

Comment on “Neutral Ecological Theory Reveals Isolation and Rapid Speciation in a Biodiversity Hot Spot”

Rampal S. Etienne,^{1*} Andrew M. Latimer,² John A. Silander Jr.,² Richard M. Cowling³

Latimer *et al.* (Reports, 9 September 2005, p. 1722) used an approximate likelihood function to estimate parameters of Hubbell’s neutral model of biodiversity. Reanalysis with the exact likelihood not only yields different estimates but also shows that two similar likelihood maxima for very different parameter combinations can occur. This reveals a limitation of using species abundance data to gain insight into speciation and dispersal.

In a recent study, Latimer *et al.* (1) used Hubbell’s (2) neutral model of biodiversity to study speciation and dispersal limitation in a Fynbos community in South Africa. This community has been known for its high speciation rates (3, 4) and low migration rates (5). By fitting the neutral model to species abundance data from this community, Latimer *et al.* (1) suggested that the neutral model can simultaneously confirm these facts. They obtained a very high value for θ , the parameter that reflects speciation rate, and a very low value for m , the immigration parameter (Table 1). However, in their estimation procedures they used a likelihood function that is based on an approximation rather than an exact derivation from Hubbell’s neutral model (6, 7). Now that the exact likelihood has been derived (7), it should be used in future studies using neutral theory to analyze species abundance data. Here, we present a reanalysis of the data presented in Latimer *et al.* (1) using the exact likelihood, and we show that the use of this likelihood function not only gives accurate parameter estimations but also leads to important new insights into the application of neutral theory to species abundance data.

First, the exact likelihood provides computational advantages over the approximation and allows more efficient searching of parameter space, which in some cases can substantially change results. In our reanalysis, we obtained maximum likelihood estimates, which are listed in Table 1; this table also contains estimates for the tropical forest data sets that Latimer *et al.* (1) used for comparison. For two of the three data sets from the Cape Floristic

region, and all the tropical moist forest data sets, the exact likelihood yields estimates of θ and m that are similar to, although not precisely the same as, the estimates provided by the approximation likelihood. However, the Cape Hangklip data yield an immigration parameter m that is two orders of magnitude larger and a corresponding θ value that is one order of magnitude lower than the respective estimates obtained by Latimer *et al.* (1). There is a local likelihood maximum near the values previously reported, but the reanalysis reveals that this is not the global maximum. For reasons provided in Latimer *et al.* (1), this high estimate of m is biologically unrealistic. The fact that the biologically more plausible parameter values are not the most likely under the neutral model demonstrates that caution is warranted in drawing ecological inferences when applying the neutral model to static species abundance data.

Second, our reanalysis demonstrates that potentially two maxima can occur, and this seems to happen only at strongly contrasting parameter values. The likelihood surfaces for all the data sets (except Zuurberg, which has an intermediate value of m) possess a secondary, local likelihood maximum (indicated by the subscript L in Table 1). Although it

has not been rigorously proven mathematically under what conditions dual maxima exist, numerical experiments unmistakably indicate the possible existence of two maxima, and there is an inductive argument that makes this plausible. For this we need to use the fundamental dispersal constant I (8) that is related to m by $I = m(J - 1) / (1 - m)$. When there is no dispersal limitation, we have $\theta = \theta_0$ and $I = \infty$. When dispersal limitation is extreme, the situation is exactly the opposite: $\theta = \infty$ and $I = I_0$. In both cases, the exact likelihood (7) reduces to the same formula—the Ewens sampling formula (9)—but in the former case the parameter is θ_0 and in the latter case the parameter is I_0 . In other words, if $\theta = \theta_0$ and $I = \infty$ together give a likelihood maximum, then $\theta = \infty$ and $I = I_0$ (with the value of I_0 equal to θ_0) give an equally large likelihood maximum. Hence, species abundance data cannot distinguish between these two (extreme) cases. Although the latter can be rejected because it is biologically unrealistic, the data themselves do not contain this information. When neither θ nor I is infinite, the complete symmetry is lost, yet it is still plausible that multiple but unequally likely maxima exist. In the data sets analyzed here, one of the maxima is more likely than the other, but the likelihood values of the maxima are relatively similar, especially considering the enormously unlikely parameter combinations around these maxima (Fig. 1).

The existence of two similar maxima reveals a previously unappreciated limitation of species abundance data. Pragmatically, the problem may be solved by using a Bayesian approach (1, 10) with priors that contain our independent knowledge about the parameters (in this example, there is little doubt that the most dispersal-limited maximum is the realistic one), but the question still remains how much the data really tell us. It appears that in data sets drawn from communities with high species diversity and/or very low migration (i.e., when θ grows large and/or m becomes very small), the neutral model will support alternative parameter combinations, and relatively slight differences in the

Table 1. Parameter estimates under neutral theory obtained with the approximation likelihood used by Latimer *et al.* (1) and obtained with the exact likelihood published recently (6, 7). There are actually two local maxima in all but one of the data sets; the parameters corresponding to the lower maximum are denoted by θ_L and m_L . The log likelihoods in these two maxima are also shown, the lower log likelihood being denoted by loglik_L .

	Old MLE estimates (1)		New MLE estimates (6, 7)					
	θ	m	θ	m	Loglik	θ_L	m_L	Loglik _L
Cederberg, South Africa	394.6	0.0055	384.1	0.0056	-255.3	55.51	0.19	-257.0
Cape Hangklip, South Africa	640.3	0.0020	44.24	0.27	-346.1	597.0	0.0021	-354.8
Zuurberg, South Africa	30.2	0.037	31.76	0.036	-261.2	none	none	none
Manu National Park, Peru	187.4	0.53	187.3	0.53	-189.5	4510	0.022	-193.0
Yasuni National Park, Ecuador	152.8	0.50	152.6	0.50	-177.7	2753	0.020	-182.3
Barro Colorado Island, Panama	47.7	0.093	47.67	0.093	-308.7	241.9	0.0030	-312.6

¹Community and Conservation Ecology Group, University of Groningen, Box 14, 9750 AA Haren, The Netherlands. ²Department of Ecology and Evolutionary Biology, University of Connecticut, 75 North Eagleville Road, Storrs, CT 06269, USA. ³Department of Botany, Terrestrial Ecology Research Unit, Nelson Mandela Metropolitan University, Box 77000, Port Elizabeth 6031, South Africa.

*To whom correspondence should be addressed. E-mail: r.s.etienne@rug.nl

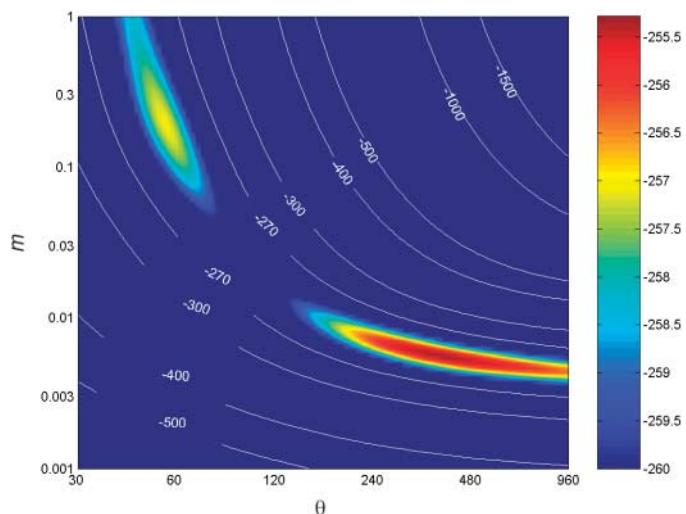


Fig. 1. Log-likelihood surface of the (θ, m) -parameter combination for the Cederberg data set. The color bar shows log likelihoods higher than -260 , so the dark blue area represents log likelihoods lower than -260 , which are indicated by contour lines (lowest value, -1523). In this light, the two local maxima have very similar likelihoods.

data sets may determine which combination is more likely. Thus, it is clear that further study is needed to explore how ecological community characteristics, sampling effects, and temporal and spatial sampling scale influence parameter estimates under the neutral model. At the same time, we need to find more ways to use information on dispersal from experimental and observational studies in general and phylogeny in particular, and paleoecological information on speciation rates in theoretical community ecology.

In sum, the neutral model is a simple, useful exploratory tool to test hypotheses about speciation and dispersal limitation using the most basic and ubiquitous community data, species' abundances. It is critical that such an analysis be done with the best tools available (6, 7), as this has far-reaching consequences for the parameter estimates and for the limitations of using species abundance data.

References

1. A. M. Latimer, J. A. Silander Jr., R. M. Cowling, *Science* **309**, 1722 (2005).
2. S. P. Hubbell, *The Unified Neutral Theory of Biodiversity and Biogeography* (Princeton Univ. Press, Princeton, 2001).
3. H. P. Linder, C. R. Hardy, *Philos. Trans. R. Soc. London Ser. B* **359**, 1623 (2004).
4. J. E. Richardson *et al.*, *Nature* **412**, 181 (2001).
5. P. Slingsby, W. J. Bond, *S. Afr. J. Bot.* **51**, 30 (1985).
6. R. S. Etienne, H. Olff, *Ecol. Lett.* **7**, 170 (2004).
7. R. S. Etienne, *Ecol. Lett.* **8**, 253 (2005).
8. R. S. Etienne, D. Alonso, *Ecol. Lett.* **8**, 1147 (2005).
9. W. J. Ewens, *Theor. Pop. Biol.* **3**, 87 (1972).
10. R. S. Etienne, H. Olff, *Ecol. Lett.* **8**, 493 (2005).

27 October 2005; accepted 28 December 2005
10.1126/science.1121914