Clustering and correlations: Inferring resilience from spatial patterns in ecosystems
SUPPLEMENTARY MATERIAL

Sumithra Sankaran,[1,*] Sabiha Majumder,[2]
Ashwin Viswanathan,[3] and Vishwesha Guttal[1]

[1]*Centre for Ecological Sciences, Indian Institute of Science, Bengaluru, India*
[2]*Institut für Integrative Biologie, ETH Zurich, Zurich, Switzerland*
[3]*Nature Conservation Foundation, Bengaluru, India*

CONTENTS

* sumithras@iisc.ac.in

Appendix A: Detailed model description

The system is modelled as a two dimensional discrete space consisting of N x N cells. Each cell can occur in any of two possible states: unoccupied (or dented by 0) and occupied (or denoted by 1). The states of each cells are updated probabilistically and the simulation sequence runs as described below.

Step 1: Select a cell at random.

Step 2: If it is unoccupied, return to step 1. If occupied, proceed to step 3.

Step 3: Select at random, one of the four nearest neighbors of the chosen cell. If the selected neighbor cell is unoccupied proceed to Step 4. Else, skip to Step 5

Step 4: Select a random value between 0 and 1. If the value is less than p, update the state of the selected neighbor cell from 0 to 1, else, update the state of the first chosen cell to 0 (the probability of this is then $1 - p$). Return to step 1 for a new iteration.

Step 5: Select a random value between 0 and 1. If the value is less than q, select one of the six possible nearest neighbors of the pair and make it occupied (if it was already occupied, nothing changes). Else, update the first chosen cell (in step 1), to 0

The cells are, thus, updated asynchronously. In each discrete time step, the update rules are iterated N^2 number of times sequentially; this procedure ensures that, on average, all sites are updated once per discrete time step.

We ran our simulations with a 1024×1024 system size at increasing resolutions of p and q to construct the steady-state phase diagrams of mean density of the landscape (Fig 3 of the main text). We then identify the critical driver values: The simulations were run for the entire parameter space of 0 to 1 for p as well as q. This was done for a base resolution of 0.01 of q and 0.001 of p. For the specific values of q ($q = 0$ and $q = 0.92$) where precise values of critical points/thresholds were required, the simulations were run at a resolution of 0.00001 for p.

The system was considered to have reached a steady-state once the mean density had saturated (typically in 10000 or less time steps away from critical points/thresholds; around 1 million time steps close to critical points/thresholds). We then compute density, i.e. proportion of occupied cells in the landscape (Fig 3 of the main text).

To compute percolation probabilities (Fig 4 of the main text), we ran another set of simulations with a system size of 256×256 and obtained 25 replicates for $q = 0, q = 0.92$ and the spatial null model. We compute the percolation probability as the fraction of these replicate snapshots had a spanning cluster.

To compute cluster-size distributions (Fig 6 of the main text) and power-spectrum functions (Fig 7 of main text), we simulate the model with a 1024×1024 system size and obtain 50 replicate spatial snapshots at steady-state. For each replicate, we use equivalence class algorithm to identify clusters (see Glossary for the definition of a cluster). Based on all replicates, we compute probability

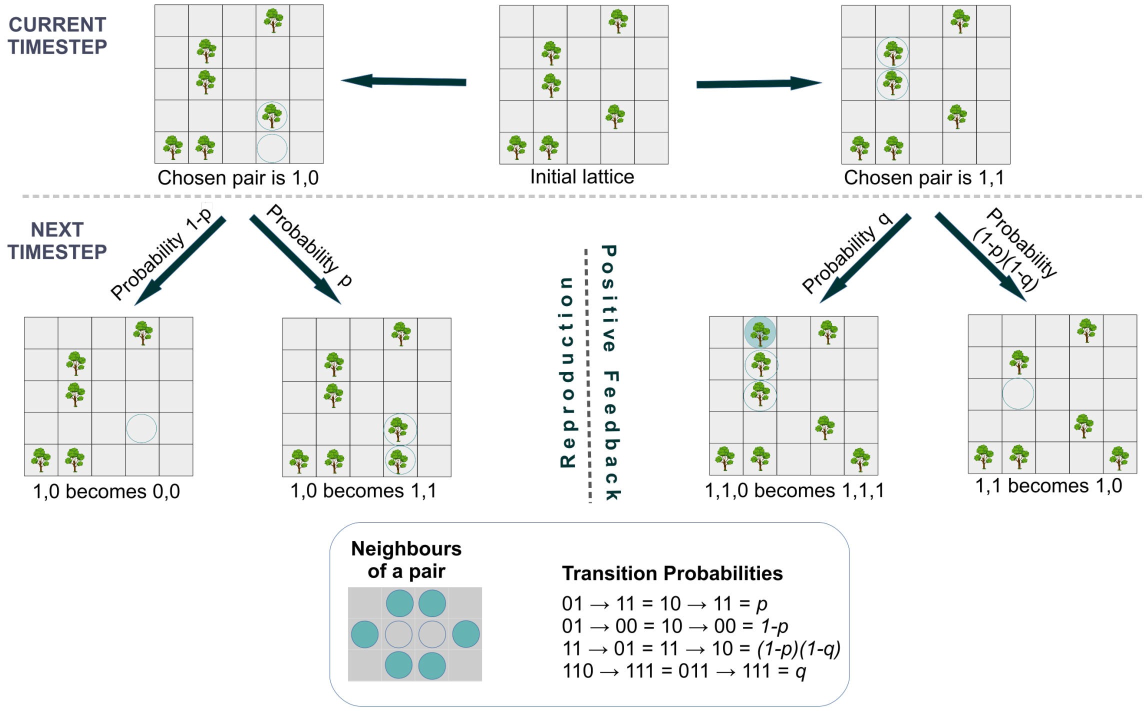


Figure 1: Schematic representation of the model and simulation procedure, for a given 'Initial lattice' shown at the centre of the top row. The parameter p represents baseline birth rate whereas q represents the strength of local positive feedback; reducing p in this model can be interpreted as increasing environmental stress. Light blue circles represent (randomly) chosen cells to update. Depending on the states of chosen cells, the update scheme results in baseline birth or death (left part of second row), or increased birth or reduced death due to positive feedback (right part of the second row). The box at the bottom shows (i) neighbours of a focal pair of cells and (ii) model update rules captured via transition probabilities.

distribution function (pdf, denoted by $P(s)$) and cumulative distribution function (cdf, denoted by $C(s)$) of cluster-sizes; we note that $C(s) = \int_{-\infty}^{s} P(s')ds'$. Inverse cdf, plotted in Figure 6 of main text amd Figure 2 in Appendix C, is defined as $1-C(s)$. See Appendix B for details of statistical fitting procedures for cluster size distributions. Also see Appendix C for cluster-size distributions for a number of values of p and q; here we used a system size of 256×256 with 25 replicates.

To compute power-spectrum (Fig 7 of the main text), we simulate the model with a 1024×1024 system size and obtain 50 replicate spatial snapshots at steady-state. For each replicate, we first take the absolute value of the Fast Fourier Transform (FFT) of the entire landscape ($N \times N$ matrix). To obtain the power at a given length of the wave number k, we perform an angular average of the resulting two dimensional FFT of the landscape; we average over all replicates. See Appendix D for the definitions related to autocovariance function and power spectrum, and see Appendix E for statistical fitting procedures of power-spectrum.

All of our code is available via github: https://zenodo.org/badge/

latestdoi/112318349

Appendix B: Statistical fitting of cluster-size distributions

Cluster size distributions were fit using methods proposed in [4]. It is worth recalling the methods of fitting power-laws have attracted much scrutiny in the literature. Therefore, we chose the statistical methods proposed by [4] which are widely accepted as rigorous to fit power-law distributions and to compare with other model candidates. We refer readers to [4] for further technical details. Readers may reproduce all of our results following the broad steps described below together with code is available on our github page: `https://github.com/ssumithra/PowerLawCriticalityPaper` (also see [5] for an R-package called *spatialwarnings*).

Is power-law a good fit?: First step in the process is to find out if power-law is even a good fit. The exponent of the distribution was estimated using Maximum Likelihood Estimation (MLE), and x_{min} was identified by minimising the Kolmogorov–Smirnov (KS) distance between the fitted model and data. We assessed goodness of fit for our power-law model by re-fitting synthetic power law distributions (which we generated) with the same estimated exponent and x_{min} values. The fraction of synthetic datasets that result in a fitted model with a KS distance larger than the KS distance calculated when fitting our dataset, was considered the p-value of our fit. As described in [4], a p-value above 0.1 represents a good fit, and only when this condition was satisfied, we proceeded to compare with alternative models of cluster size distributions. Indeed, we found p-values of 0.49 for the low positive feedback model ($q = 0$) [see Fig 6a in the main text] and 0.53 for the high positive feedback model ($q = 0.92$) [see Fig. 6b in the main text]. This suggests good power-law fit of cluster size distributions for both values of positive-feedback, but one of them is away from critical point (Fig 6a) and the other is right at the critical threshold (Fig 6b).

Is power-law the best fit?: We compared power-law (PL) fit of the cluster size distributions with three different model fits: exponential (EXP), log-normal (LN) and power-law with an exponential cut off (PLE). Each of the candidate models was fit using MLE. Since power-law is a nested model of power-law with a cut-off, these two were compared using log-likelihood ratio. The other two candidate models were compared with the power-law model using Vuong test.

We know from percolation models that as density in the system reduces, cluster size distributions typically show the following trend: a bimodal distribution, a power-law distribution, a power-law with exponential cut-off distribution and finally an exponential distribution [7, 9]. Thus, in our investigations of effects of positive feedbacks on clustering, these were obvious candidate distributions to fit to the data. In addition, we also fit and compare log-normal as a candidate function in order to make the reported results comparable with other studies discussed in [4]. However, at least to our knowledge, there exists no mechanistic processes that can yield a log-normal cluster size distribution in these models . On the other hand, based on the theory of phase transitions, there is a well-reasoned expectation of scale-free behaviour at critical points. Therefore, we do not consider log-normal distributions for our interpretations below.

Realising a true scale-free distribution requires, ideally, an infinitely large system. Our datasets consist of 50 replicates of identically sized lattices (of 1024 x 1024 cells). Even a true power-law distribution in these replicates would inevitably be best fit as a truncated power-law distribution, due to the limit imposed by the system-size. Given these finite size constraints in our data, a power-law as the best fit is inferred based on (a) the closeness of estimated parameters between the fitted power-law function and fitted power-law with exponential cut off function and (b) the range over which the power-law dominates the truncated power-law function.

For the data presented in Fig 6 (a & b) of the main text, power-law with exponential cut off was identified as the best fit model for both cases ($q = 0, p = 0.7225$ and $q = 0.92, p = 0.2852$). Given in Table 1 is the comparison statistics for the power-law (PL) fit with the other three considered models - exponential (EXP), Power-law with exponential truncation (PLE) and Log-normal (LN) fits. Based on the Table, for both datasets, log-normal could not be ruled out as a potential fit (p value given in brackets), while power-law with exponential cut-off was found to be the best fit.

Dataset	PL Vs Exp vuong test statistic	PL Vs PLE log-likelihood ratio	PL Vs LN vuong test statistic
q=0, p=0.7225	34.31***	-53.92 ***	-2.69 (p=0.996)
q=0.92, p=0.2852	41.33608 ***	-9.71**	-15.22 (p=0.999)

Table 1: Results of likelihood ratio test for fitted power-law vs other models. Positive values suggest that the power-law is a the better fit and negetive values favour the alternative model. Significance levels are as follows : '***' for $p < 0.001$, '**' for $p < 0.01$ '*' for $P < 0.1$ and '' ' for $P > 0.1$. For both datasets, log-normal could not be ruled out as a potential fit (p value given in brackets), while power-law with exponential cut-off was found to be the best fit.

Dataset	Fitted model	Exponent	Rate	X_{min}
q=0, p = 0.7225	Power law with exponential cut off	1.84	9.3×10^{-6}	17
	Power law	1.83		17
q=0.92, p = 0.2852	Power law with exponential cut off	1.75	2.9×10^{-6}	3
	Power-law	1.77		3

Table 2: Estimated parameter values of fitted power-law and power-law with exponential cut-off functions. The exponent estimates are very close for both model fits. The rate of exponential cut-off is very low, suggesting that the power-law persists over a large range of patch sizes.

We now compare parameters of the PL and PLE fits. The fitted PLE, $x^{-\beta}e^{-x/X_{max}}$, has a rate $1/X_{max}$ that is inverse of the largest patch-size X_{max}. We see from the estimated rate (Table 2), that the X_{max} is larger than the largest patch-size in the system. The power-law behaviour persists without any effect of the exponential truncation till at least four orders of magnitude in patch sizes. Hence, the observed best-fit of PLE could entirely be a consequence of finite-sizes we have used. Therefore, for all practical purposes, it is reasonable to interpret the observed pattern as a power-law distribution.

We also note the following: There is only a single point (i.e. one specific value of the driver or the density) where power-law cluster size distribution is observed. At densities higher than that, the clustering pattern shows a bimodal distribution of sizes and at lower densities, a truncated power-law distribution. Previous studies (such as [6, 7]) have dubbed the region where a bimodal distribution occurs with the point where power-law scale-free clustering emerges and have called that entire parameter space as within the power-law regime. This is because the first mode of the bimodal size distribution may indeed show power-law decay; however, the existence of bimodality in such cases may correspond to a characteristic size. Since the focus of our manuscript is only on the scale-free behaviour of the system with no characteristic size, we only consider the percolation density but not the bimodal region. Hence, only one point in our parameter space (or one density) shows scale-free power-law clustering.

Appendix C: Cluster-size distributions

Figures 2 and 3 demonstrate that trends in cluster size distributions can vary depending on the strength of positive feedback (q) (see section III-B of the main text for detailed context and explanations of the results.)

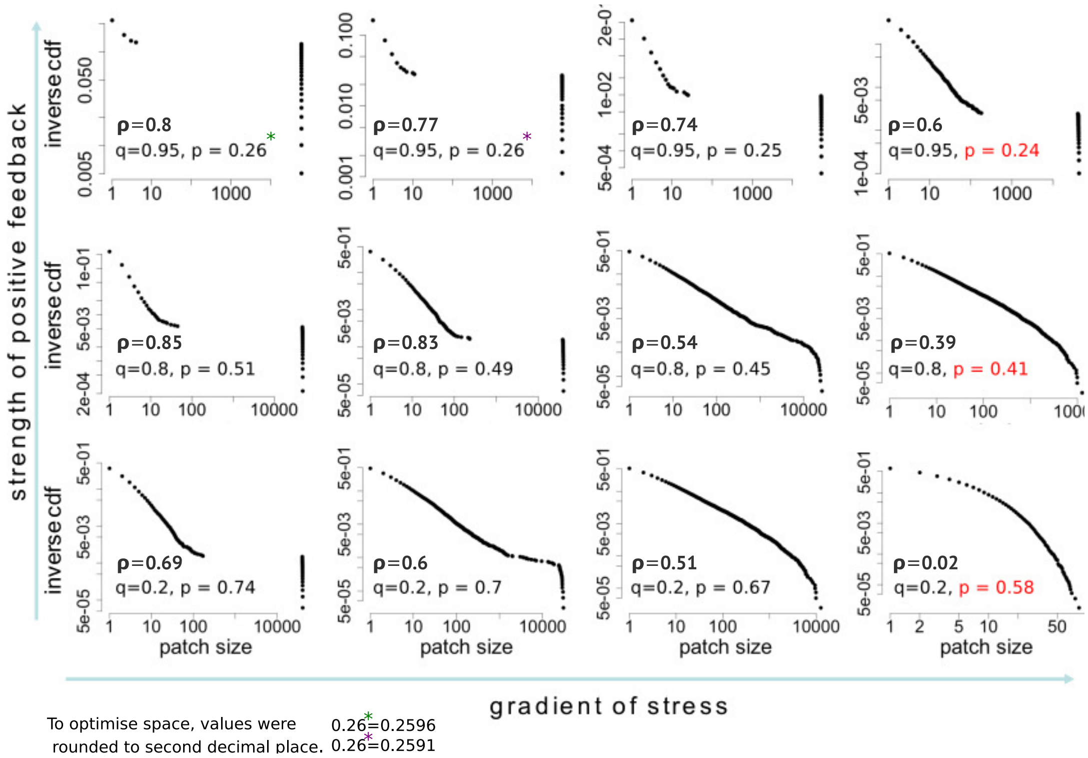


Figure 2: Trends in cluster size distributions depend on the strength of positive feedback. The right most column of plots is of systems very near/at critical points/thresholds, with the driver value shown in red. **The lower row from left to right** shows that when positive-feedback is low (($q = 0.2$), we see the entire range of cluster-size distributions: bimodal at very high density, far from the critical point to a power-law to a truncated power-law to an exponential, very near the critical point. **The middle and upper rows** show that as positive-feedback increases, 1) high densities prevail for lower p values, 2) system begins to collapse from higher density states (see Fig 3 of the main text). We then see power-law or, with very high positive-feedback, even bimodal clustering at the critical threshold of collapse. System size of 256×256 was used for these graphs

Appendix D: Effect of positive feedback on percolation and critical thresholds/points.

In the main text (section III-C) we discussed how the percolation density and density at critical points/thresholds come closer with positive feedback, till they begin to overlap. Thus with high positive feedback in the system, scale-free clustering can occur closer and closer to the critical point/threshold and early warning signals in cluster size distribution patterns may thus fail. Here we show that the same result, in terms of the driver (p). It can be seen from Fig. 4 that as the

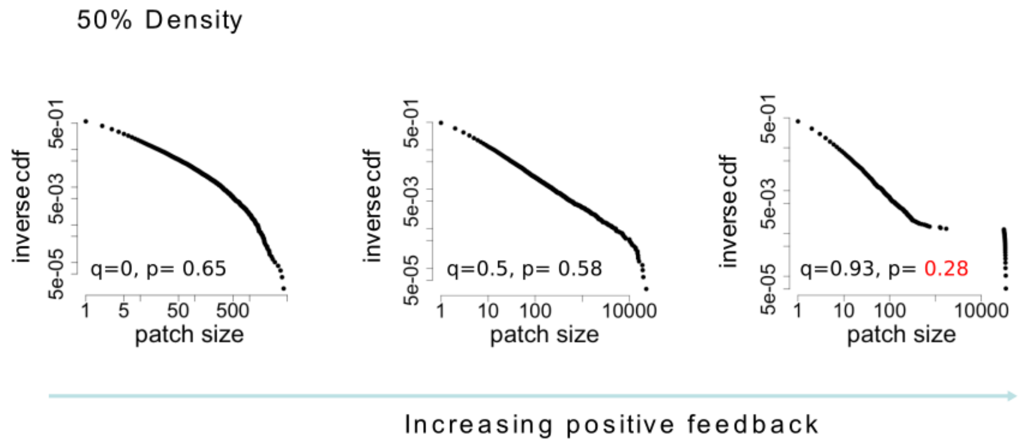


Figure 3: At the same density (50% in the above graphs), different cluster size distributions can be observed depending on positive-feedback strength (q). Since positive-feedback also results in the system collapsing from higher densities, at high positive-feedback values, fat tailed distributions occur closer to the critical threshold. System size of 256×256 was used for these graphs.

positive feedback in the system increases, the percolation point moves closer and closer and eventually overlaps with the critical point/threshold.

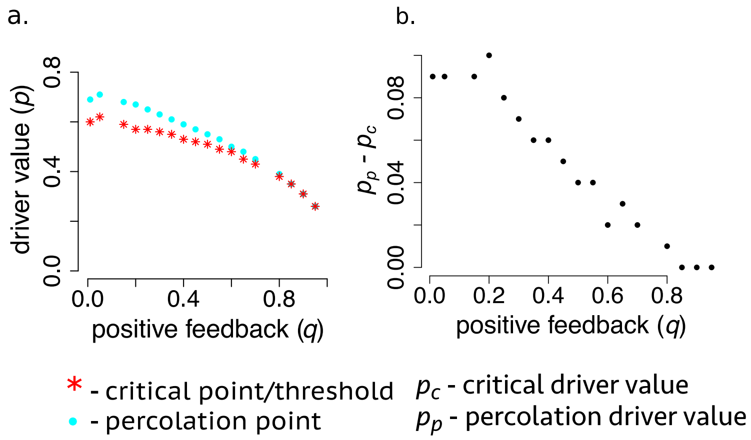


Figure 4: The distance between the percolation point p_p and the critical point/threshold p_c reduces as a function of positive feedback (q). The raw driver values at with these thresholds occur are shown in (a) and the difference between them, in (b).

Appendix E: Power-spectrum and correlations

One way to capture the spread of disturbance in a system or the length scale of spatial fluctuations, is by constructing the spatial covariance function. The *spatial autocovariance function* for local density ρ for a distance r is defined as

$$C(r) = \langle (\rho(\mathbf{x}) - \bar{\rho})(\rho(\mathbf{x}') - \bar{\rho}) \rangle \tag{E1}$$

where $\bar{\rho}$ represents mean density over the entire landscape, angular brackets denote average over all locations $\mathbf{x}$ and $\mathbf{x}'$ in the landscape that are separated by a distance r. Ecologists widely use the correlation function which is defined as

$$K(r) = \frac{\langle (\rho(\mathbf{x}) - \bar{\rho})(\rho(\mathbf{x}') - \bar{\rho}) \rangle}{\sigma^2}) \tag{E2}$$

where σ^2 is the spatial variance of densities in the ecosystem. Thus the covariance function is a product of the correlation function and the spatial variance of the data.

The *correlation length* is defined as the mean of the covariance function and can be interpreted as the average distance to which local fluctuations spread. The correlation length becomes infinite at the critical thresholds. This means that the covariance function then follows a power-law with an exponent less than two.

The *power spectrum*, denoted by $S(k)$, is the Fourier transform of its autocovariance function [1, 8]. Therefore, it can be calculated as

$$S(k) = \int C(r) e^{-ikr} dr \tag{E3}$$

At critical thresholds, we expect the spatial covariance function to exhibit a power-law relation with distance

$$C(x) = c_0 x^{-\alpha} \tag{E4}$$

where c_0 is a constant and α is an exponent less than two. The corresponding spectral function for an n-dimensional system is given by

$$|S(\mathbf{k})| \sim \mathbf{k}^{-(n-\alpha)}$$

Therefore, evidence of a power-law spectral function is also evidence of a power-law autocovariance function.

Appendix F: Power-spectrum fitting

1. Resilient systems (far from transition points)

It is well known that in systems far from transition the power-spectrum typically exhibits a Lorentzian functional form . Theoretically it is also expected that as the system reaches a critical point, its spectral function shifts to a power-law form. In our model too we found the Lorentzian function to be a good fit for resilient systems (both, with high and low positive feedback). The function was fit by running a non-linear least squared regression on the data. We present the results of the analyses below:

Dataset: low positive feedback($q = 0$)
Formula: $y = k * a/((x - x0)^2 + a^2)$
Parameters:

	Estimate	Std. Error	t value	Pr(>|t|)
k	$1.36\text{x}10^{-6}$	$1.6\text{x}10^{-8}$	87.97	<2e-16 ***
x0	-0.2308	0.0051	-45.35	<2e-16 ***
a	0.4116	0.0073	564.89	<2e-16 ***

Signif. codes: 0 ’$***$’ 0.001 ’$**$’ 0.01 ’$*$’ 0.05 ’.’ 0.1 ’ ’ 1
Residual standard error: 0.0001061 on 25647 degrees of freedom
Number of iterations to convergence: 8
Achieved convergence tolerance: 0.0000000149

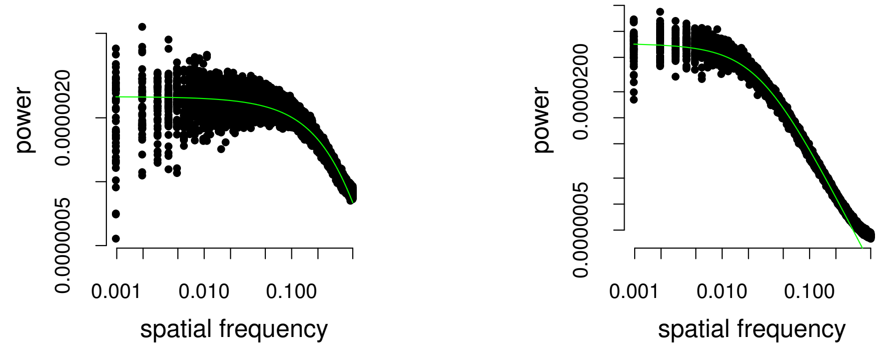
Dataset: high positive feedback ($q = 0.92$)
Formula: $y\ k * a/((x - x0)^2 + a^2)$
Parameters:

	Estimate	Std. Error	t value	Pr(>|t|)
k	$2.06\text{x}10^{-6}$	$1.13\text{x}10^{-8}$	182.99	<2e-16 ***
x0	-0.0085	0.00017	-49.67	$2.72\text{x}10^{-11}$ ***
a	0.0254	$6.79\text{x}10^{-5}$	374.72	<2e-16 ***

Signif. codes: 0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
Residual standard error: 0.0004837 on 25647 degrees of freedom
Number of iterations to convergence: 8
Achieved convergence tolerance: 0.0000000149

2. Systems near/at transition points

For systems near transition, as expected, we found that the power-law function was a good fit over a large range of the data. There are two approaches to fitting power-law relations: one is by log transforming the data and then fitting a linear model and the other, by directly fitting a non-linear model. Several studies demonstrate that log-transforming probability distribution functions (pdf) or frequency data violates assumptions underlying linear regressions such as normally distributed residuals. This, however, is not always the case - especially when the

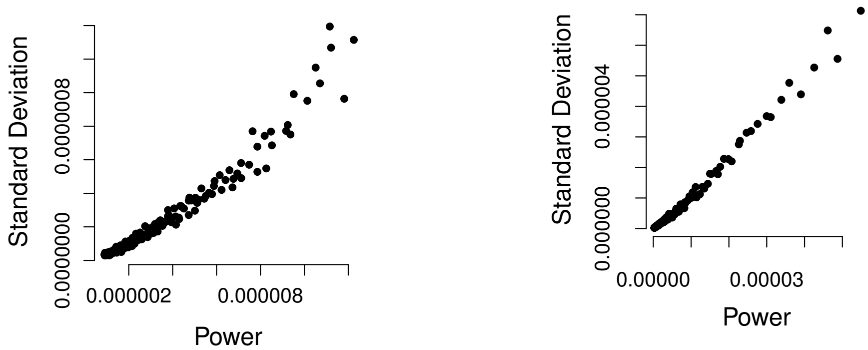


(a) Low positive feedback and close to critical point: $q = 0, p = 0.7225$

(b) High positive feedback and close to critical threshold: $q = 0.92, p = 0.2865$

Figure 5: The Lorentzian fit (green line) of power spectrum data (black dots) of resilient systems with low and high positive-feedback. The plotted data correspond to the data shown in blue in main text Figure 7.

functions are not pdfs. More precisely, fitting power-law data using non-linear least squares regressions assumes that the data have a constant standard deviation, whereas log transformation followed by linear regression is valid for data with a constant coefficent of variation [3]. Since our spectral data for systems near/at critical points, like other $1/f^{\beta}$ spectra [10], show an increasing trend of standard deviation with average power (see Fig.6), we used the linear regression method on log-transformed values. The range of the power-law function was estimated in the same way as described in the previous section, by identifying the x-value (spatial frequency) which minimises the KS distance between the predicted function and the data.



(a) Low positive feedback and close to critical point: $q = 0, p = 0.62275$

(b) High positive feedback and close to critical threshold: $q = 0.92, p = 0.2852$

Figure 6: Varying SD and constant coefficient of variation ($SD/mean$) of power-spectra of systems near transition.

Dataset: low positive feedback($q = 0$)
Formula: $y = (k * x^a)$
lm(formula = log(y)   log(x), weights = x)
Parameters:

	Estimate	Std. Error	t value	Pr(>∣t∣)
a	-1.2103	0.00096	-1265	<2e-16 ***
$\log k$	-15.4159	0.00159	-9688	<2e-16 ***

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.01954 on 15698 degrees of freedom
Multiple R-squared: 0.9932, Adjusted R-squared: 0.9932
F-statistic: 1.6x10^6 on 1 and 15698 DF, p-value: < 2.2e-16

Dataset: high positive feedback ($q = 0.92$)
Formula: $y = (k * x^a)$
lm(formula = log(y)   log(x), weights = x)
Parameters:

	Estimate	Std. Error	t value	Pr(>∣t∣)
a	-1.7819	0.00128	-1392	<2e-16 ***
$\log k$	-16.8830	0.00283	-5976	<2e-16 ***

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.1006 on 14798 degrees of freedom
Multiple R-squared: 0.9944, Adjusted R-squared: 0.9944
F-statistic: 1.938e+06 on 1 and 1 DF, p-value: < 2.2e-16

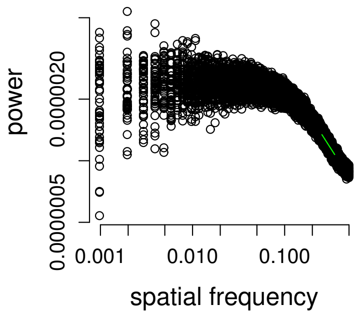
We note that, the exact numerical values of these exponents may often require simulations with much larger system sizes. Hence, we have not emphasized much about the estimation of exponents themselves and have focused only on qualitative features of the fitted functions.

3. Comparing Lorenztian and scale-free spectra

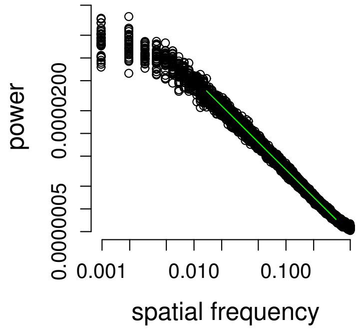
Comparison of the models (Lorentzian vs power-law) for the different datasets is impossible as the models not only have a different number of parameters, but are also not fit over the same range of data. One way to get around this is to consider the theory underlying the emergence of scale-free power-spectra in systems near critical points/thresholds. Even data that follow a Lorentzian function, will follow a power-law over some (small) range of spatial frequencies. However, as the system approaches a critical point, low frequency interactions begin to dominate, thus increasing in power and leading to a shift in the spectrum such that the extent of the power-law region increases. Thus to compare the power-spectra behaviour in systems near/at the critical point with that of resilient systems, one can examine the range over which the power-law fit extends.

To compare the range of power law (PLR), we use the method proposed by [2], who define it as:

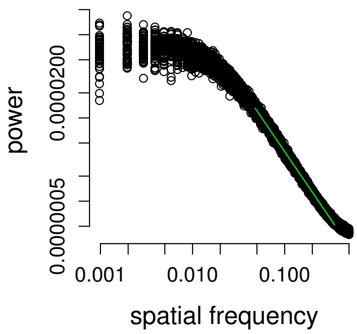
$$PLR = 1 - \frac{log[x_{min}] - log[x_{smallest}]}{log[x_{max}] - log[x_{smallest}]}$$



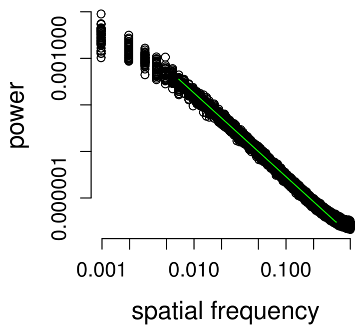
(a) Low positive feedback and far from critical point: $q = 0, p = 0.7225$, $PLR = 0.05$



(b) Low positive feedback and close to critical point: $q = 0, p = 0.26675$, $PLR = 0.56$



(c) High positive feedback and away from critical threshold: $q = 0.92, p = 0.2865$, $PLR = 0.32$



(d) High positive feedback and close to critical threshold: $q = 0.92, p = 0.2852$, $PLR = 0.67$

Figure 7: The shift in the functional form of the power-spectrum is captured by the increased range of power-law as systems approach critical point/thresholds. Data (from simulations) are shown with black open circles whereas the fitted power-law functions are shown as green lines; note that fitted region need not span the entire range, as described in the text.

PLR varies from 0 (when none of the data fall within the fitted power-law) to 1 (when all the data fall within the power-law region). We see in our power-spectra that as the system approaches a critical point/threshold, PLR increases (see Fig 6 captions). This is congruent with theoretical predictions of diverging correlation length at critical points.

---

[1] Baugh, C. and Murdin, P. 2006. Correlation function and power spectra in cosmology. Encycl. of Astronomy and Astrophysics. IOP Publishing, Bristol .

[2] Berdugo, M., Kéfi, S., Soliveres, S., and Maestre, F. T. 2017. Plant spatial patterns identify alternative ecosystem multifunctionality states in global drylands. Nature Ecology & Evolution 1:0003.

[3] Bolker, B. M., Gardner, B., Maunder, M., Berg, C. W., Brooks, M., Comita, L., Crone,

E., Cubaynes, S., Davies, T., Valpine, P., et al. 2013. Strategies for fitting nonlinear ecological models in r, ad model builder, and bugs. Methods in Ecology and Evolution 4:501–512.

[4] Clauset, A., Shalizi, C. R., and Newman, M. E. 2009. Power-law distributions in empirical data. SIAM review 51:661–703.

[5] Génin, A., Majumder, S., Sankaran, S., Danet, A., Guttal, V., Schneider, F. D., and Kéfi, S. 2018. Monitoring ecosystem degradation using spatial data and the r package spatialwarnings. Methods in Ecology and Evolution 9:2067–2075.

[6] Kéfi, S., Guttal, V., Brock, W. A., Carpenter, S. R., Ellison, A. M., Livina, V. N., Seekell, D. A., Scheffer, M., van Nes, E. H., and Dakos, V. 2014. Early warning signals of ecological transitions: methods for spatial patterns. PloS One 9:e92097.

[7] Kéfi, S., Rietkerk, M., Roy, M., Franc, A., De Ruiter, P., and Pascual, M. 2011. Robust scaling in ecosystems and the meltdown of patch size distributions before extinction. Ecology Letters 14:29–35.

[8] Reif, F., 2009. Fundamentals of statistical and thermal physics. Waveland Press.

[9] Stauffer, D. 1979. Scaling theory of percolation clusters. Physics reports 54:1–74.

[10] Van der Schaaf, v. A. and van Hateren, J. v. 1996. Modelling the power spectra of natural images: statistics and information. Vision research 36:2759–2770.