Modelling landscape variation in species richness: a hierarchical approach

Henry, M. (1); Cosson, J.-F. (2); Pons, J.-M. (3)

(1) INRA Laboratoire Abeilles & Environnement (INRA-UAPV), Site Agroparc, F-84000 Avignon, France, mickael.henry@bioguiana.org

(2) INRA CBGP (IRD-INRA-CIRAD, Montpellier Supagro), Campus International de Baillarguet, CS 30016, F-34988 Montferrier/Lez cedex, France

(3) UMR5202 Origine, structure et évolution de la biodiversité, Département Systématique et Evolution, Muséum National d'Histoire Naturelle, 55 Rue Buffon, F-75005 Paris, France

Abstract: Species richness is widely used by animal ecologists as a biodiversity metric. Modelling landscape variation in species richness is, however, subject to strong statistical constraints when reliable richness estimates are restricted to few sampling sites. In this study, we assessed the efficacy of some richness surrogates whose computation is based on the relative abundance of relevant species groups. Available from any single sample, abundance estimates are usually adequately modelled as a function of landscape configuration, and as such offer considerable advantages over the direct modelling of species richness. When applied to a complex bat assemblage in a fragmented neotropical rainforest, most candidate surrogates were tightly correlated with observed species richness (r= 0.71 to 0.88). These surrogates can be used as reliable tools to compare the efficiency of different landscape management scenarios or landscape restoration priorities with regard to biodiversity.

Keywords: forest fragmentation; Generalised Linear Model; functional diversity; higher taxa diversity; vertebrate communities

Introduction

Determining areas of biological interest for protection is one of the key objectives of conservationists. This can be achieved either by focusing on areas subject to substantial disturbances, by looking for diversity hotspots (Myers et al., 2000), or by prioritizing areas that complement the existing reserve network (Justus and Sarkar, 2002). For that purpose, conservationists often resort to spatial models of species richness for a wide range of
organisms, including plants (Wohlgemuth, 1998; Gould, 2000; Cayuela et al., 2006; Chust et al., 2006), invertebrates (Lobo and Martín-Piera, 2002) or vertebrates (Kerr and Packer, 1997; Rahbek and Graves, 2001; Gorresen and Willig, 2004). The modelling approach consists of using environmental factors like topography, climate, or any biotic or abiotic variable, as predictors to infer the number of species occurring in a given region.

More recently, Zipkin et al. (2009) have formalized a hierarchical method based on the combination of species-specific predictive models of occurrence. This promising approach has the advantage of delivering predictive maps of species richness integrating the specific characteristics (response to environment, detectability) of each individual species in an assemblage. In some circumstances, however, assemblages are too specious and/or field data too scarce to support species-specific modelling. As a consequence, local sampling effort must be intensified at the expense of the number of samples, and thus of statistical power.

Alternatively, one can use surrogates of species richness, i.e. statistics that are well correlated with species richness but simpler to estimate from field samples. Variation in species richness can be revealed by variation in the diversity of higher taxa (families, subfamilies, or genera; Mazaris et al., 2008), or in the diversity of a subset of species easy to survey (Vellend et al., 2008).

In this study, we assessed an intermediary approach whereby independent predictive models are not computed on a species-specific basis, but after species were pooled into larger species groups. Surrogates of species richness were then derived from this series of independent models. We tested this approach on a Phyllostomid bat assemblage in a fragmented forest of French Guiana (Cosson et al., 1999; Henry et al., 2007), using models predicting variation in bat abundance across landscapes. Abundance is often adequately modelled by local habitat features and landscape configuration in Neotropical bats (Gorresen and Willig, 2004; Gorresen et al., 2005; Meyer and Kalko, 2008; Klingbeil and Willig, 2009). Forest fragmentation usually leads to the local decline of some groups of bats and sometimes favours the demographic success of others. In particular, large fig-eating bats of the genus Artibeus may show little variation (Cosson et al., 1999) or even an increase in abundance in fragmented areas (Meyer and Kalko, 2008), while small understorey frugivores (Henry et al., 2007) and gleaning animalivores are expected to decrease in abundance (Medellín et al., 2000; Klingbeil and Willig, 2009). Therefore, we expected greater discrepancies in the relative abundance of these species groups would help predict the extent of species loss as a result of fragmentation. We tested this hypothesis and further assessed whether independent landscape models of bat abundance may be combined to produce a single predictive map of spatial variation in species richness (figure 1).
Figure 1. Comparison of the direct modelling approach and the abundance-based approach for mapping landscape variations in species richness. In the abundance-based approach, landscape models are computed for each species group separately, and ultimately overlaid into a single diversity model used as a surrogate of species richness.
We compared two species grouping methods, one based on taxonomy, the other on functional traits (foraging guilds). These bat species classifications lead to the computation of higher taxa diversity and functional diversity, respectively. We addressed here the two following questions: (i) are higher taxa diversity and functional diversity effective surrogates of bat species richness at the landscape scale? (ii) does the abundance-based modelling preserve the surrogate efficacy of higher taxa diversity and functional diversity? Due to their high mobility, bats may rapidly adapt their distribution to modifications of landscape, and may be able to subsist in moderately fragmented areas. Therefore, we expected the models to predict a rapid decrease and stabilization of Phyllostomid bat species richness during the few years following fragmentation, and that these changes would concern the most fragmented areas.

1. Methods

1.1. Bat survey and study design

Analyses were performed using a bat survey dataset totalling 827 individual bats, belonging to 31 species, and captured in 18 sites in a fragmented rainforest in French Guiana, using standardised mist-net protocols (Cosson et al., 1999; Henry et al. 2007). Sites included four continuous forest plots and 14 fragments ranging in size from 0.8-7.5 ha. The fragmentation occurred in 1994 by the completion of a hydroelectric dam. Bats were sampled during two time periods, termed the “recent” and “older” fragmentation periods (2-4 yrs and 9-11 yrs, respectively, after the fragmentation occurred).

1.2. Landscape variables and landscape categories

The local landscape configuration around the 18 sampling sites was described using three landscape descriptors known to be relevant for bats (Meyer et al., 2008; Meyer and Kalko, 2008; Klingbeil and Willig, 2009). One measures the extent of forest cover (%), which is the prime habitat of most species in our study. The other two describe the amount and complexity of edge habitat, as edge-sensitivity of bat species was recognized to be the proximal determinant of their sensitivity to forest fragmentation in another Neotropical land-bridge system (Meyer et al., 2008). Edge cover (%) measures the proportion of space covered by edge habitat, delineated as a 50-m wide stripe bordering the forest. Fractal dimension measures the spatial complexity of edge habitat and was calculated using the box-counting method. Forest cover, edge cover, and fractal dimension were calculated within square landscape windows centred on sampling sites, and focusing on three different spatial scales with particular relevance considering movement abilities of bats: 500 m, 1 km and 2 km in width.
1.3. Estimators of species richness

Species richness was measured using the first-order Jackknife richness estimator \((J_s)\) and the Fisher’s log-series \((\alpha)\). \(J_s\) gives an estimation of total species richness and has been successfully used in many studies (Colwell and Coddington, 1994). The log-series \(\alpha\) is the parameter of the log-series function used to fit species rank-abundance distributions, and indicates the number of species represented by a single individual in the species assemblage (Magurran, 2004). We discarded samples with capture numbers smaller than the maximum number of species reported in a single site \((n=21)\). Of the 36 bat samples \((18 \text{ sites } \times \text{ 2 study periods})\) only nine fulfilled this rule \((six \text{ and three sites in recent and older fragmentation periods, respectively})\).

To avoid basing the validation of our modelling approach on such a limited portion of the initial dataset, we recomputed richness estimates after sampling sites were grouped into larger landscape categories. The 18 sampling sites were assigned to consistent categories based on the most relevant combinations of landscape variables and spatial scales for bats, namely forest cover and edge cover measured at a 2-km scale (see results, section 2.). We clustered sites using the K-means splitting method that maximizes between-cluster variation and minimizes within-cluster variation. The most parsimonious classification clustered sites into six landscape categories that overall delineated a consistent fragmentation gradient: (i) undisturbed continuous forest \((2 \text{ sites})\), (ii) peninsulas surrounded by the aquatic matrix \((2 \text{ sites})\), (iii to v) three groups of fragments of different combinations of size and isolation degree \((2, 5 \text{ and 5 sites})\), and (vi) small, remote, fragments \((2 \text{ sites})\). After merging bat captures within each landscape category, the capture database was reshaped into 12 samples \((6 \text{ landscape categories } \times \text{ 2 study periods})\). The sampling completeness given by the ratio \{observed species richness / species richness estimated by \(J_s\}\} was satisfactory \( (> 93.6\% \text{ in any case})\).

1.4. Diversity surrogates of species richness

The surrogates of species richness we investigate in this study require we classify species into consistent groups. We compared two grouping strategies, one based on taxonomy \((n=5 \text{ subfamilies within the family Phyllostomidae})\), the other based on diet and foraging habits \((n=8 \text{ functional groups defined following Kalko, 1998; Patterson et al., 2003; Delaval et al., 2005})\). Functional groups included understory frugivores, large fig-eating bats \((\text{body mass } >36g)\), small fig-eating bats \((\text{body mass } <25g)\), nectarivores, gleaning insectivores, omnivores, carnivores and sanguivores.

The diversity of higher taxa and of functional groups was assessed using three common estimators based on relative abundances, namely the Berger-Parker dominance index \((d)\), and the Simpson \((1/D)\) and Shannon \((H')\) indices of evenness (Magurran, 2004). These indices were computed after captured individuals were summed within taxonomic groups and within functional groups to measure higher taxa diversity and functional diversity, respectively.
1.5. Modelling bat abundance from landscape descriptors

We performed a landscape modelling of the abundance of each species group separately, and then combined the predictive abundance matrices in order to compute higher taxa diversity and functional diversity as described above. Bat abundances were expressed as number of captures per 10 capture nights. We then used Generalised Linear Models (GLMs) to find the best landscape correlates of abundance values (n=36). Abundance values were treated as capture count data and modelled with a negative binomial error structure and a log-link function. The negative binomial distribution has been previously proposed as an appropriate underlying distribution for modelling biological count data, as many biological mechanisms are likely to generate distributions that would be overdispersed with respect to the Poisson distribution (White and Bennett, 1996).

The explanatory variables considered in GLMs were the three landscape variables and the study period (recent vs. older fragmentation period). To assess which combination of one or more of these factors best explained variation in bat abundance, we built various models and used the AICc (Akaike Information Criterion corrected for small sample size; Anderson et al., 2001), to select the minimal adequate model considering fit and complexity. For the sake of parsimony, we only computed models that were ecologically sound by using the following criteria: (i) we did not include more than one landscape variable at a time into models; (ii) the interaction term period × landscape variable was only considered in models where the corresponding main terms were also included; (iii) we removed correlations linking fractal dimension with forest cover at a given focal scale by using the regression residuals instead of the raw fractal values. We found it unnecessary to apply the same correction for edge cover because it was not significantly correlated with forest cover at any focal scale.

Overall, 29 models were built and compared for each taxonomic group and each functional group, including the null model taking into account the intercept only. Models with the lowest AICc receive the greatest statistical support. Models with a difference of AICc < 2 compared to the lowest AICc were considered as statistically equivalent (Burnham and Anderson, 2002).

1.6. Computing functional and higher taxa diversity from predicted abundances

We used a leave-one-out cross-validation method (Davidson and Hinkley, 1997) to produce series of independent predictions of higher taxa diversity and functional diversity. The abundance models that received the greatest statistical support were rebuilt 36 times after deleting in turn one of the 36 capture samples. Each of these 36 repeated models was used to predict the abundance expected for the corresponding excluded sample. The resulting abundance predictions for each taxonomic group and functional group were used to compute predicted higher taxa diversity and functional diversity, respectively.
2. Results

2.1. Correlations between observed diversity and species richness

At the level of landscape categories, $J_s$ ranged from 8.2 to 22.7 and $\alpha$ from 2.7 to 7.6. These variations were significantly captured by all indices of higher taxa diversity and functional diversity computed from observed bat samples (table 1), with Pearson correlation $r$ values mostly ranging from 0.70 to 0.88.

Table 1. Assessment of higher taxa diversity and functional diversity as surrogates of bat species richness ($\alpha$ and $J_s$) in a fragmented neotropical rainforest, French Guiana. Surrogate efficacy is given as the Pearson correlation coefficient $r$ between species richness and functional and higher taxa diversity, either computed from raw abundance data or abundance data predicted by the landscape modelling (*** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$, § $P < 0.1$).

<table>
<thead>
<tr>
<th>Surrogates of $\alpha$</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Functional diversity computed...</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Higher taxa diversity</td>
<td>from raw data</td>
<td>from predicted data</td>
<td>Functional diversity</td>
<td>from raw data</td>
<td>from predicted data</td>
<td></td>
</tr>
<tr>
<td>Evenness ($H'$)</td>
<td>0.83**</td>
<td>0.66*</td>
<td>0.82**</td>
<td>0.70*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evenness ($1/D$)</td>
<td>0.83**</td>
<td>0.75**</td>
<td>0.72**</td>
<td>0.66*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dominance ($d$)</td>
<td>-0.82**</td>
<td>-0.77**</td>
<td>-0.83**</td>
<td>0.76**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surrogates of $J_s$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evenness ($H'$)</td>
<td>0.70*</td>
<td>0.76**</td>
<td>0.66*</td>
<td>0.43</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evenness ($1/D$)</td>
<td>0.66*</td>
<td>0.80**</td>
<td>0.70*</td>
<td>0.23</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dominance ($d$)</td>
<td>-0.74**</td>
<td>-0.88**</td>
<td>-0.67*</td>
<td>-0.28</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2.2. Abundance-based diversity modelling

The number of best candidate models for a given species group ranged from two (nectarivores) to nine (gleaning insectivores). Fragmentation period was included as an important predictor in 50% of the best candidate models, either alone or in combination with a landscape descriptor. The most relevant landscape descriptors were forest cover and edge cover (53% and 28% of best candidate models, respectively) while fractal dimension received little support (8%). About 40% of these models favoured the 2-km scale, while the 1- and 0.5-km scales were selected in 25% and 32% of the cases.
Of the ten selected models (table 2), three suggested an effect of fragmentation period and its interaction with a landscape variable, and five others only contained a landscape variable. None of the explanatory variables appeared relevant to account for abundance variations in omnivores and large fig-eating bats that remained rather ubiquitous and stable over time. For these groups, we therefore used mean abundances to compute diversity estimates. Likewise, we used abundances averaged by study period for sanguivores. Due to an apparent population decline in older fragmentation period, their capture numbers were too scarce to compute any predictive model of abundance.

Table 2. Summary of the best candidate models for explaining landscape variations in abundance of each taxonomic and functional bat species group. Models were GLMs fitted using the negative binomial distribution, and selected according to the AIC corrected for small sample size. Models may include a landscape descriptor, the period effect, and their two-way interaction, indicated as “×”. Null models were used when none of the tested models explained a significant proportion of total deviance (likelihood ratio test).

<table>
<thead>
<tr>
<th>Level of Analysis</th>
<th>Species group</th>
<th>Best candidate model</th>
<th>Most appropriate spatial scale for landscape effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Higher taxa</td>
<td>Caroliinae</td>
<td>Forest cover (2km)</td>
<td>1 km</td>
</tr>
<tr>
<td>Level</td>
<td>Glossophaginae</td>
<td>Period × Forest cover (500m)</td>
<td>250 m</td>
</tr>
<tr>
<td></td>
<td>Phyllostominae</td>
<td>Period × Edge cover (500m)</td>
<td>250 m</td>
</tr>
<tr>
<td></td>
<td>Stenodermatiniae</td>
<td>Edge cover (0.5km)</td>
<td>250 m</td>
</tr>
<tr>
<td>Functional Level</td>
<td>Carnivores</td>
<td>Forest cover (2km)</td>
<td>1 km</td>
</tr>
<tr>
<td></td>
<td>Gleaning insectivores</td>
<td>Period × Edge cover (2km)</td>
<td>1km</td>
</tr>
<tr>
<td></td>
<td>Large fig-eating bats</td>
<td>Intercept (null model)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Nectarivores</td>
<td>Period × Forest cover (500m)</td>
<td>250 m</td>
</tr>
<tr>
<td></td>
<td>Omnivores</td>
<td>Intercept (null model)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Small fig-eating bats</td>
<td>Forest cover (0.5 km)</td>
<td>250 m</td>
</tr>
<tr>
<td></td>
<td>Understorey frugivores</td>
<td>Forest cover (1 km)</td>
<td>500 m</td>
</tr>
</tbody>
</table>
2.3. Correlations between predicted diversity and species richness

When computing diversity estimates from predicted abundance data, only higher taxa diversity remained significantly or marginally significantly correlated with species richness (table 1), while functional diversity did not. In other words, the modelling procedure preserved the efficacy of higher taxa diversity as surrogate of species richness, but not that of functional diversity.

The higher taxa Berger-Parker’s $d$ appeared to be the most convenient option (table 1). The resulting predictive maps of higher taxa $d$ produced two very different patterns for recent and older fragmentation periods (figure 2). While the diversity drop was mostly restricted to small, isolated, forest fragments in the recently fragmented system (figure 2A), the model predicts a pervasive diversity decrease affecting all forest fragments and gaining a vast portion of the continuous forest in older period (figure 2B). The decrease is most marked in areas with high edge cover (a nearly -60% decrease, figure 2C).
3. Discussion

3.1. Higher taxa diversity and functional diversity as surrogates of species richness

Higher taxa diversity and functional diversity provided a straightforward way of appraising variation in bat species richness in the St-Eugène fragmented forest. We used here common diversity indicators based on relative abundance data. The use of some of these indicators in diversity surveys, like the Shannon-Wiener $H'$, is controversial because they confound two aspects of diversity, namely species numbers and evenness (Magurran, 2004), making it difficult to isolate the real causes of diversity variation. However, when computed at higher taxa or functional levels, the numbers of species groups vary little, and diversity indicators mostly reveal variation in numerical evenness. Ecologists reluctant to use $H'$ may still find a valuable alternative with $d$ as a surrogate of species richness. In this context, $d$ refers to the relative proportion of the numerically dominant species group in the assemblage, and is inversely related to species richness.

3.2. The abundance-based diversity modelling

The abundance-based modelling of higher taxa diversity succeeded to predict major variation in bat species richness with a satisfactory accuracy (correlation coefficients up to $r=0.80$ to 0.88). This property could stem from two main characteristics of the bat assemblages. First, the fragmentation had a marked effect on bat abundances and species
composition at the spatial and temporal scale of our study, making it easier for the models to detect variation in bat assemblages among sites with contrasted landscape structure. Second, the abundance of most taxa could be adequately modelled as a function of particular landscape features.

The fact that the abundance-based modelling of functional diversity did not perform equally well does not mean that this approach should be discarded. Indeed, the functional classification was well supported from an ecological point of view because functional groups are supposed to be independent entities regarding food resources. In other contexts, the ecological dimension associated with the functional classification may offer advantages over the taxonomic approach. Even though predicted functional diversity is not tightly linked with species richness, it may still be correlated with, and used as an indicator of other ecological processes, e.g. seed and pollen dispersal by frugivorous and nectarivorous bats.

3.3. Limits and constraints on the abundance-based diversity modelling

In some circumstances, grouping species may be impossible or irrelevant because assemblages are species-poor. In this context, however, the usefulness of modelling landscape variation in species richness appears less critical. Species-specific modelling could be sufficient for most conservation purposes (Zipkin et al., 2009). Conversely, some community surveys are not limited by sample sizes (e.g. insects), and then would not gain in efficiency by applying the abundance-based diversity modelling. The method as developed herein is more appropriate for situations where diversity parameters cannot be conveniently estimated in all sampled sites due to small sample sizes, typically in complex vertebrate assemblages such as bats, birds, and reptiles. It may also be considered as an alternative to species richness when reliable species identification is tedious (e.g. in invertebrates or plants; Vellend et al., 2008) or restricted to broad morphological or taxonomic categories.

An important prerequisite for the abundance-based modelling to yield useful richness surrogates is that the abundance of the numerically dominant species group(s) should be adequately modelled as a function of landscape features or other environmental variables. Diversity indicators used in this study are highly sensitive to variations of the most abundant species groups. In the present study, the abundance of large fig-eating bats (> 42% of phyllostomid bat captures) appeared to vary independently from the landscape descriptors we used. This may partly explain the failure of functional diversity to withstand the modelling step.

The number of species groups introduced in the abundance-based modelling is another critical parameter. Splitting species into more groups reduces group sizes, and hence the predictive power of abundance models. Such an over-categorisation of species is another possible reason for the failure of functional diversity to withstand the modelling step. Bat functional groups are actually refined subsets within taxonomic groups, in that bat species
from the same functional group also belong to the same taxonomic group. Therefore, the two classification strategies bear similar information, but differ in the number of species groups (n = 5 vs. 8). This suggests that parsimony is recommended in terms of group numbers, and that an optimal classification must be evaluated a priori from expert knowledge.

Conclusions

In the neotropical bat assemblage we studied, taxonomic diversity and functional diversity were both valuable surrogates of species richness, but only the taxonomic diversity level sustained satisfactorily our abundance-based modelling method. The abundance-based modelling of higher taxa diversity appears to be a viable alternative to direct species richness modelling when the number and size of capture samples are limited, and may be used to improve the design of rapid assessment diversity surveys.

Furthermore, the abundance-based modelling better accounts for the multifaceted nature of bat assemblages. It combines different landscape models computed over various spatial scales, each model being the most appropriate for one of the various species group. Previous studies on bats (Gorresen and Willig, 2004; Gorresen et al., 2005; Meyer and Kalko, 2008) have clearly shown that landscape variation in species richness or diversity is scale-dependent. The abundance-based modelling method offers a straightforward way to account for this scale dependence. It does not completely solve the problem of summarizing the ecological complexity of species assemblages into a single metric. It may, however, provide a valuable alternative to predict landscape areas subject to substantial species loss and to compare the potential outcomes of different forest restoration strategies.

References


