JJBS

Volume 2, Number 2, June 2009
ISSN 1995-6673
Pages 83 - 88

Jordan Journal of Biological Sciences

Predicting Species Relative Abundance in Ecological Communities

Nidal Odat[a,*], Moh'd T. Moh'd Taleb Alodat[b], Riyadh Muhaidat[c], Faisal Ababneh[d], Abdel Rahman M. Al-Tawaha[a], Salem Aladaileh[a]

[a]Department of Biological Sciences, Al-Hussein Bin Talal University, Jordan

[b]Department of Statistics, [c]Department of Biological Sciences, Yarmouk University, Jordan

[d]Department of Mathematics and Statistics, Al-Hussein Bin Talal University, Jordan

Abstract

Ecologists are increasingly appreciating using statistical models to predict aspects of species ecology including their abundance and distribution due to their importance in biological conservation and management practices. The aim of this study is to propose a statistical model that allows predicting previously unknown plant species relative abundance (SRA) in an unsurveyed region based on small sub-samples of the whole community. We apply the model to a biodiversity data set which includes plant relative abundances collected from sub-samples of varied communities in central Europe. The results show that the predicted plant relative abundances in unsurveyed sites are close in value to those in the known sites, reflecting the accuracy and the predictive power of the model in estimating species relative abundance in previously unsurveyd ecological sites. The importance of our model is discussed in relation to conservation biology and management.

الملخص

غالبا ما يوظف باحثو علم البيئة نماذج احصائية تمكنهم من التنبؤ عن مفاهيم تتعلق ببيئة الكائنات الحية بما فيها الوفرة والانتشار وذلك نظرا لأهمية هذه المفاهيم في الحفاظ الحيوي والممارسات الإدارية المرافقة لها. تهدف هذه الدراسة إلى اقتراح نموذجا إحصائيا يسمح بالتنبؤ عن الوفرة النسبية الغير معروفه أصلا لأنواع نباتية معينة في منطقة غير مدروسة سابقا وذلك بالاعتماد على عينات مبسطة وممثلة لكافة أفراد المجتمع النباتي في تلك المنطقة. لقد قمنا بتطبيق هذا النموذج الإحصائي على مجموعة من البيانات تتعلق بالتنوع النباتي والتي تم جمعها من مجتمعات نباتية في أوروبا الوسطى. أظهرت نتائج الدراسة وجود تقارب بين قيم الوفرات النسبية المتنبأ عنها في المواقع الغير مدروسة وتلك القيم المحصلة في المواقع المعروفة. في النهاية, تمت مناقشة أهمية هذا النموذج الإحصائي من خلال علاقته بالحفاظ الحيوي والممارسات الإدارية.

Keywords: Species Relative Abundance; Statistical Model, Biodiversity; Ecological Communities; Conservation Biology.

1. Introduction

The species relative abundance, typically estimated by calculating the abundance of a given species divided by the total abundances of all other species in an ecological community, is a fundamental description of an ecological community (Ricklefs and Schluter 1993; Gaston and Spicer, 1998). Although estimating the species relative abundance is basic, yet it represents informative data for ecologists and conservation biologists (Gatson 2003).

Ecologists use data of species relative abundance to infer information about the mode of interaction and type of relationship among different species in a given community (Ferrier and Guisan, 2006). Moreover, species relative abundances data is used to perform population viability analysis (Possingham, *et al*., 2001), and estimate ecological disturbance (Regan, *et al*., 2003). In conservation biology and management, information on relative abundances is of great importance, as for example, to study the impact of habitat disturbances, such as fragmentation. It is well known that that disturbed and fragmented habitats are usually dominated by a very few species compared to the undisturbed sites (Guisan, *et al*. 1999; Regan *et al*., 2003; IUCN, 2001; Guisan and Thuiller, 2005).

Ecologists are often interested in employing statistical models to predict the occurrence and the distribution of species due to their importance in conservation biology of species (Franklin, 1995; Guisan and Zimmermann, 2000; Ferrier, 2002; Zhang, 2007). Generally, these statistical models utilize correlations between data on species occurrences, relative abundances, and environmental predictors (Franklin, 1995).

In the present study, we develop a statistical model that allows predicting species relative abundance across ecological communities. However, the model requires data

---

[*] Corresponding author. odatnidal@hotmail.com.

about the relative abundance of sub-samples collected from a given community. Predicting species relative abundance in unsurveyed communities is extremely likely, given a data set of species relative abundances of representative sub-samples from various communities is available. We apply the model to field data of plant species relative abundances collected in a biodiversity project in central Europe in which plant species diversity and plant productivity have been investigated (Perner, *et al*., 2005). Finally, we discuss the importance of the model in fields of management and conservation biology

2. MATERIALS and METHODS

2.1. Data

The data used in the presented model was based on a biodiversity project carried out in a plateau-like mountain range at the Thuringian/Bavarian border in central Germany with a maximum height of 870 m. Average annual temperature in the area varies between 68F and 78F and average annual precipitation varies between 950 and 1099 mm (Perner, *et al*, 2005). The studied plant communities located between 11.018º and 11.638º eastern longitudes and between 50.358º and 50.578º northern latitudes, and were covering about one hectare. For the model we developed ten plant species were chosen, namely *Dactylis glomerata, Taraxacum officinale, Trifolium repens, Veronica chamaedrys*, *Anthoxanthum odoratum, Rumex acetosa, Ranunculus acris, Holcus lanatus, Phleum pretense,* and *Campanula rotundifolia*. The species were collected from nine distinct sites that are almost comparable with regard to elevation, edaphic and climatic factors (Kahmen, *et al*., 2005).

2.2. The Proposed Model

The primary goal of our model is to predict previously undetermined relative abundance of ten different plant species in unsurveyed site. Our prediction was based on known data collected from nine different sites within each community. For the model parameters consider biological communities with a large number of species. Let $X_1, \cdots, X_n$ be the species relative abundances drawn from a random sample of n size collected from a given community (Figure 1). A species relative abundance is a description of the abundance (number of individuals observed) for each species encountered within a community. As such, it is one of the most basic descriptions of an ecological Community (Ricklefs and Schluter, 1993; Gaston and Spicer, 1998). We assume that the probability density function which represents that community is $f(x \mid \theta)$, where θ a parameter of interest is. The Bayesian statistics use the sample information in addition to our prior knowledge about θ to make statistical inferences about the population. The prior knowledge about θ is given as a probability distribution $\pi(\theta)$, say. We assume that the relative abundances for species community follow the power function density, i.e.,

$$f(x \mid \theta) = \begin{cases} \dfrac{\theta}{1-\theta} x^{\frac{2\theta-1}{1-\theta}}, & 0 < x < 1, 0 < \theta < 1; \\ 0 & otherwise, \end{cases}$$

where θ represents the mean of the distribution,.i.e., θ is the mean of relative abundances in whole community. As a prior knowledge about θ, we assume that $\pi(\theta) = \theta^{-1}(1-\theta)^{-1}$, the Jeffery's no-informative prior. Using both sources of information, we update the density to the posterior density

$$\pi(\theta \mid X_1, \cdots, X_n) = cf(X_1 \mid \theta) \cdots f(X_n \mid \theta)\pi(\theta),$$

where c is a normalizing constant. If Y is the relative abundance of an un surveyed site in this community, then we can predict Y according to the predictive density

The mean and the variance of the density $f(y \mid X_1, \cdots, X_n)$ are the prediction of Y and its uncertainty. Combining the power function observations together with the prior information leads us to the following predictive density :

$$f(y \mid t) = \frac{\int_0^1 \theta^n (1-\theta)^{-n-2} \exp(-\frac{\theta t}{1-\theta} + \frac{2\theta-1}{1-\theta} \log y) d\theta}{\Gamma(n) t^{-n}},$$

$$= \frac{nt^n}{y(t-\log y)^{n+1}}; \qquad \text{for } 0 < y < 1$$

where $t = -\sum_{i=1}^{n} \log X_i$.

The mean of this density is

$$\hat{Y} = \int_0^1 \frac{nt^n}{(t-\log y)^{n+1}} dy; \tag{1}$$

and the prediction error =

$$\int_0^1 \frac{nt^n y}{(t-\log y)^{n+1}} dy - \hat{Y} \tag{2}$$

All calculation and mathematical formula have been carried out with the help of the software package MATHEMATICA.

3. Results and Discussion

The model in this study is based on a biodiversity study conducted in Central Europe (Germany) in which plant species diversity, productivity and site characteristics have been investigated using multivariate statistics. In each study site a 2x2 m plot was established and used to determined plant cover, community biomass (productivity), and soil nutrients. The plant species within the 2x2 m plot were identified to the species level and the percent cover of each species was estimated visually (Perner, et al., 2005). This data was then used to calculate the relative abundance for each plant species as the number of plant of a particular species as a percentage of the total number of all plants of a given area or community. The plant communities in the study sites include all associated plant species that form the natural vegetation at each geographical site (Gaston, 2003).

The model provides a statistical tool for predicting species relative abundance at a given community. Table 1 shows the predictive relative abundances for the ten plant

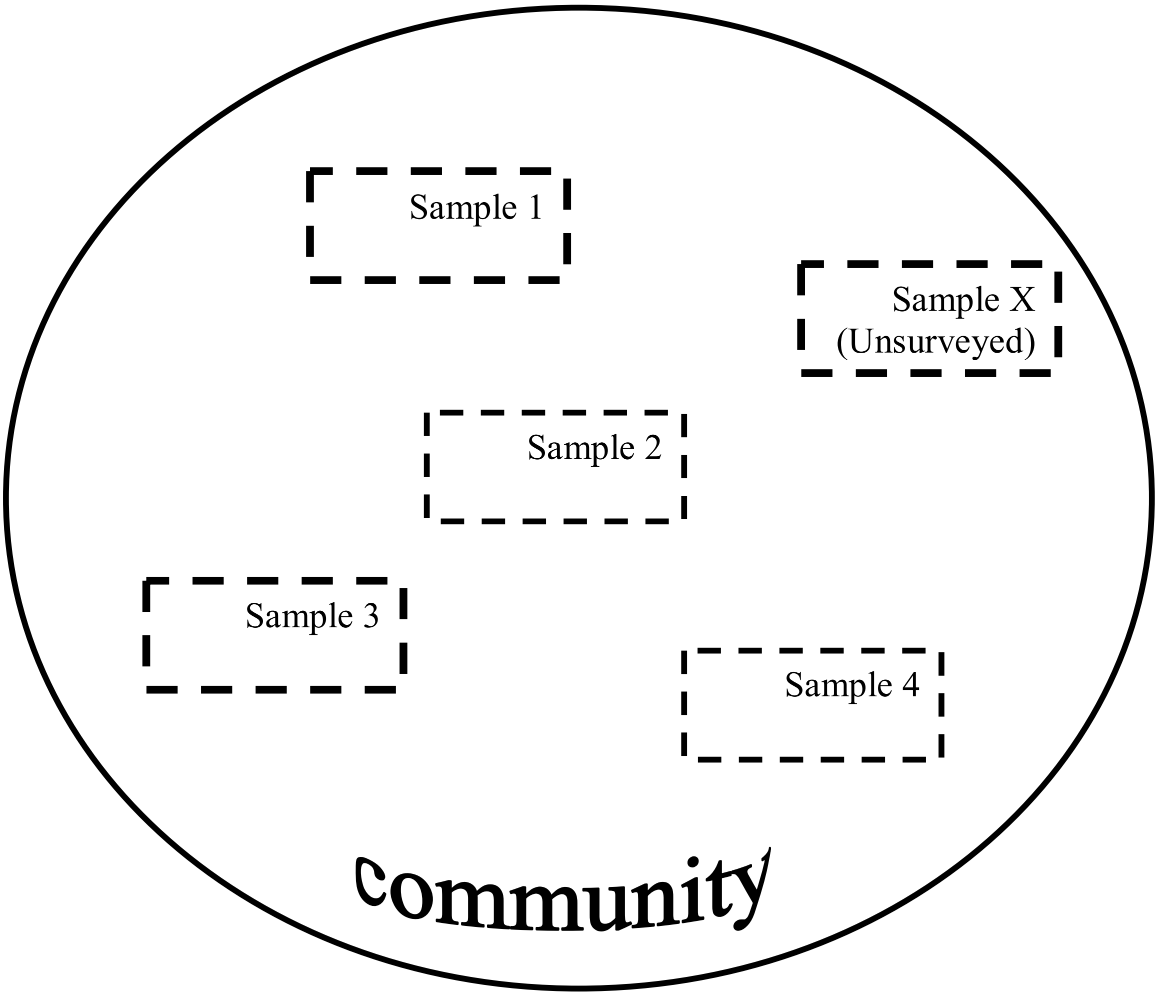

Figure 1. Hypothetical community with four known sub-samples (sample 1-4) and one unsuveyed sample.

Table 1. The predictive relative abundances for the ten plant species used in the model and their prediction error values. The prediction error was calculated according to equation 2 (see text) with the help of the software MATHEMATICA.

Plant species	Predicted relative abundances	Prediction error
Dactylis glomerata	0.252	0.081
Taraxacum officinale	0.221	0.076
Trifolium repens	0.216	0.075
Veronica chamaedrys	0.197	0.071
Anthoxanthum odoratum	0.273	0.085
Rumex acetosa	0.203	0.073
Ranunculus acris	0.153	0.060
Holcus lanatus	0.209	0.074
Phleum pretense	0.174	0.066
Campanula rotundifolia	0.150	0.059

species that are most common in the studied communities: (Dactylis glomerata, Taraxacum officinale, Trifolium repens, Veronica chamaedrys, Anthoxanthum odoratum, Rumex acetosa, Ranunculus acris, Holcus lanatus, Phleum pretense, and Campanula rotundifolia) which have been investigated in the present paper. The results show that the relative abundances for the species in the unsurveyed site are close in value to those in the known sites, reflecting the accuracy of our model in estimating species relative abundance ecological communities. Relative abundance and frequency of the predicted value for each plant species and for the whole plant community from which each species has been sampled are illustrated in Figure 2. These results indicate that the predictive values for each species are located within the density of the relative abundance calculated from the field data which may be suggest that our model is informative and comprehensive. To test the validity of the model we used simulation test to manipulate the relative abundance of five plant species drawn randomly from each community and compared the results of simulation with those obtained by our model. The results of the simulation

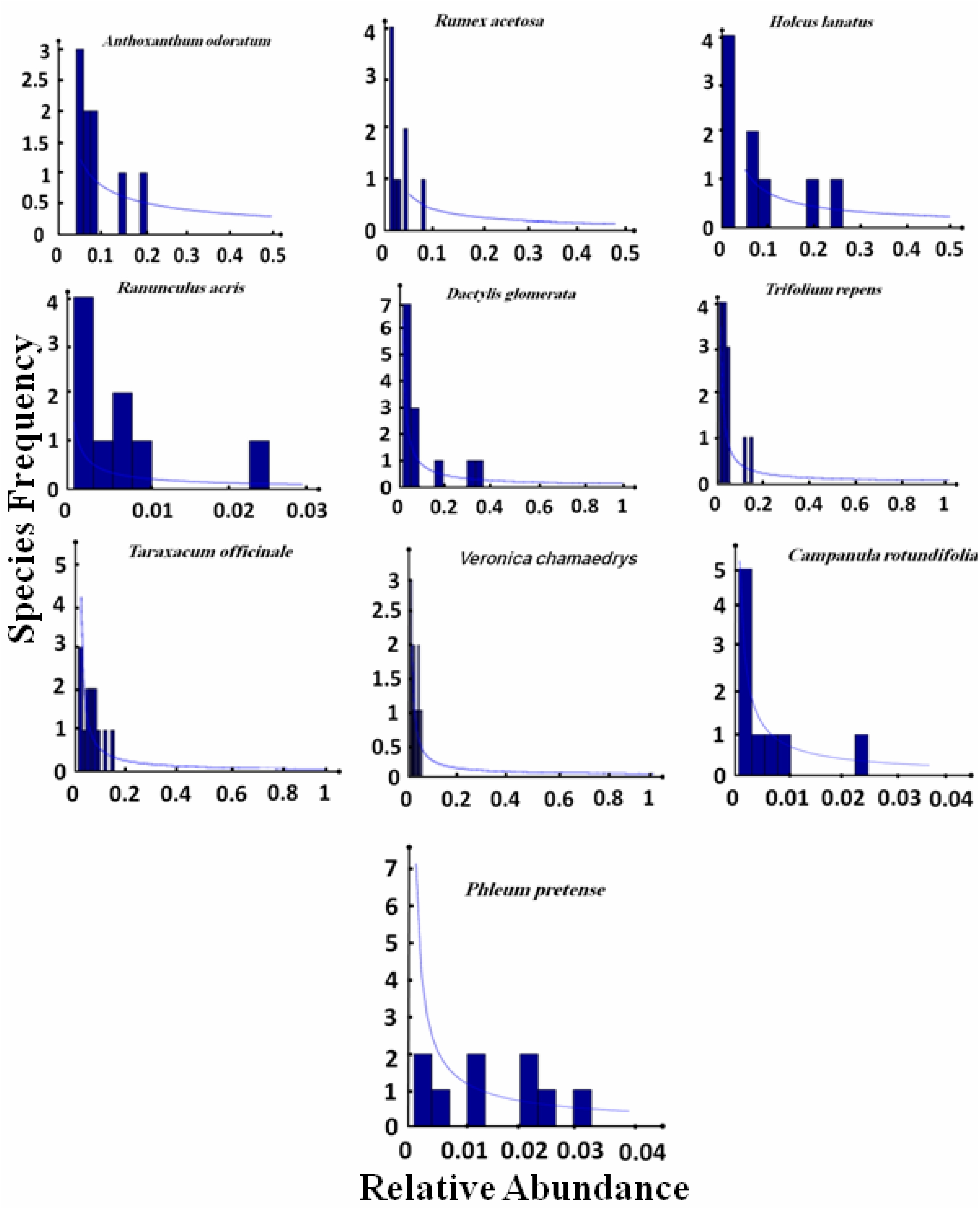

Figure 2. Histograms of the relative abundances for the ten plant species (Dactylis glomerata, Taraxacum officinale, Trifolium repens, Veronica chamaedrys, Anthoxanthum odoratum, Rumex acetosa, Ranunculus acris, Holcus lanatus, Phleum pretense, and Campanula rotundifolia). Lines are the theoretical density which represents the species relative abundance in the whole plant community.

includes two estimates for the relative abundance of each species (one estimate is based on the proposed model and one based on the simulation test). The simulation results reveal that the two values of each estimate are very close reflecting the predicated power of the model (Table 2). We propose that this model may have a wide range of applications, particularly in management practices and biological conservation. For instance, information on species relative abundance at a certain habitat can help us judge if that habitat is under human impact, such as disturbance and fragmentation. It is well known that disturbed habitats are usually dominated by a very few species compared to the undisturbed, more diverse sites (Guisan, *et al*., 1999; Guisan and Zimmermann, 2000; Regan, *et al*., 2003; Guisan and Thuiller, 2005).

The model presented in this study is rather simple and limited to habitats of similar ecological conditions, in term of species composition and abiotic factors. Nevertheless,

Table 2. The species relative abundances of five selected plant species as derived from the model and from the simulation test.

Plant species	Relative abundance based on model	Relative abundance based on simulation
Anthoxanthum odoratum	0.2796	0.2791
Rumex acetosa	0.2071	0.2062
Ranunculus acris	0.1532	0.1587
Holcus lanatus	0.2132	0.2044
Phleum pretense	0.2031	0.1820

our results maybe suggest that statistical models open the door for testing a variety of hypotheses regarding species relative abundances of plant species. The model established here can further be employed to settle other fundamental research questions related specifically to effects of population structure, species spatial distribution pattern, and habitat patchiness (homogenous vs. heterogeneous) (Lichstein, *et al.*, 2002; Holt, 2003; Reese, *et al*., 2005; Hoeting, *et al*., 2006).

References

[1] Ferrier S. 2002. Mapping spatial pattern in biodiversity for regional conservation planning: where to from here? Sys. Biol. 51: 331–363.

[2] Ferrier S and Guisan A. 2006. Spatial modeling of biodiversity at the community level. J. Appl. Ecol. 43: 393–404.

[3] Franklin J. 1995. Predictive vegetation mapping: geographic modelling of biospatial patterns in relation to environmental gradients. Prog. Phys. Geog. 19: 474–499.

[4] Gaston K. 2003. The structure and dynamics of geographic ranges. Oxford University Press, Oxford.

[5] Gaston KJ and Spicer VI. 1998. Biodiversity: an introduction. Blackwell Science, Oxford, UK.

[6] Guisan A and Zimmermann NE. 2000. Predictive habitat distribution models in ecology. Ecol. Model. 135: 147–186.

[7] [7] Guisan A and Thuiller W. 2005. Predicting species distribution: offering more than simple habitat models? Ecol. Lett. 8: 993–1009.

[8] Guisan A, Weiss SB and Weiss AD.1999. GLM versus CCA spatial modeling of plant species distribution. Plant. Ecol. 143: 107–122.

[9] Hoeting JA, Davis RA, Merton AA and Thompson SE. 2006. Model selection for geostatistical models. Ecol. Appl. 16: 87–98.

[10] Holt RD. 2003. On the evolutionary ecology of species' ranges. Evol. Ecol. Res. 5: 159–178.

[11] IUCN (2001) The IUCN Red List of Threatened Species. 3.1, Gland, Switzerland.

[12] Kahmen A, Perner J, Audorff J, Weisser W and Buchmann N. 2005. Effects of plant diversity, community composition and environmental parameters on productivity in semi-natural montane European grasslands. Oecologia 142: 606-615.

[13] Lichstein, JW, Simons TR, Shriner SA, and Franzreb KE. 2002. Spatial autocorrelation and autoregressive models in ecology. Ecol. Monogr. 72(3):445-463.

[14] Perner J, Wytrykush C, Kahmen A, Buchmann N, Egerer I, Creutzburg S, Odat N, Audorff V and Weisser WW. 2005. Effects of plant diversity, plant productivity and habitat parameters on arthropod abundance in montane European grasslands. Ecography 28: 429–442.

[15] Possingham HP, Lindenmayer DB and McCarthy MA. 2001. Population Viability Analysis, Encyclopædia of Biodiversity. Academic Press, pp 831– 843.

[16] Reese GC, Wilson KR, Hoeting JA and Flather CH. 2005. Factors affecting species distribution predictions: a simulation modeling experiment. Ecol. Appl. 15: 554–564.

[17] Regan HM, Auld TD, Keith DA and Burgman M.A. 2003. The effects of fire and predators on the long-term persistence of an endangered shrub, *Grevillea caleyi*. Biol. Conserv. 109: 73–83.

[18] Ricklefs RE and Schluter D. 1993. Species diversity in ecological communities: historical and geographical perspectives. University of Chicago Press, Chicago, Illinois, U.S.A

[19] Zhang H. 2007. Inferences on the Number of Unseen Species and the Number of Abundant/Rare Species. J. Appl. Statist. 34: 725–740.