Using wavelength and slope to infer the historical origin of semiarid vegetation bands

Citation for published version:
https://doi.org/10.1073/pnas.1420171112

Digital Object Identifier (DOI):
10.1073/pnas.1420171112

Link:
Link to publication record in Heriot-Watt Research Portal

Document Version:
Peer reviewed version

Published In:
Proceedings of the National Academy of Sciences

General rights
Copyright for the publications made accessible via Heriot-Watt Research Portal is retained by the author(s) and/or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
Heriot-Watt University has made every reasonable effort to ensure that the content in Heriot-Watt Research Portal complies with UK legislation. If you believe that the public display of this file breaches copyright please contact open.access@hw.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Using Wavelength and Slope to Infer the Historical Origin of Semi-Arid Vegetation Bands

Jonathan A. Sherratt *

*Department of Mathematics and Maxwell Institute for Mathematical Sciences, Heriot-Watt University, Edinburgh EH14 4AS, UK, j.a.sherratt@hw.ac.uk

Submitted to Proceedings of the National Academy of Sciences of the United States of America

Landscape-scale patterns of vegetation occur worldwide at interfaces between semi-arid and arid climates [1]. They are important as potential indicators of climate change and imminent regime shifts [3]. Although other mechanisms have been suggested [4,5], the patterns are widely thought to arise from positive feedback between vegetation and infiltration of rainwater. On gentle slopes the typical pattern form is bands (stripes), oriented parallel to the contours [6,7], and their wavelength is probably the most accessible statistic for vegetation patterns. Recent field studies have found an inverse correlation between pattern wavelength and slope, in apparent contradiction to the predictions of mathematical models. Here I show that this “contradiction” is based on a flawed approach to calculating the wavelength in models. When pattern generation is considered in detail, the theory is fully consistent with empirical results. For realistic parameters, degradation of uniform vegetation generates patterns whose wavelength increases with slope, while colonisation of bare ground gives the opposite trend. Therefore the empirical finding of an inverse relationship can be used, in conjunction with climate records, to infer the historical origin of the patterns. Specifically, for the African Sahel my results suggest that banded vegetation originated by the colonisation of bare ground during c. 1760-1790 or since c. 1850.

Moreover it gives valuable insights into the historical origin of these patterns.

Mathematical Modelling of Semi-Arid Vegetation

Mathematical models play a key role in understanding arid ecosystems, and a wide variety of models have been proposed over the last two decades, ranging from detailed multi-scale representations of soil-water dynamics [13] to simple models of key underlying mechanisms [14–18]. I will investigate the extent to which qualitative trends in wavelength apply irrespective of parameter values. This requires comprehensive scans across parameter space, which is only possible for very simple models. Therefore I use the Klausmeier model [14], which is one of the earliest and simplest models for vegetation patterning, and which remains in widespread use [19–23]. It is formulated in terms of vegetation biomass \( u(x,t) \) and water density \( w(x,t) \). Here \( t \) denotes time and distance \( x \) is measured in the uphill direction; I assume a uniform slope and throughout I will consider behaviour in one spatial dimension, which is sufficient for banded patterns. When suitably nondimensionalised [12,14] the model equations are:

\[
\begin{align*}
\frac{\partial u}{\partial t} &= \text{plant growth } + \text{plant loss } - \text{plant \text{\#1} uptake by plants } + \nu \frac{\partial^2 u}{\partial x^2} + D \frac{\partial}{\partial x} \frac{\partial w}{\partial x} \\
\frac{\partial w}{\partial t} &= \text{average rainfall } - \text{evaporation } - \text{plant \text{\#2} loss } + \nu \frac{\partial^2 w}{\partial x^2} + \text{flow downhill } \text{of water } \\
\end{align*}
\]

Significance

Self-organised vegetation patterns are a characteristic feature of semi-arid regions. On gentle slopes banded patterns (stripes) are typical, and their wavelength is probably the most accessible statistic for patterned vegetation. Recent data shows that on steeper slopes wavelengths are usually shorter, contradicting previous predictions of mathematical models. I resolve this “contradiction” by a detailed theoretical study of pattern generation. Moreover I show that the wavelength-slope relationship has a wholly unexpected predictive power, enabling one to infer whether the patterns arose from degradation of uniform vegetation or colonisation of bare ground. When combined with climate records, this gives new insights into the historical origin of the patterns.

Reserved for Publication Footnotes
Crucially, there are only four dimensionless parameters, which makes comprehensive scans of parameter space feasible. The key driver of pattern formation in (1) is the assumption that the per capita specific water uptake is proportional to biomass density. This is based on extensive empirical evidence that in semi-arid environments, rainfall infiltration is positively correlated with vegetation cover [24, 25], due to increasing levels of organic matter in the soil, and to the presence of root networks. The parameter \( A \) represents an average rate of rainfall, which typically occurs in discrete storm events in semi-arid regions [26, 27]. The plant loss term includes both natural death and the effects of any herbivory. Diffusion is used to model plant dispersal in the interests of mathematical simplicity; some subsequent models have used instead a nonlocal dispersal term [28, 29]. Klausmeier’s original formulation [14] did not include water diffusion but this has been added by a number of subsequent authors [19, 21–23]. Since \( D \) will typically be much larger than the (dimensionless) plant dispersal coefficient of 1, this additional term tends to enhance the pattern forming potential of the model.

When rainfall \( A \) is large, (1) predicts a stable uniformly vegetated state. As rainfall is decreased, this becomes unstable via a Turing-Hopf bifurcation, giving spatial patterns [12, 30]. To investigate the wavelengths of such patterns one considers disturbances to the uniform vegetated state with a particular spatial frequency and calculates their growth (or decay) rate. For values of \( A \) below the Turing-Hopf point, one particular frequency will give the largest growth rate; this is the “most unstable mode” (Fig. 1a), from which one can calculate the expected pattern wavelength (see Methods).

I found that whenever the rainfall \( A \) is below the critical value for patterns, the expected wavelength is positively correlated with slope (see Fig. 1b and Methods). The same trend has been found in much less systematic studies of other models [11, 12]. Moreover it is expected, because of the increased run-off on steeper gradients. However this positive correlation is the opposite of the wavelength-slope relationship found empirically [5, 8–11].

Pattern Generation from Uniform Vegetation

Calculation of wavelength using the “most unstable mode” assumes that patterns arise via disturbance of a state that is uniformly vegetated but unstable. This raises the very natural question of how the system arrived at an unstable state in the first place. In many biological contexts such as development, an unstable uniform state can arise very naturally. For example, it might be stable in an embryo that is too small to permit the destabilising frequencies, but becomes unstable as the embryo grows (e.g. [31]). Alternatively destabilisation may result from a particular gene being expressed during development (e.g. [32]). But for semi-arid vegetation there is no clear mechanism that could generate an unstable state.

It follows that the “most unstable mode” approach to the calculation of pattern wavelength and the resultant “contradiction” are not relevant to banded vegetation. Nevertheless, the pattern wavelength is possible, using the finding that wavelength is history-dependent and persistent. Studies of both (1) and other models for semi-arid vegetation have shown that as environmental parameters are gradually changed, properties of banded vegetation patterns such as net biomass and band interband ratio also change but wavelength remains constant, until a sufficiently large change in environmental parameters causes a loss of pattern stability and hence an abrupt shift in wavelength [20, 23, 33, 34]. Crucially, wavelength persistence implies that the key to its prediction lies in calculating it when patterns are first established, which can happen in two ways. Patterns might develop from uniform vegetation when changes in an environmental parameter such as rainfall cause that uniform state to become unstable. Alternatively patterns might arise via colonisation of bare ground.

For patterns arising from degradation of uniform vegetation, the wavelength-slope relationship is the result of a trade-off. The wavelength corresponding to the most unstable mode increases with slope. However steeper slopes facilitate patterning, so that pattern onset occurs at higher rainfall levels, and this tends to decrease the wavelength (see Methods). The resulting wavelength-slope relationship is non-monotonic (Fig. 1c), as found previously by Ursino [19]. Note that in Fig. 1b rainfall is fixed, and the plotted wavelength corresponds to the most unstable frequency, while in Fig. 1c rainfall varies with slope to give the pattern onset point. I found that the qualitative form of Fig. 1c applies for all relevant parameters. Thus pattern wavelength increases with slope when this is small, reaches a maximum at \( \nu = \nu_m \), say, and then decreases as slope is increased further. Although these results provide a potential explanation for the empirically observed negative correlation between wavelength and slope, all parameter estimates in the literature actually give \( \nu < \nu_m \) (see Methods and Fig. 3).

Pattern Generation from Bare Ground

I now consider banded patterns initiated by the colonisation of bare ground. Mathematically this is much more difficult because it cannot be studied via small perturbations to a uniform state: it is a fundamentally nonlinear problem. In addition to the uniformly vegetated state, (1) also has a uniform vegetated state that is stable to small perturbations. Intuitively one expects that when rainfall is sufficiently high, a localised introduction of plants will colonise this bare ground state, and I confirmed this in simulations of (1). Such an invasion involves a transition between two locally stable states, a situation that has been well studied in many contexts, for example the propagation of electrical signals along a nerve axon [35, ch. 5]. However, formulae for the invasion speed can only be calculated in the very simplest cases, and I studied colonisation for (1) using simulations.

I found that in comparison to flat terrain, vegetation has a greater propensity to invade in an uphill direction. Intuitively this is because the downhill flow of water facilitates vegetation growth at the edge of the invading front. Similarly invasion downhill is impeded in comparison to flat ground. Therefore at low levels of rainfall, both edges of a localised vegetation patch on a slope move uphill. In order to colonise bare ground, the downhill edge of a vegetation patch must invade in the downhill direction, which occurs only for rainfall levels above a critical minimum. After a drought during which vegetation has died out, recolonisation will commence when rainfall increases to this critical level, establishing a particular pattern wavelength which will then persist following subsequent moderate variations in rainfall.

Fig. 2a shows a typical plot of rainfall threshold against slope. The shaded region is where the uniformly vegetated state is stable, so that vegetation does not form bands. On sufficiently steep slopes, say \( \nu > \nu_p \), colonisation requires a rainfall level in this region; I found that \( \nu_p < \nu_m \) throughout parameter space (Fig. 4). For \( \nu < \nu_p \) colonisation leads to patterns (Fig. 5), and Fig. 2b shows that their wavelength decreases with slope, as found empirically [5, 8–11]. Simulations across parameter space confirm this to be a general result.
Discussion

To provide a specific example of the implications of my results, I consider the African Sahel, which is the transition zone between the Sahara and the Sudanian Savanna. Here banded vegetation occurs for slope gradients of about 0.2%–1% [6, 7, 10], and parameter estimates place these shallow slopes well below the smaller of the critical values, \( \nu \) [14, 19]. My results therefore suggest that wavelength would increase with slope for patterns arising from the degradation of uniform vegetation, and would decrease with slope following the colonisation of bare ground. Since there is now a large amount of data indicating the latter trend [8–10], I infer that the banded vegetation in this region has developed via colonisation of bare ground – at least in the locations providing the data, which are widespread.

This inference must be considered in conjunction with historical climate data. Rain gauge records for the Sahel are very limited before about 1920 [36], but there is considerable proxy data for the last 5 centuries [37–39]. This shows that humid conditions prevailed in the Sahel during the 16th and 17th centuries*. Evidence for this comes from three independent sources. Most quantitative are fluctuations of lake levels: for example between 1650 and 1700. Lake Chad was 4m higher than at present [41]. Secondly, historical chronologies such as those of the Bornu Empire describe prosperous conditions with famine being very rare [42, ch. 2]. Thirdly, geographical descriptions by European travellers include reports of local peoples retaining memories of markedly more humid conditions [41, pp. 229].

The approximate nature of these historical rainfall estimates make definitive conclusions impossible. However, the humid climate of the 16th and 17th centuries make it very likely that uniform vegetation was present in areas currently exhibiting patterns. Since I have concluded that vegetation will then have subsequently died out, one can expect such an event to have occurred during the most severe subsequent drought, c. 1738–1756, a period of famine being very rare [42, ch. 2].

A central concept in the understanding of desertification is the bistability between vegetated states and desert [15, 16, 43]. As rainfall is decreased, a loss in stability of a vegetated state causes a sudden transition to desert, but if rainfall is subsequently increased back above the tipping point the desert state remains, and re-establishment of vegetation requires much laborious ground-based work, and consequently older data on wavelength-slope relationships is limited. Modern elevation databases eliminate the need for \( \text{situ} \) field work. Thus remotely sensed wavelength data can easily be complemented by slope gradients. My results indicate that such combined data is far more valuable than wavelength data alone, because it may enable one to infer the historical origin of the vegetation patterns.

Methods

Calculation of the Most Unstable Mode. When \( A \geq 2B \) the model (1) has two homogeneous vegetated steady states; one is always unstable, but

\[ u = u_s = \frac{A + \sqrt{A^2 - 4B^2}}{2B} \]

\[ w = w_s = \frac{2B^2}{A + \sqrt{A^2 - 4B^2}} \]

is stable to homogeneous perturbations provided that \( B < 2 \).

This restriction on \( B \) holds for all previous parameter estimates [14, 19]; for larger \( B \) (1) can have oscillatory dynamics which are never observed in reality. To determine linear stability of \((u_s, w_s)\) I substitute \((u, w) = (u_s + \tilde{u}, w_s + \tilde{w}) e^{\lambda t + i\omega t}\); the spatial frequency of the perturbation is \(k/2\pi\). Linearising in \((\tilde{u}, \tilde{w})\) and requiring non-trivial solutions gives a quadratic for \(\lambda\) with complex coefficients, whose solution yields an explicit formula for the growth rate \(\Re \lambda\) as a function of \(k\). This formula is a small extension of previous work [12, 30]. To determine the most unstable mode I calculated \(\Re \lambda\) over a grid of \(k\) values to give an initial approximation, and then used a numerical nonlinear equation solver to refine this as a solution of \(\left(\frac{\partial}{\partial k}\right) \Re \lambda = 0\).

To investigate the correlation between the slope and the wavelength (= 2\(\pi/k\)) of the most unstable mode, I considered the parameter range 0.05 \(\leq A \leq 5\), 0.025 \(\leq B \leq 1.75\), 5 \(\leq \nu \leq 900\), 10 \(\leq D \leq 800\), which are chosen to comfortably include all reasonable estimates [14, 19, 23]. I considered 100 equally-spaced values spanning each of these ranges, giving a total of 10⁵ parameter sets. Some of these do not satisfy the constraint \(A \geq 2B\), and for some others the most unstable mode is actually stable. For the remainder (about half) I calculated the change in the frequency of the most unstable mode following small changes in \(A\) and \(\nu\). In every case, frequency increases with \(A\) and decreases with \(\nu\).

Calculation of \(\nu_m\). As rainfall \(A\) is decreased, the stability of \((u_s, w_s)\) changes at a Turing-Hopf bifurcation point, \(A = A_{TH}\). Say, for \(\nu = 0\), calculation of \(A_{TH}\) reduces to that of a standard Turing bifurcation point for reaction-diffusion equation [12, 30, 46]. From this starting point, I numerically continued \(A_{TH}\) and the corresponding spatial frequency as solutions of \(\Re \lambda = (\frac{\partial}{\partial k}) \Re \lambda = 0\), while increasing \(\nu\). I performed this procedure for the same 100 values of \(B\) and \(D\) as used in § (10^4 cases in total). For small \(D\) there is no Turing bifurcation when \(\nu = 0\); specifically this occurs when \((u_s, w_s)\) is stable for \(\nu = 0\) and \(A = 2B\), the condition for which is \(BD < 2\). This corresponds to no vegetation patterns forming on flat ground. In fact, such patterning is common, although in the absence of spatial organisation by a slope one sees labyrinthine or spotted patterns rather than bands [1, 10]. Therefore these unrealistic parameter sets (0.25% of the total) can be discounted, although for completeness I comment that patterns then exist only for \(\nu\) above a non-zero value, and pattern wavelength decreases with \(\nu\). When \(BD\) is slightly greater than 2 (about 0.7% of cases) pattern wavelength is also a decreasing function of \(\nu\). These cases will also not be relevant in applications because the rainfall range giving patterns on flat ground is so small. In all of the remaining cases

* Studies of Lake Boosambe in Ghana [40] suggest that during the same period (16th and 17th centuries) there was a severe drought near the Guinea Coast, south of the Sahel. This study has been incorrectly described as referring to the Sahel in a number of popular science articles, including (at the time of writing) the Wikipedia page on the Sahel.
(about 99%) pattern wavelength increases with $\nu$ when this is small, reaches a maximum at $\nu = \nu_m$, and then decreases (see Fig. 1c). I calculated $\nu_m$ by quadratic interpolation on my grid of $\nu$ values. Fig. 3 shows $\nu_m$ as a function of $B$ and $D$. Typical estimates for the value of $\nu$ corresponding to slopes on which banded vegetation occurs are less than 200 [14,19], while most estimates for $D$ are at least 45,000. Fig. 3 suggests that a negative correlation between wavelength and slope is restricted to unrealistic parameter values, for patterns arising from degradation of uniform vegetation.

**Investigation of Pattern Generation by Colonisation of Bare Ground.** The critical rainfall level above which colonisation of bare ground occurs is determined by the change in movement direction of the lower edge of a vegetation patch, from uphill to downhill; in the physics literature this type of transition is known as a “Maxwell point”. I ran model simulations with initial vegetation density set to $u_s$ in the right-hand (uphill) half of the domain and zero in the left-hand half, with corresponding Dirichlet boundary conditions. After initial transients have dissipated, a transition front develops, moving with a constant speed that is positive/negative for smaller/larger values of $A$. Using a nonlinear equation solver, I calculated the value $A = A_c$, at which the speed is zero: this is the threshold rainfall level for colonisation. A guide to the appropriate range of $A$ values to consider is provided by the special case $\nu = D = 0$, for which it is possible to obtain exact solutions of the ordinary differential equations satisfied by a stationary transition front, $c$, and hence of $A_c$.

I then calculated the wavelength generated by colonisation when $A = A_c$. When rainfall slowly increases, this will be the wavelength of the first patterns to be established, which will then persist following further moderate changes in rainfall [20, 23, 34]. My procedure was to solve (1) numerically with $A = A_c$ and with $u$ set to $u_s$ in the centre of the domain (in a region of width arbitrarily chosen to be 200), and zero otherwise. Fig. 5 shows a typical example of the resulting solution. The left-hand (downhill) edge of the vegetated region remains stationary because $A = A_c$, while the right-hand edge propagates uphill. The resulting vegetated region can be either uniform or patterned. The division between these cases is illustrated by plotting $A_c$ against $\nu$, and superimposing the $\nu-A$ parameter regions in which patterns do/ do not form; calculation of these regions is described in §. Fig. 2a shows one such plot: patterns develop when the slope $\nu$ is below a threshold $\nu_p$, and Fig. 2b shows that their wavelength decreases with slope. This figure is typical except that for small values of $D$ colonisation always generates uniform vegetation rather than patterns. For example when $B = 0.45$ this occurs for $D$ less than about 10.

The procedure outlined above is quite expensive in computer time, making it unfeasible to loop over a fine grid of $B$ and $D$ values such as that used in [22]. Instead I considered 25 $B$-$D$ pairs: $B = 0.05, 0.45, 0.9, 1.4, 1.9$ and $D = 50, 200, 350, 500, 650$. The slight non-uniformity in the spacing of the $B$ values is deliberate in order to include 0.45, which is the most commonly used value in the other studies using (1). Fig. 4 shows the dependence on $B$ and $D$ of $\nu_p$, and also of the ratio $\nu_p/\nu_m$; note that this ratio is always greater than 1.

**ACKNOWLEDGMENTS.** I thank Professor S.E. Nicholson (Florida State University) for helpful advice, and Professor A.R. White (Heriot-Watt University) and Dr P.L. Wiener (University of Edinburgh) for careful reading of the manuscript.
Fig. 1. (a,b) Pattern generation from pre-existing unstable vegetation, with rainfall fixed. The expected wavelength (determined by the “most unstable mode”) is positively correlated with slope. (c) Pattern generation by degradation of uniform vegetation. Rainfall is varied to give the pattern onset point, and wavelength varies non-monotonically with slope. Parameters: (a-c): $B = 0.45, D = 500$; (a,b) $A = 2$; (a) $\nu = 100$.

Fig. 2. Colonisation of bare ground gives an inverse relationship between pattern wavelength and slope. (a) The critical rainfall level above which colonisation occurs (dots). The shaded region is that in which vegetation patterns form, so that above $\nu = \nu_p$, colonisation generates uniform vegetation. (b) The wavelength of patterns generated by colonisation for $\nu < \nu_p$. Parameters: $B = 0.45, D = 500$. 
Fig. 3. The dependence of the critical slope gradient $\nu_m$ on plant loss $B$ and water diffusion $D$. For patterns generated by degradation of uniform vegetation, wavelength is negatively / positively correlated with slope for $\nu$ greater / less than $\nu_m$. The scalebar shows the colour scale for $\nu_m$, which is deliberately nonlinear to give greater visual clarity. In the grey and black regions of the parameter plane (lower left-hand corner), wavelength always decreases with $\nu$. In the black region patterns exist for all $\nu \geq 0$, while in the grey region patterns only exist for $\nu$ greater than some non-zero minimum, so that there are no patterns on flat ground. The grey and black regions are separated by the curve $BD = 2$, which is the threshold for stability of $(u_s, w_s)$ when $A = 2B$.

Fig. 4. An illustration of when colonisation of bare ground results in banded rather than uniform vegetation. This occurs when the slope $\nu$ is below $\nu_p$, whose value is indicated by the coloured circles. The wavelength of these bands is always negatively correlated with slope (see Fig. 2b for a typical result). Inside the circles I give the value of $\nu_m/\nu_p$, showing that this always exceeds 1; this implies that whenever parameters are such that colonisation generates patterns, the degradation of bare ground generates patterns whose wavelength is positively correlated with slope. For $B = 0.05$ and $D = 50$, colonisation generates uniform vegetation for all slopes $\nu \geq 0$, and also degradation of bare ground always gives a negative correlation between wavelength and slope, so that neither $\nu_p$ or $\nu_m$ are defined; however these values are significantly outside typical estimates of parameter ranges.
Fig. 5. Pattern generation at the smallest rainfall level for colonisation of bare ground. Fixing rainfall $A$ at this critical level $A_c$ causes the left hand edge of the vegetation patch to be stationary. Parameters: $B = 0.45$, $\nu = 56$, $D = 500$, which imply $A_c = 1.67$. 