Is spectral distance a proxy of beta diversity at different taxonomic ranks? A test using quantile regression

Duccio Rocchini [a,b,*], Kate S. He [c], Jianting Zhang [d]

[a] Dipartimento di Scienze Ambientali "G. Sarfatti", Università di Siena, via P.A. Mattioli 4, 53100, Siena, Italy
[b] TerraData environmetrics, Dipartimento di Scienze Ambientali "G. Sarfatti", Università di Siena, via P.A. Mattioli 4, 53100, Siena, Italy
[c] Department of Biological Sciences, Murray State University, Murray, Kentucky 42071, USA
[d] Department of Computer Science, The City College of the City University of New York, New York, NY 10031, USA

ABSTRACT

Beta diversity represents a powerful indicator of ecological conditions because of its intrinsic relation with environmental gradients. In this view, remote sensing may be profitably used to derive models characterizing or estimating species turnover over an area. While several examples exist using spectral variability to estimate species diversity at several spatial scales, most of these have relied on standard correlation or regression approaches like the common Ordinary Least Square (OLS) regression which are problematic with noisy data. Moreover, very few attempts were made to derive beta diversity characterization models at different taxonomic ranks. In this paper, we performed quantile regression to test if spectral distance represents a good proxy of beta diversity considering different data thresholds and taxonomic ranks. We used plant distribution data from the North and South Carolina including 146 counties and covering a variety of vegetation formations. The dissimilarity in species composition at different taxonomic ranks (using Sørensen distance) among pairs of counties was compared with their distance in NDVI values derived from 23 yearly MODIS images. Our results indicate that (i) spectral variability represents a good proxy of beta diversity when appropriate statistics are applied and (ii) a lower taxonomic rank is important when changes in species composition are examined spatially using remotely sensed data.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Accounting for beta diversity, i.e., the amount of turnover in composition (see Whittaker, 1960), of an area instead of simply relying on its local species richness (alpha diversity) may result in additional information for several ecological tasks allowing to: i) improve species inventorying by firstly visiting areas with higher environmental differences thus leading to a higher number of inventoried species given the same sampling effort (Rocchini et al., 2005); ii) relate the distribution of species diversity to geographical or environmental gradients at regional instead of local scale (Davidar et al., 2007; Buckley and Jetz, 2008); iii) improve conservation policies and strategies for biodiversity network design by maximizing species composition turnover (Chiarucci et al., 2008).

In this view, remote sensing may be profitably used to derive models characterizing or estimating species diversity over an area (see Gillespie et al., 2008 and references therein). As stressed by Nagendra (2001), when dealing with the design of meaningful conservation

strategies, it is nearly impossible to acquire diversity information only on the basis of field assessment. Hence, ancillary information may become crucial. As an example, knowing a-priori areas with higher diversity may help in minimizing time and costs in monitoring efforts (e.g., Bacaro et al., 2008). In recent years, considerable amount of research has been carried out to predict species rich sites by remotely sensed data based on spectral variability as summarised by the Spectral Variation Hypothesis, SVH (see Palmer et al., 2002) considering different types of organisms, such as vascular plants (see e.g., Gould, 2000; Foody and Cutler, 2003; Fairbanks and McGwire, 2004, Gillespie, 2006; Levin et al., 2007; Rocchini, 2007a, b), lichens (Waser et al., 2004), birds (Bino et al., 2008; Oindo, 2008), mammals (Oindo and Skidmore, 2002). Some examples exist about the investigation of alpha diversity with increasing local grain (Kumar et al., 2006) or about local diversity in species composition versus spectral variability (Schmidtlein and Sassin, 2004). However, most of the tests were substantially based on predictive regression models of species richness vs. spectral variability at local scale, considering only alpha diversity.

To date few efforts have been made to relate ecological variability measured by remote sensing to the other components of species diversity, such as beta diversity taking into account the species turnover over the whole extent of a study area (Tuomisto et al., 2003;

* Corresponding author. Dipartimento di Scienze Ambientali "G. Sarfatti", Università di Siena, via P.A. Mattioli 4, 53100, Siena, Italy.
E-mail address: rocchini@unisi.it (D. Rocchini).

Rocchini et al., 2005). The major assumption beyond the use of spectral distance for predicting species beta diversity is that a higher variability in the spectral response over a landscape derives from the presence of several different habitats, which are expected to host different species.

In a recent paper, He et al. (2009) firstly investigated the relation between beta diversity and spectral distance by considering different taxonomic ranks (plant species, genera and families) at the regional scale. In that paper, a Pearson correlation obtained from a Mantel test based on the whole set of data points was applied. While they provide for the quantification of the general trend in the data, using standard correlation or regression approaches, like the common Ordinary Least Square (OLS) regression in these cases, may represent a potential problem. In fact, when applying symmetric loss functions which minimize the residual sum of square, the final model will only rely on the mean of an over-dispersed cloud of points (Cade and Noon, 2003). Applying other regression techniques can overcome the problem and contribute with additional information which might be lost or underestimated by classical regression analysis. As an example, one may ask if, considering different part of the cloud of input data in a scatterplot, the pattern achieved is consistent over different thresholds (see Brown and Peet, 2003), e.g., if the slope of the regression model is still significant across the whole turnover range, from lower to upper threshold. In this view, quantile regression may represent the most efficient method for solving the problem, fitting models to any part (denominated τ) of a response distribution (Koenker and Bassett, 1978; Cade and Noon, 2003; Austin, 2007).

In this paper, extending on He et al. (2009), we aim at testing if spectral distance represents a good proxy of beta diversity considering different data thresholds (τ_s) and taxonomic ranks.

2. Methods

2.1. Ordinary least square and quantile regression fitting procedure

Let $\{y_1, y_2, ..., y_n\}$ denote the vector of the y_n values of a set of points lying within a scatterplot of a response variable Y versus a limiting factor X.

OLS regression symmetrically minimizes residuals by solving:

$$\text{residual} = min \sum \left(y_i - \hat{y}_i\right)^2 \tag{1}$$

where $\hat{y}_i$ = estimated value for each y_i.

Quantile-based fitting gives different weights to positive and negative residuals leading to an asymmetric minimization, solving:

$$\text{residual} = min \sum |y_i - \hat{y}_i| T \quad \text{with } T = \begin{cases} \tau & \text{for } (y_i - \hat{y}_i) > 0 \\ 1 - \tau & \text{for } (y_i - \hat{y}_i) < 0 \end{cases} \tag{2}$$

Notice that the multiplier T equals τ (the quantile value) for positive deviations $(y_i - \hat{y}_i)$ and $(1 - \tau)$ for negative deviations. This asymmetric minimization fits a regression model through the upper part of the response distribution for $\tau > 0.5$ and through the lower part of the distribution for $\tau < 0.5$. Quantile regression with $\tau = 0.5$ is the median regression, which can be used as a central regression line similar to the mean regression estimated with OLS regression.

Notice that the quantile minimization of residuals shown in Eq. (2) is based on absolute values rather than on squared deviations like in OLS regression, thus reducing outlier effects.

We refer to Koenker and Hallock (2001) for a more detailed dissertation on the matter and to Gotelli and Ellison (2004) for a brief summary of quantile regressions applied to ecological data.

2.2. Empirical test on the prediction of beta diversity by spectral distance

2.2.1. Study area and input data

In order to test the beta diversity in relation with spectral distance we used data from the North and South Carolina including 146 counties of the two states (Fig. 1). The vegetation patterns are quite complex in the study area. The two states cover a range of habitats including coastal lowlands, large river floodplain forests, rolling plains, forested mountains, and wetlands. According to the EPA level III ecoregion classification system (http://www.epa.gov/wed/pages/ecoregions/level_iii.htm#Ecoregions), there are four Level III ecoregions in the two states, containing Blue Ridge, Piedmont, Southeastern Plains, and Middle Atlantic Coastal Plain (See Fig. 1 in He et al., 2009 for the boundaries of these four ecoregions). More specific, the Blue Ridge is covered by the Appalachian oak forests, northern hardwoods, and southeastern spruce-fir forests at higher elevations. The Piedmont, a nonmountainous portion of the old Appalachians Highland, is dominated by successional pine and hardwood woodlands. The Southeastern Plains include a mosaic of forest woodland (oak-hickory-pine and southern mixed forest) and pasture/cropland. Lastly, the Middle Atlantic Coastal Plain contains mostly swamps and salt marshes, and some forested areas including loblolly and some shortleaf pine species, with patches of other trees such as oak, gum, and cypress, etc. Given the diverse vegetation formations found in the Carolinas, the variability in productivity is highly expected.

For species information, we used plant species data obtained from USDA plant database (http://plants.usda.gov) with a total of 3157 species in 1001 genera and 189 families and a mean number of 687 (± 171 sd) species, 375 (± 66 sd) genera and 116 (± 15 sd) families per county. The USDA plant database represents a huge effort made to provide species information across the United States (see Brown and Peet, 2003). One county did not have species data, thus reducing the set to 145 counties.

In order to calculate spectral distance we relied on Moderate Resolution Imaging Spectroradiometer (MODIS) Normalized Differential Vegetation Index (NDVI) data (spatial resolution: 250 m, temporal resolution: 16 days). Such data are freely available from the University of Maryland's Global Land Cover Facility (GLCF) project (http://glcf.umiacs.umd.edu/data/ndvi, see Tucker et al., 2004). NDVI is derived from the ratio $(\lambda_{NIR} - \lambda_R)/(\lambda_{NIR} + \lambda_R)$, where λ_{NIR} = reflectance in the near infrared part of the spectrum and λ_R = reflectance in the red part of the spectrum. NDVI is based on (i) the high reflectance by vegetation in the NIR which is linked to scattering processes at the leaf scale and (ii) the low reflectance in the Red due to the absorption by chloroplasts for photosynthesis (see Lillesand et al., 2004). NDVI varies from -1 to 1, but in this case they were stretched to match a 8-bit radiometry (ranging from 0 to 255) in order to make comparable results with other studies on the matter using Digital Numbers (DNs) in other bands (e.g., Rocchini and Cade, 2008).

There are 23 MODIS NDVI images for a whole year. We downloaded the 23 NDVI images for the year 2005 and imported them within the ESRI ArcGIS software. Since several pixels occur within each county, the mean value per county (per each image) was calculated. Therefore, each county had 23 available NDVI values. Based on the conclusions by He et al. (2009) we decided to use the whole yearly mean NDVI data. Hence, each county was provided with a single spectral value (i.e., the yearly mean NDVI) with which we calculated spectral distance between pair-wise counties as described in the next section.

2.2.2. Beta diversity versus spectral distance modeling by OLS and quantile regression

Beta diversity was measured by generating semi-matrices of pairwise distance based on the Sørensen coefficient (Sørensen, 1948; but see even Koleff et al., 2003) for the 145 counties at the three taxon
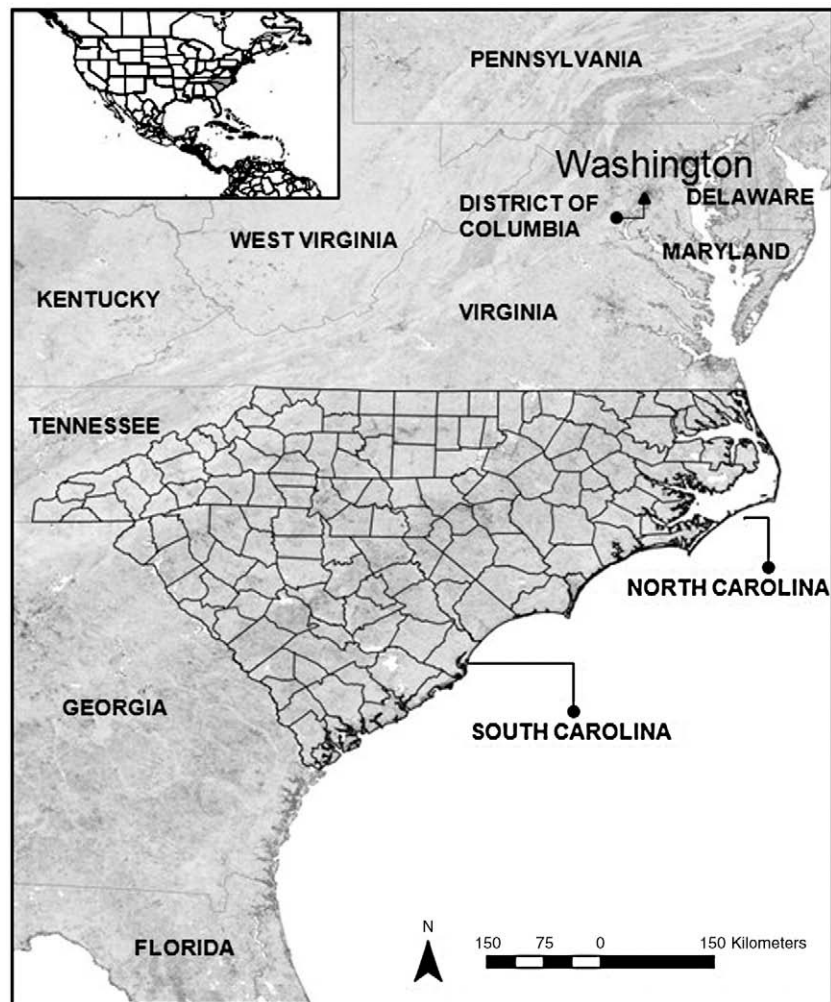
Fig. 1. The study area: North and South Carolina. In this study 145 out of 146 counties of Carolina were considered. Species data were derived from the USDA plant database while spectral data were based on MODIS imagery. In this example a MODIS NDVI image acquired in July 17th 2005 is shown. Brighter gray represents higher NDVI values. The pixels in the sea have no data values and have been excluded from mapping. We refer to the main text for major explanations.

ranks, including species, genus, and family. Sørensen coefficient (C_s) is derived as:

$$C_s = \frac{2j}{a+b} \qquad (3)$$

where $j =$ number of species shared by plot A and plot B; $a =$ total number of species in the plot A; $b =$ total number of species in the plot B, the coefficient C_s accounting for the overlap between two species lists and ranging from 0, indicating perfect dissimilarity, to 1, indicating perfect similarity.

At the same time, spectral distance was achieved by calculating NDVI distance matrices using Euclidean distance between pair-wise counties.

As a rule of thumb in quantile regression the function to be fitted is based on visual inspection of the scatterplot derived from the data being modeled (e.g., Cade and Noon, 2003; Rocchini and Vannini, in press). Thus, after visual inspection of the achieved pattern, linear models were built to minimize residuals as in Eq. (1) (OLS regression) and Eq. (2) (quantile regression). Concerning quantiles, we assessed the continuous variation of the slope of the linear models from lower ($\tau = 0.10$) to upper ($\tau = 0.90$) beta diversity values. A number of statistical software packages exist which perform quantile regression (see e.g., Blossom at USGS internet site, Cade and Richards, 2001). Instead, we used the quantreg package of *R*-software (Koenker, 2007)

on the strength of its widespread use and simple replicability of functions coded (R Development Core Team, 2008).

It has been realized that a false number of degrees of freedom can be created by distance-based models (in this case $df = 10,438$, see Legendre and Legendre, 1998; Tuomisto et al., 2003; Legendre et al., 2005; He et al., 2009). This may provoke an improper estimate of confidence intervals. Therefore, we relied on Davidar et al. (2007) who calculated confidence intervals using permutation procedures based on a bootstrapping approach. We used the boot.rq function (quantreg R-package, see Koenker, 2007) for quantile regression and the bootpred function (bootstrap package, see Efron and Tibshirani, 1994, Leisch, 2007) for OLS.

3. Results

A high degree of noise was found within the scatterplot of beta diversity versus spectral distance (Fig. 2). The slope of the relation was statistically different from zero considering all parts of the response distribution (beta diversity) conditional on the constraining factor (spectral distance, Fig. 3) despite the taxonomic rank being considered. Instead, taxonomic ranks showed marked differences in the slopes achieved by linear models, as described in the following paragraphs.

Considering the species taxonomic rank, higher quantiles showed statistically higher slopes ($p < 0.01$) than that achieved by OLS regression ($\text{slope}_{OLS} = 18.4 * 10^{-4}$), with a peak in slope at $\tau \cong 0.7$ (Fig. 3). This
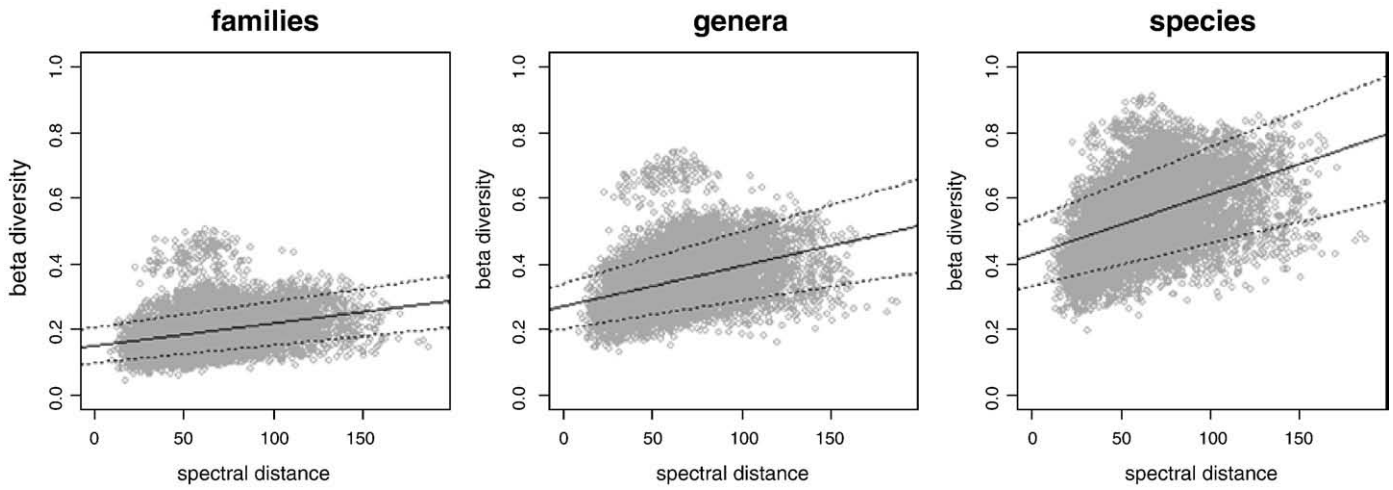
Fig. 2. Regression models of taxonomic versus spectral turnover considering three different taxonomic ranks. Solid line: Ordinary Least Square model; dashed lines: examples of quantile regressions linear models fitted at $\tau = 0.90$ (top curve) and $\tau = 0.10$ (bottom curve). Notice that taxonomic turnover measured by the Sørensen index was plotted with the same range in the y axis (0–1) thus allowing a direct comparison among the taxonomic ranks.
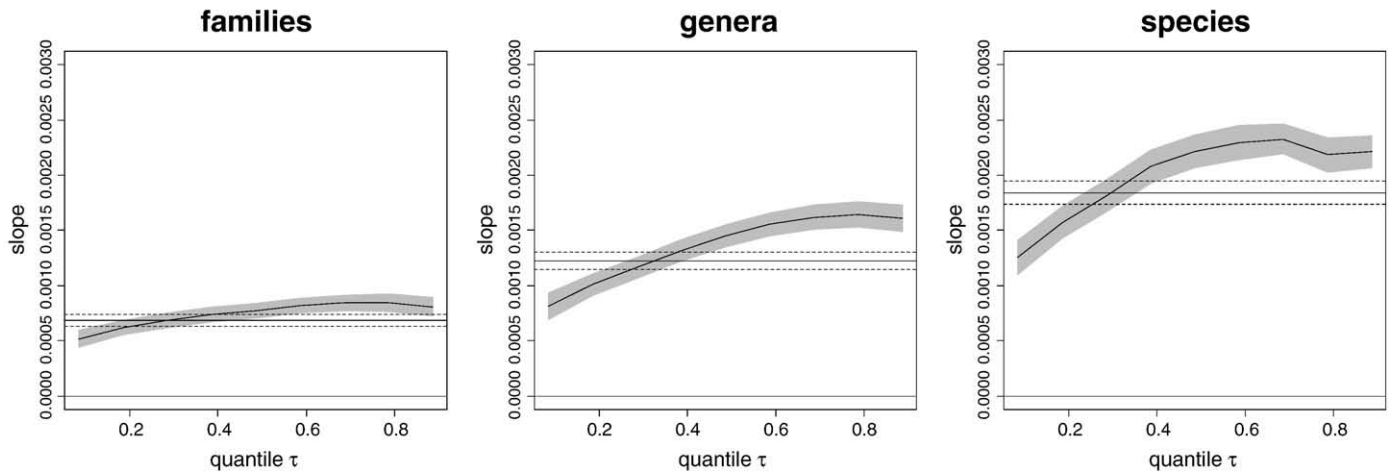

Fig. 3. Variation (solid black line) of the slope coefficient of the considered linear models. The gray area represents 99% confidence intervals; thus when they include zero (solid gray line) the coefficient is not significant at $p < 0.01$. Notice that slopes are always significant at $p < 0.01$. The horizontal solid black line represents the slope value with 99% confidence intervals (dashed black lines) obtained by OLS regression. Notice that slopes were plotted with the same range in the y axis thus allowing a direct comparison among the taxonomic ranks. A similar representation of quantile and OLS parameters is provided by Koenker and Hallock (2001), Cade and Noon (2003) and Austin (2007).

phenomenon flattened while increasing the taxonomic ranks, i.e., passing from plant species to families (Figs. 2 and 3). Passing from higher to lower taxonomic ranks, families showed lower slopes ($slope_{OLS} = 6.9 * 10^{-4}$) than those achieved for genera ($slope_{OLS} = 12.3 * 10^{-4}$) and species ($slope_{OLS} = 18.4 * 10^{-4}$).

This is true even considering slopes achieved by quantile regression. As an example, at $\tau = 0.90$ (Figs. 2 and 3), slope increased at a rate up to $13.9 * 10^{-4}$ passing from families ($slope_{\tau = 0.90} = 7.9 * 10^{-4}$) to genera ($slope_{\tau = 0.90} = 16.0 * 10^{-4}$) to species ($slope_{\tau = 0.90} = 21.8 * 10^{-4}$).

This difference among taxonomic ranks flattened at lower quantiles (lower taxonomic diversity values). For instance, at $\tau = 0.10$ (Figs. 2 and 3), slope increased at a rate up to $7.7 * 10^{-4}$ (i.e., halving that achieved at $\tau = 0.90$) passing from families ($slope_{\tau = 0.10} = 5.4 * 10^{-4}$) to genera ($slope_{\tau = 0.10} = 8.5 * 10^{-4}$) to species ($slope_{\tau = 0.10} = 13.1 * 10^{-4}$).

Hence, the range in variability in the similarity values flattened passing from species to families. As an example, concerning the difference between slopes at $\tau = 0.90$ and $\tau = 0.10$ given the same taxonomic rank this turned out to be $8.7 * 10^{-4}$ for species until $2.5 * 10^{-4}$ for families.

Since the difference between species and families was lower at lower quantiles, this range flattening effect seems to be regulated by data points situated at higher quantiles, i.e., by higher diversity values. In other words, once families instead of species were taken into account, they would not reach diversity values higher than 0.5 (Fig. 2).

4. Discussion

Very different methods have been used explicitly for modeling beta diversity by (environmental) ancillary information like the generalised dissimilarity models (GDM, e.g., Ferrier et al., 2002, 2007), boundary line regression (Rocchini, 2007a), Mantel tests (Tuomisto et al., 2003; He and Zhang, 2009), It is far beyond the aim of this paper to prove quantile regression being better than other methods. However, we claim that it directly takes into account the noise frequently found when plotting beta diversity – as measured by dissimilarity between pair-wise local or regional sites – versus spectral distance. Such noise is expected since other factors are expected to act on species variability structuring as well, such as (i) phenological variations or (ii) abiotic factors like elevation, soil, climate gradients,

Further, practical problems can arise when linking field data with satellite images, e.g., (i) the mismatch between the grain of the satellite imagery and the field data used, (ii) the difference between the time of satellite image acquisition and the field survey period, etc.

(see Gillespie et al., 2008), (iii) the distance measurement used (e.g., in this case we used the Euclidean distance among averaged NDVI which does not differentiate the spectral response of communities), (iv) the completeness of the input dataset (e.g., Palmer et al., 2002).

Concerning the last point, finding complete sets at regional to global scales is a hard issue to be solved. Plant species inventorying in a relatively large area has always been an important task for plant ecologists, given the lack of common standards in measuring the completeness of the resulting species lists and in quantifying the sampling effort (see e.g., Palmer et al., 2002; Rocchini et al., 2005). In this paper the USDA plant database was used since it has been proven to be an informative data source for testing general ecological patterns like diversity of Appalachian (East US) plant communities (Brown and Peet, 2003), exotic species invasions (Frappier and Eckert, 2003), and variability in species composition (He et al., 2009).

Of course, this dataset is not free from bias, and we suspect that no dataset is definitively free from input uncertainty. As a couple of examples, (i) one county contained no data and (ii) the outliers in the upper set of Fig. 2 may be explained if some counties had a lower number of available data thus increasing beta diversity values.

However, in this paper we were interested in testing a general pattern (i.e., the variability in species versus spectral distance) at different taxonomic ranks given the same input dataset, which is expected to contain some sort of noise in its very nature and is thus modeled using quantile coupled with OLS regression. In this view, we proved quantile regression to be a straightforward method for describing the whole pattern of beta diversity versus spectral (environmental) distance. In particular, higher quantiles, representing maximum beta diversity, seemed to rule out the differences among taxonomic ranks. In other words, the slopes of species, genera and families differed mostly in the higher part of the scatterplot. This suggests that using higher taxonomic ranks (like "families" in this case) would lead to a dramatic loss of information, in particular that related to the maximum beta diversity over an area. Our results indicate that a lower taxonomic rank is important when changes in the taxonomic composition are examined spatially using remotely sensed data.

Moreover, our findings are in line with previous studies that taxonomic rank or resolution might affect our ability to detect useful patterns in biodiversity study (Anderson et al., 2005 and references therein). Of course we agree with the issues in taxonomy demonstrating that the species taxonomic rank is more prone to taxonomic inflation which is expected to increase error variance instead of enhancing the information content (see Bacaro et al., 2009; but refer even to Knapp et al., 2005). Using higher taxonomic ranks such as families or even phyla could be effective and efficient, but critical information might be lost, as suggested by the quantile regression presented in this paper. This hampers our effort in seeking for the structure and spatial pattern of biological communities from the taxonomic data. We postulate that a lower taxonomic rank, such as species, can be sufficiently informative in local or regional studies to examine biotic interactions, facilitation, and dispersal. In contrast, higher taxonomic ranks might be useful in large-scale spatial studies relating to biogeographical and evolutionary processes (He and Zhang, 2009).

As far as we know, the only example using quantile regression in beta diversity versus spectral distance models is done by Rocchini and Cade (2008). In their test, only the species taxonomic rank was considered and at a local scale in terms of both extent (60 km) and grain of sampling units (10×10 m). As a result, a different pattern was found with lower quantiles showing higher slopes. If the grain of sampling units is small enough, a lower (than expected) number of species will be shared by two plots even if their environmental properties are the same (Nekola and White, 1999). This inevitably leads to maintain practically constant the maximum values of diversity, thus provoking a decrease in maximum slopes.

Spectral variability has proven to represent a powerful indicator of environmental heterogeneity (Palmer et al., 2002; Foody and Cutler, 2006; Rocchini, 2007b). As hypothesized by several authors (Kerr and Ostrovsky, 2003; Rocchini et al., 2005; Foody and Cutler, 2006), spectral distance represents a direct effect of environmental properties thus representing a powerful tool for gradient analysis and species diversity comparisons. As an example, Rocchini et al. (2005) demonstrated that ordering samples by their maximum spectral distance will account for a higher number of accumulated species. In other words, sites being spectrally more different might show a higher species turnover. We consider that the compositional differences of vegetation types are caused by spatial heterogeneity linked to regional climatic constraints, latitude, elevation, and historical processes. Our results suggest that remotely sensed spectral distance can be viable for detecting species compositional changes at the regional scale.

Of course, we are aware that there are costs associated with the use of remotely sensed imagery. There is still a gap in the easy availability of such data across the world (Goetz, 2007). As recently stressed by Kark et al. (2008) and by Nagendra and Rocchini (2008), this is particularly true considering developing countries. It is far beyond our aim to explore this issue and we remind to the debate made by Loarie et al. (2007) and the replies of Kark et al. (2008) and Loveland et al. (2008). However, as reported in this paper, spectral data such as MODIS NDVI are available freely on the internet and represent an undoubted source for the maximization of biodiversity monitoring. Further, the utility of such data source could facilitate biodiversity studies at the regional and even global scales, thus leading to new information for conservation planning and resource management.

5. Conclusion

In this paper, we tested the potential of using spectral distance to predict beta diversity at the regional scale. In particular, we demonstrated that: i) beta diversity has a statistically significant relation with spectral variability, since plant compositional turnover monotonically increases with increasing spectral distance among sites, and ii) the taxonomic rank being considered when performing beta diversity-based modeling is a crucial point which should be taken into account when aiming at monitoring biodiversity change over time.

These findings were attained by coupling Ordinary Least Square and quantile regression which provided a more comprehensive characterization of the process being considered than provided by estimates of the conditional mean with OLS regression. We perceive the utility in quantile regression as stated by Cade et al. (1999): 'estimating a range of regression quantiles [...] provides a comprehensive description of biological response patterns for exploratory and inferential analyses in observational studies of limiting factors, especially when sampling large spatial and temporal scales'.

Acknowledgements

References

Anderson, M.J., Connell, S.D., Gillanders, B.M., Diebel, C.E., Bloms, W.M., Saunders, J.E., Landers, T.J., 2005. Relationship between taxonomic resolution and spatial scales of multivariate variation. J. Anim. Ecol. 74, 636–646.
Austin, M., 2007. Species distribution models and ecological theory: a critical assessment and some possible new approaches. Ecol. Model. 200, 1–19.

Bacaro, G., Rocchini, D., Bonini, I., Marignani, M., Maccherini, S., Chiarucci, A., 2008. Ecological determinants of plant species richness at local scale in Mediterranean forests. Plant Biosyst. 142, 630–642.

Bacaro, G., Baragatti, E., Chiarucci, A., 2009. Using taxonomic data to assess and monitor biodiversity: are the tribes still fighting? J. Environ. Monit. 11, 798–801.

Bino, G., Levin, N., Darawshi, S., van der Hal, N., Reich-Solomon, A., Kark, S., 2008. Landsat derived NDVI and spectral unmixing accurately predict bird species richness patterns in an urban landscape. Int. J. Remote Sens. 29, 3675–3700.

Brown, R.L., Peet, R.K., 2003. Diversity and invisibility of southern Appalachian plant communities. Ecology 84, 32–39.

Buckley, L.B., Jetz, W., 2008. Linking global turnover of species and environments. PNAS 105, 17836–17841.

Cade, B.S., Noon, B.R., 2003. A gentle introduction to quantile regression for ecologists. Front. Ecol. Environ. 1, 412–420.

Cade, B.S., Richards, J.D., 2001. User manual for BLOSSOM Statistical Software. U.S. Geological Survey, Midcontinent Ecological Science Center, Fort Collins, CO.

Cade, B.S., Terrell, J.W., Schroeder, R.L., 1999. Estimating effects of limiting factors with regression quantiles. Ecology 80, 311–323.

Chiarucci, A., Bacaro, G., Rocchini, D., 2008. Quantifying plant species diversity in a Natura 2000 network: old ideas and new proposals. Biol. Conserv. 141, 2608–2618.

Davidar, P., Rajagopal, B., Mohandass, D., Puyravaud, J.P., Condit, R., Wright, S.J., Leigh, E.G., 2007. The effect of climatic gradients, topographic variation and species traits on the beta diversity of rain forest trees. Glob. Ecol. Biogeogr. 16, 510–518.

Efron, B., Tibshirani, R.J., 1994. An Introduction to the Bootstrap. Chapman and Hall, Boca Raton, USA.

Fairbanks, D.H.K., McGwire, K.C., 2004. Patterns of floristic richness in vegetation communities of California: regional scale analysis with multi-temporal NDVI. Glob. Ecol. Biogeogr. 13, 221–235.

Ferrier, S., Drielsma, M., Manion, G., Watson, G., 2002. Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales. II. Community-level modelling. Biodivers. Conserv. 11, 2309–2338.

Ferrier, S., Manion, G., Elith, J., Richardson, K., 2007. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. Divers. Distrib. 13, 252–264.

Foody, G.M., Cutler, M.E.J., 2003. Tree biodiversity in protected and logged Bornean tropical rain forests and its measurement by satellite remote sensing. J. Biogeogr. 30, 1053–1066.

Foody, G.M., Cutler, M.E.J., 2006. Mapping the species richness and composition of tropical forests from remotely sensed data with neural networks. Ecol. Model. 195, 37–42.

Frappier, B., Eckert, R.T., 2003. Utilizing the USDA PLANTS database to predict exotic woody plant invasiveness in New Hampshire. For. Ecol. Manag. 185, 207–215.

Gillespie, T.W., 2006. Predicting woody-plant species richness in tropical dry forests: a case study from South Florida, USA. Ecol. Appl. 15, 27–37.

Gillespie, T.W., Foody, G.M., Rocchini, D., Giorgi, A.P., Saatchi, S., 2008. Measuring and modelling biodiversity from space. Prog. Phys. Geogr. 32, 203–221.

Goetz, S., 2007. Crisis in earth observation. Science 315, 1767.

Gotelli, N., Ellison, A., 2004. A Primer Of Ecological Statistics. Sinauer Associates, Sunderland, Massachusets, USA.

Gould, W., 2000. Remote Sensing of vegetation, plant species richness, and regional biodiversity hot spots. Ecol. Appl. 10, 1861–1870.

He, K., Zhang, J., 2009. Testing the correlation between beta diversity and differences in productivity among global ecoregions, biomes, and biogeographical realms. Ecol. Inform. 4, 93–98.

He, K., Zhang, J., Zhang, Q., 2009. Linking variability in species composition and MODIS NDVI based on beta diversity measurements. Acta Oecol. 35, 14–21.

Kark, S., Levin, N., Phinn, S., 2008. Global environmental priorities: making sense of remote sensing: reply to TREE Letter: satellites miss environmental priorities by Loarie et al. 2007. Trends Ecol. Evol. 23, 181–182.

Kerr, J.T., Ostrovsky, M., 2003. From space to species: ecological applications for remote sensing. Trends Ecol. Evol. 18, 299–305.

Knapp, S., Lughadha, E.N., Paton, A., 2005. Taxonomic inflation, species concepts and global species lists. Trends Ecol. Evol. 20, 7–8.

Koenker, R., 2007. Quantreg: quantile regression. R package version 4.10. http://www.r-project.org.

Koenker, R., Bassett Jr., G., 1978. Regression quantiles. Econometrica 46, 33–50.

Koenker, R., Hallock, H., 2001. Quantile regression. J. Econ. Perspect. 15, 143–156.

Koleff, P., Gaston, K.J., Lennon, J.J., 2003. Measuring beta diversity for presence-absence data. J. Anim. Ecol. 72, 367–382.

Kumar, S., Stohlgren, T.J., Chong, G.W., 2006. Spatial heterogeneity influences native and nonnative plant species richness. Ecology 87, 3186–3199.

Legendre, P., Legendre, L., 1998. Numerical Ecology. Elsevier Science BV.

Legendre, P., Borcard, D., Peres-Neto, P.R., 2005. Analyzing beta diversity: partitioning the spatial variation of community composition data. Ecol. Monogr. 75, 435–450.

Leisch, F., 2007. Bootstrap: functions for the book "An Introduction to the Bootstrap". R package version 1.0-21.

Levin, N., Shmida, A., Levanoni, O., Tamari, H., Kark, S., 2007. Predicting mountain plant richness and rarity from space using satellite-derived vegetation indices. Divers. Distrib. 13, 692–703.

Lillesand, T.M., Kiefer, R.W., Chipman, J.W., 2004. Remote Sensing and Image Interpretation, 5th edn. John Wiley & Sons, New Jork, USA.

Loarie, S.R., Joppa, L.N., Pimm, S.L., 2007. Satellites miss environmental priorities. Trends Ecol. Evol. 22, 630–632.

Loveland, T.R., Cochrane, M.A., Henebry, G.M., 2008. Landsat still contributing to environmental research. Trends Ecol. Evol. 23, 182–183.

Nagendra, H., 2001. Using remote sensing to assess biodiversity. Int. J. Remote Sens. 22, 2377–2400.

Nagendra, H., Rocchini, D., 2008. High resolution satellite imagery for tropical biodiversity studies: the devil is in the detail. Biodivers. Conserv. 17, 3431–3442.

Nekola, J.C., White, P.S., 1999. The distance decay of similarity in biogeography and ecology. J. Biogeogr. 26, 867–878.

Oindo, B.O., 2008. Understanding the relationship between environmental energy availability and bird species richness in Kenya using remote sensing and ancillary data. Open Remote Sens. J. 1, 1–6.

Oindo, B.O., Skidmore, A.K., 2002. Interannual variability of NDVI and species richness in Kenya. Int. J. Remote Sens. 23, 285–298.

Palmer, M.W., Earls, P., Hoagland, B.W., White, P.S., Wohlgemuth, T., 2002. Quantitative tools for perfecting species lists. Environmetrics 13, 121–137.

R Development Core Team, 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rocchini, D., 2007a. Distance decay in spectral space in analyzing ecosystem β-diversity. Int. J. Remote Sens. 28, 2635–2644.

Rocchini, D., 2007b. Effects of spatial and spectral resolution in estimating ecosystem α-diversity by satellite imagery. Remote Sens. Environ. 111, 423–434.

Rocchini, D., Cade, B.S., 2008. Quantile regression applied to spectral distance decay. IEEE Geosci. Remote Sens. Lett. 5, 640–643.

Rocchini, D., Vannini, A., in press. What is up? Testing spectral heterogeneity vs. NDVI relationship by quantile regression. Int. J. Remote Sens.

Rocchini, D., Andreini Butini, S., Chiarucci, A., 2005. Maximizing plant species inventory efficiency by means of remotely sensed spectral distances. Glob. Ecol. Biogeogr. 14, 431–437.

Schmidtlein, S., Sassin, J., 2004. Mapping of continuous floristic gradients in grasslands using hyperspectral imagery. Remote Sens. Environ. 92, 126–138.

Sørensen, T.A., 1948. A method of establishing groups of equal amplitude in plant sociology based on similarity of species content, and its application to analyses of the vegetation on Danish commons. K. Dan. Vidensk. Selsk. Biol. Skr. 5, 1–34.

Tucker, C.J., Grant, D.M., Dykstra, J.D., 2004. NASA's global orthorectified Landsat data set. Photogramm. Eng. Remote Sensing 70, 313–322.

Tuomisto, H., Poulsen, A.D., Ruokolainen, K., Moran, R.C., Quintana, C., Celi, J., Cañas, G., 2003. Linking floristic patterns with soil heterogeneity and satellite imagery in Ecuadorian Amazonia. Ecol. Appl. 13, 352–371.

Waser, L.T., Stofer, S., Schwarz, M., Küchler, M., Ivits, E., Scheidegger, C.H., 2004. Prediction of biodiversity: regression of lichen species richness on remote sensing data. Community Ecol. 5, 121–134.

Whittaker, R.H., 1960. Vegetation of the Siskiyou Mountains, Oregon and California. Ecol. Monogr. 26, 1–80.