Journal of Entomology and Zoology Studies
Available online at www.entomoljournal.com
AkiNik
JEZS
Journal of Entomology and Zoology Studies

ISSN 2320-7078
JEZS 2013;1 (4): 73-76

Received 20-07-2013
Accepted: 04-08-2013

Youhua Chen
Department of Renewable Resources, University of Alberta, Edmonton, Canada

Correspondence:
Youhua Chen
Department of Renewable Resources, University of Alberta, Edmonton, Canada
E-Mail: youhua@ualberta.ca
haydi@126.com

Modeling species-area relationship with measurement uncertainty

Youhua Chen

ABSTRACT

In the present study, the role of measurement uncertainty of species richness has been taken into account when estimating the parameters (*c* and *z*) for the power-law species-area relationship (SAR). The nonlinear weighted estimator χ^2 is used to quantify the influence of measurement uncertainty in the values of species richness, which can be derived from the observed and estimated species richness across different areas. As a comparison, the parameters are also estimated using the conventional nonlinear least-of-square (NLOS) estimator without considering data uncertainty and only the average species richness from estimated and observed values is used. Species richness for epigean arthropods (EAR), canopy arthropods (CAJ) and ground bryophytes (BD) over different areas at the Azores, Portugal are used as empirical data sets for comparing the proposed χ^2 and conventional NLOS estimators. The results show that, both parameters *c* and *z* estimated by χ^2 estimator are significantly different from those from NLOS respectively through the paired *t*-test in all the three empirical data sets except that *c* values are not significantly different for the BD data set when comparing both estimators. Given that fact that there are significant differences on the estimated parameters for the power-law SAR model when comparing both estimators, χ^2 estimator is recommended for fitting SAR models so as to better capture the stochasticity of species richness.

Keywords: Eecological Scaling, Small island effect, Extinction risk, Measurement errors, Statistical inference.

1. Introduction

Species-area relationship (SAR) is one of the most classical ecological laws [1, 2]. SAR has been widely used to estimate and predict species' extinction risk [3–7]. A variety of SARs has been invented and applied to cope with the influences of habitat diversity, landscape heterogeneity and island age [8–13].

Limited sampling efforts have been broadly observed in ecological studies [14, 15]. As such, using the inventory data of species richness across different areas as the representative to estimate the slopes of SARs might tend to over- or under- represent the true SAR patterns. This discrepancy based on limited sampling efforts can be further exacerbated when landscape and/or dispersal complexity are taken into consideration as previously described [16–19].

Small island effect is another factor that might bring more uncertainty into the estimation of SAR parameters [11, 15, 20]. SIE is a hypothesis stating that the species richness of islands would become independent on the areal sizes of the corresponding islands [20]. As such, when plotting species richness over area sizes, data points at the zone where small areas are located would be much over dispersed based on the prediction of SIE.

As such, it seems very necessary to develop new statistical estimators to take into account of the measurement uncertainty of species richness observation over different areas so as to accurately evaluate the slopes of SAR curves. In the present study, the simple χ^2 statistic is used to control the measurement uncertainty in the species richness of a given island. The measurement uncertainty here denotes the difference of observed and estimated species richness for the focused area, which is calculated from different richness estimators.

2. Materials and Methods

2.1 Statistical methods

The nonlinear weighted estimator χ^2 has been widely applied in different disciplines including biology and physics [21]. Supposing that there is a data point *(x,y)* for constructing a SAR model *S=f(A),* we assume here that *x* denotes the size of areas, while *y* denotes the species number of that area. To incorporate the measurement uncertainty, we quantify the uncertainty of species richness as the standard deviation (δ_y) of the species richness based on the observed and estimated values using alternative richness estimators [22, 23]. Then, the χ^2 statistic is the summation of the quantity

$$\frac{(y - f(x))^2}{\delta_y^{\ 2}}$$

Over the sampling areas as follows,

$$\chi^2 = \sum_i \frac{(y_i - f(x_i))^2}{\delta_{y_i}^{\ 2}} \qquad (1)$$

Where *i* is the index of a sampling area. Minimizing the above index allows ones to obtain the estimation of parameters related to the SAR model *S=f(A)* while considering the influence of measurement uncertainty of species richness.

When $f'(x)$ is not a constant (i.e., not a linear model), equation (1) can be regarded as the nonlinear weighted least-squares estimates of the parameters of a nonlinear model. Thus, the numerical optimization technique should be applied to estimate the relevant parameters.

The conventional nonlinear least-of-square estimator (NLOS) is implemented as a comparison. The NLOS is simply to ignore the standard deviations of the data points as follows,

$$NLOS = \sum_i (y_i - f(x_i))^2 \qquad (2)$$

2.1.1 Data sets and the SAR model

Three empirical datasets are obtained from a previous study [24], which include a data set for the soil epigean arthropods at eight forest fragments in Terceira Island (named as EAR dataset), a data set for the canopy arthropods inhabiting *Juniperus brevifolia* at sixteen forest fragments of six different islands (named as CAJ dataset), and a data set for the bryophytes of seven forest fragments from Terceira and Pico islands (named as BD dataset). These data sets are built form the inventory of species diversity of arthropods and bryophytes at the Azores, Portugal [24].

In that previous study [24], the observed species richness has been provided and the associated estimated richness for each area of the data sets has been calculated using a variety of non-parametric richness estimators (ACE, ICE, Chao1, Chao2, Jackknife1, Jackknife2 and Bootstrap). The detailed introduction of these methods are not presented here for simplicity and should refer to previous studies if interested [22–27].

For the present study, the power-law SAR model is utilized for comparative studies. Here, the power-law equation is written as,

$$S = cA^z \qquad (3)$$

Where *S* is the species richness, *A* the area size, *c* and *z* free parameters required to be estimated.

The parameter estimation for *c* and *z* is carried out on the original equation (3) without any log-transformation. Transformation of the data may lead to unexpected results [28–30].

3. Results

As showed in Table 1 and Figs. 1A and 1B, for both EAR and CAJ data sets, the estimated parameters *c* and *z* values for power-law SAR models are significantly different between the NLOS and χ^2 estimators. The goodness of fit for χ^2 is slightly smaller than that for NLOS estimator because it needs to take into account of the influence of data uncertainty.

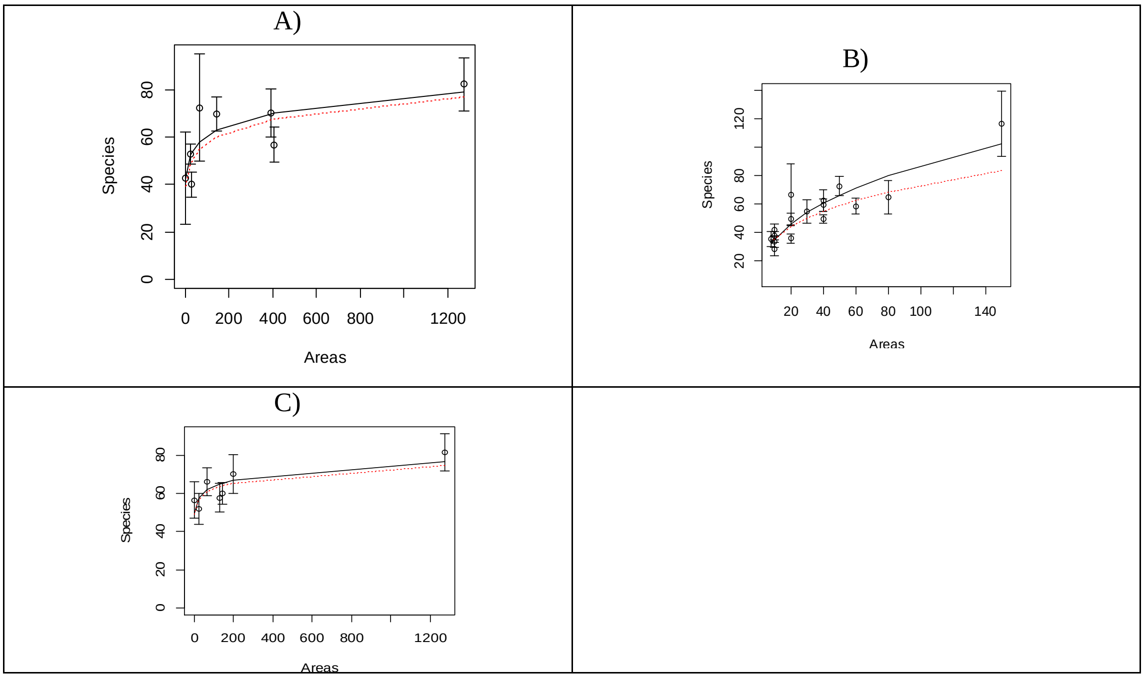
For BD data set (Fig. 1C), the results are basically similar to those for EAR and CAJ datasets. However, the estimated parameter *c* values from NLOS and χ^2 are not significantly different from each other.

Table 1: Estimation and comparison of parameters for the power-law SAR models using the conventional nonlinear fitting (NLOS) and proposed χ^2 estimators for the species diversity of epigean arthropods (EAR), canopy arthropods (CAJ) and ground bryophytes (BD) data sets at the Azores, Portugal. *c* and *z* are the parameters in the power-law SAR model. *SE* denotes the standard errors for the estimated parameters. R^2 reflects the goodness of fit of the model. *t* denotes the result from the two-tailed *t*-test, which compares the differences on simulated 1000-pair random values. These random pairs are generated from normal models with the mean=estimated value (*c* or *z*) and the standard deviation= $SE \times \sqrt{n}$ for both estimators.

Datasets	Estimators	*c*	*SE*	*t*	*z*	*SE*	*t*	R^2
EAR	NLOS	37.535	7.08		0.104	0.035		0.98
	χ^2	33.772	6.353	4.398*	0.115	0.039	-2.606*	0.98
CAJ	NLOS	13.846	2.852		0.400	0.052		0.974
	χ^2	16.952	2.904	-5.279*	0.318	0.052	8.975*	0.960
BD	NLOS	45.832	5.515		0.071	0.024		0.992
	χ^2	45.389	6.289	0.222	0.069	0.028	2.684*	0.992

* denotes a significant difference with $p<0.05$.

Fig 1: Power-law SAR curves for the species diversity of epigean arthropods (EAR, A), canopy arthropods (B) and ground bryophytes (C) data sets at the Azores, Portugal. Bars on the hollow points indicated the standard deviation of the data. The dashed red line indicated the fitting derived from the proposed χ^2 estimator using standard deviation information of species richness from estimated and observed values, while the solid back line indicated the fitting from traditional nonlinear fitting without taking into account of standard deviation of richness but only the average values of richness.



4. Discussion

The importance of data uncertainty has been growingly appraised in recent years for ecological and environmental modeling [31–36]. In the present study, the uncertainty of species richness for each area when constructing SAR models has been resolved by using the χ^2 nonlinear weighted least-squares estimator. Based on the comparison between the models with and without measurement uncertainty, it is found that the situations when uncertainty is taken into account would tend to have smaller parameter values (Table 1).

The influence of measurement uncertainty of species richness using the χ^2 estimator is dependent on the standard deviation of species richness data from observed and estimated values for each focused area. As presented in the equation (1), when the standard deviation of species richness is high, the contribution of the data point from the focused area for the overall χ^2 value will become trivial. In contrast, when the standard deviation is low, the contribution of a specific data point for the overall χ^2 value will be remarkable. Thus, the resulting estimated parameters through χ^2 and NLOS estimators are expected to be different since NLOS didn't include uncertainty information. Based on the comparative results on applying both estimators to three empirical data sets in the present study, the significant differences on the estimated parameter values indeed were observed (Table 1).

The purpose of employing nonlinear weighted least-squares estimator χ^2 by incorporating measurement errors is not because of its novelty but simplicity and straightforward understanding. It can be regarded as the extension of the ordinary least-of-square (LOS) minimization technique when considering the influence of data uncertainty. Thus, the χ^2 estimator is nothing new but a weighted version of NLOS estimator. The estimator can be feasibly implemented in numerical optimization and programming coding. Or alternatively, it can be easily implemented under the R computing environment [37] using the function "*nls*" with weights.

At last, the present results should not be directly compared to those presented in the previous study [24] for a variety of reasons. First, we take into account of all estimated and observed species richness together when reconstructing the SAR model. In contrast, the previous study estimated SAR models for each of the estimated and observed richness. Second, the present study takes into account the data uncertainty while the previous study [24] didn't do that. At last and more importantly, our estimation of SAR model parameters is carried out on the original power-law equation (3) without log-transformation. However, the previous study [24] estimated all free parameters using the log-transformed SAR model. Thus, since that there is a continuous debate on whether log-transformation of the data should be applied [28–30], it is unwise to directly compare the results presented here and the previous work.

5. Reference:

1. He F, Legendre P. On species-area relations. American Naturalist 1996; 148:719–737.
2. Arrhenius O. Species and area. Journal of Ecology 1921; 9:95–99.
3. He F, Hubbell S. Species–area relationships always overestimate extinction rates from habitat loss. Nature 2011; 473:368–371.
4. He F, Hubbell S. Geometry and scale in species-area

relationshps; Extinction and climate change: a reply. Nature 2012; 482:E5–E6.
5. He F, Hubbell S. Estimating extinction from species-area relationships: why the numbers do not add up? Ecology 2013: In press.
6. Halley J, Sgardeli V, Monokrousos N. Species-area relationships and extinction forecasts. Annals of The New York Academy of Sciences 2013; in press.
7. Kinzig A, Harte J. Implications of endemics-area relationships for estimates of species extinctions. Ecology 2000; 81:3305–3311.
8. Chen Y. Combining the species-area-habitat relationship and environmental cluster analysis to set conservation priorities: a study in the Zhoushan Archipelago, China. Conservation Biology 2009; 23:537–545.
9. Patino J, Guilhaumon F, Whittaker R. Accounting for data heterogeneity in patterns of biodiversity: an application of linear mixed effect model to the oceanic island biogeography of spore-producing plants. Ecography 2013; in press.
10. Triantis K, Mylonas M, Whittaker R. Evolutionary species-area curves as revealed by single-island endemics: insights for the interprovincial species-area relationship. Ecography 2008; 31:401–407.
11. Sfenthourakis S, Triantis K. Habitat diversity, ecological requirements of species, and the small island effect. Diversity and Distributions 2009; 15:131–140.
12. Hortal J, Triantis K, Meiri S. Island species richness increases with habitat diversity. American Naturalist 2009; 174:E205–E217.
13. Cameron R, Triantis K, Parent C. Snails on oceanic islands: testing the general dynamic model of oceanic island biogeography using linear mixed effect models. Journal of Biogeography 2013; 40:117–130.
14. Chase J, Knight T. Scale-dependent effect sizes of ecological drivers on biodiversity: why standardised sampling is not enough. Ecology Letters 2013; 16:17–26.
15. Hortal J. Escaping the trap of low sample size in island biogeography. Frontiers of biogeography 2012; 3:4.
16. Kolasa J, Manne L, Pandit S. Species-area relationships arise from interaction of habitat heterogeneity and species pool. Hydrobiologia 2012; 685:135–144.
17. Shen G, Yu M, Hu X. Species-area relationships explained by the joint effects of dispersal limitation and habitat heterogeneity. Ecology 2009; 90:3033–3041.
18. Plotkin JB, Potts MD, Leslie N. Species-area curves, spatial aggregation, and habitat specialization in tropical forests. Journal of theoretical biology 2000; 207(1):81–99.
19. Hu G, Wu J, Feeley K. The effects of landscape variables on the species-area relationship during late-stage habitat fragmentation. Plos One 2012; 7:e43894.
20. Wang Y, Zhang M, Wang S. No evidence for the small island effect in avian communities on islands of an inundated lake. Oikos 2012; 121:1945–1952.
21. Ma Y. Variable cosmological constant model: the reconstruction equations and constraints from current observational data. Nuclear Physics B 2008; 804:262–285.
22. Hortal J, Borges P, Gaspar C. Evaluating the performance of species richness estimators: sensitivity to sample grain size. Journal of Animal Ecology 2006; 75:274–287.
23. Colwell R. Estimate S: statistical estimation of species richness and shared species from samples. Version 8. 2006.
24. Borges P, Hortal J, Gabriel R, Homem N. Would species richness estimators change the observed species area relationship? Acta Oecologica 2009; 35:149–156.
25. Colwell R, Coddington J. Estimating terrestrial biodiversity through extrapolation. Philosophical transactions of the Royal Society of London. Series B, Biological sciences 1994; 345:101–118.
26. Gotelli N, Colwell R. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecology Letters 2001; 4:379–391.
27. Fattorini S. Darkling beetle communities in two geologically contrasting biotopes: testing biodiversity patterns by microsite comparisons. Biological Journal of the Linnean Society 2009; 98:787–793.
28. Xiao X, White E, Hooten M, Durham S. On the use of log-transformation vs. nonlinear regression for analyzing biological power laws. Ecology 2011; 92:1887–1894.
29. Packard G. On the use of logarithmic transformations in allometric analyses. Journal of Theoretical Biology 2009; 257:515–518.
30. Packard G. Is logarithmic transformation necessary in allometry? Biological Journal of the Linnean Society 2013; DOI: 10.1111/bij.12038.
31. Falk M. Incorporating uncertainty in environmental models informed by imagery. PhD thesis, Queensland University of Technology 2010.
32. Hui D, Jackson R. Uncertainty in allometric exponent estimation: a case study in scaling metabolic rate with body mass. Journal of Theoretical Biology 2007; 249:168–177.
33. Moore J, Semmens B. Incorporating uncertainty and prior information into stable isotope mixing models. Ecology Letters 2008; 11:470–480.
34. Cressie N, Calder C, Clark J. Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. Ecological Applications 2009; 19:553–570.
35. Jager H, King A. Spatial uncertainty and ecological models. Ecosystems 2004; 7:841–847.
36. Phillips D, Gregg J. Uncertainty in source partitioning using stable isotopes. Oecologia 2001; 127:171–179.
37. R Development Core Team. R: A Language and Environment for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org. 2011.