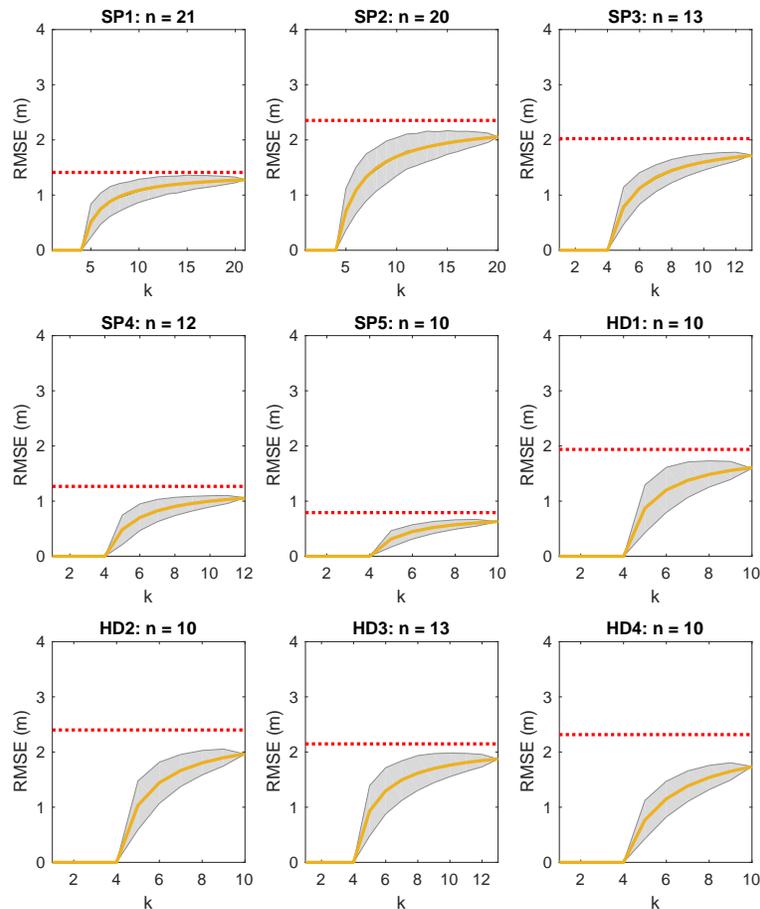
Figure C.2. Estimate of alignment error sensitivity to additional control points. For an image that was aligned using n control points, we computed the average RMSE for permutations of $5 \leq k \leq n$ control points. The resulting average RMSE curves and one standard deviation of the RMSE values are plotted. The average RMSE curves were fit to the function $ak/(1 + bk)$, and the saturating value of this function is plotted in red.

fit[2]. RMSE of this alignment is 0.94 m, which is on the order one pixel. DigitalGlobe imagery was pre-georeferenced, and was manually shifted to align more precisely with the ArcGIS World Imagery layer.

---

[2]A third-order polynomial fit is used for comparable imagery in [69]

Most recent satellite imagery used in this study contain data sensed at different frequency channels. The red, green, and blue channels were used for visualization, and the red and near infrared channels were used for computing the Soil-adjusted Vegetation Index (SAVI) [80], an index of photosynthetic activity:

$$\text{SAVI} = \frac{NIR - R}{NIR + R + L}(1 + L).$$

NIR is the near-infrared reflectance value, and R is the red reflectance value. We computed reflectances from raw pixel intensity values using radiometric calibration adjustment factors given by DigitalGlobe[3] and the imagery metadata. The parameter L is used to adjust for exposed soil surface in low-vegetation cover scenarios, and is often used in place of Normalized Difference Vegetation Index (NDVI) in dryland vegetation inference. SAVI is equivalent to NDVI for $L = 0$. For all analyses, we used a conventional value of $L = 0.5$.

C.2.2. Elevation

We used NASA Shuttle Radar Topography Mission Global 1 arc second (SRTMGL1) elevation data for our upslope migration assessment and comparison of pattern properties with slope. Datasets were obtained from the USGS website[4]. Datasets are packaged in $1°$ latitude $\times$ $1°$ longitude tiles, and were loaded, georeferenced, projected onto the WGS84 Web Mercator coordinate system (EPSG:3857) in MATLAB 2016b. This allowed for elevation data to be matched with imagery.

---

[3]https://www.digitalglobe.com/resources/technical-information
[4]https://e4ftl01.cr.usgs.gov/SRTM/

Though the SRTMGL1 dataset has a nominal resolution of 1 arcsecond ($\sim$30 m/pixel) near the equator), the true resolution is closer to 45-60 m/pixel due to the manner in which data was collected [81]. The data also contains speckle noise which is auto-correlated at a length of 1-2 pixels, and also random error, both of which together result in average vertical error of approximately 4 m in areas like the Sahara Desert. To eliminate autocorrelated errors, we subsampled the data to 3 arcsecond ($\sim$90 m/pixel) resolution.

Noise in the SRTM data presents a challenge to gradient estimation in areas of low relief, such as our regions of study, where in banded areas vertical change can be as little as 1 m per 500 m of horizontal change. To compute gradient fields, we used a second-order accuracy finite difference stencil with noise suppressing properties [82]. As an example, a 5×3 noise suppressing gradient operator as defined in [82] is

$$
f = \frac{1}{32h} \begin{bmatrix} -1 & -2 & 0 & 2 & 1 \\ -2 & -4 & 0 & 4 & 2 \\ -1 & -2 & 0 & 2 & 1 \end{bmatrix},
$$

where h is the discretization step size. Convolving this operator with the data array produces an approximation of the partial derivative in one direction. Operators with noise suppressing properties discussed in [82] can be computed for arbitrarily large stencil size.

Using a finite difference operator allows for a straightforward propagation of i.i.d. normal errors in the elevation data through the calculation of gradient and slope. We

estimated the magnitude of errors in the elevation data by computing the standard deviation of residuals from a median subtraction:

```
slidingmed = medfilt2(SRTM, [5 5]); %median-filtered data
sig = std(SRTM(:)-slidingmed(:)); %standard dev of residuals
```

The standard deviation of error propagated to each component of the gradient is then

```
sig*sqrt(sum(f(:).^2))/h
```

Slope is obtained from the magnitude of the elevation gradient vector, and to leading order the gradient error value is also equal to the error propagated to the slope calculation.

We tested the sensitivity of slope calculations to varying stencil size s (which yields a $(2s + 1) \times (2s - 1)$ operator). We note again that the truncation error of the finite difference operator is second-order for any s. Intuitively too small a stencil size will have high measurement error, and too large a stencil size will result in oversmoothing. In Figure C.3 we show the 25th, 50th, and 75th percentiles of the slope values within each study area computed over an interval of s, and indicate one standard deviation of propagated error around these values. We conclude that slope values are not sensitive to stencil size when $s \geq 15$, and we use 15 (which gives an operator of size 31×29) for all slope and gradient calculations. We confirmed by visual inspection that this stencil size produces smooth gradient fields that match hydrological features visible in the imagery (e.g., hills and channels).
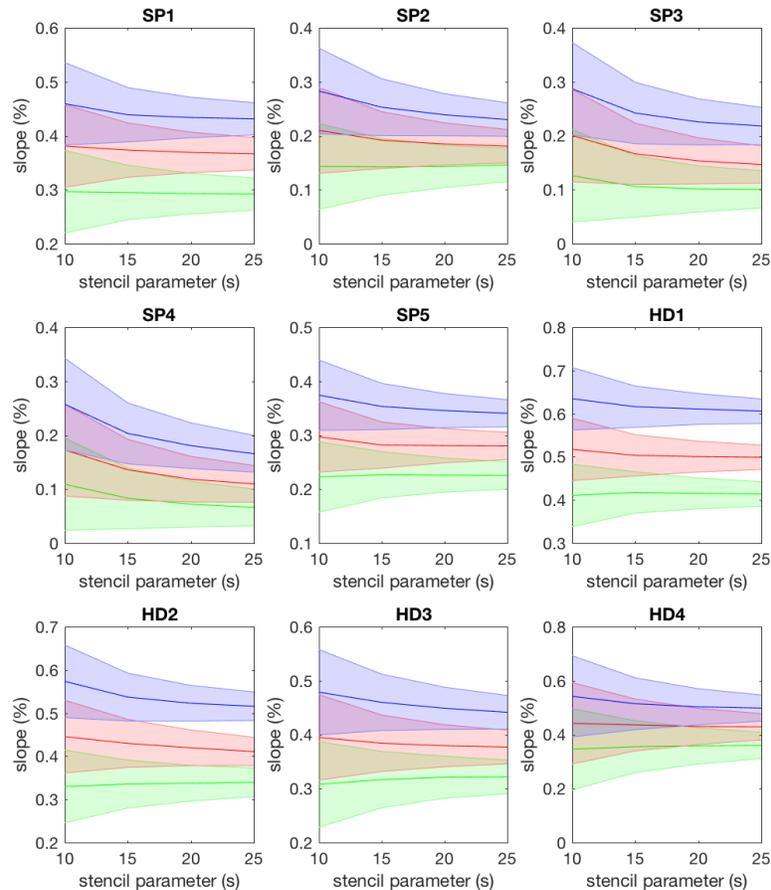
Figure C.3. Sensitivity analysis of slope calculation to stencil parameter at different study areas. The 25th (green), 50th (red), and 75th (blue) percentiles of slope values are plotted as a function of the stencil parameter, and one standard deviation of propagated error are indicated in shading. Slopes are given in units of slope percentage, which is defined as 100 times the magnitude of the elevation gradient vector.

C.3. Visual comparison

C.3.1. Protocol

We assessed changes over time at study areas via a systematic visual comparison of imagery. Roads can be visually identified in both the aerial photographs and the satellite imagery, and their presence and qualitative appearance served as our primary proxy for inferring the extent of human pressure. Vegetation in both the aerial photos and satellite imagery contrasts sharply with the light background of bare soil, and bands are clearly identifiable. Degradation was inferred through either the breakdown in regularity or disappearance of banding.

We developed a graphical user interface (GUI) in MATLAB 2016b for visually comparing images (Figure C.4). The GUI allows the user to select two imagery datasets for comparison, a georeferenced R.A.F. photograph and a more recent image. Images used for visual comparison are indicated in Table C.1. The recent image is projected onto the intrinsic coordinate system of the R.A.F. photograph, so that the data can be cleanly divided into non-overlapping 1 km $\times$ 1 km windows. The GUI simultaneously displays corresponding 1 km^2 windows of the R.A.F. photo and more recent imagery. Additionally the GUI displays a false color overlay of the two images, which was used to assess whether migration occurred. The GUI plots the local slope direction vector (computed as described in Section C.2.2) on top of the overlay, which allows the user to visually assess whether the migration is in the upslope.

For each image window, the user is prompted to enter whether regular banding is present and whether a dense settlement is present (more than 5 structures in close
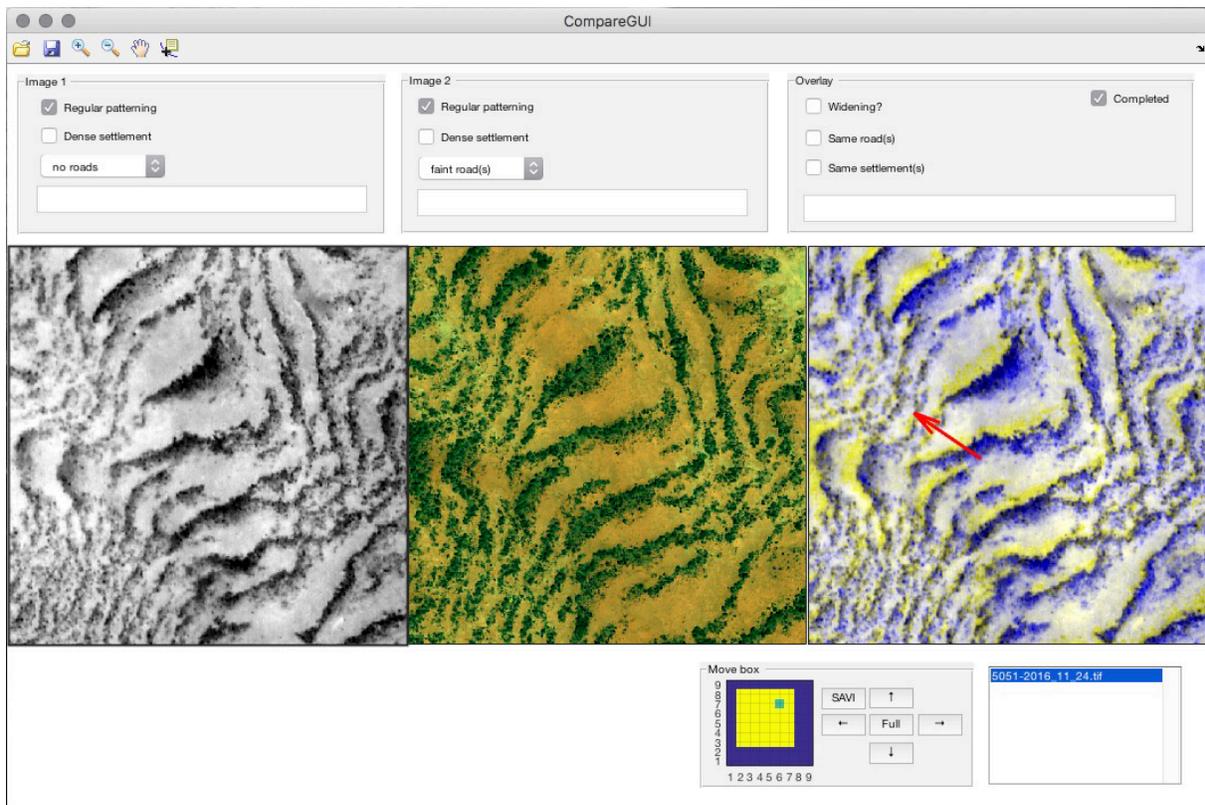
Figure C.4. Graphical user interface (GUI) used for visual comparison of images. A georeferenced R.A.F. photograph and a recent satellite image are divided into corresponding windows and shown side by side. A false color overlay of the two images is also shown, with a local slope vector overlaid to visually assess upslope migration. Blue in the false color image denotes vegetation in the R.A.F. photograph, and yellow denotes vegetation in the recent image.

proximity) using checkboxes. The user can select the extent of apparent road cover via a dropdown menu. Additionally, the user can check boxes to indicate whether band widening in the slope direction is apparent, and whether it appears that the same roads or settlements are present in both images. The user can enter comments for each image and the overlay. If the recent image contained red and near infrared channels, the Soil-adjusted Vegetation Index (SAVI) can be displayed in place of the RGB image

(see discussion of SAVI in Section C.2.1). The GUI selections are automatically saved to a MATLAB .mat file. When the user has finished assessing the window, the user can then navigate to different windows in the dataset using buttons.

The checkboxes are ternary; with banding, for example, a fully-checked state is taken to mean distinct banding, a half-checked state is taken to mean indistinct banding, and an unchecked state is taken to mean no banding. The road cover dropdown can take on one of four states: "no roads," "faint road(s)," "clear road(s)," and "clear, dense road(s)." The last state, "clear, dense road(s)," is taken to mean that a large number of well-incised, clearly visible roads cover a large portion of the window.

C.3.2. Highlighted examples

In many sites of the Haud study areas, we observed that human-made structures appeared to persist from 1952 to the present. In Figure C.5a,b, we show two examples of such structures.

In addition, we observed that in some areas of SP4, bands appeared to degrade without apparent change in wavelength. We show an example in Figure C.5c. We verified that the pictured bands have significantly lower vegetation index (SAVI) values than nearby bands in the study area, and are plausibly degraded.

C.4. Fourier analysis

C.4.1. Protocol

We quantitatively assessed changes in band wavelength using a modification of the Fourier window method by Penny *et al.* [83]. Penny *et al.* developed the method to
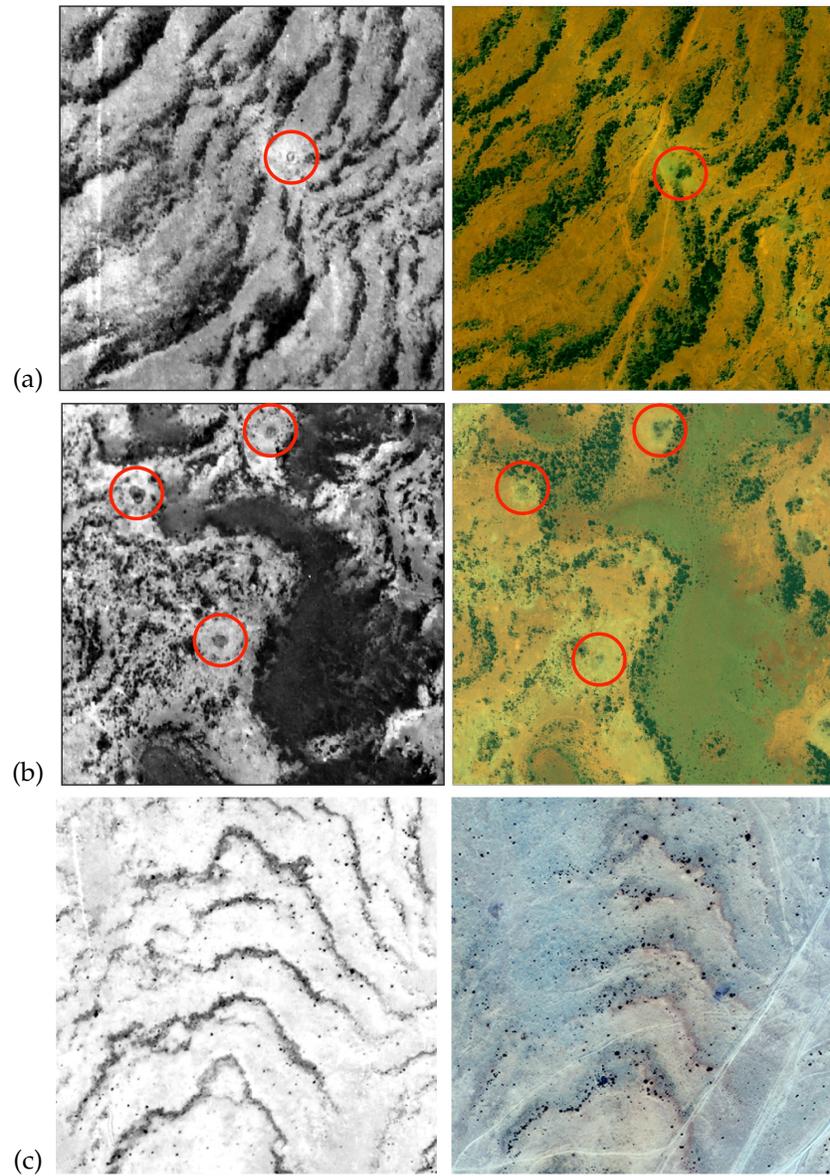
Figure C.5. Highlighted examples from visual inspection. (a) and (b) show examples at HD2 and HD4, respectively, of man-made structures that appear to persist from 1952 to 2012. (c) shows an example at SP4 where degradation appears to have occurred without an apparent history of change in band wavelength.

compute spatial maps of local wavelength and orientation from imagery over banded areas in Fort Stockton, Texas, USA. In a manner analogous to a short-time Fourier transform, the method measures wavelength and orientation in a sliding window using a 2D FFT. Vegetation banding typically contains sufficient irregularity to complicate the inference of dominant wavelength and orientation from a 2D power spectrum. The Fourier window method addresses this issue by binning power, radially for estimating wavelength and angularly for orientation, and by computing a weighted average among the contiguous bins with largest power. The method computes a uniqueness metric for both wavelength and orientation based on the distance between the maximal peak and the nearest peak with 75% of the maximal power, if present. The metric equals one if the maximal peak is the only powerful peak present, and approaches 0 as distance to the nearest powerful peak increases. In order to exclude short-wavelength noise and long wavelengths which are under-sampled for the given window size, the bins are only computed for a specified minimum and maximum wavelength interval. As presented in [83], the pattern irregularity issue is also addressed by averaging measurements over overlapping windows. We do not perform the latter step for the analyses in this study.

Penny *et al.* provide MATLAB code for their method, which we modify for our analyses. We modify the main routine to take as input two images that have been resized to the same dimensions. Computations are then performed on square windows. To reduce aperiodicity effects, we apply a 2D Hamming filter (a bell-shaped function that decays to zero away from the center) to each window. We note that this filter also

has the effect of giving more weight to the central area of the image, focusing analysis on this area. Wavelengths and orientations are then computed using Penny *et al.*'s routine.

We applied this methodology to all study areas using the image pairs indicated in Table C.1. In order to perform the windowing on a rectangularly-oriented dataset, we transformed the recent imagery onto the intrinsic coordinate system of the aerial photograph. We downscaled image pairs to a resolution of approximately 2.5 m/pixel to reduce computation time. We then applied two layers of preprocessing to emphasize the vegetation bands and de-emphasize other features in the imagery: we subtracted a coarsely gaussian-blurred version of the image to eliminate large scale variations in pixel intensity (such as darkening near the borders of the aerial photographs), and we applied a manually-tuned threshold to create a binary image of the vegetation bands.

For each binary image pair, we computed wavelength maps for the three square window sizes: 384 pixels ($\sim$1 km), 512 pixels ($\sim$1.3 km), and 768 pixels ($\sim$2 km). After applying a Hamming filter, about 4-8 vegetation bands can be sampled in the central area of a 512 pixel window (3-5 bands for a 384 pixel window, or 6-10 for a 768 pixel window). For the binning procedure, we set the minimum wavelength to 10 pixels ($\sim$25 m), and the maximum wavelength to one-fourth of the window size ($\sim$240 m for a 384 pixel window, $\sim$320 m for a 512 pixel window, and $\sim$480 m for a 768 pixel window). In order to balance computation time and even-sampling of the data, we set the step length of the sliding window to be one-fourth of the window size, resulting in adjacent windows that overlap in 75% of their area.

After wavelength maps were computed for an imagery pair, we transformed the measurements from units of pixels to units of meters. We then manually drew a mask on the imagery and applied it to the measurements in order to exclude measurements from areas without vegetation bands. Additionally we excluded measurements with wavelength uniqueness metrics smaller than 0.75. We found that in some areas, a window size of 384 pixels was too small to detect the largest wavelengths. The results of 512 and 768 pixel windows did not differ strongly, so we used the 512 pixel window computation for the results reported in Table 3.1.

C.4.2. Wavelength change

Given spatial maps of wavelength for a pair of images, we computed change maps where elements are given by

$$W_2^{i,j}/W_1^{i,j} - 1,$$

where $W_1^{i,j}$ is the wavelength in the first image at position (i, j), and $W_2^{i,j}$ is the wavelength in the second image at (i, j) (Figure C.6). Note that for computing change maps, we have manually masked out areas with no banding, and we have also excluded measurements with a wavelength uniqueness metric smaller than 0.75. We made the latter choice to reduce the incidence of falsely detected changes between the maps, reasoning that measurements in areas with multiple dominant band wavelengths are error prone. Typical change ranges between 0–10% for all study areas except SP3 and SP4, where change ranges between 0–20%.
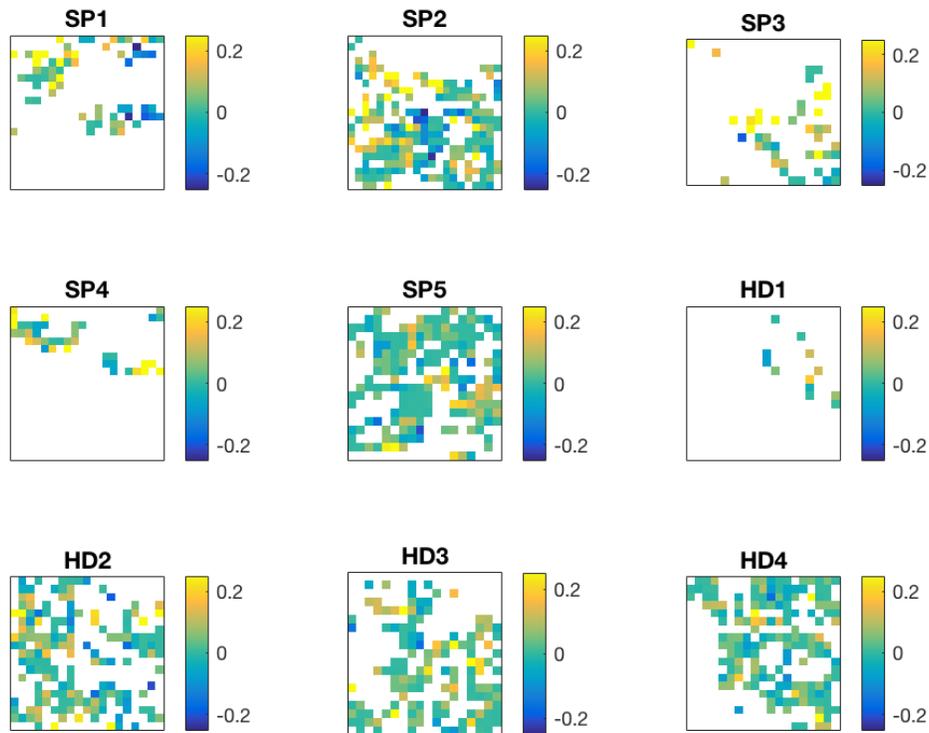
Figure C.6. Wavelength change maps for all study areas. White pixels indicate data points in areas without banding or areas with sufficiently low values of the computed wavelength uniqueness metric. All color axes are scaled to the interval [-0.25,0.25].

Our previous visual inspection suggested that there were no obvious systematic changes in wavelength at any study area, except perhaps for those associated with isolated instances of band loss in human-impacted areas. We visually reinspected areas where measured change was greater than 25% in magnitude. In some cases, it appears that these detected changes occur due to the loss of an individual band, often near evidence of human activity (Figure C.7). In Figure C.7b-e, significant road cover appears in the interband areas of the recent imagery, and are likely related to the loss of bands in

these areas. In most cases, however, we saw no clear indication for the detected wavelength changes, and attributed these false detections to wavelength measurement error that arises due to the irregularity of the banding.

C.4.3. Wavelength-slope correlations

We computed the correlation between local wavelength and slope, and the results are reported in Table 3.1). We used slope values that are closest to the center point of the window corresponding to the wavelength/migration measurement. To assess the significance of correlations, we used a paired t-test for which sample size is corrected to account for spatial autocorrelation in the data [85]. The test is implemented in the library SpatialPack for R [95]. The correlation between wavelength and slope has been empirically investigated in [83] and [69]. We found apparent correlations between slope and wavelength only at study areas SP3 ($r = -0.34, p = 0.04$) and SP5 ($r = -0.23$, $p = 0.06$).

C.5. Automated transect measurements

C.5.1. Protocol

We quantified aspects of vegetation dynamics in all study areas using automated transect measurements of individual bands. To do this, we segmented the aerial photograph to identify bands, gathered image intensity profiles along transects through the bands in direction of slope, and fit a simple step function to extract band width and
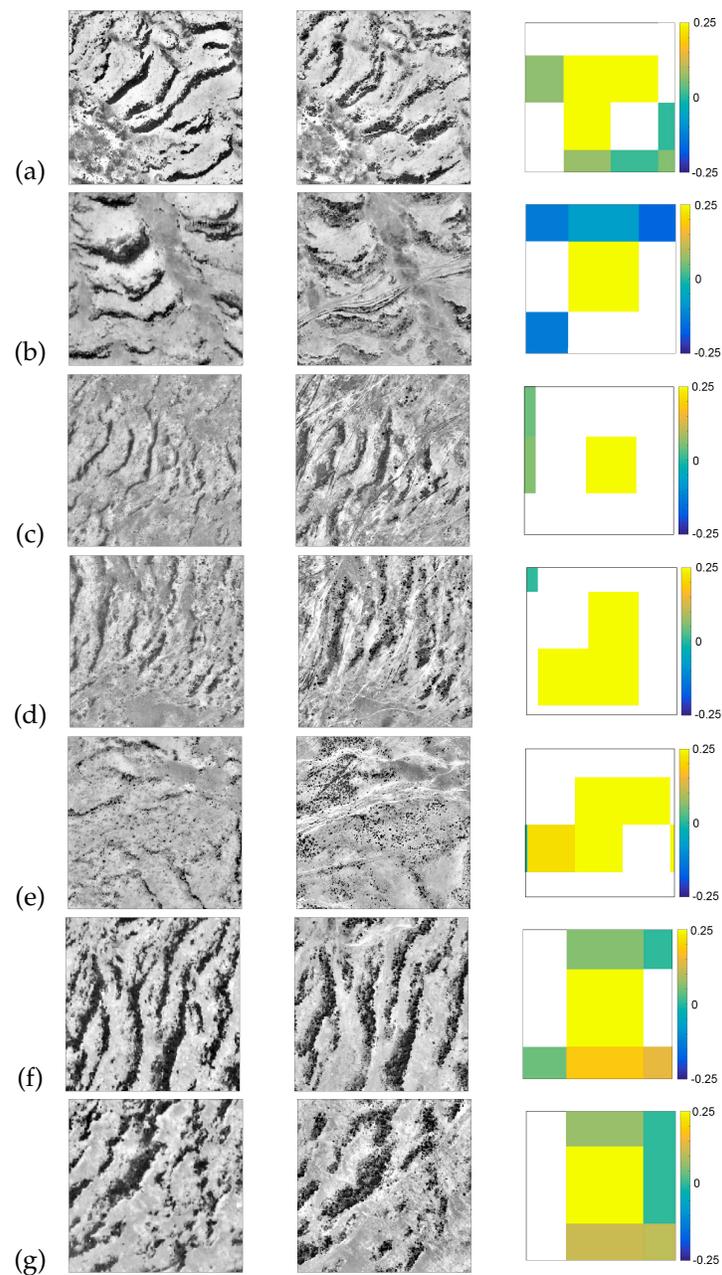
Figure C.7. Examples of areas where detected wavelength change corresponds to band loss or degradation. The first column is R.A.F. aerial photography, the second column is recent satellite imagery, and the third column is the detected wavelength change map. Locations of examples are as follows: (a) SP1, (b) SP2, (c) and (d) SP3, (e) SP4, (f) and (g) HD4.

position along the transect. We used this information to assess changes in band width over time, as well as band migration.

We eliminated the large-scale background variations in pixel intensity in the aerial photographs by subtracting a coarsely gaussian-blurred version of the image. We then applied a manually-tuned threshold to create a binary image of the vegetation bands. We passed this binary image to the `regionprops` function in MATLAB 2016b, and extracted the areas and centroids of connected components, as well as the major and minor axis lengths and orientations of ellipses fit to the connected components. We applied manually tuned criteria on the areas and ratio between major and minor axis length to isolate the vegetation bands in the binary image.

We then drew linear transects through the centroids of the stripes in the direction of the minor axis. We visually confirmed that the minor axis direction serves as an effective proxy for the slope direction. For all study areas, transect lengths are approximately 100 pixels (∼190 m). The transects were drawn so that 25% of the transect lies downslope of the centroid, and 75% lies upslope. We did this so that the same transect could be used for both the aerial photographs and the more recent imagery, accounting for band migration upslope. In order to obtain replicate measurements for estimation of error and variance, we drew eight additional transects transverse to the original (four on either side). These transects are spaced approximately 4 m apart, which precludes double-sampling of pixels by adjacent transects. We then used these transects to extract pixel intensity profiles of the aerial photography and more recent imagery. We converted color images to grayscale before extracting intensities using the `rgb2gray` MATLAB function. Whenever near-infrared channels were available in the

data, we computed the SAVI values for the image and extracted the SAVI intensities along transects as well.

We fit the intensity profile with simple plateau-like curves using MATLAB's non-linear least squares curve fitting function, `lsqcurvefit`. The curve has the form

$$f(x; \mathbf{b}) = b_1 + \frac{b_2}{2} \left[\tanh(\alpha(x - b_3)) - \tanh(\alpha(x - b_4)) \right],$$

which approaches a piecewise constant function with levels b_1 and b_2 and breakpoints at b_3 and b_4 in the limit as $\alpha \to \infty$ (Figure C.8). We used $\alpha = 500$. To fit this function, we rescaled all transects to lie along the interval $x \in [0, 1]$. For our data, the squared error cost function typically had many local minima, and so the result was sensitive to the initial guess for the parameters b_3 and b_4. We fit each intensity profile using 20 uniformly random initial guesses for b_3 and b_4 (such that $b_3 < b_4$), and chose among these the result with the minimum squared error. We then used the value $w = b_4 - b_3$ as the measured width of the band along the particular transect, and compared b_3 and b_4 values along a transect at different time points to measure migration. If the transect had an accompanying SAVI profile, then the median SAVI value s along the interval $b_3 \leq x \leq b_4$ was recorded.

Since each band was measured using multiple parallel transects, we obtained multiple measurements of w, b_3, and b_4 for each band. We used a threshold on the standard deviation ($\sigma \leq 0.2$) of the w measurements to exclude data points where bands may have substantially degraded or disappeared, or where the measurement is likely poor for some other reason. After applying the threshold, we calculated the mean w for
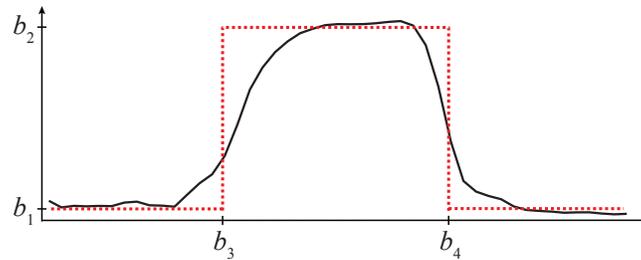
Figure C.8. Schematic example of pixel intensity profile along a transect (black) and plateau function fit to the profile (red). Levels b_1 and b_2, as well as breakpoints b_3 and b_4 are indicated.

each remaining band at each time point. When SAVI data is available, we calculate the mean s for each band as well.

C.5.2. Sool Plateau measurements

We observed appreciable increases in band width in SP1-SP4. We measured widths at multiple time points in these areas to assess when the widening may have occurred and whether it is a seasonal phenomenon. In Figure C.9, we plot the distribution of band widths in the Sool Plateau sites SP1-SP5 over time. We have reconnaissance imagery taken in 1967 for all these sites. We observe that band widths changed little between 1952 and 1967. In SP1–SP4, widths are then larger in the recent imagery (onward from 2004), and do not return to their 1952/1967 widths. In SP5, widths remain unchanged between 1967 and 2016. We conclude that band widths increased in SP1–SP4 sometime between 1967 and 2004, and that this widening is not a seasonal effect.
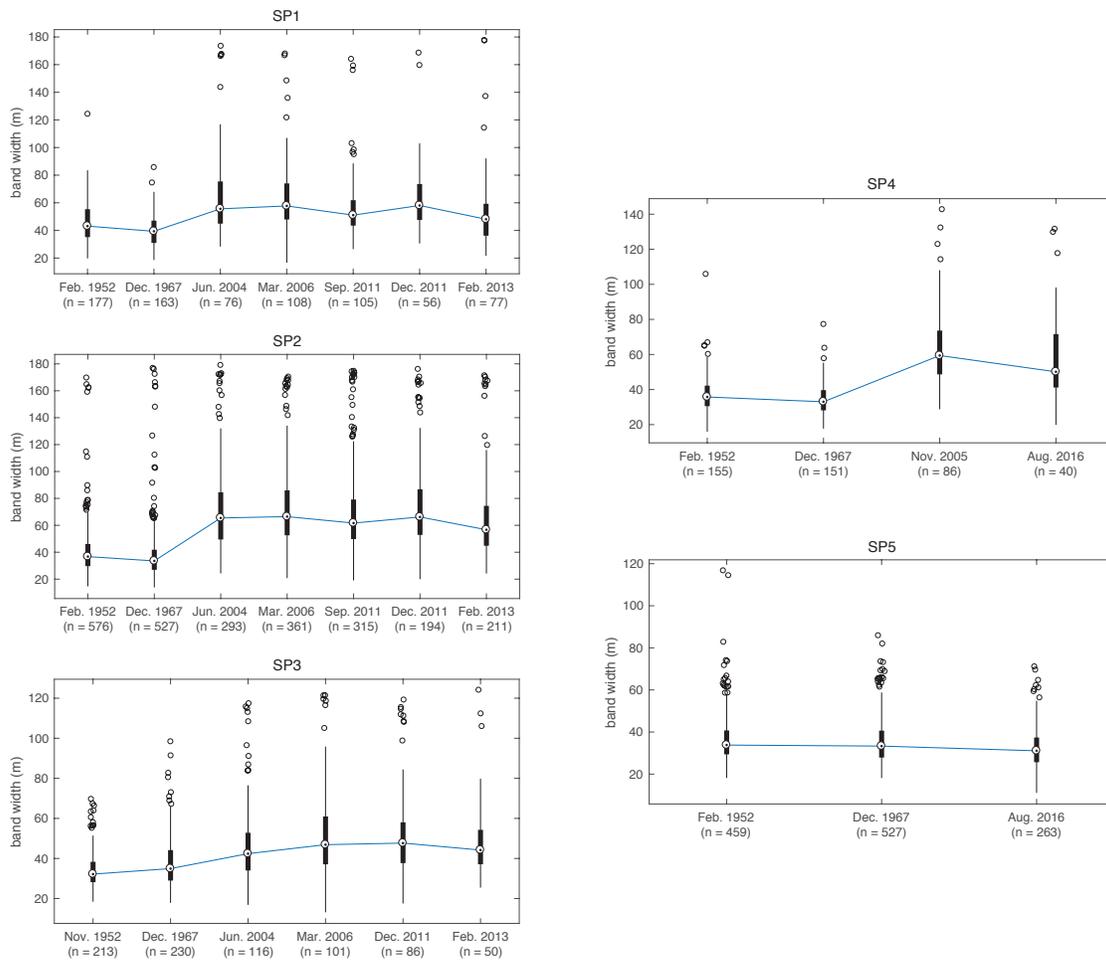
Figure C.9. Band widths measured at SP1–SP5 are shown at multiple points in time. At SP1–SP4, widths change little between 1952 and 1967, and increase between 1967 and recent imagery. Widths remain approximately constant at SP5 from 1952-2016.

C.6. Model simulations of band widening

We performed a sensitivity analysis to estimate the linear response of band width and peak band biomass in the model by Klausmeier [9] (K99) to changes in the parameter set. In dimensional form, the model is written as

(C.1)
$$\frac{\partial W}{\partial T} = \underbrace{A}_{\text{precip.}} - \underbrace{LW}_{\text{evap.}} - \underbrace{RWN^2}_{\text{transp.}} + \underbrace{V\frac{\partial W}{\partial X}}_{\text{advec.}},$$

$$\frac{\partial N}{\partial T} = \underbrace{-MN}_{\text{mort.}} + \underbrace{JRWN^2}_{\text{growth}} + \underbrace{D_N\frac{\partial^2 N}{\partial X^2}}_{\text{dispersal}}.$$

We began each simulation using the parameter set shown in Table C.2 with an initial state of the uniform equilibrium value plus small-magnitude spatial noise. We evolved the initial state to 10,000 years to obtain an equilibrium migrating patterned state. We then began a set of simulations where in each we perturb one value in the parameter set listed in Table 3.3 by a percentage between 5 and 100% that we tuned to produce a 5-10% response in width ratio. We then evolved the initial equilibrium patterned state by 50 years. The resulting patterned states are pulselike, and we measured widths by thresholding using a small value (10^{-2}) and counting the size of the resulting connected components. We stored the maximum values for each band as well. We show the width ratios and height ratios for all parameter perturbation simulations in Figure C.10, where the ratios are computed by dividing band widths or band peak values in perturbed simulations by the widths or peak values from the initial patterned state.

Parameter	Units	Description	Value
A	mm H_2O yr^{-1}	mean annual rainfall	150
L	yr^{-1}	evaporation rate	4
J	kg m^{-2} (mm H_2O)$^{-1}$	biomass yield per unit H_2O	0.0025
M	yr^{-1}	mortality rate	0.75
R	mm H_2O yr^{-1} (kg dry mass)$^{-2}$	transpiration rate	50
D_N	m^2 yr^{-1}	plant dispersal rate	1
V	yr^{-1}	water flow speed	35

Table C.2. Baseline parameter set used for sensitivity analysis of K99 based on the values given in [9]. Since bands in our regions of study are composed of a mix of grasses and shrubs, parameters which are stated in [9] to differ between grasses and trees (M, J, and R) are set at intermediate values so that the spatial scale of banding resembles the scales in our regions of study. Water flow rate V was also approximately tuned to the time scale of migration in our regions of study. We set the mean annual rainfall parameter to 150 mm, which is within the range of typical rainfall levels in our regions of study.

Increasing A, J, R, D_n, and V and decreasing L, M, and D_w results in band width increases. All parameter changes which increase band width also increase peak biomass, except for D_n and D_w.

To simulate a scenario where band species composition shifts from woody to grass biomass, we simultaneously increase J, R, and M by 10% and D_n by 50% (Figure C.10). Although increasing mortality by itself reduces the band width, the simultaneous increase of these four parameters results in band width increase.
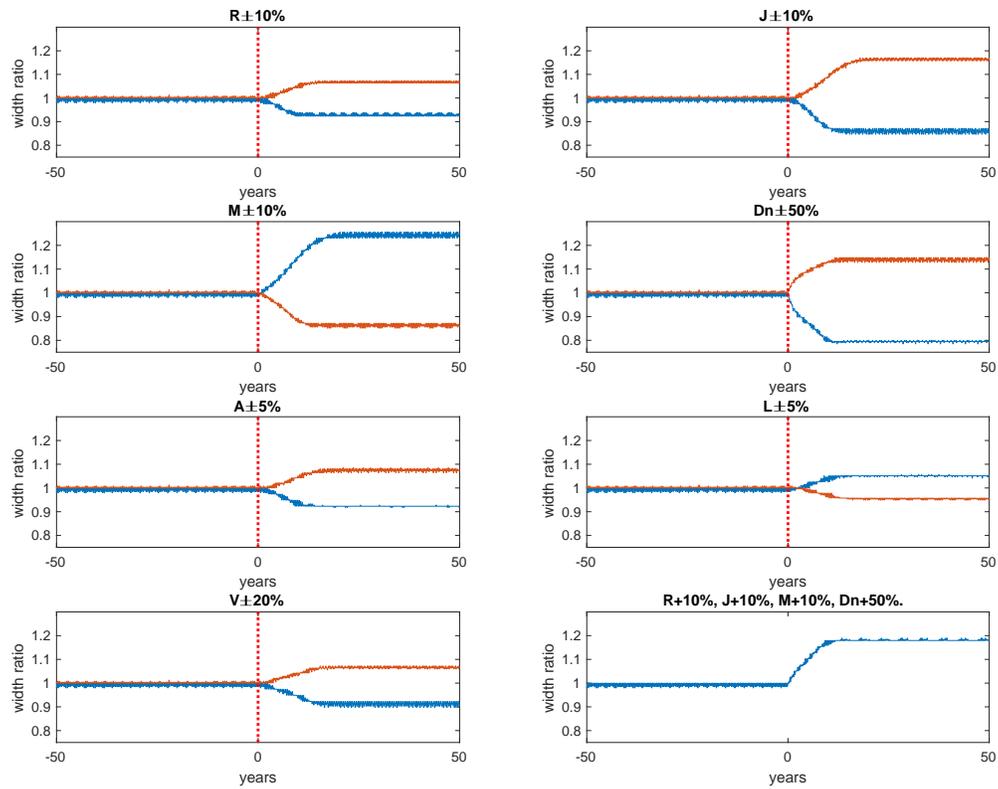
Figure C.10. K99 sensitivity analysis to width ratio. Parameters are perturbed individually by + (red) or - (blue) the percentage of the parameter indicated.

Vita

Karna Gowda was born in Bourbonnais, Illinois, and grew up watching the Bulls in the north suburbs of Chicago. He graduated from Mundelein High School in 2005, and from the University of Illinois at Urbana-Champaign with a BS in Mathematics in 2008. After spending some years teaching and working in the corporate world, Karna turned to academic research in 2011 as a student in the Department of Engineering Sciences and Applied Mathematics at Northwestern University. In 2017, Karna will transition to research in experimental ecology and evolution as a James S. McDonnell Foundation postdoctoral fellow in the Department of Physics at the University of Illinois at Urbana-Champaign.

In his off-hours, Karna can be found reading, running, watching The Simpsons with his partner, Sara, or on a long walk with his dog, Kuromatsu.