

Online Supplementary Material

Dispersal increases the resilience of tropical savanna and forest distributions

Nikunj Goel ^{1,*}, Vishweshha Guttal ², Simon A. Levin ³ and A. Carla Staver ¹

The American Naturalist

1. Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT, USA;
2. Centre for Ecological Sciences, Indian Institute of Science, Bengaluru, Karnataka, India;
3. Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA.

* Corresponding author; e-mail: nikunj.goel@yale.edu.

Table of Contents

Part 1: Models	3
Model 1: Savanna and forest model	3
Model 2: A simple catastrophic bifurcation model	3
Model 3: Harvesting model	4
Model 4: Light attenuation model	4
Part 2: Numerical methods.....	5
A numerical scheme for simulating reaction-diffusion in a 1D landscape	5
A numerical scheme for simulating reaction-diffusion in a 2D landscape	5
Part 3: Estimating Maxwell precipitation from empirical data	6
Data Sources	6
Estimating PM.....	7
Sensitivity analysis	8
Large-scale simulations	8
Part 4: Supplementary figures	10
<i>References</i>.....	12

Part 1: Models

In this paper, we use four bistable models to study the role of space in alternative stable states ecosystems. This allows us to probe the generality of our results for other bistable ecosystems and not just savannas and forests. Moreover, these models also allow us to capture ecological complexity of natural systems that has not been presented in the main text.

The first model (Eq. 1) is a simple savanna-forest model developed by Staver and Levin (2012) (see main text for details). The second (Eq. S1) and third (Eq. S2) models are single-variable dynamical models for which we can define a potential function (Strogatz 2014). The fourth model (Eqs. S3 and S4) is a two-variable coupled dynamical system with no defined potential function (Van Geest et al. 2007). We choose these models only for numerical simulations. Analytical results presented in the main text are robust to the choice of model.

Model 1: Savanna and forest model

See main text for full discussion.

Model 2: A simple catastrophic bifurcation model

The second model we consider is the so-called imperfect catastrophic bifurcation model. We use this model to simulate biome distributions for short- and long-range dispersal kernels (see Figs. 6, A4, and S4). This model has the following dynamical equation (Strogatz 2014):

$$\frac{dT}{dt} = -P + rT - T^3 = f(T, P), \quad (\text{S1})$$

where T is the state of the system (tree cover), and P is the control parameter (precipitation). The equilibria of this model are obtained by setting $f(T^*, P) = 0$. In this paper, we use $r = 3$ as the default parameter value for which the system exhibits three equilibria. We use T_S^* and T_F^* to denote stable equilibria of the system, corresponding to savanna and forest state, respectively; the roots are stable (unstable) if $df/dT|_{T=T^*} < 0$ (> 0). Furthermore, the critical precipitation (P_{FS} and P_{SF}) values of the system are obtained by simultaneously solving $f(T^*) = 0$ and $df/dT|_{T=T^*} = 0$. P_{FS} (or P_{SF}) = $\pm 2(r/3)^{3/2}$. We obtained the Maxwell precipitation (P_M) by setting $\int_{T_S^*}^{T_F^*} f(T, P) = 0$.

Using the mathematical conditions above, we obtained $P_{FS} = -2$, $P_{SF} = 2$, and $P_M = 0$. Please note that here the numbers are to be interpreted in a relative sense: Positive values of vegetation and precipitation correspond to the forest and higher precipitation regions, respectively. Similarly, negative

values of vegetation and precipitation correspond to savanna and lower precipitation regions, respectively.

Model 3: Harvesting model

$$\frac{dV}{dt} = rV \left(1 - \frac{V}{K}\right) - c \frac{V^2}{V_o^2 + V^2} + \sigma g(T, P) \eta(x, y, t) \quad (S2)$$

The third model we use in this paper is an insect outbreak population model used by Ludwig to analyze spruce budworm outbreak in eastern Canada (Ludwig et al. 1978; May 1977). This model has also been used to study vegetation dynamics in semiarid ecosystems (Guttal and Jayaprakash 2007; Guttal and Jayaprakash 2008; Noy-Meir 1975; van de Leemput et al. 2015). We use this model to study steady-state profiles of the vegetation ecotones under additive and multiplicative noise (Fig. S1). Here, V represents the state of the system, e.g., biomass. The model consists of two terms. The first term is a logistic growth function with r and K as the per-capita growth rate for small population sizes and carrying capacity of the system, respectively. The second term is the harvesting term with c as the maximum harvesting rate (control parameter) and V_o as the half-saturation constant. In this paper, we use $r = 1$, $K = 10$, and $V_o = 1$ as the parameter values. η is a standard normal random variable and σ is the standard deviation.

Using the above parameter values, we calculated the critical harvesting parameter to be $c_1 = 1.8$ (savanna to forest transition) and $c_2 = 2.6$. (forest to savanna transition). Furthermore, we found $c_M = 2.35$.

Model 4: Light attenuation model

$$\frac{dV}{dt} = r_v V \left(1 - V \frac{h_E^p}{E^p}\right) \quad (S3)$$

$$\frac{dE}{dt} = r_E E \left(1 - \frac{E}{E_o} \frac{h_v + V}{h^v}\right) \quad (S4)$$

Finally, we consider a two-variable macrophyte and turbidity model to study the role of dispersal in two-dynamical variable system that has no potential function (Fig. S2). This model has been used to describe lake-eutrophication in shallow lake water ecosystem (Van Geest et al. 2007). The two dynamical variable V and E are macrophyte and vertical light attenuation, respectively, with E_o as the control parameter of the system. The control parameter describes the vertical light attenuation in the absence of macrophytes. E_o is a proxy used in place of the nutrient loading in the lake. See Scheffer (2004) for further details on the model.

We use $h_E = 2$, $h_v = 0.2$, $P = 4$, $r_E = 0.1$ and $r_v = 0.05$ as default parameter values. Critical points of this model are calculated by setting the determinant of the Jacobian to zero (Strogatz 2014) ($E_{o1} = 5.2$ and $E_{o2} = 7.3$). This model has no potential function, and the Maxwell point is

dependent on the ratio of the diffusion coefficients (D_v/D_e ; Fig. S2) (van de Leemput et al. 2015).

Part 2: Numerical methods

All the numerical simulations were performed using R (version 3.2.2) and MATLAB (R 2015a). By default, in all the simulations, we used reflective boundary conditions. These simulation methods closely follow the methods described in Guttal and Jayaprakash (2009).

A numerical scheme for simulating reaction-diffusion in a 1D landscape

We performed numerical simulations for the 1D reaction-diffusion model using the following Euler forward-time scheme that is first-order accurate in time and second order accurate in space:

$$T_i^{t+1} = T_i^t + \Delta t [f(T_i^t, P_i) + D(T_{i-1}^t + T_{i+1}^t - 2T_i^t)], \quad (\text{S5})$$

where i represents the cell coordinate and t represents the time step. The simulations were truncated when $\sum_i |T_i^{t+1} - T_i^t| < 0.1$ because near equilibrium the change in tree density over successive time steps is very small. This numerical scheme was used to simulate Figs. 1 and 2.

A numerical scheme for simulating reaction-diffusion in a 2D landscape

Reaction-diffusion equation without noise

For 2-D reaction-diffusion model, using the following Euler forward time scheme that is first order accurate in time and second order accurate in space:

$$T_{i,j}^{t+1} = T_{i,j}^t + \Delta t [f(T_{i,j}^t, P_{i,j}) + D(T_{i-1,j}^t + T_{i+1,j}^t + T_{i,j-1}^t + T_{i,j+1}^t - 4T_{i,j}^t)], \quad (\text{S6})$$

where i and j represent the cell coordinates. Similar to the 1-D model, we truncated the simulations when $\sum_{i,j} |T_{i,j}^{t+1} - T_{i,j}^t| < 0.1$. This numerical scheme was used to simulate figures 3, 4, 6, and S2.

Reaction-diffusion equation with noise

$$\frac{\partial T}{\partial t} = f(T, P) + D \nabla^2 T + \sigma g(T, P) \eta(x, y, t) \quad (\text{S7})$$

To simulate the above stochastic reaction-diffusion equation, which is to be interpreted in an Ito-sense (Gardiner and others 1985), we used the following numerical scheme:

$$T_{i,j}^{t+1} = T_{i,j}^t + \Delta t [f(T_{i,j}^t, P_{i,j}) + D(T_{i-1,j}^t + T_{i+1,j}^t + T_{i,j-1}^t + T_{i,j+1}^t - 4T_{i,j}^t)] + \sigma \sqrt{\Delta t} g(T_{i,j}^t) \eta_{i,j}^t \quad (\text{S8})$$

where $\eta_{i,j}^t$ is a standard normal random variable and σ is the standard deviation. We truncated the simulation after $t = 10$. This numerical scheme was used to simulate figure S1.

Integro-differential equations

For a generalized kernel in equation (1) we used the following numerical scheme:

$$T_{i,j}^{t+1} = T_{i,j}^t + \Delta t (f(T_{i,j}^t, P_{i,j}) + D \sum_{i',j'} K(i-i', j-j') [T_{i,j}^t - T_{i',j'}^t]) \quad (\text{S9})$$

To reduce the computational time, we only evaluated the convolution integral (summation) for 95% confidence interval and then re-normalized the kernel to unity. The convolution was solved using *conv2* function in MATLAB. The simulations were truncated when $\sum_{i,j} |T_{i,j}^{t+1} - T_{i,j}^t| < 0.1$. This numerical scheme was used to simulate figure A4.

Part 3: Estimating Maxwell precipitation from empirical data

In the main text, we argued that, in a 1D homogeneous precipitation landscape, the boundary is stationary at a unique precipitation value—Maxwell precipitation (P_M). But this equilibrium precipitation value is unstable unless we assume a rainfall gradient. We showed analytically that if the gradient is shallow, the 1D diffusion model predicts that the boundary equilibrates at a spatial location where the rainfall value is P_M . Since P_M is an important parameter in our analysis, we present two methods of estimating P_M . All the data analyses were performed using R 3.2.2.

Data Sources

Precipitation:

We used monthly precipitation records derived from Tropical Rainfall Measuring Mission monthly precipitation product [TRMM 3B43; see Huffman and Bolvin (2013) for documentation]. To get a mean annual estimate for precipitation, we averaged the precipitation over a period of 10 years (January 1, 2000 to December 31, 2009), for Sub Saharan Africa ($20.5^\circ W - 53.75^\circ E$ and $20^\circ S - 21^\circ N$). The TRMM dataset is prepared using a combination of satellite data, precipitation gauges and modeled precipitation based on cloud cover. To estimate curvature of the savanna-forest boundary we projected the original precipitation raster of $0.25^\circ \times 0.25^\circ$ to a Lambert equal area precipitation raster (using *projectRaster* function) of resolution 25 km x 25 km (henceforth referred as low-resolution precipitation raster). For large-scale spatial simulations, we increased the resolution of the original precipitation raster ($0.25^\circ \times 0.25^\circ$) by a factor of 4 (using a *resample* function with bilinear interpolation). We then projected it to a Lambert equal-area projection (using *projectRaster* function), to get a precipitation raster of resolution 12.5 km x 12.5 km (henceforth referred to as high-resolution precipitation raster). This was done to resolve some stability issues while simulating biome distributions in Central Africa. All the cells corresponding to water bodies (like lakes and oceans) were omitted.

Vegetation:

We used a freely available global tree cover distribution product produced by Hansen et al. (2013), which reports relative tree cover distribution ranging from values between 0% (completely open canopy) to 100% (completely closed canopy) estimated from Landsat. We chose the minima between the two modes (savanna and forest) of tree cover in Sub Saharan Africa as the savanna-forest boundary, corresponding to a threshold (B_v) of 76% for the boundary, although our results were robust to variations to this threshold (Fig. S3). The vegetation raster was initially resampled to the same resolution of original precipitation raster ($0.25^\circ \times 0.25^\circ$). Then we followed the same steps described in the preparation of precipitation raster to get a low- and high-resolution vegetation raster.

Estimating P_M

To estimate the curvature of the savanna-forest boundary, we used the *contourLines* function to extract the cell coordinates of the boundary (B_v) from the low-resolution vegetation raster (a total of N pair of x and y cell coordinates). These coordinates were real values and not integers. The typical length of the boundary was the order of $10^3 - 10^4$ km. All the boundaries of length less than 150 km were ignored, as they were too small to estimate curvature. Then, for each boundary, we took a rolling window of segment length $2n + 1$ points (a pair of x and y cell coordinates) and estimated the radius of curvature (or curvature, the inverse of the radius of curvature) by fitting a circle to the cell coordinates of the segment using *lsfit.circle* function. The cell coordinates were then transformed to actual distances (1 unit of distance in cell coordinates = 25 km). To estimate the precipitation corresponding to this curvature value, we took the precipitation value corresponding to the midpoint of the segment ($(n + 1)^{th}$ point). Since the midpoint was not an integer, we rounded midpoints to closest integer values so that precipitation values could be extracted from the low-resolution precipitation raster. We then moved the rolling window by 1 point and repeated the same procedure. For the n points at the beginning (1 to n) and end ($N + 1 - n$ to N) of the boundary for which the curvature could not be estimated, we used the curvature values at $(n + 1)^{th}$ and $(N - n)^{th}$ points, respectively.

To estimate P_M , we fitted equation (A12) to the curvature and boundary precipitation data using *nls2* function (black V-shape curve in Fig. 5B). For better visualization, we grouped precipitation values into 11 precipitation intervals of 200 mm each, ranging from 800 mm to 3000 mm. For each of these precipitation intervals, we grouped the corresponding curvature values into one category and plotted the mean values of the curvature in each category against the mid-point of each precipitation interval (like 900 mm, 11000 mm and so on). We also estimated a 95% confidence interval for curvature values in each precipitation category (grey spindles in Fig. 5B).

In this paper, we provide a simple method to estimate P_M by exploiting the fact that the boundaries

Savanna-forest boundary dynamics

deviate linearly from the Maxwell contour as a function of the local curvature of the boundary (also see large-scale simulations below for an alternate method to estimate P_M). We argue that this method of estimating P_M improves upon the method proposed by Staal et al. (2016), which uses the concept of potential landscapes, for two reasons. First, their methods use only one-dimension to model space, which provides an incomplete picture of the actual dynamics of savanna and forest ecosystems. Here, we show that including a second dimension qualitatively changes the dynamics of the system. Second, they estimate P_M by reconstructing the potential function of the system using a mean-field approach. If one assumes that savanna-forest systems show strong spatial interactions, the potential function of the system is modified to (Malchow et al. 1983; Schlögl 1972)

$$U = \int \left(\frac{D}{2} (\nabla T)^2 - \int f(T, P) dT \right) dr, \quad (\text{S10})$$

and one can no longer use the mean-field potential function. Nevertheless, mean-field potential function provides valuable insights (see Appendix) and should be used, but with caution.

Sensitivity analysis

For estimating P_M , we used two parameters: the vegetation cover corresponding to the savanna-forest boundary (B_v) and the length of boundary segment (rolling window) for which the curvature values were estimated (n). To check whether our estimate of P_M was sensitive to specific values of the parameter, we computed P_M for 8 values of B_v (from 73 to 80 in steps of 1) and 20 values of n (from 6 to 25 in steps of 1). The sensitivity analysis showed that the estimate of P_M was robust to the choice parameters (B_v and n) and lay within the range of 1508 ± 84 mm MAP, with 1508 mm MAP and 1513 mm MAP as the median and mode of P_M estimates, respectively (Fig. S3).

Large-scale simulations

The model results indicate that the equilibrium position of the boundary in 2D is dependent on two independent parameters: P_M and the product of κ_{Mc} and D (Eq. 6). For a landscape with linear precipitation contours, the boundary always aligns with P_{Mc} irrespective of the choice of D and the underlying model. However, if the precipitation contours are curved, the boundary deviates from P_{Mc} , and these deviations are proportional to the product of the diffusion coefficient and a model dependent constant (Fig. 3B and Eq. 6). Exploiting this feature of the bistable reaction-diffusion model, we argue that the choice of the model will not affect the estimate of P_M , but it may not give us accurate length scales of spatial interactions.

We use imperfect catastrophic bifurcation model in our large-scale spatial simulations (Eq. S1). For a given value of P_M , we rescale the high-resolution precipitation raster (P_{old}) by the following relation

Savanna-forest boundary dynamics

$$P_{new} = \frac{P_{old} - P_{Mc}}{250}. \quad (S11)$$

We use this to rescale precipitation raster because we know from previous studies that the critical points corresponding to the forest to savanna and savanna to forest transition are around 1000 mm and 2000-2600mm, respectively (Aleman and Staver 2018; Staver et al. 2011). Furthermore, the curvature calculations indicate that P_M is around 1500mm (see the estimation of P_M from curvature and precipitation data). If we plug these estimates of critical values and Maxwell point in the above relationship (Eq. S11), we recover back the critical values and Maxwell point corresponding imperfect catastrophic bifurcation model (Eq. S1).

Next, we rescale the high-resolution vegetation raster (T_{old}) by the following relation

$$T_{new} = (T_{old} - B_v). \quad (S12)$$

In these simulations, we use $B_v = 76\%$. We use this to rescale vegetation raster because we know that the imperfect bifurcation model has a boundary at $T = 0$. In this rescaled vegetation raster (T_{new}), all the cells above (below) zero are forests (savannas). For a given value of P_M and D , we numerically simulate the distribution of savanna and forest using the rescaled precipitation raster (P_{new}) as the input (see numerical simulations). All the cells corresponding to water bodies in the simulated vegetation raster (T_{sim}) were reclassified as zero at every time step of the simulation since they neither act as savanna or forest.

Now we define a fitness function F such that $F =$ (number of forest cells in T_{sim} that are also forest in T_{new} + number of savanna cells in T_{sim} that are also savanna in T_{new}) - (number of forest cells in T_{sim} that are also savanna in T_{new} + number of savanna cells in T_{sim} that are also forest in T_{new}). In other words, the fitness function is maximum when the simulated distribution of forest and savanna exactly match the empirical data and is minimum when there is a complete mismatch. The simulated vegetation matrix T_{sim} is initialized with T_{new} (current distribution of savanna and forest).

Next, we use a Genetic Algorithm (Scrucca 2013) to find parameter values of P_M and D that maximize this fitness function. In the GA simulation, we used the following default parameter values: type = "real-valued", $D_{min} = 4$, $D_{max} = 100$, $(P_M)_{min} = 1400$ mm, $(P_M)_{max} = 1600$ mm, popSize = 10, Maxiter = 50 and parallel=T. We excluded all the areas in Western Africa and edaphic savannas of Bateke plateau from our analyses. Next, using the optimized parameter values of P_M and D , we simulated biome distributions for two other initial conditions: 'all savanna' (Fig. 6B) and 'all forest' (Fig. 6C).

Part 4: Supplementary figures

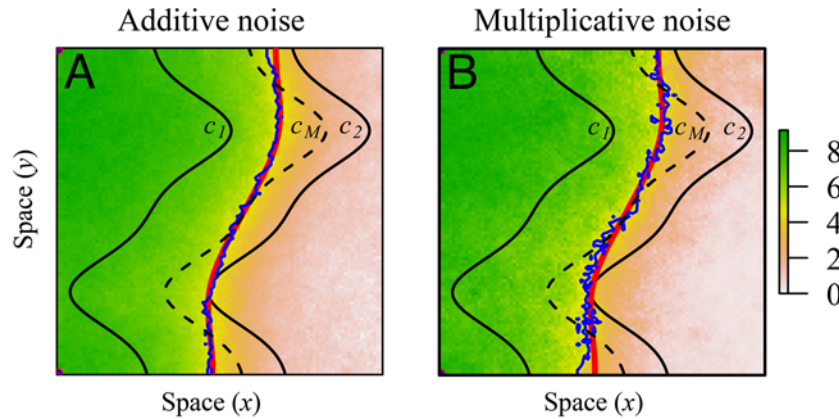


Figure S1: Boundary dynamics with (A) additive noise, $g(x) = 1$, and (B) multiplicative noise, $g(x) = N^2/(N_0^2 + N^2)$ (Harvesting model with noise term; equation S2). The blue and red line in plots represents the position of the boundary with and without the noise term, respectively. In both plots (A) and (B), we see that the addition of noise does not change the qualitative behavior of the system. Simulations were truncated after 80 time units. These simulations were performed using Harvesting model (see Models section above) on a 100x100 lattice. Here, we used $D = 50$ and $\sigma = 2$ (see Numerical Methods section above).

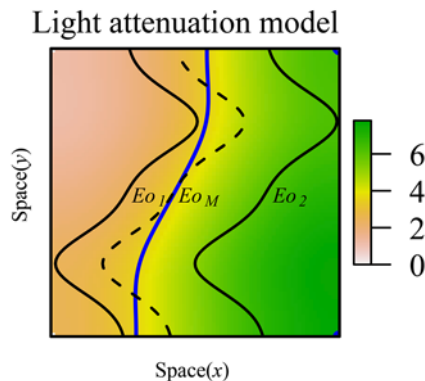


Figure S2: The simulation results from the two-variable light attenuation model (see Models section above) with no potential function is qualitatively similar to the one variable model. However, unlike the one variable model, the equilibrium position of the boundary between the two stable states is also dependent on D_v/D_E . These simulations were performed on a 100x100 lattice. Here, we used $D_E = D_v = 10$ (see Numerical Methods section above).

Savanna-forest boundary dynamics

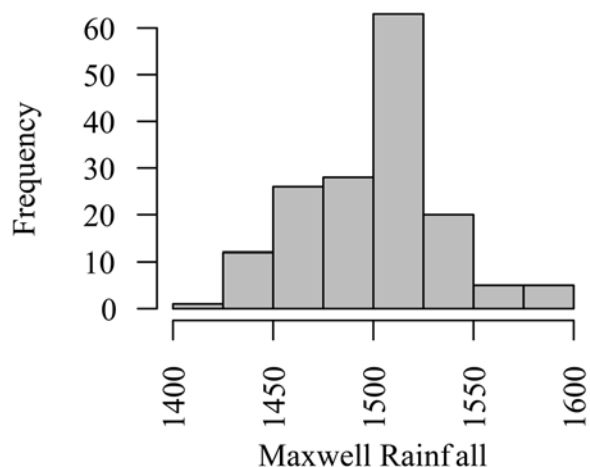


Figure S3: Sensitivity histogram of P_M estimated for various combinations of parameter values (Boundary Vegetation B_v and length of boundary segment used to calculate curvature) used in data analysis. The histogram suggests that the estimated value of P_M is robust to the changes in the parameter values and lies within 1508 ± 84 mm MAP, with 1508 mm MAP and 1513 mm MAP as the median and mode of P_M estimates (see Data Analysis section above).

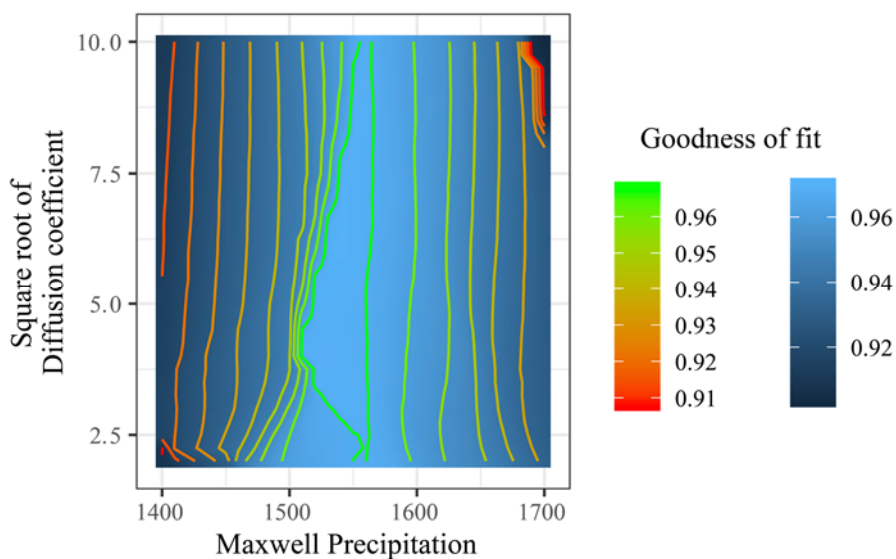


Figure S4: Comparing the simulated distribution of biomes (using a 2D reaction-diffusion model) with the current distribution of biomes in Sub Saharan Africa for various combinations of parameter values of P_M and \sqrt{D} . In these simulations, we excluded the contribution of the deforested regions in Western Africa and edaphic savannas of Bateke Plateau from the fitness function. The sensitivity analysis suggests that distribution of biomes in Sub Saharan Africa is relatively independent of D . Furthermore, the fitness function is maximum near $P_M = 1500 - 1550$ mm MAP, which is consistent with the estimate of P_M obtained from the curvature analysis in figure 5B (see Data Analysis section above).

References

- Aleman, J. C., and A. C. Staver. 2018. Spatial patterns in the global distributions of savanna and forest. *Global Ecology and Biogeography* 27:792-803.
- Gardiner, C. W., and others. 1985, Handbook of stochastic methods, v. 3, Springer Berlin.
- Guttal, V., and C. Jayaprakash. 2007. Self-organization and productivity in semi-arid ecosystems: implications of seasonality in rainfall. *Journal of theoretical biology* 248:490-500.
- . 2008. Changing skewness: an early warning signal of regime shifts in ecosystems. *Ecology Letters* 11:450-460.
- . 2009. Spatial variance and spatial skewness: leading indicators of regime shifts in spatial ecological systems. *Theoretical Ecology* 2:3-12.
- Hansen, M. C., P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau et al. 2013. High-resolution global maps of 21st-century forest cover change. *Science* 342:850-853.
- Huffman, G. J., and D. T. Bolvin. 2013. TRMM and other data precipitation data set documentation. NASA, Greenbelt, USA 28:1.
- Ludwig, D., D. D. Jones, and C. S. Holling. 1978. Qualitative analysis of insect outbreak systems: the spruce budworm and forest. *The Journal of Animal Ecology* 47:315--332.
- Malchow, H., W. Ebeling, R. Feistel, and L. Schimanskygeier. 1983. Stochastic Bifurcations in a Bistable Reaction-Diffusion System with Neumann Boundary-Conditions. *Annalen Der Physik* 40:151-160.
- May, R. M. 1977. Thresholds and Breakpoints in Ecosystems with a Multiplicity of Stable States. *Nature* 269:471-477.
- Noy-Meir, I. 1975. Stability of grazing systems: an application of predator-prey graphs. *The Journal of Ecology*:459--481.
- Scheffer, M. 2004, Ecology of shallow lakes, Springer Science & Business Media.
- Schlögl, F. 1972. Chemical reaction models for non-equilibrium phase transitions. *Zeitschrift für Physik A Hadrons and Nuclei* 253:147--161.
- Scrucca, L. 2013. GA: A Package for Genetic Algorithms in R. *Journal of Statistical Software* 53:1-37.
- Staal, A., S. C. Dekker, C. Xu, and E. H. van Nes. 2016. Bistability, Spatial Interaction, and the Distribution of Tropical Forests and Savannas. *Ecosystems* 19:1080-1091.
- Staver, A. C., S. Archibald, and S. Levin. 2011. Tree cover in sub-Saharan Africa: Rainfall and fire constrain forest and savanna as alternative stable states. *Ecology* 92:1063-1072.
- Staver, A. C., and S. A. Levin. 2012. Integrating theoretical climate and fire effects on savanna and forest systems. *The American Naturalist* 180:211-224.
- Strogatz, S. H. 2014, Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and

Savanna-forest boundary dynamics

engineering, CRC Press.

van de Leemput, I. A., E. H. van Nes, and M. Scheffer. 2015. Resilience of alternative states in spatially extended ecosystems. *PLoS One* 10:e0116859.

Van Geest, G. J., H. Coops, M. Scheffer, and E. H. Van Nes. 2007. Long transients near the ghost of a stable state in eutrophic shallow lakes with fluctuating water levels. *Ecosystems* 10:37--47.