

# Rarefaction theory applied to satellite imagery for relating spectral and species diversity

Duccio Rocchini<sup>1,2</sup>, Thomas Wohlgemuth<sup>3</sup>, Carlo Ricotta<sup>4</sup>, Sara Ghisleni<sup>1</sup>, Andrea Stefanini<sup>1</sup> and Alessandro Chiarucci<sup>1,2</sup>

<sup>1</sup>Dipartimento di Scienze Ambientali "G. Sarfatti", Università di Siena, via P.A. Mattioli 4, 53100, Siena, Italy. E-mail: rocchini@unisi.it

<sup>2</sup>Terradata srl environmetrics, Dipartimento di Scienze Ambientali "G. Sarfatti", Università di Siena, via P.A. Mattioli 4, 53100, Siena, Italy

<sup>3</sup> WSL Swiss Federal Institute for Forest, Snow and Landscape Research, Zürcherstrasse 111, CH-8903 Birmensdorf, Switzerland

<sup>4</sup> Dipartimento di Biologia Vegetale, Università "La Sapienza" di Roma, piazzale A. Moro 5, 00185 Roma, Italy

## Abstract

Species rarefaction curves have long been used for estimating the expected number of species as a function of sampling effort. Nonetheless, sampling species based on standard plant inventories represents a cost expensive approach. In this view, remotely sensed information may be straightforwardly used for predicting species rich sites on the strength of the *Spectral Variation Hypothesis* which predicts that sites with a higher spectral (environmental) variability will show a higher species diversity.

In this paper we present spectral rarefaction, i.e. the rarefaction of reflectance values derived from satellite imagery, as an effective tool for predicting bio-diverse sites.

**Keywords:** biodiversity, ecological gradients, Landsat ETM+, satellite imagery, species rarefaction, species richness, species turnover, spectral rarefaction, Spectral Variation Hypothesis.

## *Tecniche di rarefazione applicate ad immagini satellitari: possibili relazioni tra diversità spettrale e specifica*

### **Riassunto**

*Le curve di rarefazione delle specie sono state a lungo utilizzate per stimare il numero di specie atteso in funzione dello sforzo di campionamento. Tuttavia, il campionamento di specie a terra rappresenta un approccio poco efficiente in termini di tempo e costi. L'informazione remota potrebbe essere usata per individuare i siti con elevata ricchezza di specie sulla base della Spectral Variation Hypothesis che prevede che siti con maggior diversità spettrale (ambientali) presentino una maggiore diversità in termini di specie. In questo articolo viene proposta la rarefazione spettrale, intesa come rarefazione dei valori di riflettanza derivata da immagini satellitari, come metodo predittivo di siti a più alta biodiversità.*

**Parole chiave:** biodiversità, Landsat ETM+, gradienti ecologici, immagini satellitari, arefazione, ricchezza specifica, Spectral Variation Hypotesis, turnover specifico.

## Introduction

Plant species diversity assessment is an important task for ecologists [see e.g. Palmer, 1995; Palmer et al., 2002]. As stated by Palmer et al. [2002], accurately inventorying species over a large region is complicated by the fact that the botanists cannot inspect every individual plant in the region and that species composition changes through time [Robinson et al., 1994; Kirby and Thomas, 2000; McCollin et al., 2000]. Different methods have been proposed to locate those environmental gradients offering the maximum change in species richness [e.g. Gillison and Brewer, 1985; Rocchini et al., 2005]. Noteworthy, subjective sampling is likely to outperform any objective sampling in terms of maximising plant species inventories [Palmer et al., 2002]. However, replicable methods for inventorying species are strongly encouraged for improving statistical estimates of species richness [Chiarucci et al., 2001, 2003; D'Alessandro and Fattorini, 2002; Baffetta et al., 2007], large scale evaluation and comparison [Koellner et al., 2004; Chiarucci and Bonini, 2005], and multitemporal monitoring [Ferretti and Chiarucci, 2003; Kalkhan et al., 2007].

Hence, measuring heterogeneity in satellite imagery is crucial since it is expected that the higher the spectral heterogeneity of an area (i.e. the higher the number of habitats) the higher will be the diversity in terms of species living therein [*Spectral Variation Hypothesis*, see Palmer et al., 2002; Rocchini, 2007b; Gillespie et al., 2008]. The availability of data with improved both spatial and spectral resolutions represents an opportunity for the study of the relationship between spectral and ecological heterogeneity [Innes and Koch, 1998]. In this view, spectral heterogeneity has been previously demonstrated to have a sort of predictive power with respect to species richness within a given site, at different spatial scales [Gould, 2000; Oindo and Skidmore, 2002; Rocchini et al., 2004]. Further, Rocchini et al. [2005] demonstrated that species complementarity among sites ( $\beta$ -diversity), i.e. ecological gradients, can also be maximised by spectral variability. We refer to the reviews by Nagendra and Gadgil [1999], Nagendra [2001], Foody [2008], Gillespie et al. [2008], Nagendra and Rocchini [2008] for additional information.

Despite the attribute being considered, the diversity of that attribute has been proven to change as a function of scale [Whittaker, 1972]. Most measures of spectral diversity have been proposed based on the Boltzmann index [Boltzmann, 1872; Gorelick, 2006], commonly referred to as Shannon entropy index [Shannon, 1948 but see even Ricotta, 2005 and Kumar et al., 2006 for its application in remote sensing]  $H = -\sum p \times \ln(p)$ , where  $p$  is the relative abundance of each spectral reflectance value (*Digital Number*, DN). The *Shannon index* will increase if the DN values are equally distributed with no DN value being dominant with respect to the others. The *Shannon index* has been advocated as a powerful algorithm for measuring diversity. Nonetheless, it does not explicitly consider how the measure of diversity changes as a function of scale if it is applied to an entire image. It may be made to account for the variation of diversity across spatial scales if it is repeatedly calculated while increasing the sampling extent within the chosen study area. This process may be time expensive. Quoting Gorelick [Gorelick, 2006], who made a critique on diversity measured by Shannon and Simpson indices, one can never capture all aspects of diversity in a single statistic. This is true regardless of the attribute being considered.

The aim of this paper is to propose a method for measuring spectral heterogeneity at multiple scales simultaneously based on ecological theory and to show the strict relation between spectral and species based measures of diversity.

## Algorithmic foundation of spectral rarefaction

In ecology, there is a long history of studies dealing with species diversity over space or time. In particular, given  $N$  plots, i.e. sampling units with a certain dimension, three different kinds of species diversity may be recognized:

- alpha or local diversity ( $\alpha$ ), i.e. the number of species within one plot
- gamma or total diversity ( $\gamma$ ), i.e. the number of species considering  $N$  plots
- beta or between-plots diversity ( $\beta$ ), i.e. the diversity deriving from the complementarity of the species composition considering pairs of plots [Whittaker, 1972].

In this view, accumulation curves, showing the number of accumulated species given a certain number of sampled plots, have long been used for estimating the expected number of species within a study area given a specific sampling effort. Since the order that samples are added to an accumulation curve accounts for its shape [Ugland et al., 2003; Rocchini et al., 2005], an order-free curve is derived by means of (i) an analytical solution or of (ii) permutations of samples [Ricotta et al., 2002; Fattorini, 2007; Chiarucci et al., 2008]. This order-free curve is referred to as a rarefaction curve. Considering permutation (ii), once  $N$  plots have been visited across a study area and the presence of all species has been recorded (obtaining a presence/absence matrix  $M_s$  of  $N$  plots per  $S$  species), a rarefaction curve is then obtained by repeatedly resampling the pool of  $N$  plots at random without replacement and plotting the average number of species represented by 1, 2, ...,  $N$  plots [Gotelli and Colwell, 2001; Fattorini, 2007]. Thus, sample-based rarefaction generates the expected number of accumulated species as the number of sampled plots increases from 1 to  $N$ .

On the other hand, an analytical solution (i) may be formalized as:

Let  $M_s$  be a presence/absence matrix of  $N$  plots per  $S$  species, the formal estimate of the expected number of species per number of plots turns out to be:

$$E[S] = S - \frac{\sum_{i=1}^S \binom{N - N_i}{n}}{\binom{N}{n}} \quad [1]$$

where  $N_i$  = number of plots where species  $i$  is found and  $n$  = number of randomly chosen plots [see Shinozaki, 1963; Kobayashi, 1974; Koellner et al., 2004; Chiarucci et al., 2008].

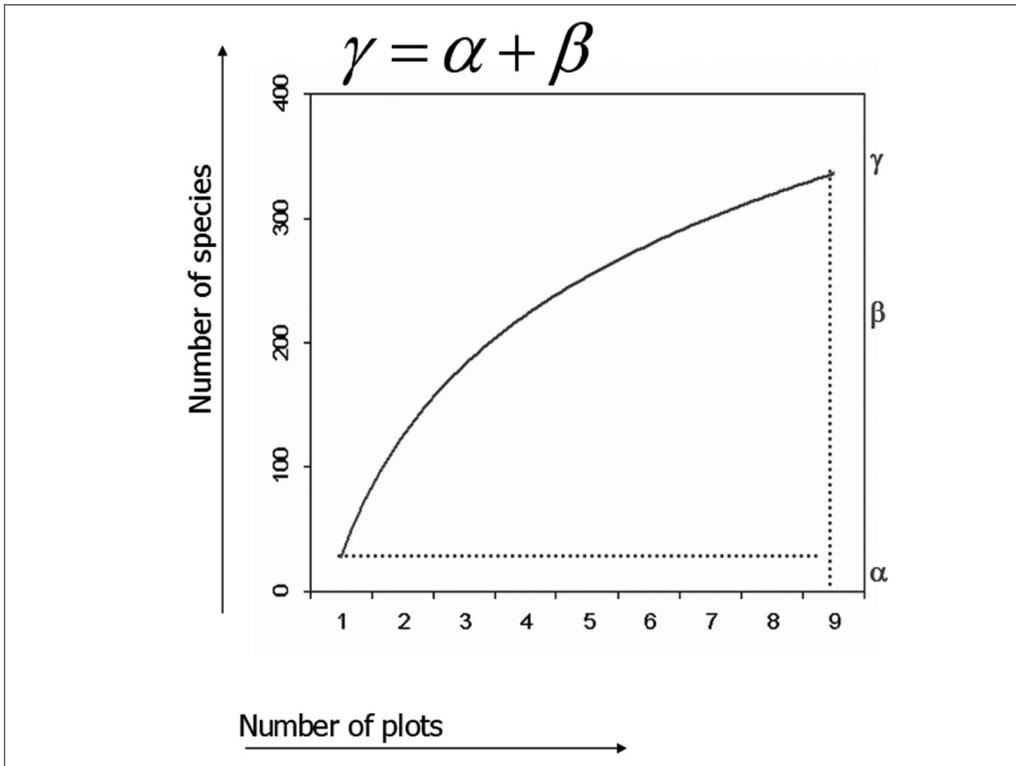
Generally, the steeper the curve, the greater the increase in species richness as the sample size increases [Kobayashi, 1974; Rocchini et al., 2005].

From a landscape perspective, rarefaction curves are directly related to the environmental heterogeneity of the area sampled. In fact, it is expected that the greater the landscape heterogeneity, the greater the species diversity, including both fine-scale and coarse-scale species richness (i.e.  $\alpha$ - and  $\gamma$ -diversity, respectively), and compositional variability, or  $\beta$ -diversity [Rocchini et al., 2005].

Computing  $\beta$ -diversity deals with looking at the difference between pairs of plots in terms of species composition [Chao et al., 2005; Bacaro et al., 2007]. Popular indices of  $\beta$ -diversity, e.g. the *Jaccard index*, are based on the intersection of the composition in species between pairs of plots with respect to their union, as  $C_j = 1 - (n / \cup)$  [Chao et al., 2005]. The higher the intersection in species composition the lower the  $\beta$ -diversity. As an example,

given two plots containing five species each ( $\alpha = 5$ ), using  $C_p$ ,  $\beta$ -diversity will range from  $\beta=0$  when the 5 species will be exactly the same, while the maximum  $\beta$ -diversity ( $\beta=1$ ) will occur when all the 5 species will be different.

An alternative definition of  $\beta$ -diversity has been provided by Whittaker [Whittaker, 1972] who expressed it as  $\beta = \gamma/\bar{\alpha}$ . This was later modified by Lande [Lande, 1996] being more consistent with the rarefaction theory. Using rarefaction curves, diversity may be partitioned by additive partitioning as  $\gamma = \alpha + \beta$  [Lande, 1996; Crist and Veech, 2006], leading to considering  $\beta$  in the same unit of measurement (i.e. number of species) of  $\alpha$  and  $\gamma$  as:  $\beta = \gamma - \alpha$  (Fig. 1).



**Figure 1 – Additive partitioning of diversity.  $\gamma$ -diversity is represented by the sum between  $\alpha$  and  $\beta$ . This leads to consider  $\beta$  in the same unit of measurement (i.e. number of species) of  $\alpha$  and  $\gamma$ .**

In this paper, different “species” will be replaced by different “DNs” (*Digital Numbers*, i.e. spectral values).

Consider a satellite image with a radiometric resolution of 8 bit. This means that the reflectance values of the pixels, i.e. the *Digital Numbers* (DNs), may range from 0 to 255. Subsampling the image by means of  $N$  plots, i.e. spatial windows with a certain dimension, will lead to a presence/absence matrix  $M_{DN}$  of  $N$  plots per  $S$  DNs.

Given the matrix  $M_{DN}$ , [1] previously introduced for species diversity can also provide a formal estimate of the number of DN per number of windows when  $N_i$  = number of plots

where the DN value  $i$  is found.

Therefore, the same concepts introduced for species diversity may thus be applied to satellite imagery diversity. Applying rarefaction theory to DNs rather than species leads to consider three different components of pixels diversity:

- alpha or local diversity ( $\alpha_{DN}$ ), i.e. the number of different DNs within one plot
- gamma or total diversity ( $\gamma_{DN}$ ), i.e. the number of different DNs considering  $N$  plots
- beta or between-plots diversity ( $\beta_{DN}$ ), i.e. the diversity deriving from  $\beta_{DN} = \gamma_{DN} - \alpha_{DN}$  [Lande, 1996; Crist and Veech, 2006].

### Theoretical example

Equation [1] only works with one-dimensional systems. In fact the dimension  $\text{Dim}(M_{DN})$  of the presence/absence matrix  $M_{DN}$  of  $N$  plots per  $S$  DN values equals  $\text{Dim}(M_{DN})=(N,S)$ , implying that:

- the plots are rows
- the DN values are columns
- the cells composing the matrix are presence/absence values, i.e. they are dummy coded as 1s and 0s.

For instance, let  $M_{DN}$  be the presence/absence matrix of  $N$  plots per  $S$  DN values derived from an 8-bit image sampled by 6 plots. In this case,  $\text{Dim}(M_{DN})=(N,S)=(6,256)$ , with DN values in one dimension ranging from 0 to 255.

Thus, before building rarefaction curves one should choose a single band to work with. A straightforward option may be based on performing data reduction with a method such as PCA and further using the first principal component explaining most of the variance [Rocchini, 2007b].

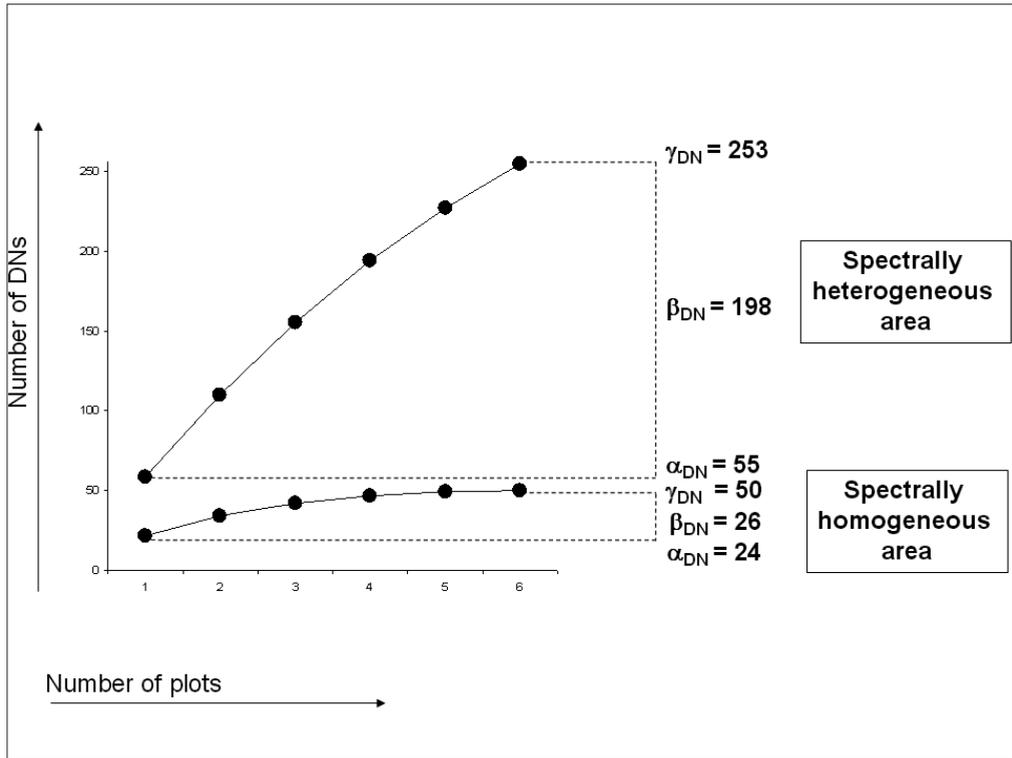
Once the rarefaction algorithm [1] has been applied to the presence absence matrix  $M_{DN}$ , different study areas sampled by the same number of plots containing the same number of inner pixels (e.g. 1000 pixels per spatial window) will possibly show very different curves. As an example, Figure 2 shows the spectrally based rarefaction curves of two areas with different levels of heterogeneity, each sampled by six spatial windows (plots).

Considering the ecologically heterogeneous area (upper curve of Fig. 2) with respect to the more homogeneous one (lower curve of Fig. 2),  $\alpha_{DN}$  equals 55 and 24 respectively, i.e. there are on average 55 and 24 distinct reflectance values for each plot (spatial window).

Meanwhile  $\gamma$  turns out to be 253 and 50, for heterogeneous vs. homogeneous area, respectively.

This means that the spectral value diversity  $\beta_{DN}$  as calculated by  $\gamma_{DN}-\alpha_{DN}$  is 198 and 26, respectively.

Notice that in this theoretical example, the rarefaction algorithm [1] allowed us to: (i) represent the variation in diversity by means of a single algorithm applied to a matrix, i.e. the spatial variation of diversity components, (ii) represent  $\beta$ -diversity by means of only one statistic without considering pair-wise distances among spatial windows (as in the case of e.g. the *Jaccard index*), and (iii) represent  $\beta$ -diversity in the same units as  $\alpha$ - and  $\gamma$ -diversity (i.e. number of different DN values).



**Figure 2** – A theoretical example of spectral rarefaction. Once differently heterogeneous areas have been sampled by the same number of plots (windows) containing the same number of inner pixels, the rarefaction curves computed by [1] provide an estimate of the number of different DNs at various spatial scales. In this case the two areas have been sampled by 6 plots each.

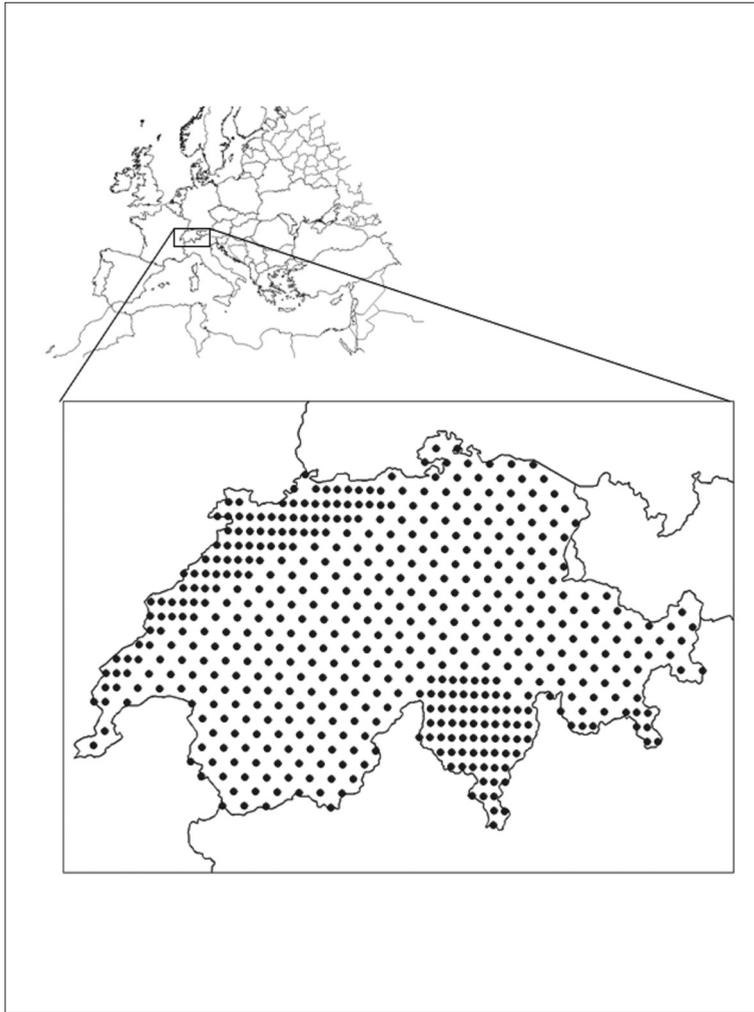
## Worked example

### Study area

The following worked example will make use of empirical data gathered within the whole Switzerland, which covers 41244 km<sup>2</sup> in central Europe and ranges in altitude from 193 to 4634 m a.s.l. (45°49'–47°48' N latitude, 5°57'–10°30' E longitude, Fig. 3). The average elevation is 1300 m a.s.l. Mountain landscapes predominate, with 60% of the country being formed by the Alps and 10% by the Jura Mountains. About 7% of the country consists of urban settlements including buildings, associated green areas, and road and rail networks [BFS, 1992/1997]. Switzerland can be easily subdivided into five biogeographical regions: Northern, Central and Southern Alps, the Jura Mountains and the Central Plateau [see Koellner et al. 2004].

### Data sampling

The Swiss 'Biodiversity Monitoring' programme (BDM) uses indicators of pressure, state and response. In total, 11 state indicators (*Z* for German 'Zustand') are repeatedly assessed. The *Z7* indicator observes the diversity of vascular plants on a landscape scale using a systematic sample of 520 1 km x 1 km quadrats (Fig. 3).



**Figure 3 – Study area (Switzerland) and the BDM systematic sampling design based on 520 quadrats of 1x1 km. Circles represent the centres of the quadrats.**

Spacing between neighbour quadrats was 19.1 km in most regions and 14.3 km in the Southern Alps and the Jura Mountains [Hintermann et al., 2000]. The BDM aims, among other objectives, to survey landscape biodiversity – from lowland to alpine zones – over a long period. In the quadrats, data were collected along 2500 m long transect routes, in buffers of 2.5 m on both sides of the transect [Plattner et al., 2004]. All sample quadrats had been visited for a first assessment by the end of 2005.

#### ***Satellite data***

Seven ortho-Landsat ETM+ images (spatial resolution: 30m, spectral range: 0.45-2.35  $\mu\text{m}$  considering bands 1-5 and 7, radiometric resolution: 8-bit) taken in the late spring

and summer period (i.e. in a similar phenological period, Tab. 1) and covering the whole Switzerland (spanning a period from 1999 to 2001) were acquired. Theoretically, a large difference between the time of satellite image acquisition and field survey could affect the relations between spectral values and species composition [Rocchini, 2007a]. However, in this case, the temporal gap between images and field data should not threaten the results, because of the large grain of the quadrats and the longer time expected for vegetation dynamics. Moreover, considering the extent of the study area, general patterns were tested in this paper, thus discarding isolated disturbance situations.

Thermal Infrared (band 6) having a spatial resolution of 60 m was not used in this paper. To reduce atmospheric effects a dark-object subtraction was applied to each image [Chavez, 1988; 1996].

**Table 1 – Information related to each Landsat ETM+ image used in this study.**

Path	Row	Acquisition date
193	027	September 13th 1999
194	027	August 24th 2001
194	028	June 21st 2001
195	027	September 11th 1999
195	028	July 30th 2001
196	027	May 15th 2000
196	028	July 21st 2001

### **Rarefaction**

Among the 520 quadrats of the previously described assessment of vascular plants in the frame of the BDM, the first available 361 quadrats were considered. Of these, 347 quadrats free from spectral noise (e.g. clouds, shadows, etc.) were retained for further analysis. The number of units per biogeographical region is reported in Table 2.

**Table 2 – Number of sampling units per biogeographical region.**

Biogeographical region	Number of sampling units
Northern Alps	83
Central Alps	73
Southern Alps	55
Jura Mountains	52
Plateau	84

Species rarefaction curves for each biogeographical region were achieved by using the R software [vegan package, Oksanen et al., 2007] according to [1]:

Similarly to species rarefaction curves, the number of accumulated spectral DNs (*Digital numbers*) for a given sample size (number of plots) was expressed by applying the rarefaction formula to the spectral DNs, rather than species counts. Since, as for species rarefaction, spectral rarefaction is based on one-dimensional values, an unstandardised PCA was applied to extract the one-dimensional data set mostly related to the original Landsat ETM+ bands (see even section “*Theoretical example*”). Accordingly, the first PCA axis (PC1), explaining 71% of the variance of the whole multispectral dataset, was retained for further analysis. Noteworthy, a PCA axis contains continuous values, and these cannot be used as classes for rarefaction purposes. For this reason PC1 was converted into a 8-bit

band splitting values into 256 equal intervals by RmcdR R-package [Fox et al., 2007]. This range was arbitrarily chosen on the strength of the input radiometric resolution (8 bit = 256 values), but other ranges could even be adopted [see Le Hégarat-Masclé et al., 1997]. This choice does not impact the analyses, since the interest is focused on relative differences among biogeographical regions. Consequently, for each quadrat, formed by 1089 pixels on average, the number of different DN values could theoretically range from 1 (homogeneous environment such as water) to 256 (heterogeneous environment composed by different land cover classes). Notice that a maximum number of DN values lower than 256 per quadrat is expected on the strength of the spatial autocorrelation of spectral values. Once spectral rarefaction curves are built, the number of DN values per quadrat can be directly calculated ( $\alpha_{DN}$ ) and rises until theoretically reaching a value of 256 while new quadrats are added to the curve. Obviously the theoretical maximum of 256 different values is reached only when the considered biogeographical region is so heterogeneous that it comprises all the 256 values obtained by the first PCA axis.

Species and spectral rarefaction curves per biogeographical region were compared considering pairs of accumulated species and DN values per region. The rarefaction curves of the five biogeographical regions differed in length because of the different size per region. Thus we chose the range of 1 to 52 quadrats, which equals the number of quadrats in the smallest region (see Tab. 2). In fact comparisons among areas need to consider the same minimum "sampling effort" (hereafter referred to as  $S_m$ , in this case equalling to 52, Tab. 2).

## Results

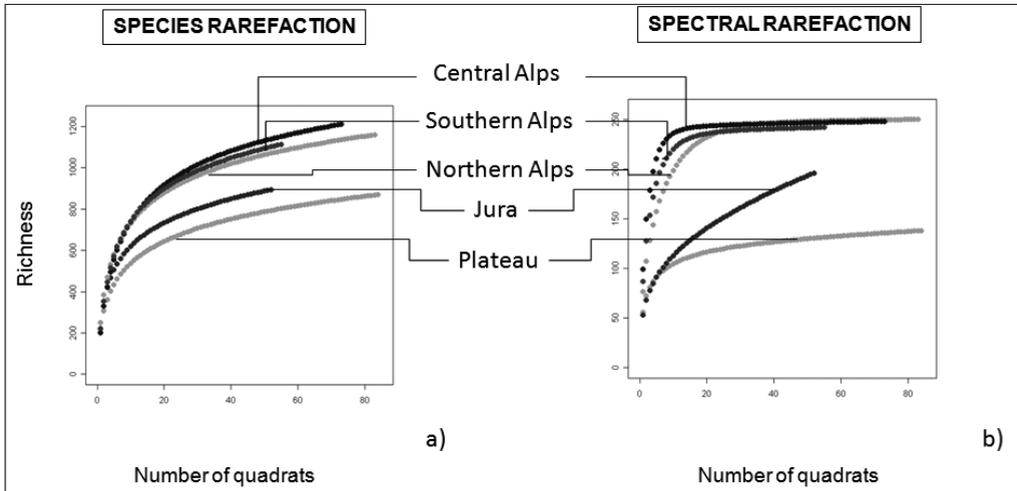
Species rarefaction curves derived from quadrats in the Alps (Northern, Central and Southern Alps) rose up more than Jura and Plateau, indicating a higher species turnover ( $\beta_{Sm}$ ) and global species richness ( $\gamma_{Sm}$ ) starting from similar values of local species richness ( $\alpha$ -diversity, Fig. 4, Tab. 3).

Considering that vascular plant species richness in the quadrats referred to the number of species recorded along transects, this does not represent the whole species richness of a given quadrat, but only a list of particularly frequent species [Stohlgren et al., 1997a]. Transects are not ideal for capturing species richness because patchy species and rare habitats are often missed [Stohlgren, 2007]. This may result in an underestimate of the real  $\alpha$ -diversity within each quadrat [Palmer, 1995] and in a consequent levelling of the differences among quadrats. We refer to Stohlgren et al. [1997a, 2000], Palmer et al. [2002], Baffetta et al. [2007] describing other sampling designs addressing the problem.

Concerning spectral rarefaction the same pattern was found, i.e. quadrats in the Alps were more diverse rather than quadrats in the Jura and Plateau areas (Fig. 4, Tab. 3). Contrary to species richness pattern, local spectral heterogeneity ( $\alpha_{DN}$ ) showed a distinctive pattern with Alps being more spectrally diverse even at local scale with respect to Jura and Plateau, reaching up to the double of spectral variability. Further, at regional scale, Alps showed a very high spectral turnover ( $\beta_{DN_Sm}$ ) which led to an asymptote in the gamma diversity ( $\gamma_{DN_Sm}$ ) reached by sampling a low number of (ca. 20) quadrats (Fig. 4, Tab. 3).

From an ecological perspective, mountainous regions in Switzerland are richer in plant species than low-elevation regions [Wohlgemuth, 1993]. In general, habitat heterogeneity at the landscape scale including species pools for lowland and mountain plants is the main reason for this pattern [Wohlgemuth, 1998]. The Alps and Jura regions, ranging from ca.

200 to 4600 m a.s.l. and from 500 to 1600 m a.s.l. respectively, are richer than the lower elevated Plateau because, in the latter region, mountain species (reflecting a different habitat type) are absent [Wohlgemuth et al., 2008].



**Figure 4 – Species (a) and spectral (b) rarefaction for each biogeographical region: Central Alps, Southern Alps, Northern Alps, Jura, Plateau. Notice that the ordering of biogeographical areas is roughly the same in both graphs, indicating a correspondence between species and spectral diversity. We refer to the main text for major explanations.**

**Table 3 – Diversity measures for each biogeographical region considering both species and spectral diversity.**

Biogeographical region	Species $\alpha$	Spectral $\alpha_{DN}$	Species $\gamma_{Sm}$	Spectral $\gamma_{DN, Sm}$	Species $\beta_{Sm}$	Spectral $\beta_{DN, Sm}$
Northern Alps	249	76	1070	248	820	172
Central Alps	199	99	1139	247	940	148
Southern Alps	220	87	1101	242	881	155
Jura Mountains	248	53	893	195	645	143
Plateau	214	56	792	130	578	75

## Discussion

The results achieved in this study indicated that spectral rarefaction is a powerful tool for detecting regionally different biodiversity.

Spectral rarefaction brings both advantages and disadvantages in its very nature. Considering the latter, it should be applied with caution mainly because of scale problems in matching ground and satellite imagery. An inappropriate matching of satellite spatial resolution and the grain of field data could hide actual spatial heterogeneity with sub-pixel variability remaining undetected [Small, 2004; Rocchini, 2007b]. In this sense, the use of hyperspectral satellite images at coarse spatial resolution like MODIS, whose pixel size approximates that of the sampling quadrats adopted here, may hide the variability of an area simply because of the high amount of mixed pixels which hampers

to detect fine grained patterns [Fisher, 1997]. It is well known that a phenomenon could be undetected only because the scale of analysis is not appropriate for study of such phenomenon [Stohlgren et al., 1997a].

Moreover, one should consider multiple scales during the analysis process from field unit grain to the whole extent of the study area to detect both accordance and possible discrepancies between species richness and spectral variability [see Stohlgren et al., 1997a; Stohlgren et al., 1997b; Wu et al., 2000 about the importance of applying multiscale approaches to diversity estimate].

Advantages of the spectral rarefaction technique mainly arise from its implicit capability in estimating landscape (ecological) heterogeneity from the local to the regional scale in such a way that it may visualise scaling up properties. This should allow to individuate heterogeneous areas a-priori during the planning phase of species inventorying or monitoring programs. Considering species inventory issues, once landscape heterogeneous areas have been identified by spectral rarefaction, sampling designs weighted on landscape heterogeneity could be built to improve species inventory efficiency [Rocchini et al., 2005]. Obviously, the achieved results should be viewed as an help to plan field survey rather than a replacement of it, limiting remote information as a driver for field sampling design strategies.

In this paper, spectral values rather than classified images were directly used which maintains the input continuous information. In most studies using remotely sensed images, predictors of species variability were mainly based on landscape metrics derived from remote sensing classification [Stohlgren et al., 1997b; Schindler et al., 2008]. Of course, image classification allows to estimate not only landscape compositional variability, like in the present paper, but even structural variability over space by applying landscape structural metrics like shape indices or interspersion [see e.g. Kumar et al., 2006]. However, as stressed by Palmer et al. [2002] and Schwarz and Zimmermann [2005] processing remote sensing data may lead to a loss of information. In fact, as long as the used classes contain a high degree of reflectance mixture, end-members (i.e. pixels occupied solely by one cover type) do not accurately represent actual ecological patterns [see Townshend et al. 2004 on the matter]. This inevitably leads to the application of several techniques based on robust theoretical background for classifying images avoiding Boolean memberships, basically relying on mixture modelling [Small, 2005; Shanmugam et al., 2006; Nichol and Wong, 2007] or on fuzzy classification [Foody, 1996; Woodcock and Gopal, 2000; Tang et al., 2005; Okeke and Karnieli, 2006; Rocchini and Ricotta, 2007]. On the contrary, the results achieved in this paper promote satellite imagery as high potential continuous ancillary data for predicting species richness.

Of course, other techniques rather than spectral rarefaction could account for the spatial variability of DN values as well, e.g. semivariograms [Curran and Atkinson, 1998; Bacaro and Ricotta, 2007]. However, spectral rarefaction coupled with additive partitioning exhibits mathematical and statistical properties which may be directly related to spectral and species  $\alpha$ -,  $\beta$ - and  $\gamma$ - diversity. **This is an enormous advantage of spectral rarefaction** as a straightforward method for (i) robustly estimating local to global diversity of an area directly relating sensor-based and field-based heterogeneity and (ii) quantitatively comparing different areas with different degrees of heterogeneity at multiple scales.

## Acknowledgements

We are particularly grateful to Giovanni Bacaro for precious insights about the models used in the present paper. We also thank Andrea Billi, Marta Giordano and Alessia Nucci for their huge work on data acquirement and analysis.

We thank Matthias Plattner as the data manager of the BDM program. A number of botanists participated to the field campaign for gathering species data: Alain Jotterand, Andrea Persico, Barbara Berner, Christian Hadorn, Christoph Bühler, Christoph Käsermann, Daniel Knecht, David Galeuchet, Dunja Al-Jabaji, Elisabeth Danner, Gabriele Carraro, Jean-François Burri, Jens Paulsen, Karin Marti, Markus Bichsel, Martin Camenisch, Martin Frei, Martin Valencak, Matthias Vust, Michael Ryf, Nils Tonascia, Philippe Druart, Regula Langenauer, Sabine Joss, Stefan Birrer, Thomas Breunig, Urs Weber, Ursula Bollens, and Ursula Kradolfer.

## References

- Bacaro G., Ricotta C. (2007) - *A spatially explicit measure of beta diversity*. *Community Ecol.*, 8: 41-46.
- Bacaro G., Ricotta C., Mazzoleni S. (2007) - *Measuring beta diversity from taxonomic similarity*. *J. Veg. Sci.*, 18: 793-798.
- Baffetta F., Bacaro G., Fattorini L., Rocchini D., Chiarucci A. (2007) - *Multi-stage cluster sampling for estimating average species richness at different spatial grains*. *Community Ecol.*, 8: 119-127.
- BFS (1992/1997) - *Arealstatistik*. Bundesamt für Statistik, Servicestelle GEOSTAT, Neuchâtel, Switzerland.
- Boltzmann L. (1872) - *Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen*. *Wiener Berichte*, 66: 275-370.
- Chao A., Chazdon R.L., Colwell R.K., Shen T.-J. (2005) - *A new statistical approach for assessing similarity of species composition with incidence and abundance data*. *Ecol. Lett.*, 8: 148-159.
- Chavez P.S. Jr. (1988) - *An improved dark object subtraction technique for atmospheric scattering correction of multispectral data*. *Remote Sens. Environ.*, 24: 459-479.
- Chavez P.S. Jr. (1996) - *Image based calibration revisited and improved*. *Photogramm. Eng. Remote Sens.*, 62: 1025-1036.
- Chiarucci A., Maccherini S., De Dominicis V. (2001) - *Evaluation and monitoring of the flora in a nature reserve by estimation methods*. *Biol. Conserv.*, 101: 305-314.
- Chiarucci A., Enright N.J., Perry G.L.W., Miller B.P., Lamont B.B. (2003) - *Performance of nonparametric species richness estimators in a high diversity plant community*. *Divers. Distrib.* 9: 283-295.
- Chiarucci A., Bonini I. (2005) - *Quantitative floristics as a tool for the assessment of plant diversity in Tuscan forests*. *Forest Ecol. Manag.*, 212: 160-170.
- Chiarucci A., Bacaro G., Rocchini D., Fattorini L. (2008) - *Discovering and ... rediscovering the rarefaction formula in ecological literature*. *Community Ecol.*, 9: 121-123.
- Crist T.O., Veech J.A. (2006) - *Additive partitioning of rarefaction curves and species-area relationships: unifying alpha-, beta- and gamma-diversity with sample size and habitat area*. *Ecol. Lett.*, 9: 923-932.

- Curran P.J., Atkinson P.M. (1998) - *Geostatistics and remote sensing*. Prog Phys Geog, 22: 61-78.
- D'Alessandro L., Fattorini L. (2002) - *Resampling estimators of species richness from presence-absence data: why they don't work*. Metron 61: 5-19.
- Fattorini L. (2007) - *Statistical inference on accumulation curves for inventorying forest diversity: a design-based critical look*. Plant Biosyst, 14: 231-242.
- Ferretti M., Chiarucci A. (2003) - *Design concepts adopted in long-term forest monitoring programs in Europe: problems for the future?* Sci. Total Environ. 310: 171-178.
- Fisher P. (1997) - *The pixel: a snare and a delusion*. Int. J. Remote Sens., 18: 679-685.
- Foody G.M. (1996) - *Fuzzy modelling of vegetation from remotely sensed imagery*. Ecol. Model., 85: 3-12.
- Foody G.M. (2008) - *GIS: biodiversity applications*. Progr Phys Geogr, 32: 223-235.
- Fox J., Ash M., Boye T., Calza S., Chang A., Grosjean P., Heiberger R., Kerns G.J., Lancelot R., Lesnoff M., Messad S., Maechler M., Putler D., Ristic M., Wolf P. (2007) - *Rcmdr: R Commander*. R package version 1.2-9. <http://www.r-project.org>, <http://socserv.socsci.mcmaster.ca/jfox/Misc/Rcmdr/>.
- Gillespie T.W., Foody G.M., Rocchini D., Giorgi A.P., Saatchi S. (2008) - *Measuring and modeling biodiversity from space*. Progr Phys Geogr, 32: 203-221.
- Gillison A.N., Brewer K.R.W. (1985) - *The use of Gradient Directed Transects or Gradsects in Natural Resource surveys*. J. Environ. Manag. 20: 103-127.
- Gorelick R. (2006) - *Combining richness and abundance into a single diversity index using matrix analogues of Shannon's and Simpson's indices*. Ecography, 29: 525-530.
- Gotelli N.J., Colwell R.K. (2001) - *Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness*. Ecol. Lett., 4: 379-391.
- Gould W. (2000) - *Remote Sensing of vegetation, plant species richness, and regional biodiversity hot spots*. Ecol. Appl., 10: 1861-1870.
- Hintermann U., Weber D., Zangger A. (2000) - *Biodiversity monitoring in Switzerland*. Schriftenreihe für Landschaftspflege und Naturschutz, 62: 47-58.
- Innes J.L., Kock B. (1998) - *Forest biodiversity and its assessment by remote sensing*. Global Ecol. Biogeogr. Letters, 7: 397-419.
- Kalkhan M.A., Stafford E.J., Stohlgren T.J. (2007) - *Rapid plant diversity assessment using a pixel nested plot design: A case study in Beaver Meadows, Rocky Mountain National Park, Colorado, USA*. Divers. Distrib. 13: 379-388.
- Kirby K.J., Thomas R.C. (2000) - *Changes in the ground flora in Wytham Woods, southern England from 1974 to 1991 – implications for nature conservation*. J. Veg. Sci., 11: 871-880.
- Kobayashi S. (1974) - *The species-area relation I. A model for discrete sampling*. Res. Popul. Ecol. 15: 223-237.
- Koellner T., Hersperger A.M., Wohlgemuth T. (2004) - *Rarefaction method for assessing plant species diversity on a regional scale*. Ecography, 27: 532-544.
- Kumar S., Stohlgren T.J., Chong J.W. (2006) - *Spatial heterogeneity influences native and nonnative plant species richness*. Ecology, 87: 3186-3199.
- Lande R. (1996) - *Statistics and partitioning of species diversity, and similarity among multiple communities*. Oikos, 76: 5-13.

- Le Hégarat-Masclé S., Vidal-Madjar D., Taconet O., Zribi M. (1997) - *Application of Shannon information theory to a comparison between L- and C-band SIR-C polarimetric data versus incidence angle*. Remote Sens. Environ., 60: 121-130.
- McCollin D., Moore L., Sparks T. (2000) - *The flora of a cultural landscape: environmental determinants of change using archival sources*. Biol. Conserv., 92: 249-263.
- Nagendra H. (2001) - *Using remote sensing to assess biodiversity*, Int. J. Remote Sens., 22: 2377-2400.
- Nagendra H., Gadgil M. (1999) - *Satellite imagery as a tool for monitoring species diversity: an assessment*, J. Appl. Ecol., 36: 388-397.
- Nagendra H., Rocchini D. (2008) - *Satellite imagery applied to biodiversity study in the tropics: the devil is in the detail*. Biodiv. Conserv., 17: 3431-3442.
- Nichol J., Wong M.S. (2007) - *Remote sensing of urban vegetation life form by spectral mixture analysis of high-resolution IKONOS satellite images*. Int. J. Remote Sens., 28: 985-1000.
- Oindo B.O., Skidmore A.K. (2002) - *Interannual variability of NDVI and species richness in Kenya*. Int. J. Remote Sens., 23: 285-298.
- Okeke F., Karnieli A. (2006) - *Methods for fuzzy classification and accuracy assessment of historical aerial photographs for vegetation change analyses. Part I: Algorithm development*. Int. J. Remote Sens., 27: 153-176.
- Oksanen J., Kindt R., Legendre P., O'Hara R.B. (2007) - *vegan: Community Ecology Package version 1.8-6*. <http://cran.r-project.org/>.
- Palmer M.W. (1995) - *How should one count species?* Nat. Area J., 15: 124-135.
- Palmer M.W., Earls P., Hoagland B.W., White P.S., Wohlgemuth T., (2002) - *Quantitative tools for perfecting species lists*. Environmetrics, 13: 121-137.
- Plattner M., Birrer S., Weber D. (2004) - *Data quality in monitoring plant species richness in Switzerland*. Community Ecol., 5: 135-143.
- Ricotta C. (2005) - *On possible measures for evaluating the degree of uncertainty of fuzzy thematic maps*. Int. J. Remote Sens., 26: 5573-5583.
- Ricotta C., Carranza M.L., Avena G.C. (2002) - *Computing  $\beta$ -diversity from species area curves*. Basic Appl. Ecol., 3: 15-18.
- Robinson G.R., Yurlina M.E., Handel S.N. (1994) - *A century of change in the Staten island flora: ecological correlates of species losses and invasions*. Bulletin Torrey Bot. Club, 121: 119-129.
- Rocchini D. (2007a) - *Distance decay in spectral space in analysing ecosystem  $\beta$ -diversity*. Int. J. Remote Sens., 28: 2635-2644.
- Rocchini D. (2007b) - *Effects of spatial and spectral resolution in estimating ecosystem  $\alpha$ -diversity by satellite imagery*. Remote Sens. Environ., 111: 423-434.
- Rocchini D., Chiarucci A., Loiselle S.A. (2004) - *Testing the spectral variation hypothesis by using satellite multispectral images*. Acta Oecol., 26: 117-120.
- Rocchini D., Andreini Butini S., Chiarucci A. (2005) - *Maximizing plant species inventory efficiency by means of remotely sensed spectral distances*. Global Ecol. Biogeogr., 14: 431-437.
- Rocchini D., Ricotta C. (2007) - *Are landscapes as crisp as we may think?* Ecol. Model., 204: 535-539.

- Schindler S., Poirazidis K., Wrbka T. (2008) - *Towards a core set of landscape metrics for biodiversity assessments: A case study from Dadia National Park, Greece*. *Ecol. Indic.*, 8: 502-514
- Schwarz M., Zimmermann N.E. (2005) - *A new GLM-based method for mapping tree cover continuous fields using regional MODIS reflectance data*. *Remote Sens. Environ.*, 95: 428-443.
- Shanmugam P., Ahn Y.H., Sanjeevi S. (2006) - *A comparison of the classification of wetland characteristics by linear spectral mixture modelling and traditional hard classifiers on multispectral remotely sensed imagery in southern India*. *Ecol. Model.*, 194: 379-394.
- Shannon, C. (1948) - *A mathematical theory of communication*. *Bell Systems Tech. J.*, 27: 379-423.
- Shinozaki K. (1963) - *Note on the species area curve*. *Proceedings of the 10th Annual Meeting of Ecological Society of Japan*, 5 (in Japanese).
- Small C. (2004) - *The Landsat ETM+ spectral mixing space*. *Remote Sens. Environ.*, 93: 1-17.
- Small C. (2005) - *A global analysis of urban reflectance*. *Int. J. Remote Sens.*, 26: 661-681.
- Stohlgren T.J. (2007) - *Measuring plant diversity: lessons from the field*. Oxford University Press, New York, USA.
- Stohlgren T.J., Chong G.W., Kalkhan M.A., Schell L.D. (1997a) - *Multiscale sampling of plant diversity: effects of minimum mapping unit size*. *Ecol. Appl.*, 7: 1064-1074.
- Stohlgren T.J., Coughenour M.B., Chong G.W., Binkley D., Kalkhan M.A., Schell L.D., Buckley D.J., Berry J.K. (1997b) - *Landscape analysis of plant diversity*. *Landscape Ecol.*, 12: 155-170.
- Stohlgren T.J., Owen A.J., Lee M. (2000) - *Monitoring shifts in plant diversity in response to climate change: a method for landscapes*. *Biodiv. Conserv.*, 9: 65-86.
- Tang X.M., Kainz W., Fang Y. (2005) - *Reasoning about changes of land covers with fuzzy settings*. *Int. J. Remote Sens.*, 26: 3025-3046.
- Townshend J.R.G., Huang C., Kalluri S.N.V., Defries R.S., Liang S., Yang K. (2004) - *Beware of per-pixel characterization of land cover*. *Int. J. Remote Sens.*, 21: 839-843.
- Ugland K.I., Gray J.S., Ellingen K.E. (2003) - *The species-accumulation curve and estimation of species richness*. *J. Anim. Ecol.*, 72: 888-897.
- Whittaker R. (1972) - *Evolution and measurement of species diversity*. *Taxon*, 21: 213-251.
- Wohlgemuth T. (1993) - *The distribution atlas of pteridophytes and phanerogams of Switzerland (Welten and Sutter 1982) in a relational database – species number per mapping unit and its dependence on various factors*. *Bot. Helv.*, 103: 55-71.
- Wohlgemuth T. (1998) - *Modelling floristic species richness on a regional scale: A case study in Switzerland*. *Biodivers. Conserv.*, 7: 159-177.
- Wohlgemuth T., Nobis M., Kienast F., Plattner M. (2008) - *Modelling vascular plant diversity at the landscape scale using systematic samples*. *J. Biogeogr.*, 35: 1226-1240.
- Woodcock C.E., Gopal S. (2000) - *Fuzzy set theory and thematic maps: accuracy assessment and area estimation*. *Int. J. Geogr. Inf. Sci.*, 14: 153-172.
- Wu J., Jelinski D.E., Luck M., Tueller P.T. (2000) - *Multiscale Analysis of Landscape Heterogeneity: Scale Variance and Pattern Metrics*. *Geogr. Inf. Sci.*, 6: 6-19.

**Received 26/03/2009, accepted 21/04/2009.**