From local spectral species to global spectral communities: A benchmark for ecosystem diversity estimate by remote sensing

Duccio Rocchini, Nicole Salvatori, Carl Beierkuhnlein, Alessandro Chiarucci, Florian de Boissieu, Michael W Förster, Carol Garzon-Lopez, Thomas Gillespie, Heidi Hauffe, Kate He, et al.

Duccio Rocchini, Nicole Salvatori, Carl Beierkuhnlein, Alessandro Chiarucci, Florian de Boissieu, et al.. From local spectral species to global spectral communities: A benchmark for ecosystem diversity estimate by remote sensing. Ecological Informatics, 2021, 61, pp.101195. 10.1016/j.ecoinf.2020.101195 . hal-03122931

HAL Id: hal-03122931
https://hal.inrae.fr/hal-03122931v1

Submitted on 3 Nov 2021

From local spectral species to global spectral communities: a benchmark for ecosystem diversity estimate by remote sensing

Duccio Rocchini[1,2]*, Nicole Salvatori[3,4], Carl Beierkuhnlein[5], Alessandro Chiarucci[1], Florian de Boissieu[6], Michael Förster[7], Carol X. Garzon-Lopez[8], Thomas W. Gillespie[9], Heidi C. Hauffe[10], Kate S. He[11], David Kleijn[12], Birgit Kleinschmit[7], Jonathan Lenoir[13], Marco Malavasi[2], Vítězslav Moudrý[2], Harini Nagendra[14], Davnah Payne[15], Petra Šímová[2], Michele Torresani[16], Martin Wegmann[17], Jean-Baptiste Féret[6]

October 3, 2020

[1] Alma Mater Studiorum University of Bologna, Department of Biological, Geological and Environmental Sciences, via Irnerio 42, 40126, Bologna, Italy

[2] Czech University of Life Sciences Prague, Faculty of Environmental Sciences, Department of Applied Geoinformatics and Spatial Planning, Kamýcka 129, Praha - Suchdol, 16500, Czech Republic

[3] University of Udine, Department of Agri-Food, Animal and Environmental Sciences (DI4A), via delle scienze 206, 33100 Udine (UD), Italy

[4] University of Trieste, Department of Life Sciences, via Giorgieri 5, 34100 Trieste (TS), Italy

[5] Biogeography, BayCEER, University of Bayreuth, Universitaetsstraße 30, 95440 Bayreuth, Germany

[6] UMR-TETIS, IRSTEA Montpellier, Maison de la Télédétection, 500 rue JF Breton, 34093, Montpellier Cedex 5, France

[7] Department of Geoinformation for Environmental Planning, Technical University of Berlin, Straße des 17. Juni 145, 10623 Berlin, Germany

[8] Ecology and Vegetation Physiology Group (EcoFiv), Universidad de los Andes, Cr. 1E No 18A, Bogotá, Colombia

[9] Department of Geography, University of California Los Angeles, Los Angeles, California 90095-1524, USA

[10] Fondazione Edmund Mach, Research and Innovation Centre, Department of Biodiversity and Molecular Ecology, Via E. Mach 1, 38010 S. Michele all'Adige (TN), Italy

[11] Department of Biological Sciences, Murray State University, Murray, KY 42071, USA

[12] Plant Ecology and Nature Conservation Group, Wageningen University, Droevendaalsesteeg 3a, 6708 PB Wageningen, The Netherlands

[13] UR "Ecologie et dynamique des systèmes anthropisées" (EDYSAN, UMR 7058 CNRS-UPJV), Université de Picardie Jules Verne, 1 Rue des Louvels, 80037 Amiens Cedex 1, France

[14] Azim Premji University, PES Institute of Technology Campus, Pixel Park, B Block, Electronics City, Hosur Road, Bangalore 560100, India

[15] University of Bern, Institute of Plant Sciences, GMBA office Altenbergrain 21, 3013 Bern, Switzerland

[16] Free University of Bolzano/Bozen, Faculty of Science and Technology, Piazza Universitá/Universitätsplatz 1, 39100, Bolzano, Bozen, Italy

[17] Department of Remote Sensing, Remote Sensing and Biodiversity Research Group, University of Wuerzburg, Wuerzburg, Germany

Abstract

In the light of unprecedented change in global biodiversity, real-time and accurate ecosystem and biodiversity assessment are becoming increasingly essential. Nevertheless, estimation of biodiversity using ecological field data can be difficult for several reasons. In particular, for larger extents, it is challenging to collect data to provide reliable information. Some of these restrictions in Earth observation can be avoided through the use of remote sensing approaches. Different studies have estimated biodiversity on the basis of the Spectral Variation Hypothesis (SVH). According to this hypothesis, spectral heterogeneity over the different pixel units of a spatial grid reflects a higher niche heterogeneity, allowing more organisms to coexist. Recently, the spectral species concept has been derived, following the consideration that the spectral heterogeneity at a landscape scale corresponds to a combination of subspaces sharing a similar spectral signature. At a local scale, with the use of high resolution remote sensing data, the different subspaces can be identified as different spectral entities, the so called "spectral species". Our approach extends this concept over wide spatial extents and to a higher level of biological organisation.

We applied this method to MODIS imagery data over Europe. Obviously, in our case, a spectral species from MODIS is not related to a single plant species in the field but rather to a species assemblage, habitat, or ecosystem. Based on such spectral information, we propose a straightforward method to derive α- (local relative abundance and richness of spectral species) and β-diversity (turnover of spectral species) maps at wide geographical extents.

Keywords: biodiversity; ecological informatics; modelling; remote sensing; satellite imagery.

1 Introduction

1.1 A quest for robust and reproducible α- and β-diversity measurement

The variability of life on Earth is heterogeneously distributed across the planet; ecologists and biogeographers have long been questioning about the potential causes of biodiversity distribution. Nowadays, the speed of changes and the uncertainty about consequences thereof is urging the whole scientific community worldwide. The perception of these processes translates into the need to use standardized methods for biodiversity assessment and monitoring in order to gain a better understanding and identify generalities.

A diatribe is still open about what are the most reliable metrics to assess biodiversity (see Jurasinski et al. (2009); Tuomisto (2010). Until now, no consistent definition exists and also the definition of the CBD (1992, `https://www.cbd.int/convention/text/`) is more confusing than clear: "Biological Diversity means the variability among living organisms from all sources, including, inter alia, terrestrial, marine and other aquatic ecosystems and the ecological complexes of which they are part; this includes diversity within species, between species and of ecosystems." Biodiversity obviously includes quantitative (number of species, alpha-diversity, gamma-diversity), qualitative (turnover, composition, beta-diversity) and functional (complexity, trophic levels, ecosystem services) aspects. To sum up our understanding on the term biodiversity (i.e. biological diversity) and to base our study on a more general and consistent concept, "biodiversity characterizes qualitative, quantitative and functional aspects of biotic units at various levels of organization in a concrete or abstract context, and at a given temporal or/and spatial scale" (Beierkuhnlein, 2003). In consequence, species richness and metrics that are based on it are very important, but they represent just one aspect of biodiversity. The total number of species co-occurring in a given community (α-diversity) is nested within the total number of a species pools occurring for instance at the landscape level (γ-diversity). But the reduction of biodiversity to the perspective of inventory and proportion would not cover spatial gradients in composition and species turnover (differentiation, β-diversity) (Jurasinski et al., 2009; Baselga, 2012) and also ignores functional diversity (e.g. functional traits), which is the main driver of ecosystem functioning.

In general, β-diversity is a crucial measure, since, given the same local richness of different sites (α-diversity), it directly considers the turnover among them. As an example, let A and B be two sampling sites with 10 different species each. If all 10 species are fully shared, the total γ-diversity would equal 10 species, while if all 10 species are completely different from one site to the other (high turnover, high β-diversity) the total diversity of the whole area based on the two focal sites would double.

It would then be interesting to understand how β-diversity originates, investigating how species composition differs among sites. In fact, species composition could be related to environmental conditions, or it could randomly fluctuate. A generally accepted hypothesis suggests that β-diversity might change as a function of species types living in a certain community. For instance, β-diversity should be small when communities are dominated by a limited number of competitive species; this is recognized as the null hypothesis and it entails a uniform distribution in species composition (Legendre et al., 2005).

The β-diversity concept is mainly reflecting the environmental heterogeneity between sites and thus within a given larger area that contains several of the focal study sites. Heterogeneity is in fact highly associated with a high degree of biological diversity since heterogeneous sites offer a diversity of ecological niches (sensu Elton (1958)) that can be occupied if the species pool offers the respective ecological diversity to address these niches (Gaston, 2000; Rocchini et al., 2010). Furthermore, since β-diversity can be described as the spatial turnover among sites within a given region, it captures a fundamental feature of the spatial pattern of biodiversity.

In some cases, spatial turnover can result from local extinction processes that hit certain species more than others and enhance the dissimilarity between sites without dispersal (Steinitz et al., 2006). This is the case in highly fragmented landscapes where dispersal is limited (Hobbs et al., 2006). Even stochastic processes (sensu Moran (1950) and Clark (2008)) may enhance β-diversity in previously homogeneous ecosystems. For instance, sudden fragmentation (Alados et al., 2009) can lead to disfunctional source-sink metapopulations with intrinsic influences on the degree of spatial (and genetic) connectivity of organisms (Waples and Gaggiotti, 2006), resulting in the local loss of sink populations. However, in most situations, the spatial turnover and therefore the dispersal of species between sites (metapopulation and metacommunity dynamics) is linked to the distance among sites. Strictly speaking, the similarity between two sites decays with increasing distance between them (Rocchini, 2007), a process also known as the distance decay in similarity or the Tobler's first law of geography (Tobler, 1970).

Hence, modelling the distribution of β-diversity in space is based on softening the role of individual species, which are not even completely described at wide geographical scales, for the benefit of having a more abstract proxy for ecosystem patterns and processes. Such process can lead to efficiently monitor and further preserve entire ecosystems, as it has been established by the Aichi Biodiversity Targets (`https://www.cbd.int/sp/`).

In general, an enormous effort from ecologists is necessary in order to collect reliable data on biodiversity. A pioneering example is the public database of the Global Biodiversity Information Facility (GBIF, `https://www.gbif.org/`). GBIF is a network funded by the world's governments which contains almost 41.000 databases of species occurrences spread out across the world. The huge amount of data accessible and the different techniques now available to analyze them should facilitate biodiversity assessment and monitoring. Unfortunately, although it would be possible, in principle, to use these data to make reasonable assumptions on the biodiversity at wide spatial extents, there are several limitations due to their quality (Maldonado et al., 2015). The errors that usually rise are due to: (i) poor quality in getting exact geographic coordinates of the sampling sites; (ii) wrong taxonomic identification with poor quality control; and (iii) difficulties in proving a reliable random sampling with large areas being poorly covered. Furthermore, these data appear as point data and usually grids are used in order to synthesize diversity metrics. These data mostly come from presence-only data without any link to relative abundance, dominance, biomass or cover, which would be reflected in remote sensing. Finally, GBIF data are inadequate for local estimates of biodiversity as they do not consider co-occurrence data. Indeed, and contrary to recent databases at the community level such as the European Vegetation Archive (EVA) (Chytry et al., 2016) or the sPlot initiative (Bruelheide et al., 2019), GBIF does not provide information on species co-occurrence which is very problematic for biodiversity assessment and monitoring. Despite the disadvantages that come from the use of public databases, there is some benefits in the use of such data. First of all, there is a huge amount of data collected and provided by citizens and research institutions available in the GBIF database when compared to the data that could be collected locally, resulting in a huge saving of time and costs. Moreover, GBIF data are standardized to the same format and therefore ready to use.

To overcome the issues due to the collection and availability of in situ ecological data, remote sensing imagery has become more and more important and is now considered a reliable tool to assess and monitor biodiversity (Tuanmu and Jetz, 2015).

1.2 The spectral species concept

Remote sensing based approaches have proven to be useful modelling techniques to detect the variability of biodiversity in space and time across scales of biological organisation, at different grains (spatial resolutions) and extents (Rocchini et al., 2013). Airborne sensors have even been used to detect and map single species distributions (Skorownek et al., 2017a), even the most tiny and inconspicuous ones such as *Campylopus introflexus*, a moss species which is highly invasive in Europe (Skorownek et al., 2017b).

Remote sensing techniques have been used to study the impact of landscape and environment on biodiversity, and to explore and visualize spatial data and biodiversity change. Therefore, remote sensing data have become among the most time and cost effective tools, allowing to make relevant conservation actions in a relatively short period of time. Furthermore, also the contrary holds true: remote sensing demonstrated the impact of biodiversity (including non-native invasive species) on ecosystem functioning (Ewald et al., 2018).

In general, vegetation absorbs the blue and the red light, for photosynthesis, while it reflects NIR radiation due to the physical structure of the cells composing the leaf mesophyllum (Wegmann et al., 2016). The bands relative to RED and NIR are used as proxies for photosynthetic activity of the vegetation. These bands are usually incorporated in a widely used index, the normalized difference vegetation index (NDVI), which is calculated as NDVI=(NIR-RED)/(NIR+RED). The higher the relative abundance of photosynthetic vegetation, the higher would be the reflectance in the NIR band and the absorption in the RED band. NDVI ranges from -1 to 1, with 0 values usually associated with non vegetated areas and negative values associated with water surfaces or snow.

This index has widely been used to discriminate different vegetation types over an area. In fact, in several studies, NDVI is positively correlated to the net primary productivity (NPP, e.g. Gillespie et al. (2008)). Therefore, it can be used as a proxy to quantify species richness and diversity, based on the species-energy theory, proposed by Currie (1991), namely a relation between species richness and energy, that would depend mainly on annual potentiafl evapotranspiration and actual evapotranspiration. Another hypothesis related to the variability in space of the spectral signal has been proposed by Palmer at al. (2002). The so called spectral variation hypothesis (SVH) states that the higher the environmental heterogeneity the higher would be the species diversity of an area, due to a higher amount of ecological niches available.

Hence, based on the SVH, spectral variability can effectively be related to environmental heterogeneity and therefore it could be used to assess species biodiversity of an area. In this sense, since the spectral variability is derived from the information present in the pixels of an acquired image, it is important that such pixels, describing the area of study, would have a spatial resolution coherent with the ecological assumptions taken into account and such that it would allow to make some predictions on biodiversity.

Among the most novel methods to estimate diversity by remote sensing, described in Rocchini et al. (2018), the spectral species concept (Féret and Asner, 2014) is one of the most powerful, since it allows to couple k-means approaches to the gridded data obtained from remote sensing technologies as a mean to derive α- and β-diversity 2D-matrices. The spectral species algorithm allows the separation of the spectral space in subunits identified as spectral species. Its root theory is built upon two major founding principles. The first is the aforementioned Spectral Variation Hypothesis, relating spectral to environmental heterogeneity. The second is based on the plant optical types proposed by Ustin and Gamon (2010). This concept is mainly related to the use of particular sensors providing high spatial resolution images and able to measure different signals about the phenology, the biochemistry and the structure of vegetation. Such sensors can obtain information at the individual plant scale level.

The method is based on an unsupervised clustering algorithm, first relying on dimensionality reduction obtained after running a principal component analysis (PCA) and then on the actual clustering of the pixels, with the subsequent assignment to spectral species, based on a k-means approach. PCA and similar clustering methods have already been shown to reliably reduce the multidimensional spectral sets for models on species and biodiversity distribution (Rocchini et al., 2010). Furthermore, the method provides an interesting visual inspection of diversity building α- and β-diversity maps.

As far as we know, the spectral species concept has been applied so far only at the local scale (Féret and Asner, 2014). Hence, the aim of this manuscript is to extend this concept over wider spatial extents passing to a spectral community concept, by generating a heterogeneity map at a wide geographical scale to estimate α- and β-diversity across Europe.

2 The algorithm

The spectral species algorithm has been originally developed to map tropical forest canopy diversity using imaging spectroscopy with a spatial resolution up to 2 meters (Féret and Asner (2014), Figure 1). Following the hypothesis that species are spectrally separable (Asner and Martin, 2009), the approach is based on the segmentation of the spectral space defined by the remote sensing data. In fact the spectral space is assumed to be a combination of several subspaces, reflecting the "signature" of one or several species. Therefore these subspaces would be the expression of a more general "spectral species". From the resultant "spectral community", it would be possible to derive the diversity of an area. The output of this algorithm will not be a list of the species of the area, but rather a map of the distribution of the spectral communities available within the area from which it might be possible to calculate several diversity indices. In particular we focused our attention on α- and β-diversity metrics. Both introduced by Whittaker (1972), the first reflects the mean species diversity in sites at a local scale whereas the second is an indicator of the spatial (or temporal) heterogeneity at a relatively larger scale. In the algorithm, α-diversity is calculated in a neighbourhood (plot) of $n \times n$ pixels by the Shannon diversity index (Shannon, 1948) calculated as follow:

$$H' = -\sum_{s=1}^{N} p_{\mathrm{s}} \ln p_{\mathrm{s}} \tag{1}$$

where p_{s} is the proportion of each spectral species s in each plot.

The β-diversity indicator is instead computed by the Bray-Curtis (hereafter BC) dissimilarity metric (Bray and Curtis, 1957):

$$BC_{\mathrm{ij}} = \frac{\sum_{s=1}^{N} |x_{\mathrm{is}} - x_{\mathrm{js}}|}{\sum_{s=1}^{N} (x_{\mathrm{is}} + x_{\mathrm{js}})} \tag{2}$$

where BC_{ij} is the dissimilarity between plots i and j and x_{is} and x_{js} are the abundances of spectral species s in plots i and j.

In the spectral species algorithm, once the BC dissimilarity matrix between all pairs of plots is computed, a multidimensional scaling is performed in order to translate information about the pairwise dissimilarity among P plots into a configuration of P points mapped in a 3-dimensional Cartesian space such as NMDS or PCoA (Mead, 1992). This simplified translation of the BC dissimilarity matrix can then be displayed as a colored map. More details can be found in Féret and de Boissieu (in press).

While the Shannon index has a theoretical maximum limit corresponding to the $ln(richness)$, the Bray-Curtis index ranges from 0 to 1, where 0 is indicating that the two sites are identical whereas 1 indicates that the two sites do not share species. Hence, BC can be considered as an estimate of the heterogeneity of a certain area. The final aim of the method was to generate an heterogeneity map across the study region. Strictly speaking, the method is a clustering approach which (i) divides the subspaces in spectral units and (ii) assigns it to spectral species from which (iii) different diversity maps can be obtained. Box 1 focuses in detail on the main steps of the algorithm, while the dedicated R package `biodivMapR` is now available (https://github.com/jbferet/biodivMapR) and fully described in Féret and de Boissieu (in press).

2.1 Application of the algorithm

Remote sensing data are usually provided as raster objects with a geographic coordinate system information, namely regular grids (matrices) or stacks of raster layers (e.g. one raster layer per band for multispectral or hyperspectral data), in which each cell represents a pixel with the corresponding reflectance value associated to a specific band. Such data have been manipulated with the Software R Development Core Team (2019). R can be used for remote sensing data analysis since it includes spatial functionalities throughout a suite of R packages like the `rgdal` and `raster` packages (see Box 2 for more information).

Our main purpose was to apply the spectral species algorithm to a continental-scale geographical region such as Europe. Hence, Moderate Resolution Imaging Spectroradiometer (MODIS) data, with a spatial resolution of 500m covering Europe, were downloaded at the United States Geological Survey (USGS) site (`https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod09a1_v006`). After a visual check of the images, in order to guarantee i) the coverage of a complete phenological period and to ii) avoid noise related to clouds, we referred to the RED and NIR bands from 2018 from January to December, to calculate NDVI, by generating a sample set of 12 NDVI images (Figure 2).

In this case, due to the input spatial resolution of MODIS (500m), mixed pixels would occur, by smoothing the reflectance related to single plant species. In other words, the direct relationship between the spectral species detected in the spectral space versus plant species does not hold true. This said, from the diversity measurement perspective this is just a matter of terms being used, with spectral species being more related to field plant communities, habitats or other ecological entities.

For the derivation of spectral species, in order to define the number of clusters, we relied on the highest number of clusters with stable results after a trial and error procedure, reaching 200 clusters, i.e. spectral species. Once pixels with similar NDVI values in 12 dimensions were clumped together, Shannon's H' was calculated with a window size of 10x10 pixels and an output resolution of 5km. The attained α-diversity map quantitatively showed the local spectral diversity distribution over Europe (Figure 3), with a higher heterogeneity found in i) more topographically complex regions, mainly due to strong local differences induced by elevation gradients (passing from forests to grasslands, to rocks and snow), and/or differences in terms of seasonality in relation with elevation, as in Rocchini et al. (2019), and in ii) more contrasted agricultural areas in both the spatial and temporal dimensions (Hobbs et al., 2006; Vihervaara et al., 2017).

β-diversity (Figure 3) showed a clear differentiation among different areas over Europe. The attained map was in line with the European Environmental Agency map of ecoregions (Figure 4, see Mucher et al. (2009)). The correspondence of the achieved patterns in the two maps was apparent, with a similar contour of the major ecoregions like the Mediterranean, the Atlantic, the Continental, the Boreal and the Alpin regions. This demonstrates an intrinsic ability of the spectral species approach to catch differences in the physiological and functional properties of vegetation even at wide spatial scales, starting from spectral reflectance or spectral indices. Minor differences were mainly related to the biogegraphical (i.e., purely spatial) differentiation of ecoregions in the EEA map. As an example, Alpine ecoregions could not be distinguished between North and South Alpine regions by the spectral species approach, since, having very similar conifer species composition, they show the same physiological, phenological and thus spectral pattern.

3 Discussion

In this paper, for the first time, the spectral species concept has been extended from the consideration of a single species to an entire community. We demonstrated that the combined use of the novel unsupervised clustering method proposed by Féret and Asner (2014) with NDVI time series at European scale, allows the derivation of local (α) diversity and turnover (β) relying on free to use and operationally available satellite data.

With regards to a potential validation with in-situ data, the uncertainty of wide-scale datasets hampers a spatial overlap. In this case, in-situ datasets meet all five major concerns recently raised by Hobohm et al. (2019), i.e.: i) there is insufficient data coverage across Europe to make an unbiased comparison between predicted and actual distributions, ii) taxonomic standards differ across sampled regions, iii) there are generally different shapes of areas being sampled, iv) political borders often define sampling areas and aggregated sampling areas, and v) data are not aggregated in the same way in all areas. Furthermore, spatial information has an intrinsic varying degree of relevance mainly due to the fact that, rather than species lists, it is made up of geometrical precision, attributes robustness and temporal consistency (Hobona et al., 2006). Finally, different models and approaches to measuring diversity inevitably provide different outputs, as pointed out in the generalised entropy theory put forward by Rényi (1961). Given the above validation difficulties, we decided to qualitatively compare our generated output, in particular the β-diversity map, with existing ecoregion maps, which are expected to discriminate different spatial areas based on natural borders defined by biological diversity (https://ecoregions2017.appspot.com/) and thus are intrinsically related to differences in the species and spectral turnover of communities.

Since the output of the algorithm represents the variation of the pixel values in space and time, the most diverse pixels were those with the highest turnover among the neighborhood areas and most affected by seasonality. The importance of accounting for turnover instead of simple richness has been widely discussed in the ecological literature (Tuomisto, 2010), since environmental variability over spatial gradients is one of the major drivers of the structure and composition of diversity (Legendre et al., 2005). In this view, the use of the "spectral species concept", defined as the variation of clustered pixel values, represents a powerful approach for the investigation of gradient variation of diversity in space and, potentially, in time.

In general, the measure of variability in space has been demonstrated to follow scale-based differentiation. In other words, results are expected to change with spatial scale in terms of both grain (spatial resolution) and extent (extent of geographical area of interest, Palmer at al. (2002)). Regarding extent, one of the major weaknesses of the proposed algorithm in β-diversity quantification (although this applies in general to all measurements of turnover) is that by increasing the extent of an observation area, the estimated values for an individual comparison between sites are modified by the increasing spectral species pool.

Additional drawbacks at the current stage of the algorithm include: i) the use of remotely sensed data which are not necessarily related to the main drivers of species distributions and of diversity, ii) the general multicollinearity found in most of the remotely sensed sets, iii) the unsupervised clustering process being adopted.

Concerning climate, a solution might be found in the use of remotely sensed derived climate data adding climate change as an additional layer of complexity as in Rocchini et al. (2015a) and in (Zellweger et al., 2019). Also in this case multicollinearity of climate variables should be seriously taken into account, as we did for the original remote sensing data, by applying a PCA to reduce the noise in the data and detect potential artifacts; then, PCA components might also be visualised to find potential congruence between spectral species and real species patterns. Finally, the process for grouping pixels in spectral species is based on an unsupervised clustering, where the definition of the number of clusters should be done a-priori. In this case, we hypothesized that the diversity of types of landscapes and gradient of climates across Europe may require a large number of clusters to correctly differentiate among them, relying on a fuzzy view of ecosystems (Rocchini and Ricotta, 2007). Hence, we decided to adopt a trial and error procedure until reaching a sort of threshold in which no significant changes were further found. Such threshold was 200 clusters, namely the final used number. In the near future, it might be interesting to make a sensitivity analysis to show the impact of the number of clusters on the final analysis.

Considering the use of remote sensing for species diversity estimates, correlation and determination coefficients are generally statistically significant but low, hampering the direct use of remotely sensed diversity in simple univariate models (Rocchini et al., 2018). In fact, the relationship between α- or β-diversity and habitat heterogeneity, which is the founding principle of the use of remote sensing data for these analyses, is rarely linear (Ferrier et al., 2007), mainly because of variation in the rate of species turnover along an environmental gradient. However, remote sensing variables are generally well suited in more complex multivariate models accounting for part of the diversity explained for species communities (Rocchini et al., 2018). This is especially true considering that environmental turnover generally explains more variation in species diversity rather than mere spatial structure (Hernandez-Stefanoni et al., 2012). Moreover, based on their high temporal resolution, remote sensing data might be useful to detect drastic changes of diversity in space and time, e.g. related to catastrophic events, overall considering the intrinsic difficulties in relying on in-situ data for wide geographical scales (Cord and Rödder, 2011; Hobohm et al., 2019).

From an ecological perspective, remote sensing imagery bands (dimensions) show a high affinity with the hypervolume axes proposed by Hutchinson (1957) for modelling species niches. In the Hutchinson's theory, an hypervolume is represented by a space defined by a set of n independent axes which could be related to the final variables driving the realised niche of a species (see also Blonder (2017) and Ricotta et al. (2010) on the niche differentiation concept). In our case, such axes would be the original satellite sensor bands being strictly related to the identification of a spectral species and the resulting spectral community in a site, instead of a niche. From this point of view, spectral species and communities are in line with joint species distribution models (JSDMs), which explicitly take into account biotic interactions among species in a community, while in our model the "interaction" among pixel values is ruled out in general by their proximity both from a spatial and from a spectral point of view. In this paper, the final aim was not to model single spectral species or spectral communities but rather to estimate diversity and its change over space and time, following the mathematical principles described in Liu et al. (2014) and Rocchini et al. (2015b), for which the distribution of diversity over space is actually a particular case of the so-called switched systems, i.e. hybrid systems resulting from both continuous and discrete dynamics with a high number of different potential variables acting as main drivers of diversity response. In practice, in our view we succeeded to fill a previous gap in spatio-ecological analysis, i.e. the translation of what in remote sensing science is known as "spectral mixture modeling" (Jensen, 2015) into an ecological diversity theory perspective. In fact, in spectral mixture modeling the measured spectral reflectance is decomposed as a mixture of endmembers. In our case, such mixture was used to directly compute alpha- and overall beta-diversity over wide spatial areas in few time.

4 Conclusion

Predicting and mapping α- and β-diversity using remotely sensed images acquired over large areas is currently a key topic in ecology, and could provide landscape managers with effective tools to confront global change. In this paper, we proposed a novel method based on preliminary unsupervised clustering of spectral data (NDVI time series derived from MODIS data), assigning each pixel to a "spectral species" and then calculating diversity based on a dissimilarity metric. At the scale of this study, the one-to-one relationship between spectral species and in-situ plant species is not achieved, but the spectral species concept still holds true once considering that the detected spectral species in the spectral space are related to higher-order plant hierarchies (assemblages, entire habitats, etc.). That is, from an algorithmic point of view, the bulk of the calculations are unaltered.

Based on the results presented here, the use of the spectral species and communities concept would appear to promote more effective planning and policies related to the conservation of wild species, by improving our understanding of the dynamics of local and global biodiversity at different spatial and temporal scales.

5 Acknowledgements

DR and DK were supported by the H2020 project SHOWCASE. DR was also supported by the H2020 COST Action CA17134 “Optical synergies for spatiotemporal sensing of scalable ecophysiological traits (SENSECO)”. This paper has been partially developed during a Short Term Scientific Mission supported by the H2020 COST Action CA15212 “Citizen Science to promote creativity, scientific literacy, and innovation throughout Europe”.

JBF and FdB acknowledge financial support from Agence Nationale de la Recherche (BioCop project—ANR-17-CE32-0001) and TOSCA program grant of the French Space Agency (CNES) (HyperTropik/HyperBIO project).

References

Alados, C.L., Pueyo, Y., Escos, J., Andujar, A. (2009). Effects of the spatial pattern of disturbance on the patch-occupancy dynamics of juniper–pine open woodland. Ecological Modelling, 220: 1544-1550.

Asner, G.P. and Martin, R.E. (2009). Airborne spectranomics: mapping canopy chemical and taxonomic diversity in tropical forests. Frontiers in Ecology and the Environment, 7: 269-276.

Baselga, A. (2012).The relationship between species replacement, dissimilarity derived from nestedness, and nestedness. Global Ecology and Biogeography 21: 1223-1232.

Beierkuhnlein, C. (2003). Der Begriff Biodiversität. Nova Acta Leopoldina N.F., 87: 51-72.

Blonder, B. (2017). Hypervolume concepts in niche- and trait-based ecology. Ecography 41: 1441-1445.

Borg, I., Groenen, P. (2005). Modern Multidimensional Scaling: theory and applications (2nd edition). Springer-Verlag, New York.

Bray, J.R. and Curtis J.T. (1957). An ordination of the upland forest communities of southern Wisconsin. Ecological monographs 27: 325-349.

Bruelheide, H., Dengler, J., Jimenez-Alfaro, B., Purschke, O., Hennekens, S.M., Chytry, M., Pillar, V.D., Jansen, F., Kattge, J., Sandel, B., Aubin, I., Biurrun, I., Field, R., Haider, S., Jandt, U., Lenoir, J., Peet, R.K., Peyre, G., Sabatini, F.M., Schmidt, M., Schrodt, F., Winter, M., Acic, S., Agrillo, E., Alvarez, M., Ambarlı, D., Angelini, P., Apostolova, I., Arfin Khan, M.A.S., Arnst, E., Attorre, F., Baraloto, C., Beckmann, M., Berg, C., Bergeron, Y., Bergmeier, E., Bjorkman, A.D., Bondareva, V., Borchardt, P., Botta-Dukat, Z., Boyle, B., Breen, A., Brisse, H., Byun, C., Cabido, M.R., Casella, L., Cayuela, L., Cerny, T., Chepinoga, V., Csiky, J., Curran, M., Custerevska, R., Dajic, Z., De Bie, S.E., de Ruffray, P., De Sanctis, M., Dimopoulos, P., Dressler, S., Ejrnaes, R., El-Rouf, M.A., El-Sheikh, M., Enquist, B., Ewald, J., Fagundez, J., Finckh, M., Font, X., Forey, E., Fotiadis, G., Garcia-Mijangos, I.., Luis de Gasper, A., Golub, V., Gutierrez, A.G., Hatim, M.Z., He, T., Higuchi, P., Holubova, D., Holzel, N., Homeier, J., Indreica, A., Isık Gursoy, D., Jansen, S., Janssen, J., Jedrzejek, B., Jirousek, M., Jurgens, N., Kacki, Z., Kavgacı, A., Kearsley, E., Kessler, M., Knollova, I., Kolomiychuk, V., Korolyuk, A., Kozhevnikova, M., Kozub, L., Krstonosic, D., Kuhl, H., Kuhn, I., Kuzemko, A., Kuzmic, F., Landucci, F., Lee, M.T., Levesley, A., Li, C.-F., Liu, H., Lopez-Gonzalez, G., Lysenko, T., Macanovic, A., Mahdavi, P., Manning, P., Marceno, C., Martynenko, V., Mencuccini, M., Minden, V., Moeslund, J.E., Moretti, M., Muller, J.V., Munzinger, J., Niinemets, U., Nobis, M., Noroozi, J., Nowak, A., Onyshchenko, V., Overbeck, G.E., Ozinga, W.A., Pauchard, A., Pedashenko, H., Penuelas, J., Perez-Haase, A., Peterka, T., Pettrk, P., Phillips, O.L., Prokhorov, V., Rasomavicius, V., Revermann, R., Rodwell, J., Ruprecht, E., Rusina, S., Samimi, C., Schaminee, J.H.J., Schmiedel, U., Sibik, J., Silc, U., Skvorc, Z., Smyth, A., Sop, T., Sopotlieva, D., Sparrow, B., Stancic, Z., Svenning, J.-C., Swacha, G., Tang, Z., Tsiripidis, I., Turtureanu, P.D., Ugurlu, E., Uogintas, D., Valachovic, M., Andre, K., Vashenyak, V.Y., Vassilev, K., Velez-Martin, E., Venanzoni, R., Vibrans, A.C., Violle, C., Virtanen, R., von Wehrden, H., Wagner, V., Walker, D.A., Wana, D., Weiher. E., Wesche, K., Whitfeld, T., Willner, W., Wiser, S., Wohlgemuth, T., Yamalov, S., Zizka, G., Zverev, A. (2019). sPlot – a new tool for global vegetation analyses. Journal of Vegetation Science, 30: 161-186.

Chytry, M. , Hennekens, S. M., Jimenez-Alfaro, B. , Knollova, I. , Dengler, J. , Jansen, F. , Landucci, F. , Schaminee, J. H., Acic, S. , Agrillo, E. , Ambarlı, D. , Angelini, P. , Apostolova, I. , Attorre, F. , Berg, C. , Bergmeier, E. , Biurrun, I. , Botta-Dukat, Z. , Brisse, H. , Campos, J. A., Carlon, L. , Carni, A. , Casella, L. , Csiky, J. , Custerevska, R. , Dajic Stevanovic, Z. , Danihelka, J. , De Bie, E. , Ruffray, P. , De Sanctis, M. , Dickore, W. B., Dimopoulos, P. , Dubyna, D. , Dziuba, T. , Ejrnæs, R. , Ermakov, N. , Ewald, J. , Fanelli, G. , Fernandez-Gonzalez, F. , FitzPatrick, U. , Font, X. , Garcia-Mijangos, I. , Gavilan, R. G., Golub, V. , Guarino, R. , Haveman, R. , Indreica, A. , Isık Gursoy, D. , Jandt, U. , Janssen, J. A., Jirousek, M. , Kacki, Z. , Kavgacı, A. , Kleikamp, M. , Kolomiychuk, V. , Krstivojevic Cuk, M. , Krstonosic, D. , Kuzemko, A. , Lenoir, J. , Lysenko, T. , Marcenò, C. , Martynenko, V. , Michalcova, D. , Moeslund, J. E., Onyshchenko, V. , Pedashenko, H. , Perez-Haase, A. , Peterka, T. , Prokhorov, V. , Rasomavicius, V. , Rodriguez-Rojo, M. P., Rodwell, J. S., Rogova, T. , Ruprecht, E. , Rusina, S. , Seidler, G. , Sibik, J. , Silc, U. , Skvorc, Z. , Sopotlieva, D. , Stancic, Z. , Svenning, J. , Swacha, G. , Tsiripidis, I. , Turtureanu, P. D., Ugurlu, E. , Uogintas, D. , Valachovic, M. , Vashenyak, Y. , Vassilev, K. , Venanzoni, R. , Virtanen, R. , Weekes, L. , Willner, W. , Wohlgemuth, T., Yamalov, S. (2016), European Vegetation Archive (EVA): an integrated database of European vegetation plots. European Vegetation Archive (EVA): an integrated database of European vegetation plots. Applied Vegetation Science 19: 173–180.

Cord, A., Rödder; D. (2011). Inclusion of habitat availability in species distribution models through multi-temporal remote-sensing data? Ecological Applications, 21: 3285-3298.

Clark, J.S. (2008). Beyond neutral science. Trends in Ecology and Evolution, 24: 8-15.

Currie, D.J. (1991). Energy and large-scale patterns of animal-and plant-species richness. The American Naturalist 137: 27-49.

Dinerstein, E., Olson, D., Joshi, A., Vynne, C., Burgess, N.D., Wikramanayake, E., Hahn, N., Palminteri, S., Hedao, P., Noss, R., Hansen, M., Locke, H., Ellis, E.C., Jones, B., Barber, C.V., Hayes, R., Kormos, C., Martin, V., Crist, E., Sechrest, W., Price, L., Baillie, J.E.M., Weeden, D., Suckling, K., Davis, C., Sizer, N., Moore, R., Thau, D., Birch, T., Potapov, P., Turubanova, S., Tyukavina, A., de Souza, N., Pintea, L., Brito, J.C., Llewellyn, O.A., Miller, A.G., Patzelt, A., Ghazanfar, S.A., Timberlake, J., Kloser, H., Shennan-Farpon, Y., Kindt, R., Lilleso, J.B., van Breugel, P., Graudal, L., Voge, M., Al-Shammari, K.F., Saleem, M.A. (2017). Ecoregion-based approach to protecting half the terrestrial realm. Bioscience, 67: 534-545.

Elton, C.S. (1958). The Ecology of invasions by animals and plants. Methuen, London.

Ewald., M., Skowronek, S., Aerts, R., Dolos, K., Lenoir, J., Nicolas, M., Warrie, J., Hattab, T., Feilhauer, H., Honnay, O., Garzon-Lopez, C.X., Decocq, G., Van De Kerchove, R., Somers, B., Rocchini, D., Schmidtlein, S. (2018). Analyzing remotely sensed structural and chemical canopy traits of a forest invaded by *Prunus serotina* over multiple spatial scales. Biological Invasions, 20: 2257-2271.

Féret, J.-B., Asner, G.P. (2014). Mapping tropical forest canopy diversity using high-fidelity imaging spectroscopy. Ecological Applications 24: 1289-1296.

Féret, J.-B., de Boissieu, F. (in press). `biodivMapR`: An `R` package for $\alpha-$ and β-diversity mapping using remotely sensed images. Methods in Ecology and Evolution.

Ferrier, S., Manion, G., Elith, J. and Richardson, K. (2007). Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. Diversity and Distributions 13: 252-264.

Foody, G.M. (2008). GIS: biodiversity applications. Progress in Physical Geography 32: 223-235.

Gaston, K.J. (2000). Global patterns in biodiversity. Nature 405: 220.

Gillespie, T.W., Foody, G.M., Rocchini, D., Giorgi, A.P. and Saatchi, S. (2008). Measuring and modelling biodiversity from space. Progress in Physical Geography 32: 203-221.

Hernandez-Stefanoni, J.L., Gallardo-Cruz, J.A., Meave, J.A., Rocchini, D., Bello-Pineda, J., Lopez-Martinez, J.O. (2012). Modeling alpha- and beta-diversity in a tropical forest from remotely sensed and spatial data. International Journal of Applied Earth Observation and Geoinformation, 19: 359-368.

Hobbs, R.J., Arico, S., Aronson, J., Baron, J.S., Bridgewater, P., Cramer, V.A., Epstein, P.R., Ewel, J.J., Klink, C.A., Lugo, A.E., Norton, D., Ojima, D., Richardson, D.M., Sanderson, E.W., Valladares, F., Vila, M., Zamora, R., Zobel, M. (2006). Novel ecosystems: theoretical and management aspects of the new ecological world order. Global Ecology and Biogeography, 15: 1-7.

Hobohm, C., Janisova, M., Steinbauer, M., Landi, S., Field, R., Vanderplank, S., Beierkuhnlein, C., Grytnes, J.-A., Vetaas, R.O., Fidelis, A., de Nascimento, L., Clark, V.P., Fernandez-Palacios, J.M., Franklin, S., Guarino, R., Huang, J., Krestov, P., Ma, K., Onipchenko, V., Palmer, M.W., Fragomeni Simon, M., Stolz, C., Chiarucci, A. (2019). Global endemics-area relationships of vascular plants. Perspectives in Ecology and Conservation, 17, 41-49.

Hobona, G., James, P., Fairbairn, D. (2006). Multidimensional visualisation of degrees of relevance of geographic data. International Journal of Geographical Information Science. 20: 469-490.

Hutchinson, G. (1957). Concluding remarks. Cold Spring Harb. Symp. Quant. Biol. 22: 415-427.

Jurasinski, G., Retzer, V., Beierkuhnlein, C. (2009). Inventory, differentiation, and proportional diversity: a consistent terminology for quantifying species diversity. Oecologia, 159: 15-26.

Jensen, R.J. (2015). Introductory Digital Image Processing: A Remote Sensing Perspective. Pearson Publishing.

Kohonen, T. (1982). Analysis of a simple self-organizing process. Biological cybernetics 44: 135-140.

Lande, R. (1996). Statistics and partitioning of species diversity, and similarity among multiple communities. Oikos 76: 5-13.

Legendre, P., Borcard, D., Peres-Neto, P.R. (2005). Analyzing beta diversity: partitioning the spatial variation of community composition data. Ecological Monographs 75: 435-450.

Liu, X., Zhong, S., Ding, X. (2014). A Razumikhin approach to exponential admissibility of switched descriptor delayed systems. Applied Mathematical Modelling 38, 1647–1659.

Maldonado, C., Molina, C.I., Zizka, A., Persson, C., Taylor, C.M., Albán, J., Chilquillo, E., Ronsted, N. and Antonelli, A. (2015). Estimating species diversity and distribution in the era of Big Data: to what extent can we trust public databases?. Global Ecology and Biogeography 24: 973-984.

Mead, A. (1992). Review of the Development of Multidimensional Scaling Methods. The Statistician, 41: 27-39.

Moran, P.A.P. (1950). Notes on continuous stochastic phenomena. Biometrika, 37: 17-23.

Mouchet, M., Levers, C., Zupan, L., Kuemmerle, T., Plutzar, C., Erb, K., Lavorel, S., Thuiller, W., Haberl, H. (2015). Testing the effectiveness of environmental variables to explain European terrestrial vertebrate species richness across biogeographical scales. PLoS ONE, 10: e0131924.

Mucher, C.A., Hennekens, S.M., Bunce, R.G.H., Schaminee, J.H.J., Schaepman, M.E. (2009). Modelling the spatial distribution of Natura 2000 habitats across Europe. Landscape and Urban Planning, 92: 148-159.

Palmer, M.W., Earls, P.G., Hoagland, B.W., White, P.S. and Wohlgemuth, T. (2002). Quantitative tools for perfecting species lists. Environmetrics, 13: 121-137.

Peuquet, D.J. (1992). An algorithm for calculating minimum euclidean distance between two geographic features. Computers & Geosciences 18: 989-1001.

R Development Core Team (2019). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. `https://www.R-project.org/`

Rényi, A. (1961). On measures of information and entropy. Proceedings of the fourth Berkeley Symposium on Mathematics, Statistics and Probability 1960: 547-561.

Ricotta, C., Godefroid, S., Rocchini, D. (2010). Patterns of native and exotic species richness in the urban flora of Brussels: rejecting the 'rich get richer' model. Biological Invasions, 12. 233-240.

Rocchini, D. (2007). Distance decay in spectral space in analysing ecosystem β-diversity. International Journal of Remote Sensing, 28: 2635-2644.

Rocchini, D., Andreo, V., Förster, M., Garzon-Lopez, C.X., Gutierrez, A.P., Gillespie ,T.W., Hauffe, H.C., He, K.S., Kleinschmit, B., Mairota, P., Marcantonio, M., Metz, M., Nagendra, H., Pareeth, S., Ponti, L., Ricotta, C., Rizzoli, A., Schaab, G., Zebisch, M., Zorer, R., Neteler, M. (2015a). Potential of remote sensing to predict species invasions: A modelling perspective. Progress in Physical Geography, 39: 283-309.

Rocchini, D., Balkenhol, N., Carter, G.A., Foody, G.M. Gillespie, T.W., He, K.S., Kark, S., Levin, N., Lucas, K., Luoto, M., Nagendra, H., Oldeland, J., Ricotta, C., Southworth, J. and Neteler, M. (2010). Remotely sensed spectral heterogeneity as a proxy of species diversity: recent advances and open challenges. Ecological Informatics, 5: 318-329.

Rocchini, D., Delucchi, L., Bacaro, G., Cavallini, P., Feilhauer, H., Foody, G.M., He, K.S., Nagendra, H., Porta, C., Ricotta, C., Schmidtlein, S., Spano, L.D., Wegmann, M., Neteler, M. (2013). Calculating landscape diversity with information-theory based indices: A GRASS GIS solution. Ecological Informatics, 17: 82-93.

Rocchini, D., Luque, S., Pettorelli, N., Bastin, L., Doktor, D., Faedi, N., Feilhauer, H., Féret, J.B., Foody, G.M., Gavish, Y., Godinho, S., Kunin, W.E., Lausch, A., Leitao, P.J., Marcantonio, M., Neteler, M., Ricotta, C., Schmidtlein, S., Vihervaara, P., Wegmann, M., Nagendra, H. (2018). β-diversity by remote sensing: A challenge for biodiversity monitoring. Methods in Ecology and Evolution 9: 1787-1798.

Rocchini, D., Hernandez Stefanoni, J.L., He, K.S. (2015b). Advancing species diversity estimate by remotely sensed proxies: a conceptual review. Ecological Informatics, 25: 22-28.

Rocchini, D., McGlinn, D., Ricotta, C., Neteler, M., Wohlgemuth, T. (2011). Landscape complexity and spatial scale influence the relationship between remotely sensed spectral diversity and survey based plant species richness. Journal of Vegetation Science, 22: 688-698.

Rocchini, D., Marcantonio, M., Da Re, D., Chirici, G., Galluzzi, M., Lenoir, J., Ricotta, C., Torresani, M., Ziv, G. (2019). Time-lapsing biodiversity: an open source method for measuring diversity changes by remote sensing. Remote Sensing of Environment, 231: 111192.

Rocchini, D., Ricotta, C. (2007). Are landscapes as crisp as we may think? Ecological Modelling, 204: 535-539.

Skowronek, S., Asner, G.P., Feilhauer, H. (2017). Performance of one-class classifiers for invasive species mapping using airborne imaging spectroscopy. Ecological Informatics, 37: 66-76.

Skowronek, S., Ewald, M., Isermann, M., Van De Kerchove, R., Lenoir, J., Aerts, R., Warrie, J., Hattab, T., Honnay, O., Schmidtlein, S., Rocchini, D., Somers, B., Feilhauer, H. (2017). Mapping an invasive bryophyte species using hyperspectral remote sensing data. Biological Invasions, 19: 239-254.

Shannon, C.E. (1948). A mathematical theory of communication. ACM SIGMOBILE mobile computing and communications review 5: 3-55.

Steinitz, O., Heller, J., Tsoar, A., Rotem, D., Kadmon, R. (2006). Environment, dispersal and patterns of species similarity. Journal of Biogeography: 33, 1044-1054.

Sun, L., Mi, X., Wei, J., Wang, J., Tian, X., Yu, H. and Gan, P. (2017). A cloud detection algorithm-generating method for remote sensing data at visible to short-wave infrared wavelengths. ISPRS Journal of Photogrammetry and Remote Sensing 124: 70-88.

Tobler, W., (1970). A computer movie simulating urban growth in the Detroit region. Economic Geography, 46: 234-240.

Tuanmu, M.N., Jetz, W. (2015). A global, remote sensing-based characterization of terrestrial habitat heterogeneity for biodiversity and ecosystem modelling. Global Ecology and Biogeography, 24: 1329-1339.

Tuomisto, H. (2010). A consistent terminology for quantifying species diversity? Yes, it does exist. Oecologia 164: 853-860.

Ustin, S.L. and Gamon, J.A. (2010). Remote sensing of plant functional types. New Phytologist 186.4: 795-816.

Vihervaara, P., Auvinen, A.P., Mononen, L., Torma, M., Ahlroth, P., Anttila, S., Bottcher, K., Forsius, M., Heino, J., Heliola, J., Koskelainen, M., Kuussaari, M., Meissner, K., Ojala, O., Tuominen, S., Viitasalo, M., Virkkala, R. (2017). How Essential Biodiversity Variables and remote sensing can help national biodiversity monitoring. Global Ecology and Conservation, 10: 43-59.

Waples, R., Gaggiotti, O. (2006). What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Molecular Ecology, 15: 1419-1439.

Wegmann, M., Leutner, B. and Dech, S. eds. Remote sensing and GIS for ecologists: using open source software. (2016). Pelagic Publishing Ltd, 2016.

Whittaker, R.H. (1972). Evolution and measurement of species diversity. Taxon 21: 213-251.

Zellweger, F., De Frenne, P., Lenoir, J., Rocchini, D., Coomes, D. (2019). Advances in microclimate ecology arising from remote sensing. Trends in Ecology & Evolution, 34: 327-341.

Figures

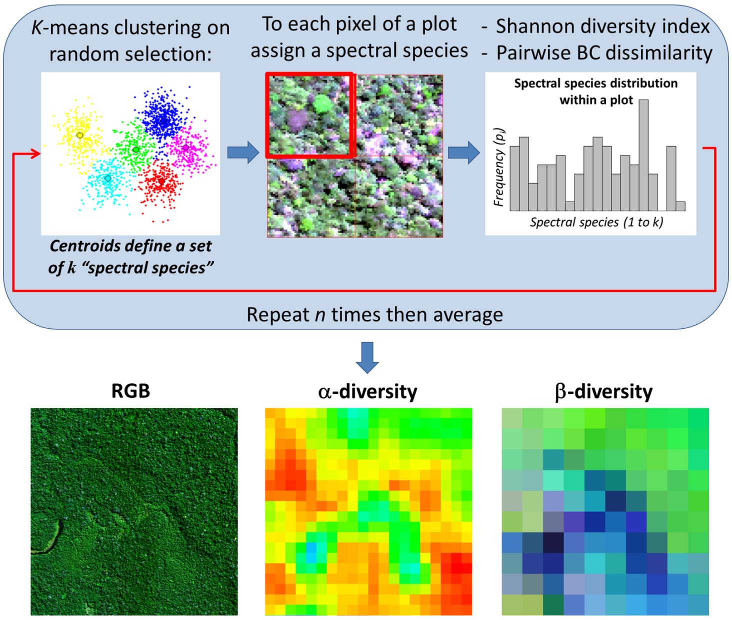

Figure 1: Diagrammatic representation of the steps of the algorithm used to achieve α- and β-diversities, redrawn from (Féret and Asner, 2014). Pixels are clumped in a spectral species and spectral community diversity is calculated. We refer to the main text and to Box 1 for additional information.

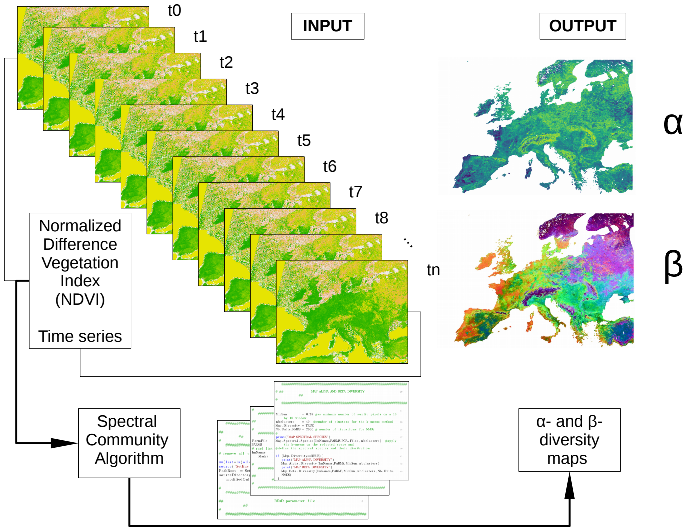

Figure 2: An input set of n images can be handled to create a time series and use the stack to further calculate the spectral community diversity. In our paper, a stack of 12 NDVI images of 2018 from the MODIS sensor was processed by the spectral species algorithm, by producing α- and β-diversity maps.

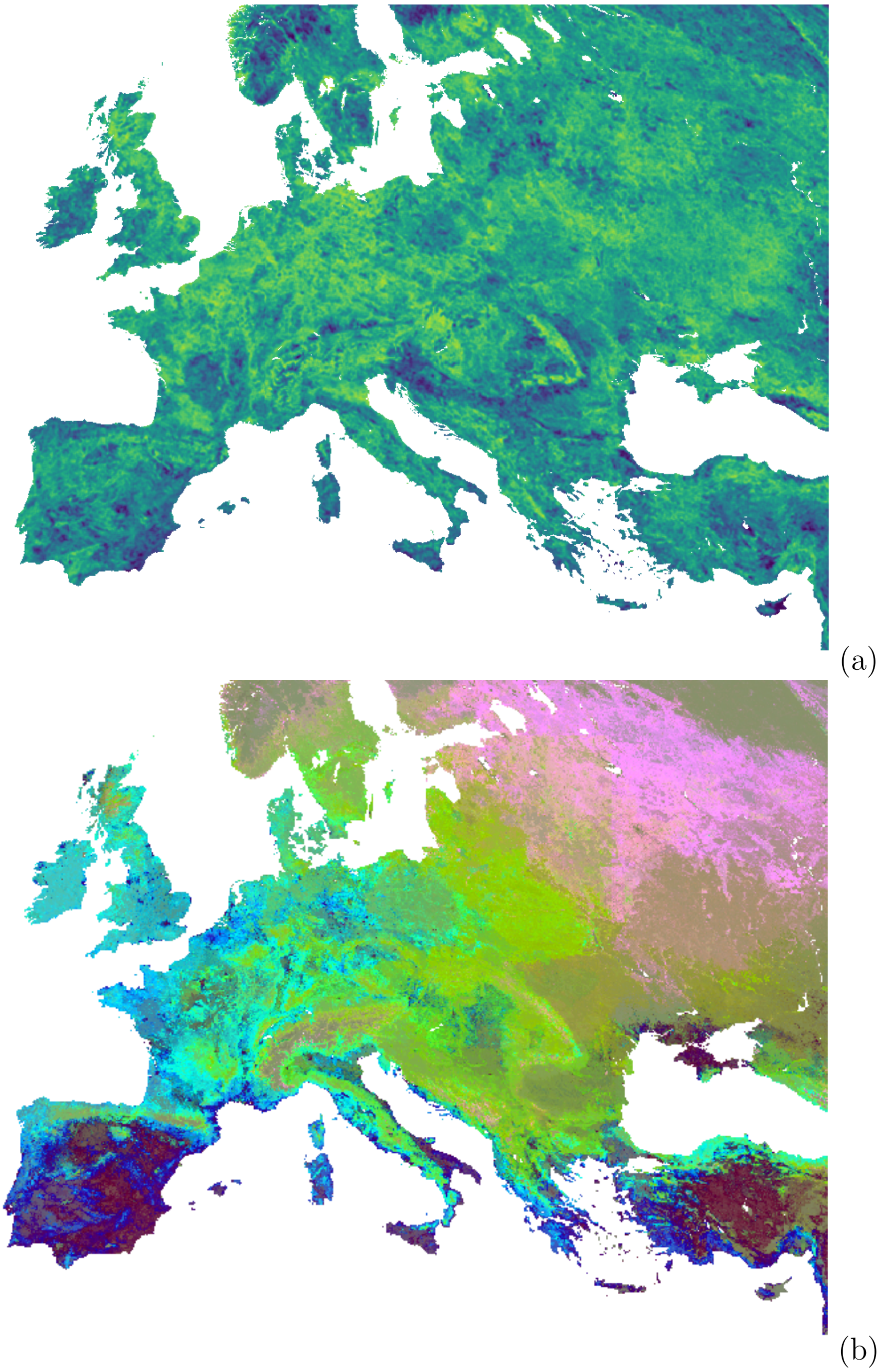

Figure 3: α- (a) and β-diversity (b) maps obtained by the spectral species algorithm. (a) The α-diversity map, based on Shannon's H' index (ranging from blue [low values] to light green [high values]) calculated in a 10x10 pixels local neighbourhood, corresponds to the local entropy of clusters, so that each location is independent from the others; (b) The β-diversity map - Bray-Curtis dissimilarity reduced to 3 dimensions with NMDS - provides information about the dissimilarity among any location in the image. Here, the distance between pairs of spatial units is expressed as a 3 colour code.

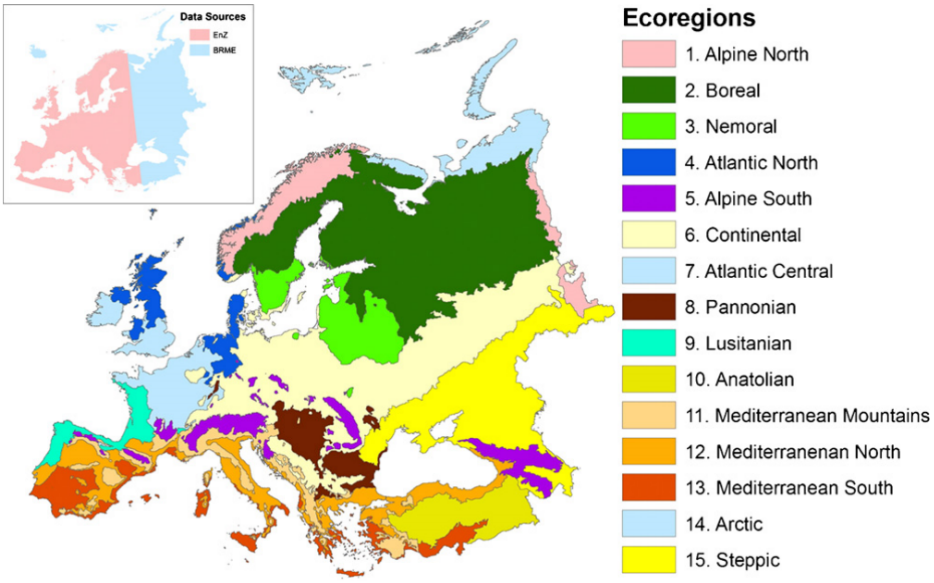

Figure 4: European Environmental Agency ecoregions map redrawn by Mucher et al. (2009). Similar maps at a coarser grain are provided by Mouchet et al. (2015) and Dinerstein et al. (2017).

Box 1 - Steps composing the spectral species algorithm

1. A Principal Component Analysis (PCA) is applied to the spectral data. PCA is not performed on the whole image, but only on a large subset of pixels randomly selected from the image. Due to the high dimensionality of the data, the reduction of the dataset is not altering the result. Those principal components explaining most of the variance of the original set are then retained for further steps.

2. A subset of pixels is then randomly selected across the entire map and the spectral space containing such a subset is partitioned into spectral species using k-means clustering with the number of k clusters being decided a priori. Then the centroids defining the spectral species are located.

3. The spectral dataset is divided into final mapping units. Each pixel is assigned to a given spectral species based on the minimal Euclidean distance between pixels (Peuquet, 1992) and the previously defined centroids.

4. A spectral species distribution is obtained for each mapping unit from which the α- and β-diversity indices are computed as previously stated.

5. Since the spectral species distribution is obtained by a subset of pixels, in order to avoid under-representation of some small-scaled ecological classes (e.g. small scale vegetation patterns), steps 4 and 5 are repeated 100 times, and the indicators obtained for each repetition are averaged. In particular the Bray-Curtis dissimilarity matrix is computed for each pair of spatial units, based on their spectral species distribution at each iteration; then the final matrix corresponds to the BC dissimilarity averaged over all the iterations.

6. Non metric Multidimensional Scaling (NMDS) (e.g. Borg and Groenen (2005)) is applied to the matrices in order to obtain a visual representation of the results. NMDS is an ordination technique usually applied in ecology that differs from other ordination techniques as PCA, since in NMDS a small number of axes are chosen prior to the analysis and then the data are fitted into the chosen dimensions. Furthermore, NMDS

is not an analytical but numerical technique, seeking for the right solution (convergence) iteratively. Finally, NMDS is not an eigenvector-eigenvalue technique, hence a NMDS ordination can be rotated among the axes. NMDS is mostly used in ecology for its versatility since it accepts any distance measure of the samples. In this case the Bray-Curtis matrix was used. In the applied NMDS approach, the first step is generally to decide the number of reduced dimensions; in this case 3 dimensions were chosen. The algorithm starts with the construction of initial random arrangements of the pixels. Then the Euclidean distances among the samples is calculated in this first configuration; those distances are regressed against the original distance matrix, and the predicted ordination distances are calculated. Finally, the regression is fitted by the least-squares method. The goodness of fit is measured by the sum of squared differences between ordination-based distances and the predicted distances. The goodness of fit is calculated through the Kruskal's Stress index:

$$Stress = \sqrt{\frac{[r]\sum_{h,i}(d_{\mathrm{hi}} - \widehat{d}_{\mathrm{hi}})^2}{[r]\sum_{h,i}{d^2}_{\mathrm{hi}}}} \tag{3}$$

where d_{hi} is the ordinated distance between pixels h and i, and $\widehat{d}_{\mathrm{hi}}$ is the distance predicted from the regression. Then, a new configuration is computed moving in the direction in which stress changes most rapidly. The entire procedure is repeated until convergence. A *Stress* value that provides an excellent representation in the reduced dimensions is considered to be lower than 0.05; nevertheless a value of $Stress < 0.2$ is still considered a good representation Borg and Groenen (2005).

Basically, the algorithm provides both single spectral species maps and the α- and β-diversity maps. The algorithm input file needs to be in ENVI binary format with the corresponding header file. The file should be in Band Interleave by Line (BIL) format and 2-byte signed integer, and should not have extension. A further masking file in the same format is necessary in order to mask clouds and water surfaces.

Box 2 - Packages used in this manuscript to handle and analyse spatial data in R

- `raster`: It provides classes and functions to manipulate geographic data in raster format. Raster data divides space into cells (as pixels) of equal size (in units of the coordinate reference system). Along with the `raster` package, the `sp` package is also loaded, which provides spatial object classes and methods to retrieve coordinates.
- `rgdal`: It provides functions to import ad export spatial data in different formats.
- `RStoolbox`: A toolbox for remote sensing image processing and analysis.