



HAL
open science

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.

► To cite this version:

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-03122931>

Submitted on 3 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

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

4 Duccio Rocchini^{1,2*}, Nicole Salvatori^{3,4}, Carl Beierkuhnlein⁵,
Alessandro Chiarucci¹, Florian de Boissieu⁶, Michael Förster⁷,
Carol X. Garzon-Lopez⁸, Thomas W. Gillespie⁹, Heidi C. Hauffe¹⁰,
Kate S. He¹¹, David Kleijn¹², Birgit Kleinschmit⁷,
Jonathan Lenoir¹³, Marco Malavasi², Vítězslav Moudrý²,
Harini Nagendra¹⁴, Davnah Payne¹⁵, Petra Šímová²,
Michele Torresani¹⁶, Martin Wegmann¹⁷, Jean-Baptiste Féret⁶

5 October 3, 2020

6 ¹ Alma Mater Studiorum University of Bologna, Department of Biologi-
7 cal, Geological and Environmental Sciences, via Irnerio 42, 40126, Bologna,
8 Italy

9 ² Czech University of Life Sciences Prague, Faculty of Environmental Sci-
10 ences, Department of Applied Geoinformatics and Spatial Planning, Kamýcka
11 129, Praha - Suchbátka, 16500, Czech Republic

12 ³ University of Udine, Department of Agri-Food, Animal and Environ-
13 mental Sciences (DI4A), via delle scienze 206, 33100 Udine (UD), Italy

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

16 ⁵ Biogeography, BayCEER, University of Bayreuth, Universitätsstraße
17 30, 95440 Bayreuth, Germany

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

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

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

24 ⁹ Department of Geography, University of California Los Angeles, Los
25 Angeles, California 90095-1524, USA

26 ¹⁰ Fondazione Edmund Mach, Research and Innovation Centre, Depart-
27 ment of Biodiversity and Molecular Ecology, Via E. Mach 1, 38010 S. Michele
28 all'Adige (TN), Italy

29 ¹¹ Department of Biological Sciences, Murray State University, Murray,
30 KY 42071, USA

31 ¹² Plant Ecology and Nature Conservation Group, Wageningen Univer-
32 sity, Droevendaalsesteeg 3a, 6708 PB Wageningen, The Netherlands

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

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

38 ¹⁵ University of Bern, Institute of Plant Sciences, GMBA office Altenber-
39 grain 21, 3013 Bern, Switzerland

40 ¹⁶ Free University of Bolzano/Bozen, Faculty of Science and Technology,
41 Piazza Università/Universitätsplatz 1, 39100, Bolzano, Bozen, Italy

42 ¹⁷ Department of Remote Sensing, Remote Sensing and Biodiversity Re-
43 search Group, University of Wuerzburg, Wuerzburg, Germany

44 **Abstract**

45 In the light of unprecedented change in global biodiversity, real-
46 time and accurate ecosystem and biodiversity assessment are becom-
47 ing increasingly essential. Nevertheless, estimation of biodiversity us-
48 ing ecological field data can be difficult for several reasons. In partic-
49 ular, for larger extents, it is challenging to collect data to provide reli-
50 able information. Some of these restrictions in Earth observation can
51 be avoided through the use of remote sensing approaches. Different
52 studies have estimated biodiversity on the basis of the Spectral Varia-
53 tion Hypothesis (SVH). According to this hypothesis, spectral hetero-
54 geneity over the different pixel units of a spatial grid reflects a higher
55 niche heterogeneity, allowing more organisms to coexist. Recently, the
56 spectral species concept has been derived, following the consideration
57 that the spectral heterogeneity at a landscape scale corresponds to a
58 combination of subspaces sharing a similar spectral signature. At a
59 local scale, with the use of high resolution remote sensing data, the
60 different subspaces can be identified as different spectral entities, the
61 so called "spectral species". Our approach extends this concept over
62 wide spatial extents and to a higher level of biological organisation.

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

70 *Keywords:* biodiversity; ecological informatics; modelling; remote sens-
71 ing; satellite imagery.

72 1 Introduction

73 1.1 A quest for robust and reproducible α - and β - 74 diversity measurement

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

82 A diatribe is still open about what are the most reliable metrics to as-
83 sess biodiversity (see Jurasinski et al. (2009); Tuomisto (2010)). Until now,
84 no consistent definition exists and also the definition of the CBD (1992,
85 <https://www.cbd.int/convention/text/>) is more confusing than clear:
86 “Biological Diversity means the variability among living organisms from all
87 sources, including, inter alia, terrestrial, marine and other aquatic ecosystems
88 and the ecological complexes of which they are part; this includes diversity
89 within species, between species and of ecosystems.” Biodiversity obviously
90 includes quantitative (number of species, alpha-diversity, gamma-diversity),
91 qualitative (turnover, composition, beta-diversity) and functional (complex-
92 ity, trophic levels, ecosystem services) aspects. To sum up our understanding
93 on the term biodiversity (i.e. biological diversity) and to base our study on a
94 more general and consistent concept, “biodiversity characterizes qualitative,
95 quantitative and functional aspects of biotic units at various levels of orga-
96 nization in a concrete or abstract context, and at a given temporal or/and
97 spatial scale” (Beierkuhnlein, 2003). In consequence, species richness and
98 metrics that are based on it are very important, but they represent just one

99 aspect of biodiversity. The total number of species co-occurring in a given
100 community (α -diversity) is nested within the total number of a species pools
101 occurring for instance at the landscape level (γ -diversity). But the reduction
102 of biodiversity to the perspective of inventory and proportion would not
103 cover spatial gradients in composition and species turnover (differentiation,
104 β -diversity) (Jurasinski et al., 2009; Baselga, 2012) and also ignores functional
105 diversity (e.g. functional traits), which is the main driver of ecosystem
106 functioning.

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

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

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

132 In some cases, spatial turnover can result from local extinction processes
133 that hit certain species more than others and enhance the dissimilarity between
134 sites without dispersal (Steinitz et al., 2006). This is the case in
135 highly fragmented landscapes where dispersal is limited (Hobbs et al., 2006).
136 Even stochastic processes (sensu Moran (1950) and Clark (2008)) may enhance
137 β -diversity in previously homogeneous ecosystems. For instance, sudden
138 fragmentation (Alados et al., 2009) can lead to dysfunctional source-sink
139 metapopulations with intrinsic influences on the degree of spatial (and ge-

140 netic) connectivity of organisms (Waples and Gaggiotti, 2006), resulting in
141 the local loss of sink populations. However, in most situations, the spatial
142 turnover and therefore the dispersal of species between sites (metapopula-
143 tion and metacommunity dynamics) is linked to the distance among sites.
144 [Strictly speaking, the similarity between two sites decays with increasing dis-
145 tance between them](#) (Rocchini, 2007), a process also known as the distance
146 decay in similarity or the Tobler’s first law of geography (Tobler, 1970).

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

153 In general, an enormous effort from ecologists is necessary in order to
154 collect reliable data on biodiversity. A pioneering example is the public
155 database of the Global Biodiversity Information Facility (GBIF, <https://www.gbif.org/>). GBIF is a network funded by the world’s governments
156 which contains almost 41.000 databases of species occurrences spread out
157 across the world. The huge amount of data accessible and the different tech-
158 niques now available to analyze them should facilitate biodiversity assessment
159 and monitoring. Unfortunately, although it would be possible, in principle,
160 to use these data to make reasonable assumptions on the biodiversity at wide
161 spatial extents, there are several limitations due to their quality (Maldonado
162 et al., 2015). The errors that usually rise are due to: (i) poor quality in get-
163 ting exact geographic coordinates of the sampling sites; (ii) wrong taxonomic
164 identification with poor quality control; and (iii) difficulties in proving a re-
165 liable random sampling with large areas being poorly covered. Furthermore,
166 these data appear as point data and usually grids are used in order to syn-
167 thesize diversity metrics. These data mostly come from presence-only data
168 without any link to relative abundance, dominance, biomass or cover, which
169 would be reflected in remote sensing. Finally, GBIF data are [inadequate](#) for
170 local estimates of biodiversity as they do not consider co-occurrence data.
171 Indeed, and contrary to recent databases at the community level such as
172 the European Vegetation Archive (EVA) (Chytry et al., 2016) or the sPlot
173 initiative (Bruehlheide et al., 2019), GBIF does not provide information on
174 species co-occurrence which is very problematic for biodiversity assessment
175 and monitoring. Despite the disadvantages that come from the use of public
176 databases, there is some benefits in the use of such data. First of all, there
177 is a huge amount of data collected and provided by citizens and research
178 institutions available in the GBIF database when compared to the data that
179 could be collected locally, resulting in a huge saving of time and costs. More-
180

181 over, GBIF data are standardized to the same format and therefore ready to
182 use.

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

187 **1.2 The spectral species concept**

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

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

203 In general, vegetation absorbs the blue and the red light, for photosyn-
204 thesis, while it reflects NIR radiation due to the physical structure of the
205 cells composing the leaf mesophyllum (Wegmann et al., 2016). The bands
206 relative to RED and NIR are used as proxies for photosynthetic activity of
207 the vegetation. These bands are usually incorporated in a widely used in-
208 dex, the normalized difference vegetation index (NDVI), which is calculated
209 as $NDVI = (NIR - RED) / (NIR + RED)$. The higher the relative abundance of
210 photosynthetic vegetation, the higher would be the reflectance in the NIR
211 band and the absorption in the RED band. NDVI ranges from -1 to 1, with
212 0 values usually associated with non vegetated areas and negative values
213 associated with water surfaces or snow.

214 This index has widely been used to discriminate different vegetation types
215 over an area. In fact, in several studies, NDVI is positively correlated to the
216 net primary productivity (NPP, e.g. Gillespie et al. (2008)). Therefore, it
217 can be used as a proxy to quantify species richness and diversity, based on the
218 species-energy theory, proposed by Currie (1991), namely a relation between
219 species richness and energy, that would depend mainly on annual potenti-

220 afl evapotranspiration and actual evapotranspiration. Another hypothesis
221 related to the variability in space of the spectral signal has been proposed
222 by Palmer et al. (2002). The so called spectral variation hypothesis (SVH)
223 states that the higher the environmental heterogeneity the higher would be
224 the species diversity of an area, due to a higher amount of ecological niches
225 available.

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

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

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

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

260 **2 The algorithm**

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

$$H' = -\sum_{s=1}^N p_s \ln p_s \quad (1)$$

280 where p_s is the proportion of each spectral species s in each plot.

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

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

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

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

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

304 2.1 Application of the algorithm

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

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

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

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

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

358 **3 Discussion**

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

365 With regards to a potential validation with in-situ data, the uncertainty
366 of wide-scale datasets hampers a spatial overlap. In this case, in-situ datasets
367 meet all five major concerns recently raised by Hobohm et al. (2019), i.e.: i)
368 there is insufficient data coverage across Europe to make an unbiased com-

369 parison between predicted and actual distributions, ii) taxonomic standards
370 differ across sampled regions, iii) there are generally different shapes of areas
371 being sampled, iv) political borders often define sampling areas and aggregated
372 sampling areas, and v) data are not aggregated in the same way in
373 all areas. Furthermore, spatial information has an intrinsic varying degree
374 of relevance mainly due to the fact that, rather than species lists, it is made
375 up of geometrical precision, attributes robustness and temporal consistency
376 (Hobona et al., 2006). Finally, different models and approaches to measuring
377 diversity inevitably provide different outputs, as pointed out in the gener-
378 alised entropy theory put forward by Rényi (1961). Given the above valida-
379 tion difficulties, we decided to qualitatively compare our generated output,
380 in particular the β -diversity map, with existing ecoregion maps, which are
381 expected to discriminate different spatial areas based on natural borders de-
382 fined by biological diversity (<https://ecoregions2017.appspot.com/>) and thus
383 are intrinsically related to differences in the species and spectral turnover of
384 communities.

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

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

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

409 Concerning climate, a solution might be found in the use of remotely

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

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

444 From an ecological perspective, remote sensing imagery bands (dimen-
445 sions) show a high affinity with the hypervolume axes proposed by Hutchin-
446 son (1957) for modelling species niches. In the Hutchinson's theory, an hy-
447 pervolume is represented by a space defined by a set of n independent axes
448 which could be related to the final variables driving the realised niche of a
449 species (see also Blonder (2017) and Ricotta et al. (2010) on the niche dif-
450 ferentiation concept). In our case, such axes would be the original satellite

451 sensor bands being strictly related to the identification of a spectral species
452 and the resulting spectral community in a site, instead of a niche. From this
453 point of view, spectral species and communities are in line with joint species
454 distribution models (JSDMs), which explicitly take into account biotic inter-
455 actions among species in a community, while in our model the “interaction”
456 among pixel values is ruled out in general by their proximity both from a
457 spatial and from a spectral point of view. In this paper, the final aim was
458 not to model single spectral species or spectral communities but rather to
459 estimate diversity and its change over space and time, following the mathe-
460 matical principles described in Liu et al. (2014) and Rocchini et al. (2015b),
461 for which the distribution of diversity over space is actually a particular case
462 of the so-called switched systems, i.e. hybrid systems resulting from both
463 continuous and discrete dynamics with a high number of different potential
464 variables acting as main drivers of diversity response. In practice, in our
465 view we succeeded to fill a previous gap in spatio-ecological analysis, i.e. the
466 translation of what in remote sensing science is known as “spectral mixture
467 modeling” (Jensen, 2015) into an ecological diversity theory perspective. In
468 fact, in spectral mixture modeling the measured spectral reflectance is de-
469 composed as a mixture of endmembers. In our case, such mixture was used
470 to directly compute alpha- and overall beta-diversity over wide spatial areas
471 in few time.

472 4 Conclusion

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

485 Based on the results presented here, the use of the spectral species and
486 communities concept would appear to promote more effective planning and
487 policies related to the conservation of wild species, by improving our un-
488 derstanding of the dynamics of local and global biodiversity at different spatial

489 and temporal scales.

490 **5 Acknowledgements**

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

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

500 **References**

- 501 Alados, C.L., Pueyo, Y., Escos, J., Andujar, A. (2009). Effects of the spatial
502 pattern of disturbance on the patch-occupancy dynamics of juniper–pine
503 open woodland. *Ecological Modelling*, 220: 1544-1550.
- 504 Asner, G.P. and Martin, R.E. (2009). Airborne spectranomics: mapping
505 canopy chemical and taxonomic diversity in tropical forests. *Frontiers in
506 Ecology and the Environment*, 7: 269-276.
- 507 Baselga, A. (2012). The relationship between species replacement, dissimilar-
508 ity derived from nestedness, and nestedness. *Global Ecology and Biogeog-
509 raphy* 21: 1223-1232.
- 510 Beierkuhnlein, C. (2003). Der Begriff Biodiversität. *Nova Acta Leopoldina
511 N.F.*, 87: 51-72.
- 512 Blonder, B. (2017). Hypervolume concepts in niche- and trait-based ecology.
513 *Ecography* 41: 1441-1445.
- 514 Borg, I., Groenen, P. (2005). *Modern Multidimensional Scaling: theory and
515 applications* (2nd edition). Springer-Verlag, New York.
- 516 Bray, J.R. and Curtis J.T. (1957). An ordination of the upland forest com-
517 munities of southern Wisconsin. *Ecological monographs* 27: 325-349.
- 518 Bruelheide, H., Dengler, J., Jimenez-Alfaro, B., Purschke, O., Hennekens,
519 S.M., Chytrý, M., Pillar, V.D., Jansen, F., Kattge, J., Sandel, B., Aubin,
520 I., Biurrun, I., Field, R., Haider, S., Jandt, U., Lenoir, J., Peet, R.K.,
521 Peyre, G., Sabatini, F.M., Schmidt, M., Schrodte, F., Winter, M., Acic,
522 S., Agrillo, E., Alvarez, M., Ambarlı, D., Angelini, P., Apostolova, I.,
523 Arfin Khan, M.A.S., Arnst, E., Attorre, F., Baraloto, C., Beckmann, M.,
524 Berg, C., Bergeron, Y., Bergmeier, E., Bjorkman, A.D., Bondareva, V.,
525 Borchardt, P., Botta-Dukat, Z., Boyle, B., Breen, A., Brisse, H., Byun,
526 C., Cabido, M.R., Casella, L., Cayuela, L., Cerny, T., Chepinoga, V.,
527 Csiky, J., Curran, M., Custerovska, R., Dajic, Z., De Bie, S.E., de Ruffray,
528 P., De Sanctis, M., Dimopoulos, P., Dressler, S., Ejrnaes, R., El-Rouf,
529 M.A., El-Sheikh, M., Enquist, B., Ewald, J., Fagundes, J., Finckh, M.,
530 Font, X., Forey, E., Fotiadis, G., Garcia-Mijangos, I., Luis de Gasper,
531 A., Golub, V., Gutierrez, A.G., Hatim, M.Z., He, T., Higuchi, P., Hol-
532 ubova, D., Holzner, N., Homeier, J., Indreica, A., Isik Gursoy, D., Jansen,
533 S., Janssen, J., Jedrzejek, B., Jirousek, M., Jurgens, N., Kacki, Z., Kav-
534 gacı, A., Kearsley, E., Kessler, M., Knollova, I., Kolomiychuk, V., Ko-
535 rolyuk, A., Kozhevnikova, M., Kozub, L., Krstonosic, D., Kuhl, H., Kuhn,

536 I., Kuzemko, A., Kuzmic, F., Landucci, F., Lee, M.T., Levesley, A., Li,
537 C.-F., Liu, H., Lopez-Gonzalez, G., Lysenko, T., Macanovic, A., Mahdavi,
538 P., Manning, P., Marceno, C., Martynenko, V., Mencuccini, M., Minden,
539 V., Moeslund, J.E., Moretti, M., Muller, J.V., Munzinger, J., Niinemets,
540 U., Nobis, M., Noroozi, J., Nowak, A., Onyshchenko, V., Overbeck, G.E.,
541 Ozinga, W.A., Pauchard, A., Pedashenko, H., Penuelas, J., Perez-Haase,
542 A., Peterka, T., Pettrk, P., Phillips, O.L., Prokhorov, V., Rasomavicius,
543 V., Revermann, R., Rodwell, J., Ruprecht, E., Rusina, S., Samimi, C.,
544 Schaminee, J.H.J., Schmiedel, U., Sibik, J., Silc, U., Skvorc, Z., Smyth, A.,
545 Sop, T., Sopotlieva, D., Sparrow, B., Stancic, Z., Svenning, J.-C., Swacha,
546 G., Tang, Z., Tsiripidis, I., Turtureanu, P.D., Ugurlu, E., Uogintas, D.,
547 Valachovic, M., Andre, K., Vashenyak, V.Y., Vassilev, K., Velez-Martin,
548 E., Venanzoni, R., Vibrans, A.C., Violle, C., Virtanen, R., von Wehrden,
549 H., Wagner, V., Walker, D.A., Wana, D., Weiher, E., Wesche, K., Whit-
550 field, T., Willner, W., Wiser, S., Wohlgemuth, T., Yamalov, S., Zizka,
551 G., Zverev, A. (2019). sPlot – a new tool for global vegetation analyses.
552 *Journal of Vegetation Science*, 30: 161-186.

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

- 576 Cord, A., Rödder, D. (2011). Inclusion of habitat availability in species dis-
577 tribution models through multi-temporal remote-sensing data? Ecological
578 Applications, 21: 3285-3298.
- 579 Clark, J.S. (2008). Beyond neutral science. Trends in Ecology and Evolution,
580 24: 8-15.
- 581 Currie, D.J. (1991). Energy and large-scale patterns of animal-and plant-
582 species richness. The American Naturalist 137: 27-49.
- 583 Dinerstein, E., Olson, D., Joshi, A., Vynne, C., Burgess, N.D., Wikra-
584 manayake, E., Hahn, N., Palminteri, S., Hedao, P., Noss, R., Hansen,
585 M., Locke, H., Ellis, E.C., Jones, B., Barber, C.V., Hayes, R., Kormos,
586 C., Martin, V., Crist, E., Sechrest, W., Price, L., Baillie, J.E.M., Wee-
587 den, D., Suckling, K., Davis, C., Sizer, N., Moore, R., Thau, D., Birch,
588 T., Potapov, P., Turubanova, S., Tyukavina, A., de Souza, N., Pintea, L.,
589 Brito, J.C., Llewellyn, O.A., Miller, A.G., Patzelt, A., Ghazanfar, S.A.,
590 Timberlake, J., Kloser, H., Shennan-Farpon, Y., Kindt, R., Lilleso, J.B.,
591 van Breugel, P., Graudal, L., Voge, M., Al-Shammari, K.F., Saleem, M.A.
592 (2017). Ecoregion-based approach to protecting half the terrestrial realm.
593 Bioscience, 67: 534-545.
- 594 Elton, C.S. (1958). The Ecology of invasions by animals and plants. Methuen,
595 London.
- 596 Ewald., M., Skowronek, S., Aerts, R., Dolos, K., Lenoir, J., Nicolas, M.,
597 Warrie, J., Hattab, T., Feilhauer, H., Honnay, O., Garzon-Lopez, C.X.,
598 Decocq, G., Van De Kerchove, R., Somers, B., Rocchini, D., Schmidlein,
599 S. (2018). Analyzing remotely sensed structural and chemical canopy traits
600 of a forest invaded by *Prunus serotina* over multiple spatial scales. Biologi-
601 cal Invasions, 20: 2257-2271.
- 602 Féret, J.-B., Asner, G.P. (2014). Mapping tropical forest canopy diversity
603 using high-fidelity imaging spectroscopy. Ecological Applications 24: 1289-
604 1296.
- 605 Féret, J.-B., de Boissieu, F. (in press). **biodivMapR**: An R package for α - and
606 β -diversity mapping using remotely sensed images. Methods in Ecology
607 and Evolution.
- 608 Ferrier, S., Manion, G., Elith, J. and Richardson, K. (2007). Using gen-
609 eralized dissimilarity modelling to analyse and predict patterns of beta
610 diversity in regional biodiversity assessment. Diversity and Distributions
611 13: 252-264.

- 612 Foody, G.M. (2008). GIS: biodiversity applications. *Progress in Physical Ge-*
613 *ography* 32: 223-235.
- 614 Gaston, K.J. (2000). Global patterns in biodiversity. *Nature* 405: 220.
- 615 Gillespie, T.W., Foody, G.M., Rocchini, D., Giorgi, A.P. and Saatchi, S.
616 (2008). Measuring and modelling biodiversity from space. *Progress in Phys-*
617 *ical Geography* 32: 203-221.
- 618 Hernandez-Stefanoni, J.L., Gallardo-Cruz, J.A., Meave, J.A., Rocchini, D.,
619 Bello-Pineda, J., Lopez-Martinez, J.O. (2012). Modeling alpha- and beta-
620 diversity in a tropical forest from remotely sensed and spatial data. *Inter-*
621 *national Journal of Applied Earth Observation and Geoinformation*, 19:
622 359-368.
- 623 Hobbs, R.J., Arico, S., Aronson, J., Baron, J.S., Bridgewater, P., Cramer,
624 V.A., Epstein, P.R., Ewel, J.J., Klink, C.A., Lugo, A.E., Norton, D.,
625 Ojima, D., Richardson, D.M., Sanderson, E.W., Valladares, F., Vila, M.,
626 Zamora, R., Zobel, M. (2006). Novel ecosystems: theoretical and man-
627 agement aspects of the new ecological world order. *Global Ecology and*
628 *Biogeography*, 15: 1-7.
- 629 Hobohm, C., Janisova, M., Steinbauer, M., Landi, S., Field, R., Vander-
630 plank, S., Beierkuhnlein, C., Grytnes, J.-A., Vetaas, R.O., Fidelis, A., de
631 Nascimento, L., Clark, V.P., Fernandez-Palacios, J.M., Franklin, S., Guar-
632 ino, R., Huang, J., Krestov, P., Ma, K., Onipchenko, V., Palmer, M.W.,
633 Fragomeni Simon, M., Stolz, C., Chiarucci, A. (2019). Global endemics-
634 area relationships of vascular plants. *Perspectives in Ecology and Conser-*
635 *vation*, 17, 41-49.
- 636 Hobona, G., James, P., Fairbairn, D. (2006). Multidimensional visualisation
637 of degrees of relevance of geographic data. *International Journal of Geo-*
638 *graphical Information Science*. 20: 469-490.
- 639 Hutchinson, G. (1957). Concluding remarks. *Cold Spring Harb. Symp.*
640 *Quant. Biol.* 22: 415-427.
- 641 Jurasinski, G., Retzer, V., Beierkuhnlein, C. (2009). Inventory, differentia-
642 tion, and proportional diversity: a consistent terminology for quantifying
643 species diversity. *Oecologia*, 159: 15-26.
- 644 [Jensen, R.J. \(2015\). *Introductory Digital Image Processing: A Remote Sens-*](#)
645 [ing Perspective. Pearson Publishing.](#)

- 646 Kohonen, T. (1982). Analysis of a simple self-organizing process. *Biological*
647 *cybernetics* 44: 135-140.
- 648 Lande, R. (1996). Statistics and partitioning of species diversity, and simi-
649 larity among multiple communities. *Oikos* 76: 5-13.
- 650 Legendre, P., Borcard, D., Peres-Neto, P.R. (2005). Analyzing beta diver-
651 sity: partitioning the spatial variation of community composition data.
652 *Ecological Monographs* 75: 435-450.
- 653 Liu, X., Zhong, S., Ding, X. (2014). A Razumikhin approach to exponential
654 admissibility of switched descriptor delayed systems. *Applied Mathemati-*
655 *cal Modelling* 38, 1647–1659.
- 656 Maldonado, C., Molina, C.I., Zizka, A., Persson, C., Taylor, C.M., Albán, J.,
657 Chilquillo, E., Ronsted, N. and Antonelli, A. (2015). Estimating species
658 diversity and distribution in the era of Big Data: to what extent can we
659 trust public databases?. *Global Ecology and Biogeography* 24: 973-984.
- 660 Mead, A. (1992). [Review of the Development of Multidimensional Scaling](#)
661 [Methods. *The Statistician*, 41: 27-39.](#)
- 662 Moran, P.A.P. (1950). Notes on continuous stochastic phenomena.
663 *Biometrika*, 37: 17-23.
- 664 Mouchet, M., Levers, C., Zupan, L., Kuemmerle, T., Plutzer, C., Erb, K.,
665 Lavorel, S., Thuiller, W., Haberl, H. (2015). Testing the effectiveness of
666 environmental variables to explain European terrestrial vertebrate species
667 richness across biogeographical scales. *PLoS ONE*, 10: e0131924.
- 668 Mucher, C.A., Hennekens, S.M., Bunce, R.G.H., Schaminee, J.H.J., Schaep-
669 man, M.E. (2009). Modelling the spatial distribution of Natura 2000 habi-
670 tats across Europe. *Landscape and Urban Planning*, 92: 148-159.
- 671 Palmer, M.W., Earls, P.G., Hoagland, B.W., White, P.S. and Wohlgemuth,
672 T. (2002). Quantitative tools for perfecting species lists. *Environmetrics*,
673 13: 121-137.
- 674 Peuquet, D.J. (1992). An algorithm for calculating minimum euclidean dis-
675 tance between two geographic features. *Computers & Geosciences* 18: 989-
676 1001.
- 677 R Development Core Team (2019). R: A Language and Environment for
678 Statistical Computing. R Foundation for Statistical Computing, Vienna,
679 Austria. <https://www.R-project.org/>

- 680 Rényi, A. (1961). On measures of information and entropy. Proceedings of
681 the fourth Berkeley Symposium on Mathematics, Statistics and Probability
682 1960: 547-561.
- 683 Ricotta, C., Godefroid, S., Rocchini, D. (2010). Patterns of native and exotic
684 species richness in the urban flora of Brussels: rejecting the 'rich get richer'
685 model. *Biological Invasions*, 12: 233-240.
- 686 Rocchini, D. (2007). Distance decay in spectral space in analysing ecosystem
687 β -diversity. *International Journal of Remote Sensing*, 28: 2635-2644.
- 688 Rocchini, D., Andreo, V., Förster, M., Garzon-Lopez, C.X., Gutierrez, A.P.,
689 Gillespie, T.W., Hauffe, H.C., He, K.S., Kleinschmit, B., Mairota, P., Mar-
690 cantonio, M., Metz, M., Nagendra, H., Pareeth, S., Ponti, L., Ricotta, C.,
691 Rizzoli, A., Schaab, G., Zebisch, M., Zorer, R., Neteler, M. (2015a). Potential
692 of remote sensing to predict species invasions: A modelling perspective.
693 *Progress in Physical Geography*, 39: 283-309.
- 694 Rocchini, D., Balkenhol, N., Carter, G.A., Foody, G.M. Gillespie, T.W., He,
695 K.S., Kark, S., Levin, N., Lucas, K., Luoto, M., Nagendra, H., Oldeland,
696 J., Ricotta, C., Southworth, J. and Neteler, M. (2010). Remotely sensed
697 spectral heterogeneity as a proxy of species diversity: recent advances and
698 open challenges. *Ecological Informatics*, 5: 318-329.
- 699 Rocchini, D., Delucchi, L., Bacaro, G., Cavallini, P., Feilhauer, H., Foody,
700 G.M., He, K.S., Nagendra, H., Porta, C., Ricotta, C., Schmidtlein, S.,
701 Spano, L.D., Wegmann, M., Neteler, M. (2013). Calculating landscape
702 diversity with information-theory based indices: A GRASS GIS solution.
703 *Ecological Informatics*, 17: 82-93.
- 704 Rocchini, D., Luque, S., Pettorelli, N., Bastin, L., Doktor, D., Faedi, N.,
705 Feilhauer, H., Féret, J.B., Foody, G.M., Gavish, Y., Godinho, S., Kunin,
706 W.E., Lausch, A., Leitao, P.J., Marcantonio, M., Neteler, M., Ricotta,
707 C., Schmidtlein, S., Vihervaara, P., Wegmann, M., Nagendra, H. (2018).
708 β -diversity by remote sensing: A challenge for biodiversity monitoring.
709 *Methods in Ecology and Evolution* 9: 1787-1798.
- 710 Rocchini, D., Hernandez Stefanoni, J.L., He, K.S. (2015b). Advancing species
711 diversity estimate by remotely sensed proxies: a conceptual review. *Eco-
712 logical Informatics*, 25: 22-28.
- 713 Rocchini, D., McGlenn, D., Ricotta, C., Neteler, M., Wohlgemuth, T. (2011).
714 Landscape complexity and spatial scale influence the relationship between

- 715 remotely sensed spectral diversity and survey based plant species richness.
716 *Journal of Vegetation Science*, 22: 688-698.
- 717 Rocchini, D., Marcantonio, M., Da Re, D., Chirici, G., Galluzzi, M., Lenoir,
718 J., Ricotta, C., Torresani, M., Ziv, G. (2019). Time-lapsing biodiversity:
719 an open source method for measuring diversity changes by remote sensing.
720 *Remote Sensing of Environment*, 231: 111192.
- 721 Rocchini, D., Ricotta, C. (2007). Are landscapes as crisp as we may think?
722 *Ecological Modelling*, 204: 535-539.
- 723 Skowronek, S., Asner, G.P., Feilhauer, H. (2017). Performance of one-
724 class classifiers for invasive species mapping using airborne imaging spec-
725 troscopy. *Ecological Informatics*, 37: 66-76.
- 726 Skowronek, S., Ewald, M., Isermann, M., Van De Kerchove, R., Lenoir,
727 J., Aerts, R., Warrie, J., Hattab, T., Honnay, O., Schmidtlein, S., Roc-
728 chini, D., Somers, B., Feilhauer, H. (2017). Mapping an invasive bryophyte
729 species using hyperspectral remote sensing data. *Biological Invasions*, 19:
730 239-254.
- 731 Shannon, C.E. (1948). A mathematical theory of communication. *ACM SIG-*
732 *MOBILE mobile computing and communications review* 5: 3-55.
- 733 Steinitz, O., Heller, J., Tsoar, A., Rotem, D., Kadmon, R. (2006). Environ-
734 ment, dispersal and patterns of species similarity. *Journal of Biogeography*:
735 33, 1044-1054.
- 736 Sun, L., Mi, X., Wei, J., Wang, J., Tian, X., Yu, H. and Gan, P. (2017). A
737 cloud detection algorithm-generating method for remote sensing data at
738 visible to short-wave infrared wavelengths. *ISPRS Journal of Photogram-*
739 *metry and Remote Sensing* 124: 70-88.
- 740 Tobler, W., (1970). A computer movie simulating urban growth in the Detroit
741 region. *Economic Geography*, 46: 234-240.
- 742 Tuanmu, M.N., Jetz, W. (2015). A global, remote sensing-based character-
743 ization of terrestrial habitat heterogeneity for biodiversity and ecosystem
744 modelling. *Global Ecology and Biogeography*, 24: 1329-1339.
- 745 Tuomisto, H. (2010). A consistent terminology for quantifying species diver-
746 sity? Yes, it does exist. *Oecologia* 164: 853-860.
- 747 Ustin, S.L. and Gamon, J.A. (2010). Remote sensing of plant functional
748 types. *New Phytologist* 186.4: 795-816.

- 749 Vihervaara, P., Auvinen, A.P., Mononen, L., Torma, M., Ahlroth, P., Anttila,
750 S., Bottcher, K., Forsius, M., Heino, J., Heliola, J., Koskelainen, M., Kuus-
751 saari, M., Meissner, K., Ojala, O., Tuominen, S., Viitasalo, M., Virkkala,
752 R. (2017). How Essential Biodiversity Variables and remote sensing can
753 help national biodiversity monitoring. *Global Ecology and Conservation*,
754 10: 43-59.
- 755 Waples, R., Gaggiotti, O. (2006). What is a population? An empirical eval-
756 uation of some genetic methods for identifying the number of gene pools
757 and their degree of connectivity. *Molecular Ecology*, 15: 1419-1439.
- 758 Wegmann, M., Leutner, B. and Dech, S. eds. Remote sensing and GIS for
759 ecologists: using open source software. (2016). Pelagic Publishing Ltd,
760 2016.
- 761 Whittaker, R.H. (1972). Evolution and measurement of species diversity.
762 *Taxon* 21: 213-251.
- 763 Zellweger, F., De Frenne, P., Lenoir, J., Rocchini, D., Coomes, D. (2019).
764 Advances in microclimate ecology arising from remote sensing. *Trends in*
765 *Ecology & Evolution*, 34: 327-341.

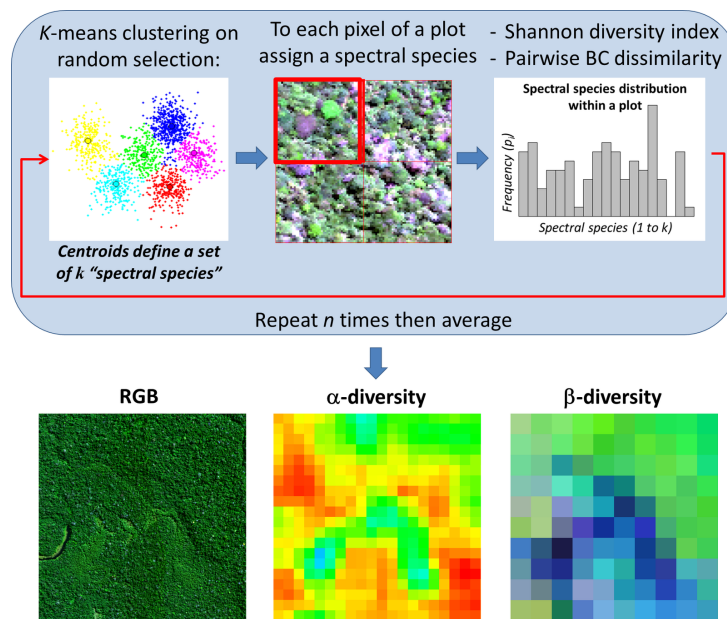


Figure 1: Diagrammatic representation of the steps of the algorithm used to achieve α - and β -diversities, redrawn from (Féret and Asner, 2014). Pixels are clumped in a spectral species and spectral community diversity is calculated. We refer to the main text and to Box 1 for additional information.

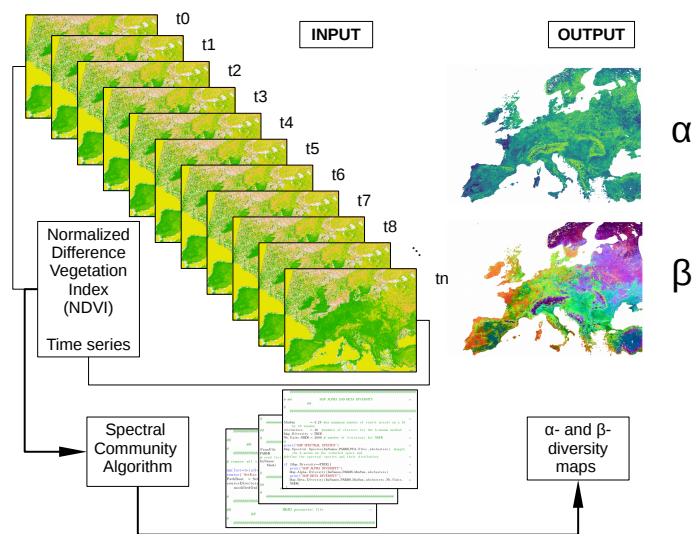


Figure 2: An input set of n images can be handled to create a time series and use the stack to further calculate the spectral community diversity. In our paper, a stack of 12 NDVI images of 2018 from the MODIS sensor was processed by the spectral species algorithm, by producing α - and β -diversity maps.

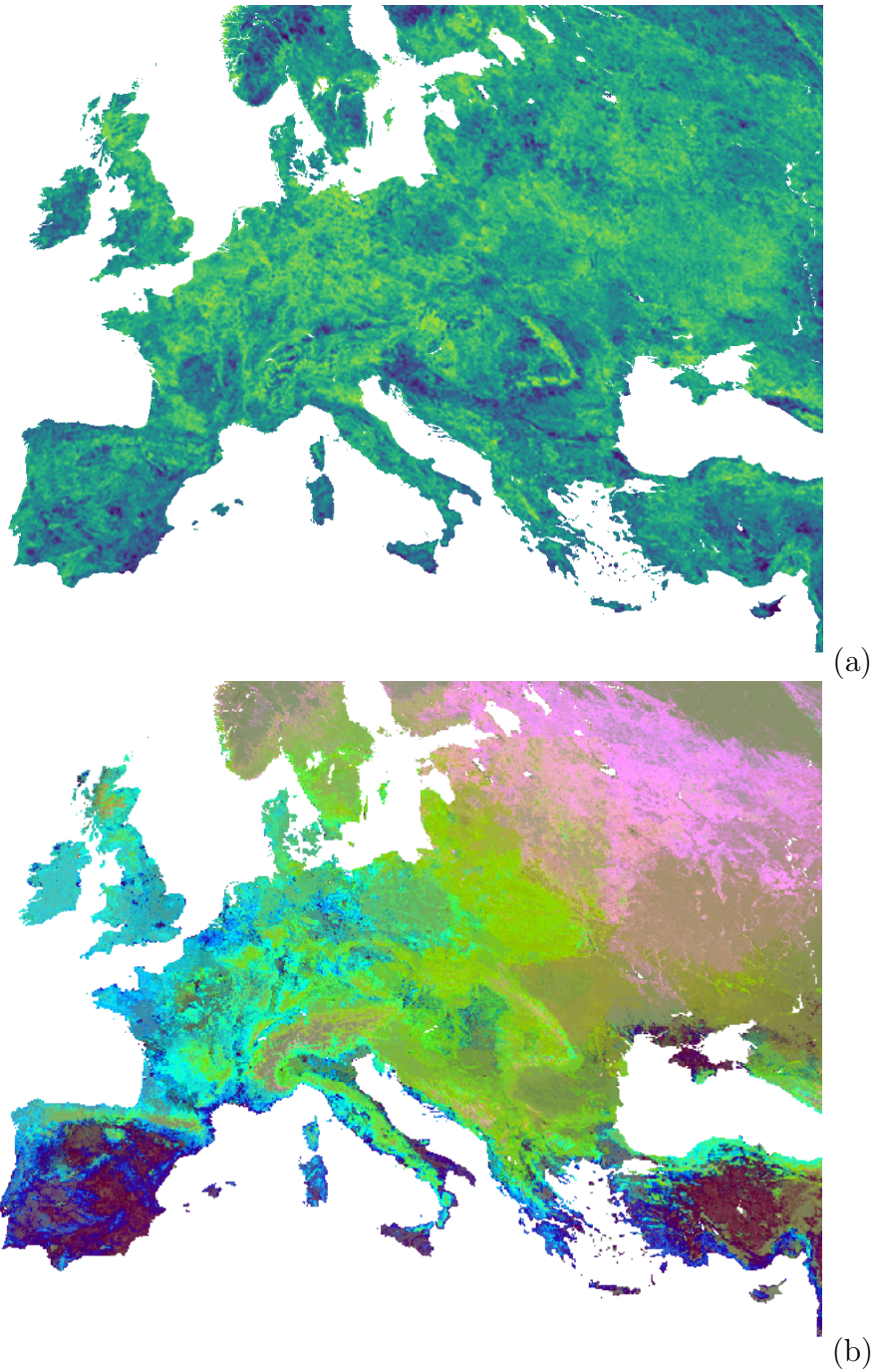


Figure 3: α - (a) and β -diversity (b) maps obtained by the spectral species algorithm. (a) The α -diversity map, based on Shannon's H' index (ranging from blue [low values] to light green [high values]) calculated in a 10x10 pixels local neighbourhood, corresponds to the local entropy of clusters, so 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 along any location in the image. Here, the distance between pairs of spatial units is expressed as a 3 colour code.

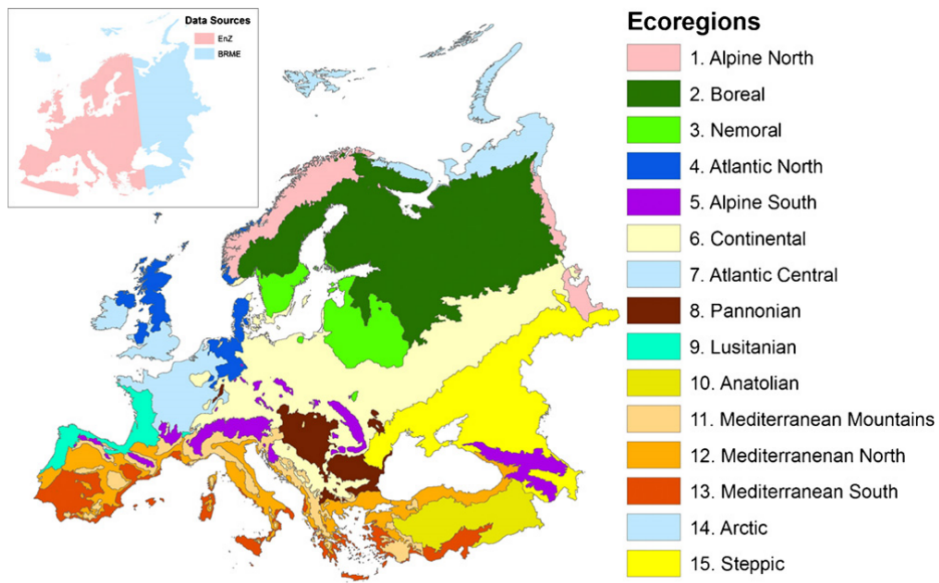


Figure 4: European Environmental Agency ecoregions map redrawn by Mucher et al. (2009). Similar maps at a coarser grain are provided by Mouchet et al. (2015) and Dinerstein et al. (2017).

767 **Box 1 - Steps composing the spectral species**
768 **algorithm**

- 769 1. A Principal Component Analysis (PCA) is applied to the spectral data.
770 PCA is not performed on the whole image, but only on a large subset
771 of pixels randomly selected from the image. Due to the high dimen-
772 sionality of the data, the reduction of the dataset is not altering the
773 result. Those principal components explaining most of the variance of
774 the original set are then retained for further steps.
- 775 2. A subset of pixels is then randomly selected across the entire map and
776 the spectral space containing such a subset is partitioned into spectral
777 species using k-means clustering with the number of k clusters being
778 decided a priori. Then the centroids defining the spectral species are
779 located.
- 780 3. The spectral dataset is divided into final mapping units. Each pixel
781 is assigned to a given spectral species based on the minimal Euclidean
782 distance between pixels (Peuquet, 1992) and the previously defined
783 centroids.
- 784 4. A spectral species distribution is obtained for each mapping unit from
785 which the α - and β -diversity indices are computed as previously stated.
- 786 5. Since the spectral species distribution is obtained by a subset of pixels,
787 in order to avoid under-representation of some small-scaled ecological
788 classes (e.g. small scale vegetation patterns), steps 4 and 5 are repeated
789 100 times, and the indicators obtained for each repetition are averaged.
790 In particular the Bray-Curtis dissimilarity matrix is computed for each
791 pair of spatial units, based on their spectral species distribution at each
792 iteration; then the final matrix corresponds to the BC dissimilarity
793 averaged over all the iterations.
- 794 6. Non metric Multidimensional Scaling (NMDS) (e.g. Borg and Groenen
795 (2005)) is applied to the matrices in order to obtain a visual representa-
796 tion of the results. NMDS is an ordination technique usually applied in
797 ecology that differs from other ordination techniques as PCA, since in
798 NMDS a small number of axes are chosen prior to the analysis and then
799 the data are fitted into the chosen dimensions. Furthermore, NMDS

800 is not an analytical but numerical technique, seeking for the right so-
801 lution (convergence) iteratively. Finally, NMDS is not an eigenvector-
802 eigenvalue technique, hence a NMDS ordination can be rotated among
803 the axes. NMDS is mostly used in ecology for its versatility since it
804 accepts any distance measure of the samples. In this case the Bray-
805 Curtis matrix was used. In the applied NMDS approach, the first step
806 is generally to decide the number of reduced dimensions; in this case
807 3 dimensions were chosen. The algorithm starts with the construction
808 of initial random arrangements of the pixels. Then the Euclidean dis-
809 tances among the samples is calculated in this first configuration; those
810 distances are regressed against the original distance matrix, and the
811 predicted ordination distances are calculated. Finally, the regression is
812 fitted by the least-squares method. The goodness of fit is measured by
813 the sum of squared differences between ordination-based distances and
814 the predicted distances. The goodness of fit is calculated through the
815 Kruskal's Stress index:

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

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

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

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

- 831 • **raster**: It provides classes and functions to manipulate geographic
832 data in raster format. Raster data divides space into cells (as pixels) of
833 equal size (in units of the coordinate reference system). Along with the
834 **raster** package, the **sp** package is also loaded, which provides spatial
835 object classes and methods to retrieve coordinates.
 - 836 • **rgdal**: It provides functions to import and export spatial data in differ-
837 ent formats.
 - 838 • **RStoolbox**: A toolbox for remote sensing image processing and analy-
839 sis.
-