

SCIENTIFIC REPORTS



OPEN

Cross-scale neutral ecology and the maintenance of biodiversity

James P. O'Dwyer¹ & Stephen J. Cornell²

One of the first successes of neutral ecology was to predict realistically-broad distributions of rare and abundant species. However, it has remained an outstanding theoretical challenge to describe how this distribution of abundances changes with spatial scale, and this gap has hampered attempts to use observed species abundances as a way to quantify what non-neutral processes are needed to fully explain observed patterns. To address this, we introduce a new formulation of spatial neutral biodiversity theory and derive analytical predictions for the way abundance distributions change with scale. For tropical forest data where neutrality has been extensively tested before now, we apply this approach and identify an incompatibility between neutral fits at regional and local scales. We use this approach to derive a sharp quantification of what remains to be explained by non-neutral processes at the local scale, setting a quantitative target for more general models for the maintenance of biodiversity.

Neutral biodiversity theory has become one of the most tested paradigms of macroecology^{1–4}. It combines the ecological mechanisms of birth, death, competition, speciation, and spatial dispersal to make predictions for ecological patterns, and makes manifest the belief that the many differences between species may not be critical for successfully predicting large-scale, aggregated phenomena. Subsequent studies have expanded on the original neutral approach^{5,6}, generalizing the theory to include life history^{7,8}, fitness differences^{9–11} and multiple modes of speciation^{12,13}. But at the core of this theory there is a missing link: we lack a complete picture of how neutral predictions change with spatial scale.

Censusing of species abundances at the scale of hectares has provided a tractable approach to test neutral predictions. At these local scales we would expect dispersal limitation to significantly affect the distribution of species abundances alongside birth, death, and competition, leading to departures from the log series distribution of abundances that neutral theory predicts at the largest scales^{1,14,15}. But existing neutral predictions for local community abundances are based on spatially-implicit models, meaning that the process of dispersal is not modeled explicitly, and is instead treated by fitting a parameter in a sampling process. Because there are differences between this spatially-implicit approximation and the full, spatially-explicit neutral processes¹⁶, it has been unclear whether neutrality is consistent with observed local community species abundances or not. Spatially explicit neutral approaches have been developed by modelling dispersal with a dispersal kernel, including numerical simulations^{1,17–20}; hybrid approaches where non-spatial parameters are fitted to a spatially-explicit simulations^{16,21}; the limit of very short-scale dispersal²²; a focus on predicting pairwise correlations in species composition^{23–25}; phenomenological models²⁶; and analytical approaches that make statistical assumptions which are violated in real communities^{27,28}. In short, each has some drawbacks, approximations, or limitations in applicability, so that while spatially explicit models make more realistic assumptions, only the spatially implicit theory has so far been compared exhaustively to empirical abundance patterns.

In this paper, we address this gap by introducing a new mathematical formulation of the spatial theory of neutral biodiversity, derived using the backward equation formulation of stochastic processes. While an exact solution of these equations is not available due to non-linearities, we have identified an accurate approximation scheme which we test extensively using spatially-explicit numerical simulations. These new results allow us to generate a set of predictions that connect both local observations and regional data, in a way that is faithful to neutral model assumptions across this range of scales. We subsequently parametrize the neutral model using sparse, regional and continental-scale observations, and go on to test whether it is then consistent with distributions of abundance at the local scale. We focus on data that has already been fitted using spatially-implicit models, to see whether our spatially-explicit approach deviates from these earlier results.

¹Department of Plant Biology, University of Illinois, Urbana, IL, USA. ²Institute of Integrative Biology, University of Liverpool, Liverpool, L69 7ZB, UK. Correspondence and requests for materials should be addressed to S.J.C. (email: Stephen.Cornell@liverpool.ac.uk)

Combining our modeling approach with data from these multiple scales, we find that neutrality alone significantly underestimates local species diversity, and also deviates from the observed distribution of rare and abundant species. Our intuition might have been that dispersal and neutrality would lead to many rare, transient species, which disperse into a local community and quickly drop out before proliferating. In fact, our spatial neutral prediction dramatically under-predicts the observed number of rare species. This indicates that local stabilizing mechanisms are likely important to understand and accurately predict local patterns of biodiversity^{29–31}, and precisely quantifies what remains for these approaches to explain.

Model

Our model is based around the neutral assumptions of intrinsic birth and mortality rates that are identical across all species, in addition to symmetric competition for a single resource, which we approximate using the mean field approach¹⁰. This is also known as a non-zero sum formulation³² because the total community size is allowed to fluctuate around an average value. We assume that individuals are sessile and dispersal takes place at birth, though as we show in the Supplementary Information the same species-area curve and species-abundance distribution would occur if the organism moved during its lifetime. The resulting model is an assemblage of ecologically identical species, with a constant, total density across space and time when in steady state. New species enter the community via speciation, which occurs at a fixed per capita rate, and hence a fixed rate per unit time and area. All species eventually leave the community due to extinction. So the model reduces to a set of independent populations, beginning their existence with a single individual, and proliferating transiently across space. Meanwhile, we would like to predict the probability that a focal species has a given number of individuals in our sample location in the present day.

In our Supplementary Information we derive the following backward equation (so-called because we look ‘backwards’ from the present day, as explained in our Supplementary materials) to characterize these dynamics and this observable:

$$\begin{aligned} \frac{\partial P(k, A, x, y, t)}{\partial t} = & (b - \nu) \sum_{m=0}^k P(k - m, A, x, y, t) P(m, A, x, y, t) \\ & + b\sigma^2 \sum_{m=0}^k P(k - m, A, x, y, t) \\ & \times \left\{ \frac{\partial^2}{\partial x^2} P(m, A, x, y, t) + \frac{\partial^2}{\partial y^2} P(m, A, x, y, t) \right\} \\ & - (2b - \nu) P(k, A, x, y, t) + b\delta_{k,0}. \end{aligned} \quad (1)$$

In this equation, $P(k, A, x, y, t)$ is the probability that a single individual at time 0, at a spatial location defined by coordinates x and y , will have k conspecific descendants in a sample region of area A , after a time t has elapsed. We can derive patterns of species abundance from this quantity because, if the individual at time 0 is the founding individual of the species, then $P(k, A, x, y, t)$ is the total number of members of this species in the sampling region at time t . This equation describes how P changes in time and space due to the processes of birth (with associated dispersal) and death, and is derived by considering how the lineage starting from a single individual can change in a short time interval. The first term represents birth of a conspecific (which takes place at rate $(b - \nu)$), and is bilinear in P because birth leads to two independent lineages which can contribute descendants to the sampling region. The sum over m represents the different combinations of lineages that lead to a total k descendants in the focal region. Offspring of a different species are produced at rate ν , but these do not contribute to $P(k, A, x, y, t)$. The second term contains a Laplacian to represent the fact that offspring are distributed at a random distance, proportional to the parameter σ , from their parent. The term $b\delta_{k,0}$ (using the Kronecker δ symbol, which is equal to one if $k = 0$, and zero otherwise) represents the fact that a death event (occurring at rate b) causes a lineage containing a single individual to go extinct (in which case the only possibility is $k = 0$). The linear term, whose prefactor $(2b - \nu)$ is the sum of the birth and death rates, represents the fact that either a birth or a death event will change the number of descendants.

From the definition of the quantity $P(k, A, x, y, 0)$, the appropriate initial condition for this equation is simply that at $t = 0$, if the location (x, y) is inside the sample area A , then $P(k, A, x, y, 0) = \delta_{k,1}$, again using the Kronecker δ symbol. Conversely, if (x, y) is outside the sample area, $P(k, A, x, y, 0) = \delta_{k,0}$. We have chosen the sample area to be circular, and centered at the point 0,0 in two-dimensional space. Assuming for the time being that we can solve Eq. (1) for $P(k, A, x, y, t)$, then we immediately have a community level prediction for the average number of species with exactly abundance k in a sample area A :

$$S(k, A) = \nu\rho \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \int_0^{\infty} dt P(k, A, x, y, t), \quad (2)$$

where ρ is the constant average total density across space. However, solving Eq. (1) with the appropriate initial condition is non-trivial, due to the quadratic terms in P , which derive from the birth process, and we do not know of any closed-form solution. This non-linearity is the essence of why this is a difficult problem, and is also reflected in the challenge of finding exact solutions in the corresponding forward-in-time, field theory version of this model²⁷.

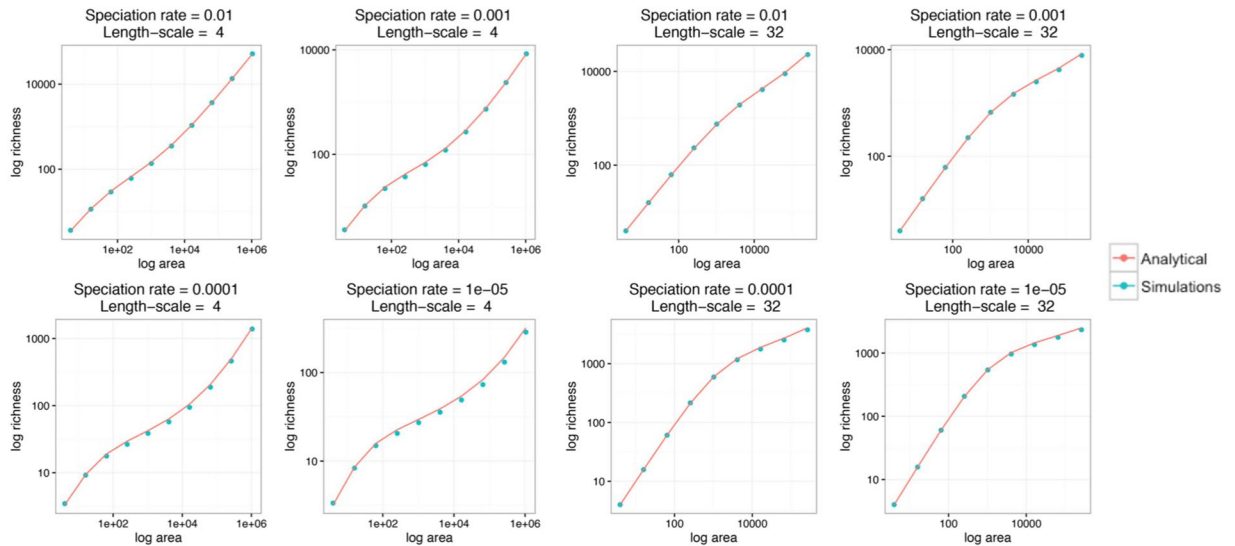


Figure 1. The species-area curve. We show a comparison between species richness as a function of sampled area for our analytical approximation to spatial neutral theory, compared with numerical simulations¹⁷. Over this range of values of speciation rate ($\frac{\nu}{b}$) and dispersal length-scale (σ in the main text), we see quantitative agreement between our approximation method and these earlier numerical results.

Model Solutions

Species-area curve. In our Supplementary Information we introduce an approximation scheme to linearize Eq. (1), with different linearizations applying in different regions of the landscape. As a special case of Eq. (2), we first focus on solutions for the Species-area curve, which counts the total number of distinct species (with any value of $k > 0$) as A increases, in this case for a circular sample region. We find the following approximate solution for this relationship:

$$S(A) = \rho\nu_{\text{eff}}A + \frac{2\rho\sqrt{A}\pi\sigma^2(1 - \nu_{\text{eff}})I_1\left(\frac{\sqrt{A}}{\sqrt{\pi\sigma^2}}\right)}{\frac{1}{\sqrt{\nu_{\text{eff}}}}I_1\left(\frac{\sqrt{A}}{\sqrt{\pi\sigma^2}}\right)\frac{K_0\left(\sqrt{A\nu_{\text{eff}}/\pi\sigma^2}\right)}{K_1\left(\sqrt{A\nu_{\text{eff}}/\pi\sigma^2}\right)} + I_0\left(\frac{\sqrt{A}}{\sqrt{\pi\sigma^2}}\right)} \quad (3)$$

In this solution, we have used the short-hand $\nu_{\text{eff}} = \frac{\nu}{b - \nu} \log(b/\nu)$, but no new parameters have been introduced, while I_n and K_n are modified Bessel functions. Note that only the per capita, per generation speciation rate, ν/b enters this solution, and so the rates b and ν do not independently affect the Species-area curve. How well does this approximation work? In Fig. 1 we demonstrate the agreement between theoretical and simulated curves over a range of speciation rates and values of σ .

At small areas, with $A \ll \pi\sigma^2$, both simulations and theoretical results give $S(A) \simeq \rho A$, i.e. where most new individuals belong to distinct species as the sample area is increased. At large areas, $A \gg \pi\sigma^2/\nu_{\text{eff}}$, both simulations and theoretical results approach $S(A) \simeq \rho\nu_{\text{eff}}A$, so that richness again increases linearly with area, but with a smaller overall coefficient. In between these extremes, we also see good agreement between the simulated and theoretical curves. The transition between large and intermediate scales has been modeled before, by making various phenomenological assumptions about species range shapes and distributions^{33,34}. Here we can see that explicitly the first correction to large-scale linear behavior is proportional to \sqrt{A} , identical to these earlier results³³, so that at intermediate to large scales:

$$S(A) \simeq \rho\nu_{\text{eff}}A + \frac{2\rho\sqrt{A\nu_{\text{eff}}\pi\sigma^2}}{\sqrt{\nu_{\text{eff}}} + 1}, \quad (4)$$

again only valid when $A\nu_{\text{eff}}/\pi\sigma^2 \gg 1$. This agreement is non-trivial, given that the shape of any given neutral species range will not satisfy the simplifying assumptions (of circularity or convexity) made in the phenomenological approaches. Finally, the intermediate region as a whole has been fitted to empirical data drawn from across many taxa and environments using a power law³⁵, and our results show that in neutral theory the power law SAC can only ever be an approximate description.

Spatial Scaling of the Species Abundance Distribution. We now apply the same approximation method to solve for the species abundance distribution, $S(k, A)$, given by Eq. (2). Our solution is expressed in

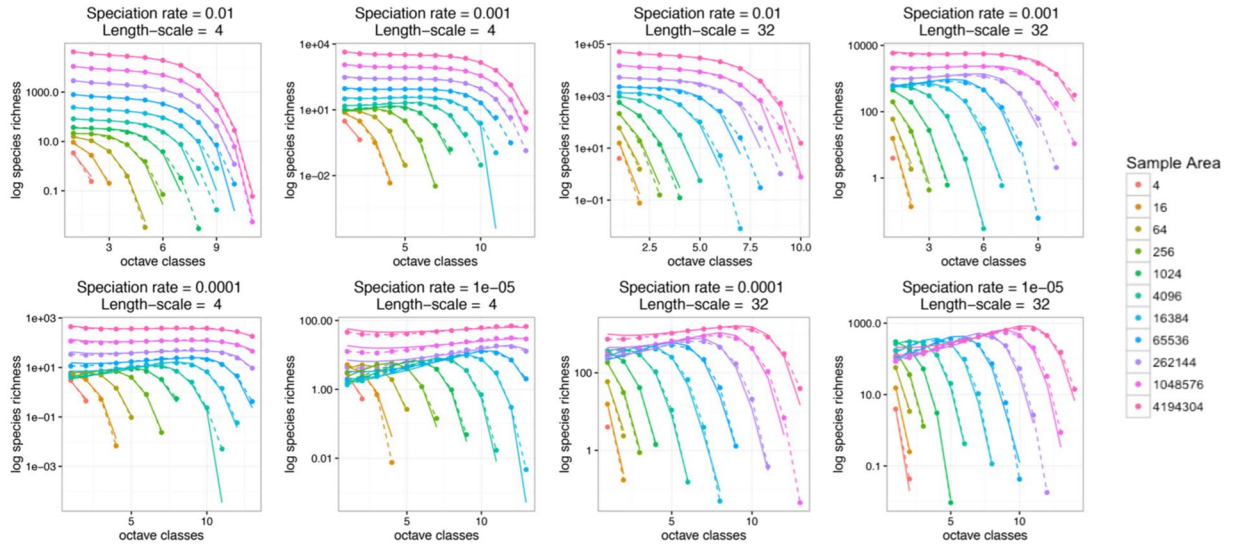


Figure 2. The species-abundance distribution. We test our approximation over a range of speciation rates, ν/b (and two different dispersal length-scale, σ), by comparing the predictions using Eq. (5) with numerical simulations¹⁹. Our results show good (though not perfect, due to both our approximations and the details of the numerical simulation) agreement over this range of parameter values.

terms of the generating function, $\Psi(z, A) = \sum_{k=1}^{\infty} S(k, A)z^k$, and we supply R code to extract the SAD itself from this generating function, following the method of³⁶. Our solution is given by:

$$\Psi(z, A) = S(A) - \rho f(z)A + \frac{\frac{2(f(z) - (1-z))}{\sqrt{h(z)}} \rho \sqrt{A\pi\sigma^2} I_1\left(\sqrt{\frac{h(z)A}{\pi\sigma^2}}\right)}{I_0\left(\sqrt{\frac{Ah(z)}{\pi\sigma^2}}\right) + \sqrt{\frac{1-z}{f(z)}} \frac{K_0\left(\sqrt{\frac{Af(z)h(z)}{(1-z)\pi\sigma^2}}\right)}{K_1\left(\sqrt{\frac{Af(z)h(z)}{(1-z)\pi\sigma^2}}\right)} I_1\left(\sqrt{\frac{Ah(z)}{\pi\sigma^2}}\right)}$$

where we have defined the functions $f(z) = \frac{\nu}{b-\nu} \log\left[\frac{b-(b-\nu)z}{\nu}\right]$ and $h(z) = 1 - z(1 - \nu/b)$ for ease of notation, and $S(A)$ is given by Eq. (3). While finding the Species-area curve is already a promising step, matching the full species abundance distribution as a function of area is a much sterner test for our approximation scheme. In Fig. 2, we show that our solution closely matches earlier numerical simulations over a range of speciation rates ν , values of dispersal length-scale, σ , and sample areas.

This expression for Ψ displays the properties of species abundance distributions that have previously been found by simulations of spatial neutral models¹⁹. First, when A is very large, the third term becomes much smaller than the second term, so the generating function is approximately $\rho f(z)A$. Expanding in powers of z , we find in this limit

$$S(k, A) \propto \frac{\left(1 - \frac{\nu}{b}\right)^n}{n},$$

which is a Fisher logseries with diversity parameter $\alpha = 1 - \frac{\nu}{b}$.

Second, the species-abundance distributions display the “universality” noted by Rosindell and Cornell¹⁹. While the expression for Ψ depends on all four quantities z , A , ν/b , and σ , in Appendix 2.3 we show in that, when the speciation rate is small ($\nu/b \rightarrow 0$), it reduces to an expression that depends only on the two combinations $Z = (1 - z)b/\nu$ and $Y = A\nu/(b\sigma^2)$. We also show in Appendix 2.3 that this is not limited to our approximation, but is also a property of the exact solution to the backward equation. We further show in Appendix 2.3 that this is equivalent to the species abundance distribution taking the scaling form $S(k, A) = \nu \tilde{S}(k\nu, A\nu/\sigma^2)$. This confirms analytically that species abundance distributions for spatial neutral models form a single-parameter family of curves, which extends the universality described by Storch *et al.*³⁴ for species-area curves and endemics-area curves.

Application to Tropical Forest Communities

Now armed with a spatially-explicit prediction for the species abundance distribution, we test whether the observed distribution of tree species abundances at the Barro Colorado Island 50 ha plot (BCI) is consistent with a neutral model where parameters are fixed independently of the plot-scale counts. Due to its high diversity and regular and comprehensive census, this plot has often been a testing ground for theoretical explanations of

Comparison of Observed and Predicted Abundances

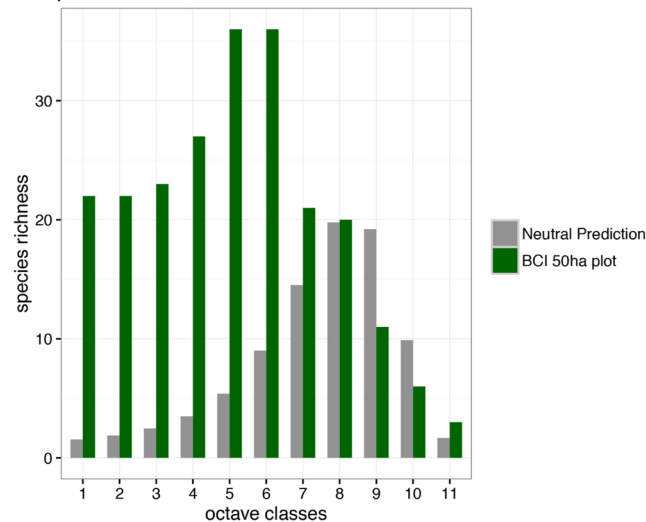


Figure 3. Neutral predictions at BCI. This comparison demonstrates the discrepancies between neutral predictions and the observed data at the 50 ha plot on Barro Colorado Island^{43–45}. Neutral predictions are generated by fitting our spatial neutral model using large-scale data reported and analyzed in²⁴. The results show that these large-scale fits produce a local-scale prediction for species abundances that both underestimates local species richness, compared with observed data, and also skews abundances from rare to more abundant species.

biodiversity patterns. It has also been extensively compared to the spatially-implicit neutral predictions, which have closely matched the observed abundance distribution^{37–39}, although even early on it was emphasized that it may be difficult to distinguish neutral fits from alternatives with the same number of parameters^{40,41}. Taking our alternative route, how should we determine the parameters of our spatially-explicit model? Density ρ is straightforward to estimate, and we could conceivably match the dispersal length-scale σ using inverse modeling and seed-trap data⁴². However, the speciation rate ν would be extremely challenging to measure directly, even to the extent that it is well-defined^{12,13}.

Here we take a different approach, leveraging the methods and results of earlier studies focusing on large-scale spatial correlation functions^{23,24}. These papers focus on the two-point spatial correlation function, known as $F(r)$, the probability that two trees sampled at a separation r from each other are conspecifics. For spatial neutral theory, it has already been shown^{23,27} that this function takes the following form at large spatial separations:

$$F(r) = \frac{1}{\rho\pi\sigma^2} K_0\left(\frac{r\sqrt{2\nu/b}}{\sigma}\right). \quad (7)$$

Using this result, and data from trees with diameter >10 cm in 34 1 ha plots in Panama (separated by values of r between ~ 0.5 and ~ 50 km), Condit *et al.*²⁴ obtained parameter fits of $\sigma = 40.2$ m and $\nu/b = 5.10^{-8}$. These fitted values used the observed density of $\rho \simeq 0.04$ m^{-2} . While speciation would be difficult to estimate independently of this fit, this value of σ is similar to those obtained from seed-trap data⁴².

With these parameters fixed, we can test whether these large-scale data are consistent with the observed distribution of species abundance at the 50 ha plot scale^{43–45}. Figure 3 demonstrates that the spatial neutral model severely underestimates diversity at the 50 ha scale, by approximately a factor of two. It also skews the distribution of species abundances towards more dominant species, with singleton species (those with just one stem >10 cm in the plot) underestimated by a factor of around twenty compared with observed counts. We also acknowledge the huge variation in fitted speciation rates across different regions using large-scale plot data—for example, South American plots fitted in²⁴ led to much lower speciation rates than the data from Panama. To allow for the possibility that the large-scale Panama data is skewed in some way, in Fig. S1 we use the parameters fitted using Ecuadorian forest data, but find the same general pattern: an underestimate of overall local species richness at the 50 ha scale, alongside a skew away from rare species. We emphasize that while we are only looking at one plot, but this is a data set where spatially-implicit neutral predictions had already passed a series of tests, and so it is important to see whether these hold up when the neutral model is spatially-explicit. In summary, the formulation of spatial neutral theory we have considered here allows us to show that local abundances are not consistent with the parameters inferred from large-scale data.

Discussion

Neutral theory has most often been formulated in a spatially-implicit way, so that local species abundance distributions depend on two free parameters characterizing the influx from a larger (but unmeasured) regional community^{1,37}. These parameters, known as the fundamental diversity parameter, θ , and dispersal limitation parameter, m , respectively determine the richness of this larger community and then the rate of immigration

from the regional to the local scale. It is certainly difficult to estimate the richness of this larger community, and hence θ , and while some model approaches have attempted to connect the m to explicit mechanisms of dispersal^{21,46}, this matching only works in certain idealized limits. It has therefore been difficult to know the values of these two fitted parameters θ and m are biologically reasonable or not, even when the neutral theory successfully matches the distribution of species abundances in a local community.

We have introduced a new formulation of spatially-explicit, stochastic biodiversity theory that complements and extends the predictions of earlier approaches^{17,23,27}. Making predictions from our model reduces to the solution of a non-linear partial differential equation, and while it is unlikely that this equation has a closed form solution, we identified an approximation scheme which closely matches the quantitative results of numerical simulations reported in previous work. The underlying model is essentially the same as considered in²⁷, but the new formulation leads to a much better approximation of macroecological patterns (see Supplementary Information for a further discussion of the unrealistic approximations used in the ‘forward-in-time’ approach of²⁷). We focused here on predictions for the species-area curve, and for the distribution of species abundances as a function of spatial scale. The latter prediction is a key advance over earlier formulations of neutral theory, as it allows us to test whether neutral theory matches observed abundance distributions without tuning parameters to fit this data.

It is uncontroversial to say that neutrality is an incomplete description of any given natural system. Instead, neutrality provides a starting point from which we might hope to infer the importance of non-neutral processes. The species abundance distribution has been largely written off as an approach to achieving this, in part because spatially-implicit neutral models are flexible enough to fit a vast range of different local abundance distributions. In applying our spatially-explicit methodology to Panamanian tropical forest data, we in part rehabilitate the species abundance distribution as a diagnostic for what is missing from the neutral explanation, in an approach consistent with previous calls to test multiple patterns simultaneously, rather than just species abundances alone⁴⁷. Specifically, by fitting neutral parameters using large-scale pairwise correlation data, we identified a mismatch between those fitted parameters and local community abundances, under the neutral assumption: by fitting neutral parameters using large-scale data for the pairwise-similarity of widely separated plots, we were able to show that the corresponding neutral prediction for species abundances underestimates diversity at the 50 ha scale, and dramatically skews the distribution of abundances away from rare species. Conversely, we could in principle tune the neutral parameters to more closely match the 50 ha scale data, but at the expense of explaining large-scale spatial correlations poorly. Overall, this shows that neutral demographic processes, combined with Gaussian dispersal limitation alone, are unlikely to explain the maintenance of diversity at the both plot scale and regional scale.

Our goal in this study was not to identify what specific mechanisms could be added to the neutral dynamics to explain the maintenance of observed distributions of species abundances. However, there are several likely ways to resolve this mismatch, and our analysis now opens up the possibility of quantifying what kinds of additional ecological mechanisms provide the best explanation. Very generally, the skew towards rare species in the empirical data can be explained by the presence of stabilizing mechanisms at the local scale. Stabilization can arise from density-dependent interactions, perhaps in turn driven by plant-soil feedbacks^{29,30}, which act to reduce both local dominance and extirpation. An alternative is that neutral models can still explain the presence of these rare species, but that we need to consider so-called “fat-tailed” dispersal, where the probability of dispersing a given distance from a parent tree drops off relatively slowly with distance^{18,23,48}. Our results raise a challenge to either of these explanations for rare diversity in tropical forests. For example, if plant-soil feedbacks explain this combination of patterns, can we quantify exactly how strong and at what spatial scales these mechanisms must act? Similarly, can we quantify exactly what type of long-distance dispersal, if any, can explain the same patterns? Building on the development of this spatial model to include more general processes will provide a sharp, quantitative test of whether a given proposed mechanism is consistent with observations.

Neutrality has perhaps been tested more than any other single theory of biodiversity. This scrutiny has ranged across decadal fluctuations^{49–51} and evolutionary timescales^{52–56}, and across taxonomic groups and environments^{37,55,57–59}. In this manuscript, we show that in terms of the patterns where it has seen greatest success, species abundance distributions, we are seeing discrepancies between the theoretical predictions and observed data. On the other hand, the precise formulation of the neutral theory is exactly what makes it possible to perform these quantitative tests. While the presence of species differences and local niche structure has also been extensively tested, it has rarely been possible to translate the existence of these mechanisms into quantitative predictions for biogeographical patterns, like the distribution of species abundances as a function of spatial scale. The approach we have taken and discrepancies we have identified may therefore serve to motivate new, and more accurate, models of biodiversity, taking us a step closer to identifying precisely what mechanisms do and do not matter for the prediction of biodiversity patterns^{60,61}.

References

- Hubbell, S. P. *The Unified Neutral Theory of Biodiversity and Biogeography* (Princeton Univ. Press, Princeton, 2001).
- Rosindell, J., Hubbell, S. & Etienne, R. The unified neutral theory of biodiversity and biogeography at age ten. *Trends Ecol Evol* **26**, 340–348 (2011).
- O’Dwyer, J. & Chisholm, R. Neutral Theory and Beyond. In *Encyclopedia of Biodiversity* (Elsevier, 2013).
- Azaele, S. *et al.* Statistical mechanics of ecological systems: Neutral theory and beyond. *Reviews of Modern Physics* **88**, 035003 (2016).
- Etienne, R. S. A new sampling formula for neutral biodiversity. *Ecol Lett* **8**, 253–260 (2005).
- Etienne, R. Maximum likelihood estimation of neutral model parameters for multiple samples with different degrees of dispersal limitation. *Journal of Theoretical Biology* **257**, 510–514 (2009).
- O’Dwyer, J., Lake, J., Ostling, A., Savage, V. & Green, J. An integrative framework for stochastic, size-structured community assembly. *Proc Natl Acad Sci* **106**, 6170–6175 (2009).
- Xiao, X., O’Dwyer, J. P. & White, E. P. Comparing process-based and constraint-based approaches for modeling macroecological patterns. *Ecology* (2016).

9. Haegeman, B. & Loreau, M. A mathematical synthesis of niche and neutral theories in community ecology. *J Theor Biol* **269**, 150–165 (2011).
10. O'Dwyer, J. & Chisholm, R. A mean field model for competition: From neutral ecology to the red queen. *Ecology Letters* **17**, 961–969 (2014).
11. Kessler, D. A. & Shnerb, N. M. Neutral-like abundance distributions in the presence of selection in a continuous fitness landscape. *Journal of theoretical biology* **345**, 1–11 (2014).
12. Rosindell, J., Cornell, S. J., Hubbell, S. P. & Etienne, R. S. Protracted speciation revitalizes the neutral theory of biodiversity. *Ecology Letters* **13**, 716–727 (2010).
13. Etienne, R. S. & Haegeman, B. The neutral theory of biodiversity with random fission speciation. *Theoretical Ecology* **4**, 87–109 (2011).
14. Fisher, R. A., Corbet, A. S. & Williams, C. B. The relation between the number of species and the number of individuals in a random sample of an animal population. *The Journal of Animal Ecology* **12**, 42–58 (1943).
15. White, E. P., Thibault, K. M. & Xiao, X. Characterizing species abundance distributions across taxa and ecosystems using a simple maximum entropy model. *Ecology* **93**, 1772–1778 (2012).
16. Etienne, R. S. & Rosindell, J. The spatial limitations of current neutral models of biodiversity. *PloS One* **6**, e14717 (2011).
17. Rosindell, J. & Cornell, S. J. Species-area relationships from a spatially explicit neutral model in an infinite landscape. *Ecol Lett* **10**, 586–595 (2007).
18. Rosindell, J. & Cornell, S. J. Species–area curves, neutral models, and long-distance dispersal. *Ecology* **90**, 1743–1750 (2009).
19. Rosindell, J. & Cornell, S. J. Universal scaling of species-abundance distributions across multiple scales. *Oikos* **122**, 1101–1111 (2013).
20. Shem-Tov, Y., Danino, M. & Shnerb, N. M. Solution of the spatial neutral model yields new bounds on the amazonian species richness. *Scientific Reports* **7** (2017).
21. Chisholm, R. & Lichstein, J. Linking dispersal, immigration and scale in the neutral theory of biodiversity. *Ecol Lett* **12**, 1385–1393 (2009).
22. Durrett, R. & Levin, S. Spatial models for species-area curves. *Journal of Theoretical Biology* **179**, 119–127 (1996).
23. Chave, J. & Leigh, E. G. A spatially explicit neutral model of beta-diversity in tropical forests. *Theor Pop Biol* **62**, 153–168 (2002).
24. Condit, R. *et al.* Beta-diversity in tropical forest trees. *Science* **295**, 666–669 (2002).
25. Houchmandzadeh, B. & Vallade, M. Clustering in neutral ecology. *Physical Review E* **68**, 061912 (2003).
26. Peruzzo, F. & Azaele, S. A phenomenological spatial model for macro-ecological patterns in species-rich ecosystems. *arXiv preprint arXiv:1609.02721* (2016).
27. O'Dwyer, J. & Green, J. Field theory for biogeography: a spatially-explicit model for predicting patterns of biodiversity. *Ecology Letters* **13**, 87–95 (2010).
28. Grilli, J., Azaele, S., Banavar, J. R. & Maritan, A. Absence of detailed balance in ecology. *EPL (Europhysics Letters)* **100**, 38002 (2012).
29. Mangan, S. *et al.* Negative plant-soil feedback predicts tree-species relative abundance in a tropical forest. *Nature* **466**, 752–755 (2010).
30. Comita, L., Muller-Landau, H., Aguilar, S. & Hubbell, S. Asymmetric density dependence shapes species abundances in a tropical tree community. *Science* **329**, 330–332 (2010).
31. Chisholm, R. A. & Muller-Landau, H. C. A theoretical model linking interspecific variation in density dependence to species abundances. *Theoretical Ecology* **4**, 241–253 (2011).
32. Etienne, R. S., Alonso, D. & McKane, A. J. The zero-sum assumption in neutral biodiversity theory. *J Theor Biol* **248**, 522–536 (2007).
33. Allen, A. P. & White, E. P. Effects of range size on species–area relationships. *Evolutionary Ecology Research* **5**, 493–499 (2003).
34. Storch, D., Keil, P. & Jetz, W. Universal species–area and endemics–area relationships at continental scales. *Nature* **488**, 78–81 (2012).
35. Arrhenius, O. Species and area. *J Ecol* **9**, 95–99 (1921).
36. Bornemann, F. Accuracy and stability of computing high-order derivatives of analytic functions by cauchy integrals. *Foundations of Computational Mathematics* **11**, 1–63 (2011).
37. Volkov, I., Banavar, J. R., Hubbell, S. P. & Maritan, A. Neutral theory and relative species abundance in ecology. *Nature* **424**, 1035–1037 (2003).
38. Chave, J. Neutral theory and community ecology. *Ecol Lett* **7**, 241–253 (2004).
39. Volkov, I., Banavar, J. R., Hubbell, S. P. & Maritan, A. Patterns of relative species abundance in rainforests and coral reefs. *Nature* **450**, 45–49 (2007).
40. McGill, B. J. A test of the unified neutral theory of biodiversity. *Nature* **422**, 881–885 (2003).
41. May, R. Patterns of species abundance and diversity. In *Ecology and Evolution of Communities* (Belknap Press, 1975).
42. Clark, J. S., Silman, M., Kern, R., Macklin, E. & HilleRisLambers, J. Seed dispersal near and far: patterns across temperate and tropical forests. *Ecology* **80**, 1475–1494 (1999).
43. Hubbell, S. P., Condit, R. & Foster, R. B. Barro Colorado forest census plot data, <https://ctfs.arnarb.harvard.edu/webatlas/datasets/bci> (2010).
44. Condit, R. *Tropical forest census plots* (Springer-Verlag and R. G. Landes Company, Berlin, Germany, and Georgetown, Texas, 1998).
45. Hubbell, S. P. *et al.* Light-gap disturbances, recruitment limitation, and tree diversity in a neotropical forest. *Science* **283**, 554–557, <https://doi.org/10.1126/science.283.5401.554> (1999).
46. Chisholm, R. A. & Levin, S. A. Linking dispersal and immigration in multidimensional environments. *Bulletin of mathematical biology* **74**, 1754–1763 (2012).
47. McGill, B. J. *et al.* Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecology letters* **10**, 995–1015 (2007).
48. Nathan, R. Long-distance dispersal of plants. *Science* **313**, 786–788 (2006).
49. Chisholm, R. A. *et al.* Temporal variability of forest communities: empirical estimates of population change in 4000 tree species. *Ecology Letters* **17**, 855–865 (2014).
50. Fung, T., O'Dwyer, J. P., Rahman, K. A., Fletcher, C. D. & Chisholm, R. A. Reproducing static and dynamic biodiversity patterns in tropical forests: the critical role of environmental variance. *Ecology* **97**, 1207–1217 (2016).
51. Fung, T., O'Dwyer, J. P. & Chisholm, R. A. Species-abundance distributions under colored environmental noise. *Journal of mathematical biology* **74**, 289–311 (2017).
52. Nee, S. The neutral theory of biodiversity: Do the numbers add up? *Functional Ecology* **19**, 173–176, <http://www.jstor.org/stable/3599286> (2005).
53. Chisholm, R. & O'Dwyer, J. Species ages in neutral biodiversity models. *Theoretical Population Biology* **93**, 85–94 (2014).
54. Wang, S., Chen, A., Fang, J. & Pacala, S. Why abundant tropical tree species are phylogenetically old. *Proc Natl Acad Sci* **110**, 39–43 (2013).
55. O'Dwyer, J., Sharpton, T. & Kembel, S. Backbones of Evolutionary History Test Biodiversity Theory in Microbial Communities. *Proc Natl Acad Sci* **112**, 8356–8361 (2015).
56. Houchmandzadeh, B. Neutral aggregation in finite-length genotype space. *Physical Review E* **95**, 012402 (2017).
57. Muneepeerakul, R. *et al.* Neutral metacommunity models predict fish diversity patterns in mississippi–missouri basin. *Nature* **453**, 220–222 (2008).

58. Ofiteiru, I. D. *et al.* Combined niche and neutral effects in a microbial wastewater treatment community. *Proceedings of the National Academy of Sciences* **107**, 15345–15350 (2010).
59. Woodcock, S. *et al.* Neutral assembly of bacterial communities. *FEMS Microbiology Ecology* **62**, 171–180 (2007).
60. Hubbell, S. P. Neutral theory in community ecology and the hypothesis of functional equivalence. *Functional ecology* **19**, 166–172 (2005).
61. McGill, B. J. Towards a unification of unified theories of biodiversity. *Ecology Letters* **13**, 627–642 (2010).

Acknowledgements

J.O.D. acknowledges the Simons Foundation Grant #376199, McDonnell Foundation Grant #220020439, and Templeton World Charity Foundation Grant #TWCF0079/AB47. This work was supported by the Natural Environment Research Council grant number NE/H007458/1. We thank Ryan Chisholm for comments on an earlier draft, and gratefully acknowledge James Rosindell for permission to use simulated neutral model data generated in collaboration with S.J.C. as a means to test our analytical approximations. The BCI forest dynamics research project was made possible by NSF grants to S.P. Hubbell: DEB #0640386, DEB #0425651, DEB #0346488, DEB #0129874, DEB #00753102, DEB #9909347, DEB #9615226, DEB #9405933, DEB #9221033, DEB #-9100058, DEB #8906869, DEB #8605042, DEB #8206992, DEB #7922197, support from CTFS, the Smithsonian Tropical Research Institute, the John D. and Catherine T. MacArthur Foundation, the Mellon Foundation, the Small World Institute Fund, and numerous private individuals, and through the hard work of over 100 people from 10 countries over the past two decades. The plot project is part the Center for Tropical Forest Science, a global network of large-scale demographic tree plots.

Author Contributions

J.O.D. and S.J.C. designed the study, derived the model equations and solutions, wrote the paper and prepared the figures.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-27712-7>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018