

Inferring ecosystem parameters from observation of vegetation patterns

S. E. Thompson^{1,2} and G. G. Katul^{1,3}

Received 4 August 2011; revised 19 September 2011; accepted 20 September 2011; published 19 October 2011.

[1] Spatial organization of vegetation into periodic, coherent patterns arises from the interaction of positive and negative ecological feedbacks. Naturally, the patterns reflect the characteristics of the ecological processes that underlie their formation. Direct inference of the parameters describing these ecological processes from observations of vegetation spatial patterns has not been attempted. If successful, such inference can facilitate the parameterization and predictive use of vegetation pattern models. An inference technique based on nonlinear filtering is proposed here and applied to estimate the parameters of a single-equation phenomenological model of vegetation biomass patterning. Results derived from modeled biomass data indicate that for sufficiently accurate biomass observations (signal-to-noise ratios >4), and spatial resolution of better than 10% of the pattern wavelength, nonlinear filtering techniques recovered model parameters with high fidelity. When applied to real-world imagery, reasonable parameters within the pattern-forming regime were inferred. The study demonstrates, for the first time, the feasibility of inferring quantitative ecological information from spatial observations of vegetation distributions. **Citation:** Thompson, S. E., and G. G. Katul (2011), Inferring ecosystem parameters from observation of vegetation patterns, *Geophys. Res. Lett.*, 38, L20401, doi:10.1029/2011GL049182.

1. Introduction

[2] Periodic coherent vegetation patterns have a global distribution at desert margins, and arise from the superposition of facilitative interactions between plants at small scales with competitive interactions at large scales [Borgogno *et al.*, 2009]. Numerous theoretical models of these systems [Lefever and Lejeune, 1997; Rietkerk *et al.*, 2002; Gilad *et al.*, 2004], and an increasing cohort of field and remote sensing studies [Barbier *et al.*, 2006; Kefi *et al.*, 2007], suggest that they represent the vegetated state of a bistable region in which loss of vegetation cover could lead to locally irreversible desertification. Inferring the risk of such ecological collapse from observations of vegetation patterns is an emerging scientific imperative, but to date this inference has been qualitative and diagnostic, and reliant on observing changes in the morphology of the pattern as exogenous factors such as rainfall or grazing are altered [Barbier *et al.*, 2006; Kefi *et al.*, 2007;

Deblauwe *et al.*, 2011]. The main novelty in this study is to demonstrate the feasibility of directly estimating ecosystem parameters from the vegetation pattern using nonlinear filtering theory coupled to a suitable model of pattern formation. The results provide a needed link between field observations and theoretical models of vegetation patterns, setting the stage for improved data-model assimilation for ecological forecasting [Clark *et al.*, 2001]. The proposed approach is complementary to field-based work needed to assist with model selection and to constrain parameter estimates. To provide a tractable starting point, the biomass-only phenomenological model proposed by Lefever and Lejeune [1997, hereinafter LL97] is employed as a case study. The model arises from the sum of two kernels: one representing the spatial variation in facilitation, and one representing the spatial variation in competition between plants. The LL97 model offers (i) a tractable parameter space, (ii) directly observable state variables (biomass, P), (iii) a general formulation that is broadly consistent with a range of proposed pattern forming mechanisms, and (iv) predicts nondimensional biomass $P \in (0,1)$ thereby simplifying the interpretation of vegetation biomass observations from various remote sensing platforms.

2. Methods

2.1. Model and Modeled Data

[3] The parameter estimation can be simplified by applying two assumptions: (i) the perennial vegetation pattern exists in a pseudo-steady condition, and (ii) parameters estimated from 1D transects are representative of the full system. The first assumption is supported by the slow rates of change of vegetation patterns already reported in the literature [Barbier *et al.*, 2006; Deblauwe *et al.*, 2011], and the second by quantitative similarities between LL97 model predictions in 1 and 2 dimensions. With these assumptions, the LL97 model simplifies to an ordinary differential equation in space:

$$\frac{\partial^4 P}{\partial x^4} = 8(1 - \mu + (\Lambda - 1)P - P^2) + 4\left(\frac{L^2}{P} - 1\right) \frac{\partial^2 P}{\partial x^2} \quad (1)$$

where μ is a ratio of biomass growth to mortality rate; Λ is a ratio of facilitation to competition; L is a ratio of facilitative length scales to inhibiting length scales (approximately canopy radius to root zone radius); P a ratio of biomass to the potential peak biomass given local edaphic conditions, and x the spatial coordinate normalized by the inhibition length scale. All these parameters are nondimensional. The unsteady LL97 model was solved with $\mu = 1.01$, $\Lambda = 1.2$, $L = 0.1$, conditions that lie in the center of the model's pattern-forming regime. Zero-mean Gaussian noise with standard deviation (σ) set to 1% of the peak biomass (P_{\max}) was added to the solution to represent model process error. Three kinds of synthetic

¹Nicholas School of the Environment, Levine Science Research Center, Duke University, Durham, North Carolina, USA.

²Department of Civil and Environmental Engineering, University of California, Berkeley, California, USA.

³Department of Civil and Environmental Engineering, Pratt School of Engineering, Duke University, Durham, North Carolina, USA.

observations were crafted from the model output and used for parameter estimation: (i) additive zero-mean Gaussian error with $0.01 < \sigma < 1$, (ii) multiplicative log-normal errors with mean 1 and $0.01 < \sigma < 1$, and (iii) coarse-grained data formed by applying a moving-average window over 1% to 400% of the pattern wavelength (ω) before adding observation errors ($\sigma = 0.01$ and 0.1) to the coarse-grained data.

2.2. Inference Techniques

[4] Two methods were used to estimate the non-dimensional parameters. The first method, referred to as the ‘search algorithm’, solved equation (1) using factorial combinations of parameters, and identified the parameter set that minimized residuals between model predictions and observations. To improve the numerical stability, the model expression was linearized for three cases as:

$$\begin{aligned} \frac{\partial^4 P}{\partial x^4} &= 8(1 - \mu + (\Lambda - 1)P - P^2) - 4\frac{\partial^2 P}{\partial x^2}, & P \geq P_2 \gg L^2 \\ \frac{\partial^4 P}{\partial x^4} &= 8(1 - \mu + (\Lambda - 1)P - P^2), & P \geq P_1 \sim L^2 \\ P &= 0, & P < P_1 \ll L^2 \end{aligned} \quad (2)$$

Equation (2) was numerically integrated using a 4th order Runge-Kutta method, and $P_1 = 2.5L^2$ and $P_2 = 9.5L^2$. Root mean squared error (rmse) between this numerical integration and the full LL97 solution was < 0.02 . The rmse increased monotonically as parameters varied from their ‘true’ values. Confidence intervals around the minima were computed using a Chi-square test on the sum-of-square-deviations between observations and equation (2) predictions, allowing 95% confidence intervals to be independently determined for Λ , μ and L .

[5] The second estimation method was based on the unscented Kalman filter (UKF), which provides unbiased and efficient state estimates for highly nonlinear systems. The UKF has been extensively reviewed by *Julier and Uhlmann* [2004] and *Van der Merwe and Wan* [2001] and its details are not repeated here. Briefly, it differs from standard Kalman filtering approaches by approximating the nonlinear transformation of the first and second order moments of the state distribution via a set of ‘sigma points’ chosen to reproduce the moments of the state distribution. To use the UKF for parameter estimation, the parameters to be estimated were included in the state vector and its nonlinear update:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \\ \dot{x}_6 \\ \dot{x}_7 \end{bmatrix} = \frac{d}{dx} \begin{bmatrix} P \\ dP/dx \\ d^2P/dx^2 \\ d^3P/dx^3 \\ \Lambda \\ \mu \\ L \end{bmatrix} = \begin{bmatrix} x_2 + \epsilon_1 \\ x_3 + \epsilon_2 \\ x_4 + \epsilon_3 \\ f(x_1, x_3) + \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \\ \epsilon_7 \end{bmatrix} \quad (3)$$

where $f(x_1, x_3) = 8(1 - \mu + (\Lambda - 1)x_1 - x_1^2) + 4(\frac{L^2}{x_1} - 1)x_3$ and ϵ is the model process error. The sigma point generation algorithm proposed by *Van der Merwe and Wan* [2001], optimized for Gaussian model error, was used. Filtering the

series with additive error used a direct application of the method outlined by *Julier and Uhlmann* [2004]. The multiplicative noise case used a log transformation to convert the observation noise to additive Gaussian noise. The parameter estimates made by the UKF depend on the initial conditions for the parameters. The UKF was therefore applied to equation (3) over an ensemble of 200 initial parameter estimates. The parameter estimates converged over the first 50% of the biomass series, so the mean and standard deviation of the estimates for the last half of the observed series were used to compute the parameters. Additional details about the implementation of the method are provided as auxiliary material.¹

[6] The two inference techniques were used to explore three issues: (i) quality of parameter estimation as the signal-to-noise ratio (P_{\max}/σ) decreased; (ii) quality of parameter estimation with decreasing observation resolution, and (iii) ability of the filter to estimate consistent parameters from a $1 \text{ km} \times 0.75 \text{ km}$ aerial photograph of vegetation patterning. The aerial photograph used was taken in an area to the north of Parc W Niger, $12^\circ 56' 09.80'' \text{N}$, $3^\circ 19' 59.72'' \text{E}$. Details of the processing of the aerial photograph to generate parameters for analysis are also provided as auxiliary material. The observations were normalized by the peak absorbance in the image, and biomass was assumed to scale linearly with this normalized value (see auxiliary material). The spatial scaling was provided by the $8m$ root zone extent measured in vegetation patterns in Parc W, which was assumed to provide a reasonable estimate [*Barbier et al.*, 2008].

3. Results

3.1. Effect of Decreasing Signal-to-Noise Ratio and Spatial Resolution

[7] The performance of the UKF and the search algorithm approaches in estimating the parameters used to generate the synthetic data is shown in Figure 1 for a range of signal-to-noise ratios for both multiplicative and additive noise. It is evident that the UKF performed well for a wide range of signal-to-noise ratios, and for both kinds of observation errors. Confidence intervals on the predictions (omitted for clarity) were small ($< 5\%$) for all signal-to-noise ratios here. The Λ and μ were estimated with greater reliability than L , although in part this reflects the small value of L . By comparison, the search algorithm performed well for multiplicative noise at high signal-to-noise ratios (> 20), below which the error in predicted L increased dramatically, exceeding 50% at a signal-to-noise ratio of 15. Predictions of Λ and μ remained reasonable for signal-to-noise ratios > 20 , below which they became biased, erratic, and imprecise (confidence intervals of $\sim 50\%$). The search algorithm performance with additive noise was poor in all but the least noisy cases, and again became biased, erratic and imprecise at signal-to-noise ratios of < 20 . The analysis was repeated for contrasting parameter choices, and produced similar results (not shown).

[8] The coarse-grained model data ranged in resolution from 0.01ω to 4ω . The UKF performed well when applied to coarse grained datasets contaminated with multiplicative noise, generating errors of less than 5% for μ and Λ until the averaging window approached ω . The estimation of L was

¹Auxiliary materials are available in the HTML. doi:10.1029/2011GL049182.

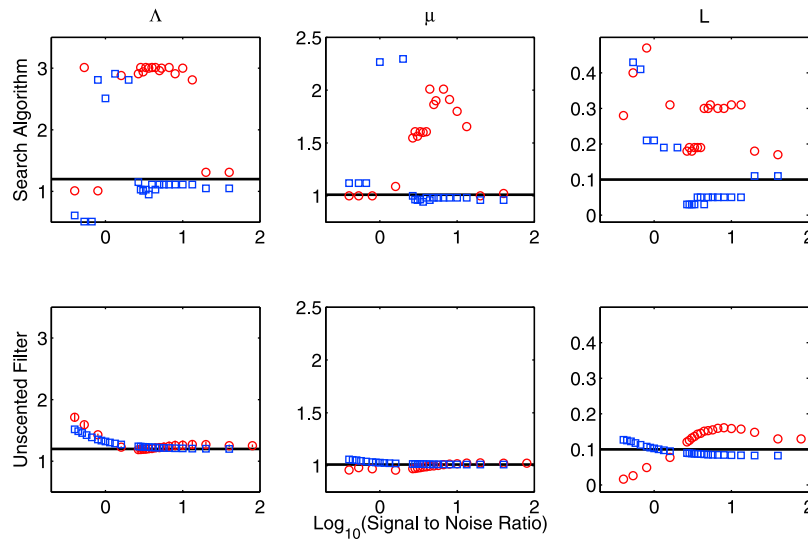


Figure 1. Performance of (top) the search algorithm and (bottom) the UKF for estimating the parameters (left) Λ , (middle) μ and (right) L , as a function of the signal-to-noise ratio in the observations (shown in log scale). Red circles indicate the parameter estimates assuming additive noise, while blue squares indicate the estimates made with multiplicative noise. The black line indicates the true parameter value. The UKF produced reasonable parameter estimates for signal-to-noise ratios greater than ~ 2 , below which the estimates became biased. The search algorithm performed reasonably well for multiplicative noise with signal-to-noise ratios greater than ~ 4 , below which predictions became erratic. Except on very clean datasets, the search algorithm performed poorly for additive noise.

sensitive to the coarse-graining, with errors increasing as the averaging window exceeded 0.1ω . Results were comparable for the clean and noisy datasets. The additive noise filter performed worse for the noisy dataset than the clean dataset at a given resolution. The sensitivity of the filter to coarse graining was similar for the two datasets, with parameter estimation becoming biased when the smoothing window approached 0.25ω . The results are shown in the auxiliary material.

3.2. Estimation of Parameters From Aerial Photograph

[9] The normalized absorbance (biomass) data, parameter estimates and an example of the LL97 model predictions using the estimated parameters (assuming multiplicative noise) are shown in Figure 2. The UKF output was consistent

across multiple transects, taken in multiple directions through the patterned image, as evidenced by the relatively constrained standard deviations of the parameter estimates: $\Lambda = 1.70 \pm 0.05$, $\mu = 0.108 \pm 0.04$ and $L = 0.02 \pm 0.02$. The estimated parameters lie within the pattern forming regime of the LL97 model (a necessary condition for their formation). However, the parameters in this case lie close to the pattern forming bifurcation so that the lower limit of the estimated value of μ lies in a region where homogeneous vegetation cover is stable. The Parc W vegetation patterns are the southernmost occurrence of vegetation patterns in the Sahelian region [Barbier *et al.*, 2006], and since the site studied here is likely to be somewhat drier and more impacted by grazing, these results are consistent with the ecological context of the study site. In contrast, repeating the analysis with the UKF optimized for

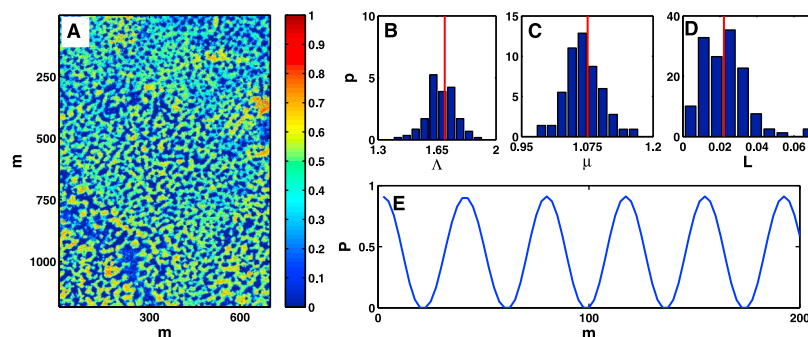


Figure 2. Parameter estimation from an aerial photograph of vegetation patterns in Niger. (a) Normalized biomass spatial distribution, determined from the total image intensity in a Google Earth image. (b–d) Estimates of parameters Λ , μ , and L obtained from sampling multiple horizontal and vertical 1D transects from Figure 2a. Mean parameter estimates shown in red. (e) LL97 1D model output with the mean parameter estimates from Figures 2b–2d, demonstrating that the estimated parameters lie in the pattern forming regime.

additive noise generated inconsistent and unrealistic estimates of the LL97 parameters.

4. Discussion and Conclusions

[10] These results provide a proof of concept of the feasibility of extracting quantitative information about small-scale ecosystem properties from observation of large-scale spatial biomass distributions via spatial patterning. Necessary conditions for successful parameter estimation were resolutions of better than 0.1ω , and observation errors of less than $\sim 25\%$ of P_{\max} , suggesting that high resolution imagery is needed in such estimation exercises. The form of the vegetation model used was reflected in the quality of the parameter estimates. For instance, L was the most poorly estimated parameter in all cases. This can be attributed to two sources: (i) direct observation of the effects of L is confined to cases when P is small, and is therefore challenging to observe, and (ii) the effects of L largely influence the biomass peaks by determining the spatial gradients of P at the edges of vegetation patches. Observation of these gradients is sensitive to small errors. The improved estimation for the multiplicative noise case was partly due to this sensitivity because multiplicative noise reduces the absolute error at small biomass values relative to additive noise. Multiplicative noise also avoids the bias associated with truncating estimates of $P < 0$ ($P < 0$ arises when assuming additive noise for small P), again improving the filter performance at small P and the estimates of L . These factors contribute to the improved precision and reliability of the estimates from the multiplicative filter applied to the aerial photograph. However, the estimate $L = 0.02$ for the Parc W patterns suggests that the length-scale over which facilitative interactions occurs is on the order of only 15 cm (40 cm at most given the error in L), smaller than tree canopy radii. Whether this indicates a systematic under-estimation of L , or the possibility that facilitative interactions are restricted to scales smaller than the canopy radius, is unclear.

[11] Extending this approach from well-studied sites like Parc W to other locations is straightforward provided the challenges associated with normalizing biomass and the spatial terms in the LL97 model can be overcome. Biomass normalization may benefit from coupling larger-scale estimates (e.g., MODIS imagery) with high resolution photography to ensure that a reasonable estimation of the saturating biomass condition (e.g., NDVI) can be made. Research to evaluate the NDVI - biomass relationships in patterned sites would be valuable. Normalizing the space terms by the inhibition lengthscale (root zone) may be challenging for areas where root measurements are unavailable, although literature estimates are available for many pattern forming species. Alternatively, a phenomenological model has been proposed in which length scales are normalized by the plant

canopy [Lefever *et al.*, 2009]. Similar filtering methods can also be applied to physically based models in which all parameters are explicit. For instance, the UKF performed well in estimating the parameters of a coupled water - biomass model [Rietkerk *et al.*, 2002] from synthetic data provided observations of soil moisture were incorporated in the filter (not shown here). Obtaining such observations, both in terms of a mean soil moisture estimate for the root zone or an average soil moisture condition through time, however, remains at the limits of contemporary remote sensing technology.

[12] **Acknowledgments.** The National Science Foundation (NSF-AGS-1102227, NSF-EAR-10-13339, and NSF-CBET-103347), the U.S. Department of Energy through the Office of Biological and Environmental Research (BER), and the U.S. Department of Agriculture (2011-67003-30222) supported this work.

[13] The Editor wishes to thank two anonymous reviewers for their assistance evaluating this paper.

References

- Barbier, N., P. Couteron, J. Lejoly, V. Deblauwe, and O. Lejeune (2006), Self organized vegetation patterning as a fingerprint of climate and human impact on semi-arid ecosystems, *J. Ecol.*, *94*, 537–547.
- Barbier, N., P. Couteron, R. Lefever, V. Deblauwe, and O. Lejeune (2008), Spatial decoupling of facilitation and competition at the origin of gapped vegetation patterns, *Ecology*, *89*(6), 1521–1531.
- Borgogno, F., P. D'Odorico, F. Laio, and L. Ridolfi (2009), Mathematical models of vegetation pattern formation in ecohydrology, *Rev. Geophys.*, *47*, RG1005, doi:10.1029/2007RG000256.
- Clark, J. S., *et al.* (2001), Ecological forecasts: An emerging imperative, *Science*, *293*(5530), 657–660, doi:10.1126/science.293.5540.657.
- Deblauwe, V., N. Barbier, P. Couteron, O. Lejeune, and J. Bogaert (2011), Environmental modulation of self-organized periodic vegetation patterns in Sudan, *Ecography*, doi:10.1111/j.1600-0587.2010.06684.x, in press.
- Gilad, E., J. von Hardenberg, A. Provenzale, M. Shachak, and E. Meron (2004), Ecosystem engineers: From pattern formation to habitat creation, *Phys. Rev. Lett.*, *93*(9), doi:10.1103/PhysRevLett.93.098105.
- Julier, S., and J. Uhlmann (2004), Unscented filtering and nonlinear estimation, *Proc. IEEE*, *92*(3), 401–422.
- Kefi, S., M. Rietkerk, C. Alados, Y. Pueyo, V. Papanastasis, A. ElAich, and P. de Ruiter (2007), Spatial vegetation patterns and imminent desertification in Mediterranean arid ecosystems, *Nature*, *449*(7159), 213–218.
- Lefever, R., and O. Lejeune (1997), On the origin of tiger bush, *Bull. Math. Biol.*, *59*, 263–294.
- Lefever, R., N. Barbier, P. Couteron, and O. Lejeune (2009), Deeply gapped vegetation patterns: On crown/root allometry, criticality and desertification, *J. Theor. Biol.*, *261*(2), 194–209, doi:10.1016/j.jtbi.2009.07.030.
- Rietkerk, M., M. C. Boerlijst, F. van Langevelde, R. HilleRisLambers, J. van de Koppel, L. Kumar, H. H. T. Prins, and A. M. de Roos (2002), Self-organization of vegetation in arid ecosystems, *Am. Nat.*, *160*(4), 524–530.
- Van der Merwe, R., and E. Wan (2001), The square-root unscented Kalman filter for state and parameter estimation, in *International Conference on Acoustics, Speech and Signal Processing*, vol. 6, pp. 3461–3464, Inst. of Electron. and Electr. Eng., New York.

G. G. Katul and S. E. Thompson, Nicholas School of the Environment, Levine Science Research Center, Duke University, Durham, NC 27707, USA. (gaby@duke.edu; set8@duke.edu)