The spectral species concept in living color

Duccio Rocchini[1,2], Maria J. Santos[3], Susan L. Ustin[4],
Jean-Baptiste Féret[5], Gregory P. Asner[6], Carl Beierkuhnlein[7],
Michele Dalponte[8], Hannes Feilhauer[9], Giles M. Foody[10], Gary
N. Geller[11], Thomas W. Gillespie[12], Kate S. He[13], David
Kleijn[14], Pedro J. Leitão[15,16], Marco Malavasi[2], Vítězslav
Moudrý[2], Jana Müllerová[17], Harini Nagendra[18], Signe
Normand[19,20], Carlo Ricotta[21], Michael E. Schaepman[22],
Sebastian Schmidtlein[23], Andrew K. Skidmore[24,25], Petra
Šímová[2], Michele Torresani[1], Philip A. Townsend[26], Woody
Turner[27], Petteri Vihervaara[28], Martin Wegmann[29], and
Jonathan Lenoir[30]

[1] *BIOME Lab, Department of Biological, Geological and Environmental Sciences, Alma Mater Studiorum University of Bologna, via Irnerio 42, 40126, Bologna, Italy*
[2] *Czech University of Life Sciences Prague, Faculty of Environmental Sciences, Department of Spatial Sciences, Kamýcka 129, Praha - Suchdol, 16500, Czech Republic*
[3] *University of Zurich, Department of Geography, Winterthurerstrasse 190, 8057 Zurich, Switzerland*
[4] *Department of Land, Air, and Water Resources, University of California Davis, Davis, CA 95616, USA*
[5] *UMR-TETIS, IRSTEA Montpellier, Maison de la Télédétection, 500 rue JF Breton, 34093, Montpellier Cedex 5, France*
[6] *Center for Global Discovery and Conservation Science, Arizona State University, Tempe, Arizona, USA*
[7] *Biogeography, BayCEER, University of Bayreuth, Universitaetsstraße 30, 95440 Bayreuth, Germany*
[8] *Sustainable Ecosystems and Bioresources Department, Research and Innovation Centre, Fondazione Edmund Mach, Via E. Mach 1, 38010 San Michele all'Adige (TN)*
[9] *Remote Sensing Center for Earth System Research, University of Leipzig, Talstr. 35, 04103 Leipzig, Germany*
[10] *School of Geography, University of Nottingham, University Park, Nottingham NG7 2RD, UK*
[11] *NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109 USA*
[12] *Department of Geography, University of California Los Angeles, Los Angeles, California 90095-1524, USA*

1

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

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

[15] *Department Landscape Ecology and Environmental System Analysis, Technische Universität Braunschweig, Braunschweig, Germany*

[16] *Geography Department, Humboldt-Universität zu Berlin, Berlin, Germany*

[17] *Department of GIS and Remote Sensing, Institute of Botany, The Czech Acad. Sciences, 252 43 Průhonice, Czech Republic*

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

[19] *Aarhus University, Department of Biology, Ecoinformatics and Biodiversity, Ny Munkegade 116, building 1540, 8000 Aarhus C, Denmark*

[20] *Center for Biodiversity Dynamics in a Changing World (BIOCHANGE), Department of Biology, Aarhus University, Ny Munkegade 116, building 1540, 8000 Aarhus C, Denmark*

[21] *Department of Environmental Biology, University of Rome "La Sapienza", Rome 00185, Italy*

[22] *University of Zurich, Department of Geography, Remote Sensing Laboratories, Winterthurerstrasse 190, 8057 Zurich, Switzerland*

[23] *Institute of Geography and Geoecology, Karlsruhe Institute of Technology, Kaiserstr. 12, 76131 Karlsruhe, Germany*

[24] *Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, The Netherlands*

[25] *Department of Earth and Environmental Science, Macquarie University, Sydney, NSW, Australia*

[26] *Department of Forest and Wildlife Ecology, University of Wisconsin, Madison, WI, USA*

[27] *Earth Science Division, NASA Headquarters, Washington, DC, 20546*

[28] *Natural Environment Centre, Finnish Environment Institute (SYKE), Helsinki, Finland*

[29] *University of Wuerzburg, Department of Remote Sensing, Oswal-Kuelpe Weg 86, Wuerzburg 97074, Germany*

[30] *UMR CNRS 7058 "Ecologie et Dynamique des Systèmes Anthropisés" (EDYSAN), Université de Picardie Jules Verne, 1 Rue des Louvels, 80000 Amiens, France*

August 10, 2022

Abstract

Biodiversity monitoring is an almost inconceivable challenge at the scale of the entire Earth. The current (and soon to be flown) generation of spaceborne and airborne optical sensors (i.e., imaging spectrometers) can collect detailed information at unprecedented spatial, temporal, and spectral resolutions. These new data streams are preceded by a revolution in modelling and analytics that can utilize the richness of these datasets to measure a wide range of plant traits, community composition, and ecosystem functions. At the heart of this framework for monitoring plant biodiversity is the idea of remotely identifying species by making use of the 'spectral species' concept. In theory, the spectral species concept can be defined as a species characterized by a unique spectral signature and thus remotely detectable within pixel units of a spectral image. In reality, depending on spatial resolution, pixels may contain several species which renders species-specific assignment of spectral information more challenging. The aim of this paper is to review the spectral species concept and relate it to underlying ecological principles, while also discussing the complexities, challenges and opportunities to apply this concept given current and future scientific advances in remote sensing.

Keywords: airborne sensors; biodiversity; ecoinformatics; hyperspectral images; plant optical types; remote sensing; satellite imagery; vegetation communities.

Plain language summary (PLS)

Biodiversity monitoring based on field data in almost inconceivable at the scale of the entire Earth. Over the past decades, remote sensing has opened possibilities for Earth observation from air and space, allowing us to monitor ecological change, primarily expressed by changes in vegetation cover, distribution and functioning, which can be subsequently linked to drivers of change in space and time, from local to global scale. Recently, the spectral species concept - an algorithm that clusterizes pixels from spectral images with a similar spectral signal (referred to as 'spectral species') - has brought attention . The aim of this paper is to review the ecological functioning principles of the spectral species concept and to refine its definition by a better linkage with field observations of plant species distribution data (i.e. presence-absence data) available from vegetation surveys.

3

Key Points (KP)

1. Remote sensing has opened possibilities for Earth observation from air and space, allowing us to monitor ecological change.

2. Biodiversity monitoring based on field data in almost inconceivable at the scale of the entire Earth.

3. The spectral species concept, relating field to remotely sensed data, can open news ways to measure diversity from space.

4

1   Background

Rapid environmental changes are occurring across the globe at small to large spatial extents due to the combined effects of climate change, land-use changes, and biological invasions (Kreft & Jetz, 2007). Therefore, there is an urgent pressing demand for operational biodiversity monitoring systems that will improve our understanding of the repercussions of these drivers of changes on ecosystem functioning and lead to better ecosystem management (Skidmore et al., 2021). New approaches are required to obtain timely biodiversity data that are consistently and routinely measured across the Earth surface. Some of these needs are fulfilled by remote sensing information (Schweiger and Laliberté, 2022). Over the past decades, remote sensing has opened possibilities for Earth observation from air and space, allowing us to monitor ecological change, primarily expressed by changes in vegetation cover, distribution and functioning, which can be subsequently linked to drivers of change in space and time, from local to global scale (Skidmore et al., 2015; Asner et al., 2017). Recent technological advances in remote sensing data acquisition and processing now open new perspectives for monitoring changes in biodiversity at unprecedented details over large geographic areas, and ultimately over the entire Earth (Luque et al., 2018; Randin et al., 2020). Furthermore, missions like the Surface Biology and Geology (SBG) by NASA (`https://sbg.jpl.nasa.gov/`, Cawse-Nicholson et al., 2021) have been implemented to support the development of algorithms for exploiting spaceborne remotely sensed data and providing a relatively fast but accurate estimate of ecological properties in vast areas over time.

Spaceborne and airborne passive optical sensors relying on imaging spectroscopy (i.e., spectral remote sensing including multispectral and hyperspectral imaging) are a good example of the recent remote-sensing revolution in ecology (Kwok, 2018). By measuring information from most of the electromagnetic spectrum operable for Earth observation, imaging spectroscopy has demonstrated significant capabilities to detect and monitor the spatial distribution of plant communities, species and traits (Asner & Martin, 2008; Schaaf et al., 2013; Schweiger et al., 2017; Skowronek et al., 2017). The pixel reflectance in an optical image results from the integration of multiple interactions between light and matter, including vegetation and the surrounding environment (soil, atmosphere). Intrinsic properties of vegetation influencing this remotely sensed information correspond to biophysical and biochemical properties (i.e., traits) of leaves and the canopy that can be related to levels of ecological organizations such as ecosystems, communities, species, and potentially to the intraspecific trait of plant genotypes (Madritch et al., 2014; Blonder et al., 2020). Spectroscopy has long been used to capture the

5

characteristic absorption features of biochemical compounds of plants, which biologically corresponds to the phenotypic expression of some of the genes that describe individuals belonging to a given species (e.g., Jacquemoud & Ustin, 2019). These biochemical traits and their dynamics are linked to functional traits, opening access to the monitoring of ecosystem functions, processes and services. Besides, imaging spectroscopy has already demonstrated capabilities for species discrimination in various types of ecosystems (Féret & Asner, 2014; Fassnacht et al., 2016; Skowronek et al., 2017). At the sub-organism level, biochemical properties estimated from observations at leaf and canopy scales (Baret et al., 1994; Kokaly et al., 2009; Ollinger, 2011; Serbin et al., 2012) are described in commonly used radiative transfer models (Jacquemoud et al., 1996; Féret et al., 2008; Torresani et al., 2021). The fine spectral resolution and spectral sampling interval on imaging spectrometers provide information to quantify key biochemical properties of vegetation such as leaf pigment (e.g., chlorophylls, carotenoids, anthocyanins), water, cellulose, lignin, nitrogen, phosphorous and protein contents based on their specific light absorption characteristics (Ewald et al., 2018).

The link between biochemical properties, functional attributes (morphological, physiological and phenological traits), and taxonomic information is often implicitly assumed when performing spectroscopic analysis for biodiversity monitoring. Recently, the spectral species concept - an algorithm that clusters pixels from spectral images with a similar spectral signal (referred to as 'spectral species') - has brought attention (Féret & Asner, 2014; Rocchini et al., 2021b). However, in reality, the automatic detection of pixel units sharing a similar spectral signature in a remotely-sensed spectral image does not necessarily match with the actual distribution pattern of a given species but may rather reflect the spatial distribution pattern of a group of species sharing similar biochemical properties (Woodcock and Strahler, 1987).

The capacity to identify species is often explained by similarities in spectral signatures between individuals of the same species, and dissimilarities in spectral signatures between individuals of different species. There is no taxonomic marker in spectroscopy, but individuals of the same species are characterized by a limited set of biophysical and biochemical properties, allowing differentiation from individuals from other species. Despite significant scientific advances, development of automated retrievals of plant biochemistry, traits and species identification from satellites across the globe and over time remains aspirational and more work is needed to accomplish this goal, especially for devising a global monitoring of biodiversity change. Here, we aim at introducing the ecological functioning principles of the spectral species concept and refining its definition by a better linkage with field observations of plant species distribution data (i.e., presence-absence data)

6

available from vegetation surveys. Finally, we will deeply face complexities, challenges and opportunities associated with the use of this concept for remotely-sensed biodiversity monitoring, including species richness and evenness (alpha-diversity) as well as composition turnover (beta-diversity).

1.1 A constellation of optical sensors to explore the spectral domain of the Earth surface

The current generation of Earth observing spaceborne and airborne sensors acquires highly-resolved images over a wide range of wavelengths in the solar (optical) and microwave domains of the electromagnetic spectrum (Ustin & Middleton, 2021). The information measured by optical sensors operating in the visible (380-750 nm) and near to shortwave infrared (750-2500 nm) regions largely corresponds to the region of solar reflectance from the Earth's surface, and are acquired for each individual pixel of an image. Image acquisition can occur at fine temporal resolutions of days to weeks and in a few cases even at multiple revisits during the day, but data availability and quality are strongly dependent on atmospheric conditions. Optical remote sensing systems include those that measure a few discrete spectral bands (multispectral imaging), such as: (i) the Thematic Mapper on USGS Landsat; (ii) MERIS (Medium Resolution Imaging Spectrometer) on the ESA Envisat; (iii) MODIS (Moderate Resolution Imaging Spectrometer) on the NASA Terra and Aqua platforms; and more recently (iv) the Sentinel-2 satellites of the ESA Copernicus program. These instruments provide open access data describing the Earth surface at frequent intervals and are complemented by commercial satellites that have high spatial resolution like IKONOS, SPOT, Quickbird, WorldView and more recently the Planet constellation.

Some spaceborne sensors provide enhanced spectral capabilities and measure surface reflectance over hundreds of narrow contiguous spectral bands covering the solar radiation spectrum. These hyperspectral satellites are also called imaging spectrometers. The term hyperspectral emphasizes instruments measuring a large number of spectral bands while imaging spectrometer refers to the type of instrument used, i.e., a spectrometer that measures bands across a spectral wavelength interval (spectrum) and produces a 2D array of pixels and spectra. Hyperspectral satellites have demonstrated strong potential for characterizing the chemical and physical structure of the Earth surface, with applications in mineralogy, soil sciences, and vegetation sciences (Plaza et al., 2009). Such images reveal details that improve estimates of key vegetation properties and better discriminate between vegetation types, species or even between genotypes of a given species when the spatial and

7

spectral resolutions are sufficient to match individuals (Madritch et al., 2014; Blonder et al., 2020). Several satellites hosting imaging spectrometers were recently launched, and a multitude of satellite missions are in preparation (see the review paper by Ustin & Middleton, 2021). These include DESIS and HISUI on the International Space Station (ISS), the free flying platforms PRISMA and EnMAP (expected launch in 2022), along with NASA's and ESA's global monitoring imaging spectrometer missions SBG and CHIME (expected launches in the late 2020s).

1.2 Remote sensing tools for monitoring biodiversity

Species identification from the measurement of the absorptive and reflective characteristics of plants is based on the hypothesis that individuals from the same species share similar biochemical properties, leading to similar spectral characteristics measured at the pixel scale. However, individuals from the same species may also share similar biochemical properties with individuals from another species which may limit our abilities to assign a given spectral signature to a given species. Besides, one needs to also consider phenological patterns, since remote sensing data is generally only acquired during specific periods of time (e.g. seasonal), or under specific conditions (e.g. health status), or with specific sensor information (e.g. spatial, spectral and radiometric resolutions across all wavelengths), all of which may complicate species identification or discrimination. Thus, there are several challenges to identifying a taxonomically identified species with a unique spectral signature. Previous research has shown that minor shifts in plant development induced by the environment and its interaction with different plant genotypes (i.e., phenotypic variability) may result in co-location between vegetative, flowering, fruiting, or senescent stages, each of which expresses different biophysical and biochemical properties. Hence, multiple spectral signatures corresponding to multiple biochemical traits may exist for the same species. Although adding complexity, these problems have solutions. For example, fast growing annual species and facultative annuals (e.g., the invasive Water Hyacinth (*Eichhornia crassipes*) are often found in different growth stages and detection requires several spectral signatures to account for the different growth stages that are later combined for mapping the focal species of concern (Khanna et al., 2011). Further, spectral discrimination among species is difficult, for example when several species in an ecosystem may share a suite of traits (i.e., strong overlap in the trait space) due to climate constraints and environmental filtering. This happens when trait combinations make individual species phenotypically similar (e.g., grass species in grassland habitats, as shown in diversity studies by Gholizadeh et al., 2019, 2020), or when dif-

8

ferent combinations of traits result in similar spectral signatures (Kokaly et al., 2009; Ollinger, 2011). Additional empirical evidence is thus necessary to contextualize and resolve these questions (Andrew and Ustin, 2008), such as the data provided by field data collections, experimental studies, and modeling, to achieve full ecological understanding of information from the satellite data to monitor biodiversity.

1.3 From optical types to the spectral species concept

Just as traditional biodiversity theory focuses on differences between individuals to assign individuals to different species in order to assess species richness, diversity measured with imaging spectrometer data from spaceborne and airborne optical sensors is based on pixels, each with its own spectral information (i.e., reflectance) (Rocchini et al., 2021a). The term spectral signature is more specific than spectral information and usually applied to a specific type of surface (soil, vegetation, water), to a specific material (i.e., traits like chlorophyll concentration), or to a specific level of biological organization (genotype, population, species, stand, community, ecosystem). Spectral diversity corresponds to the spatial variation of spectral information: it is tightly related to the notion of (multivariate) variation among species traits (Rocchini et al., 2018) which is the basis of functional diversity in classical ecological theory. In other terms, spectral diversity reflects - at least to a large extent - diversity in community functioning based on the assortment of functional traits in the community, irrespective of the species/individuals that possess these traits (Petchey & Gaston , 2002; Matson et al., 2005). Therefore, spectral diversity is conceptually closer to the notion of functional diversity than to the concept of species or taxonomic diversity. For this reason, it is important to keep in mind that the spectral signature of a given pixel unit in a spectral image cannot directly be assigned to a given individual belonging to a specific taxonomic entity (i.e., the spectral species concept). First of all, the pixel size of the image may not match with an actual individual in the field but may contain several individuals belonging to the same species (i.e., a population) or several individuals belong to different species (i.e., a community) such that the same species could be involved in different spectral signatures or 'optical types' (Figure 1). Additionally, the same suite of biochemical traits (e.g., foliar nutrient contents) may show extremely high spectral variability depending on the health status of the focal plant species such that different optical types could be assigned to the same species depending on its status: healthy vs. stressed (Figure 1, panels f,g,h). Hence, prior to defining the spectral species concept, one needs to identify the different optical types occurring within a given spectral

9

image and then assess how each optical type overlaps with the concept of taxonomic species so that one can assign an optical type to a specific species (i.e., the spectral species concept) or a group of species. To confirm that the spatial distribution pattern of pixels sharing a similar spectral signature (i.e., an optical type) can be directly assigned to the spatial distribution pattern of individuals from a single species (i.e., the spectral species), one needs ground-truth data on species distribution (i.e., presence-absence data from vegetation surveys) from within the spectral image.

Depending on pixel size and size of individuals, pixels can cover multiple individuals or parts of an individual, or stands of one species or mixed stands (Figure 1). In forests, where the spectral species concept was born, a match between spectral and biological species was furthered by the fact that pixels have a higher chance to belong to one species alone (Figure 1, panels a to e, with small pixel size). This does not apply to all situations. Grassland communities, for example, pose a greater challenge because pixels regularly contain several species and optical types can at best be linked to entire communities (Figure 1, panels f to h with any pixel size). These, however, are even less clearly delimited objects than species.

1.4 Spectral species translating spatial distributions of optical types into diversity metrics

Various analytical approaches have been developed to take advantage of the spectral information in the optical domain and investigate different dimensions of biodiversity (see the review paper by Rocchini et al., 2018). These approaches have demonstrated that it is possible to map and understand functional and taxonomic diversity through space and time, achieved through methodological approaches that differ in their focus, either statistical or process-based. They include data transformation (Rocchini et al., 2017), feature selection and dimensionality reduction (Feilhauer et al., 2011), and machine learning techniques (Kitzes et al., 2021).

The number of plant optical types proposed by Ustin & Gamon (2010) is consistent with the Spectral Variation Hypothesis (SVH) (Palmer et al., 2002; Rocchini et al., 2010) which states that ecosystem heterogeneity is associated with high spectral variability. In other words, increased environmental heterogeneity provides more niches for species to co-occur in geographical space with an expected increase in local species richness (Palmer et al., 2002). For optical types, we therefore expect greater diversity where greater spatial environmental heterogeneity occurs.

Spectral species - i.e., the number of spectrally distinct classes that ap-

10

proximate species - are based on the hypothesis that proper processing of images allows discrimination among species, groups of species, or functional groups. Spectral species aim to discretize remotely-sensed information into groups of pixels through unsupervised clustering. The notion of spectral species was successfully applied for mapping tropical biodiversity using airborne high spatial resolution imaging spectroscopy (Féret & Asner, 2014), under the assumption that the majority of pixels did not contain plant mixtures and that individual pixels could be assigned to a species, given that spectral variance among pixels meet statistical criteria and that the average tree crown size approximates the pixel size. Underwood et al. (2007) and Thorp al. (2013) evaluated information content for species mapping from different spatial and spectral resolutions, concluding that high spectral resolution contributed substantially more information for species mapping than higher spatial resolution. Hence, spectral species and optical types are conceptually equivalent in their purpose to discriminate among optical entities acquired from an image, but the appropriateness of the terminology may vary with the type of ecosystem and spatial resolution of the sensor.

A spectral species is then a set of pixels having similar spectral properties that can be used as a proxy for ecologically relevant taxonomic or functional groups, and eventually inventoried to calculate biodiversity metrics like alpha-diversity (e.g., Shannon's H, Simpson's D, Rényi's H) and beta-diversity (e.g., Bray-Curtis dissimilarity or Rao's quadratic entropy) across a landscape (**Box 1**). Clustering approaches similar to the spectral species framework have been successfully used to map plant species diversity over very different habitats and geographical regions: from African savannas (Baldeck et al., 2014) to grasslands in the Platte River ecosystem near Wood River, Nebraska, USA (Gholizadeh et al., 2020), and from the Peruvian Andes-Amazon tropical forests (Féret & Asner, 2014) to old-growth, secondary and artificial forests of the Shennongjia National Forest Natural Reserve in China (Zhao et al., 2018).

Spectral species combined with functional trait estimation contribute to biodiversity understanding by characterizing the morphological (canopy architecture, gap fraction, etc.) and functional traits, that define the functional role of a species in an ecosystem, e.g. carbon capture strategies by resource acquisition or resource conservation strategies (like drought responses), that then provide a basis to connect spectral species with the taxonomic species that exhibit these trait assemblages. Therefore, spectral species provide a basis to connect with biological species and eventually explore spatial distribution of trait assemblages, their evolution in time, and their linkages to environmental or human factors.

11

2   Complexities and challenges

Many of the uncertainties of the spectral species concept apply generally to remotely sensed classified images, which we will not review in detail because they are extensively addressed elsewhere (e.g., Lu & Weng, 2005; Kun et al., 2011). The spectral species method of identifying spectral species with distinctive traits facilitates identification of rare species that occupy few pixels in an image. Rare and endemic plants by definition do not grow in large colonies of many individuals and generally have restricted distributions. But they are of importance to biodiversity mapping because there are far more rare species than frequent species and they are often concentrated in areas of high biodiversity, especially in hotspots (Griggs, 1940; Medail and Quezel, 1997; Ricotta et al., 2010). In terms of biodiversity they may represent species with unique morphological features and genetic richness (Myers et al., 2000; Joppa et al., 2011). Other aspects of uncertainty around the operationalization of the spectral species concept are similar to many problems encountered in vegetation classification systems, such as whether rare species are retained in the classified data, whether the classes are realistic in terms of field measured data, and how well the method deals with mixed pixels (multispecies pixels, species with multiple phenotypes).

Some of the complications that can make the spectral species concept difficult to handle are shown in the aforementioned Figure 1, where differences in reflection do not necessarily adhere to species boundaries. In fact, several factors determine whether the spectral species concept predicts the current number of species in the area under study. First, more fundamentally it depends on how the taxonomic species are determined, e.g., whether related taxa are characterized by 'splitting' or 'lumping' criteria, leading to more or fewer species. There are also cases of cryptic diversity, i.e., species that are not morphologically or phenotypically distinct, but having a distinct genetic make-up. Thus, the actual number of detected species approximates the true number. Second, even if species are clearly assigned, the spectral species concept requires the measurement of optically active traits, which may not be the traits that differentiate species, or - because of trade-offs (based on the species functional strategy) or selective pressures (resulting from competition, human use and other types of ecological interactions) - they may be not expressed at the time of measurement. Third, community composition and abundance interact with what is detectable – as the sensor primarily measures top of canopy dominant species, which may identify important ecosystem functions but miss much of the diversity.

The spectral species concept, as with most remotely sensed measures, is dependent on the temporal dimension to which it is applied. For instance,

12

measuring the phenological differences among co-occurring plant species throughout the season would improve species detection. Including seasonal changes in the analysis may improve detection of understory species when the deciduous overstory is dormant. Multitemporal datasets can form multi-seasonal spectral signatures to account for seasonal changes in a species (e.g., Somers & Asner, 2014) and multitemporal spectral libraries are under development (Dudley et al., 2015). The spectral species concept has shown consistency in identifying taxonomic species and traits from other taxonomic species and their traits. Detecting species in mixed pixels depends on the sensor traits listed above and the magnitude of the differences among the spectral signatures of the species in the pixel. Species with different functional strategies that are distinguishable in optically active traits will be easier to identify than species with more similar trait assemblies, but which are not optically active. Differences in traits due to leaf types (evergreen or deciduous, needle leaf or broadleaf) create considerable differences in leaf reflectance and timing the date of data acquisition to maximize phenological differences generally always improves the ability to detect sub pixel species. The spectral species concept has yet to be widely tested over the global range of ecosystems and environmental conditions so there is a need to determine its performance and limitations at these scales.

2.1   Ecological issues behind spectral species

The spectral species concept was initially developed to allow computation of diversity indices usually computed from species inventories, but with pixels used instead of species. It did not intend to directly estimate 'absolute' species richness from the spectral clustering, but rather hypothesized that relative species richness and higher-level diversity indices integrating richness and abundance or dissimilarity over space, could be estimated and compared within a limited spatial extent corresponding to an airborne imaging spectroscopy acquisition. Even in a very unlikely situation of perfect spectral discrimination among species in a remotely sensed acquisition, the predicted number of species may be smaller than the actual species count due to problems related to the potential similarity of traits among multiple species, especially when the growth forms and phenology are similar. The predicted number of species may also be higher than the actual species count if individuals from the same species show phenological shifts. The naming of spectral classes requires matching the taxonomic species information with the spectral information. If the taxonomic naming was based on extreme clumping of groups into fewer taxa or the splitting of taxa into increased number of taxa, the predicted spectral species may differ from the number of

13

biological taxa identified at a site. Perhaps knowledge of the predicted number of spectral species might result in rethinking the criteria for recognition and taxonomic revisions, or similarly knowledge of the species may require rethinking of how spectroscopy data were analyzed.

While ecological communities have a variety of either dominant, co-dominant and rare species, the latter contribute a lot to the highest diversity measurable on Earth. This is a problem for remote sensing approaches based on clustering, as these algorithms often delete clusters with just a few data points (minimum class size or minimum/maximum number of classes may be selected in the setup), thus there is a tendency to lose rare and endemic species likely fundamental for biological diversity and ecosystem functions. A 'continuous surface' analytical method that directly addresses this problem is the widely used 'Multiple Endmember Mixture Model' (MESMA, Roberts et al., 1998), which models each pixel from a range of 'endmembers' (also referred to as 'pure spectra') identified statistically from the convex hull of the data hypervolume. The model can retain classes with small numbers of endmembers depending on user criteria but may have the difficulty in assigning class labels when the variation is continuous (e.g., across a natural continuum or from one habitat type to another, Schmidtlein et al., 2007).

Environmental heterogeneity may also affect the accuracy of the spectral species methodology (Schmidtlein & Fassnacht, 2017). Disturbances may increase or decrease biodiversity, with a net zero effect on heterogeneity. There are several examples (e.g., forest gap openings, fire spread, and urban development) that can increase heterogeneity without an increase overall biological diversity. Spectral species should detect these changes as differences in the assemblage and range of traits being measured, although there is uncertainty about the least detectable magnitude of change. Therefore, we expect remotely sensed spectral species to provide a first-order exploratory tool to, at least detect areas that are suspected of hosting a high number of species or, alternatively, identify areas that should host larger numbers of species (based on independent criteria) but don't. From this point of view, in-situ data would greatly improve the classification, validate its accuracy, and provide a basis to investigate and identify sites that do not match a priori expectations (Foody et al., 2016).

The R package (`biodivMapR`) dedicated to applications of the spectral species algorithm (Féret and de Boissieu, 2020) on raster data requires the number of spectral species be defined by the user. The method is not intended to estimate the absolute number of species from remote sensing, as this problem is highly scale and context dependent as explained earlier. The definition of the optimal number of clusters in a dataset is a specific problem

14

that the spectral species framework was also not intended to solve. Adding a method to determine the number of clusters corresponding to species richness over a full dataset would definitely contribute to the generalization of the spectral species framework over various ecosystems. The number of species included in a remote sensing acquisition is not generally known a priori for most of the Earth's ecosystems. However, there are statistical approaches to estimate the number with enough accuracy that realistic values could be produced (Chang and Du, 2004; Gholizadeh et al., 2020) , as well as deeply rooted ecological principles based on species dispersal, biogeography, landscape ecology (Rocchini et al., 2021a).

There are questions on whether this method will continue to be successful at the spatial scales of satellites, where multiple species within pixels cause spectral mixing, and under conditions of sparse vegetation where mixed pixels can include plants, soils, plant litter, geological minerals, water, ice and man-made materials. This will be resolved with the use and applications of the new generation of hyperspectral satellites. The various spaceborne imaging spectrometers (see the review paper by Ustin & Middleton, 2021) that are available now (DESIS, EnMAP, HISUI, PRISMA) or in the next few years (e.g., CHARM, EMIT, SBG) all have 30m pixels. These imaging spectrometers will provide ample opportunities to examine how well the spectral species concept scales up and across ecosystems.

In addition to passive optical sensors, reviewed in this manuscript, LiDAR and radar active sensors might complement observations with the capacity to penetrate the uppermost layer of vegetation and provide information on canopy structure, while being less sensitive to atmospheric perturbations than passive optical sensors, Asner et al., 2008; Bergen et al., 2009; Zhao et al., 2018; Mulatu et al., 2019; Simonson et al., 2021; Lenoir et al., in press). For instance, the GEDI LiDAR mission on the ISS is measuring at high vertical and spatial resolutions the distribution and height of global woody vegetation to provide structural information, including understory data, for biodiversity estimates. Besides, the NISAR and Biomass radar missions (launches in 2023) will monitor global patterns of biomass, disturbances, and impacts on biodiversity. Together, advances in the spatial, spectral and temporal dimensions of imagery offer immense data streams that can be harnessed to better understand the processes of biological functioning, and to systematically map and monitor ecosystem changes from local to regional to global scale.

15

3 Outlook

In this manuscript we focused on the need to model biodiversity from hyperspectral remote sensing and field data to calibrate such models. A huge amount of field (free) data is actually available (e.g., sPlotOpen, GBIF), but creativity for mapping species diversity from hyperspectral imagery should be reinforced. From this point of view, the spectral species concept is expected to make a major contribution in mapping and analysis of hyperspectral satellite data to produce remote sensing based essential biodiversity variables (RS-EBVs, see Skidmore et al., 2015, 2021). Even if the predicted spectral species do not precisely identify the number and identity of taxonomic species in an image, it will provide a first-order exploratory tool to detect areas of low to high species diversity. Such approaches will require a better understanding of the range of conditions under which the spectral species concept can operate. In section 2 we highlight some limitations for applications of the concept, but many of these limitations should be seen as new opportunities for research. These include, among others, identification of additional spectral characteristics that can be captured and what additional traits, and thus more species will be detected. We are aware that very small absorption features are present in plant spectra and are statistically detectable but we do not know the physical basis of the biochemical/biophysical material that is absorbing this energy, highlighting an important avenue of research for the future. There is an ever-growing number of new traits reported from imaging spectroscopy research, ranging from phenolic and isoprene compounds, non-structural carbohydrates, fiber content (Serbin et al., 2014; Singh et al., 2015; Ely et al., 2019), essential nutrients including potassium, phosphorous, and calcium (Asner et al., 2015; Ely et al., 2019) to RuBP carboxylation (Vcmax) and regeneration (Jmax) (Serbin et al., 2012, 2014; Rogers et al., 2017; Wu et al., 2019). Such information is retrieved from measurements enabled by today's imaging spectrometers, which typically have spectral resolutions of 10 nanometer (nm) spectral bands cross the solar spectrum for a total of around 200-250 bands, some have 3-5nm in the visible-near-infrared wavelengths and 5-10nm in the shortwave infrared region. New satellite sensors under development (Flora on the Flex platform and GeoCarb) are designed to detect chlorophyll fluorescence with 1-3nm narrow bands in the wavelength regions where chlorophyll fluoresces near the oxygen bands.

The spectral species concept is based on the principle that the variability in a spectral data cube is sufficient to identify a suite of traits in a pixel. In these cases, pixels are composed of multiple species with an assortment of different traits. The spectral species concept has been broadened to that of "spectranomics", which is "an approach to conceptually and geographically

16

link plant canopy species and their functional traits to their spectral-optical properties" (Asner & Martin, 2016). Broadening "spectranomics" has benefit for biodiversity assessment across different organizational levels, another area in which the spectral species concept could contribute.

While differences in traits identify different species, they can also describe patterns at different levels of ecological organization: habitat, community, ecosystem. In Ustin & Gamon (2010), the concept of "plant optical types" is based on their spectral attributes, without any direct reference to species or traits. By relating these optical types to specific traits (e.g. leaf and canopy resources allocated to productivity) these optical types acquire the definition of plant functional types. Just as actual plant species change, the spectral species concept must be dynamic and change. As the range of applications increase, the range of spectral patterns will increase (e.g., including phenological events like mass flowering or senescent states) (Poyry et al., 2018).

Understanding the interactions between biodiversity and ecological/environmental drivers is difficult (Kreft & Jetz, 2007). From this point of view, collecting exploratory remote sensing data on environmental heterogeneity across large geographical extents is relatively simple and combining that information with changing patterns of functional traits could improve species identification, mapping and monitoring of potential diversity hotspots (Skidmore et al., 2015; Asner et al., 2017).

As biodiversity research expands to global ecosystems, there are many questions unresolved. Both ecological sciences and Earth observation technologies are still in maturing phases of development. In this paper, we clarified the links between the spectral species concept, optical types and optical traits, and their analogy with ecological dimensions including species, functional types and functional traits, as well as the mechanistic link between biophysical properties of vegetation and what is usually expressed as spectral signatures, corresponding to species. The relevance of discrete approaches, and their complementarity with continuous approaches was also highlighted.

As stressed in this paper, important advances have been made in understanding how the spectral signature relates to biodiversity and where there is untapped potential to further disentangle these connections. While one should not overestimate the capacity of remote sensing to directly estimate biodiversity from space (Skidmore et al., 2021), remote sensing captures patterns of reflected or emitted electromagnetic radiation that are driven by biophysical and biochemical properties of vegetation (i.e., patterns dependent on the optical properties of plants that are the phenotypic expressions of their fitness strategies). There are a wide range of scientific disciplines that include

17

all ecological subdisciplines from the population to the global biosphere, biology, soils, hydrology, evolution and phylogenetics, optics, mathematics, statistics and informatics and engineering, that are necessary to understand, interpret, refine and improve remote sensing research. To actually develop and verify methods to identify and monitor global patterns of biodiversity, it will truly take "more than a village" but instead an engaged and committed international contingent of scientists, social scientists, citizen scientists, engineers, policy makers, land owners, farmers, and more, to shorten the long road ahead.

4   Authors' contributions

DR, MJS, SLU, JBF and JL led the writing of this manuscript. DR, JBF, GPA, MD, MJS and SS analyzed the data providing the output for the figures of this manuscript. SS conceived and crafted Figure 1. All authors contributed critically to the development of the concept of this manuscript and to the writing of the draft, and they gave final approval for publication.

5   Acknowledgments

18

References

Andrew, M.E., Ustin, S.L. (2008). The role of environmental context in mapping invasive plants with hyperspectral image data. Remote Sensing of Environment, 112, 4301-4317.

Asner, G.P., Knapp, D.E., Kennedy-Bowdoin, T., Jones, M.O., Martin, R.E., Boardman, J. & Hughes, R.F. (2008a) Invasive species detection in Hawaiian rainforests using airborne imaging spectroscopy and LiDAR. Remote Sensing of Environment, 112, 1942-1955.

Asner, G.P. & Martin, R.E. (2008) Spectral and chemical analysis of tropical forests: scaling from leaf to canopy levels. Remote Sensing of Environment, 112, 3958-3970.

Asner, G.P., Martin, R.E. (2016) Spectranomics: Emerging science and conservation opportunities at the interface of biodiversity and remote sensing. Global Ecology and Conservation, 8, 212-219.

Asner, G.P., Martin, R.E., Anderson, C.B., Knapp, D.E. (2015) Quantifying forest canopy traits: Imaging spectroscopy versus field survey. Remote Sensing of Environment, 158, 15-27.

Asner, G.P., Martin, R.E., Knapp, D.E., Tupayachi, R., Anderson, C.B., Sinca, F., Vaughn, N.R. & Llactayo, W. (2017) Airborne laser-guided imaging spectroscopy to map forest trait diversity and guide conservation. Science, 355, 385-389.

Baldeck, C.A., Colgan, M.S., Féret, J.B., Levick, S.R., Martin, R.E. & Asner, G.P. (2014) Landsape-scale variation in plant community composition of an African savanna from airborne species mapping. Ecological Applications, 24, 84-93.

Baret, F., Vanderbilt, V.C., Steven, M.D. & Jacquemoud, S. (1994) Use of spectral analogy to evaluate canopy reflectance sensitivity to leaf optical properties. Remote Sensing of Environment, 48, 253-260.

Bergen, K.M., Goetz, S.J., Dubayah, R.O., Henebry, G.M. et al. (2009) Remote sensing of vegetation 3-D structure for biodiversity and habitat: Review and implications for lidar and radar spaceborne missions. Journal of Geophysical Research: Biogeosciences, 114, G00E06.

Blonder, B., Graae, B.J., Greer, B., et al. (2020). Remote sensing of ploidy level in quaking aspen (*Populus tremuloides* Michx.). Journal of Ecology, 108, 175-188.

19

Cawse-Nicholson, K., Townsend, P.A., Schimel, D., Assiri, A.M., Blake, P.L. et al. (2021). NASA's surface biology and geology designated observable: a perspective on surface imaging algorithms. Remote Sensing of Environment, 257, 112349.

Chang, C.I., Du, Q. (2004) Estimation of number of spectrally distinct signal sources in hyperspectral imagery. IEEE Transactions on Geoscience and Remote Sensing 42: 608-619.

Dudley, K.L., Dennison, P.E., Roth, K.L., Roberts, D.A., Coates, A.R. (2015). A multi-temporal spectral library approach for mapping vegetation species across spatial and temporal phenological gradients. Remote Sensing of Environment, 167, 121-134.

Ely, K.S., Burnett, A.C., Lieberman-Cribbin, W., Serbin, S.P. & Rogers, A. (2019) Spectroscopy can predict key leaf traits associated with source-sink balance and carbon - nitrogen status. Journal of Experimental Botany, 70, 1789-1799.

Ewald, M., Skowronek, S., Aerts, R., Dolos, K., Lenoir, J., Nicolas, M., Warrie, J., et al. (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.

Fassnacht, F.E., Latifi, H., Stereńczak, K., Modzelewska, A., Lefsky, M., Waser, L.T., Straub, C. & Ghosh, A. (2016). Review of studies on tree species classification from remotely sensed data. Remote Sensing of Environment, 186, 64-87.

Feilhauer, H., Faude, U. & Schmidtlein, S. (2011) Combining Isomap ordination and imaging spectroscopy to map continuous floristic gradients in a heterogeneous landscape. Remote Sensing of Environment, 115, 2513-2524.

Feilhauer, H., Somers, B., van der Linden, S. (2017) Optical trait indicators for remote sensing of plant species composition: Predictive power and seasonal variability. Ecological Indicators, 73, 825-833.

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. (2020) biodivMapR: An R package for α- and β-diversity mapping using remotely sensed images. Methods in Ecology and Evolution, 11, 64- 70.

20

727 Féret, J.-B., François, C., Asner, G.P., Gitelson, A.A., Martin, R.E., Bidel,
728 L.P.R., Ustin, S.L. & le Maire, G., Jacquemoud, S. (2008) PROSPECT-4
729 and 5: Advances in the leaf optical properties model separating photosyn-
730 thetic pigments. Remote Sensing of Environment, 112, 3030-3043.

731 Foody, G.M., Pal, M., Rocchini, D., Garzon-Lopez, C.X. Bastin, L. (2016)
732 The sensitivity of mapping methods to reference data quality: Training
733 supervised image classifications with imperfect reference data. ISPRS In-
734 ternational Journal of Geo-Information 5: 199.

735 Gholizadeh, H., Gamon, J.A., Helzer, C. J. & Cavender-Bares, J. (2020)
736 Multi-temporal assessment of grassland α- and β-diversity using hyper-
737 spectral imaging. Ecological Applications 30: e02145.

738 Gholizadeh, H., Gamon, J.A., Townsend, P.A., Zygielbaum, A.I., Helzer,
739 C.J., Hmimina, G.Y., Yu, R., Moore, R.M., Schweiger, A.K., Cavender-
740 Bares, J. (2019) Detecting prairie biodiversity with airborne remote sens-
741 ing. Remote Sensing of Environment 221: 38-49.

742 Griggs, R.F. (1940) The ecology of rare plants. Bulletin of the Torrey Botan-
743 ical Club, 67, 575-594.

744 Khanna, S., Santos, M.J., Ustin, S.L. & Haverkamp, P.J. (2011) An in-
745 tegrated approach to a biophysiologically based classification of floating
746 aquatic macrophytes. International Journal of Remote Sensing, 32, 1067-
747 1094.

748 Kitzes, J., Blake, R., Bombaci, S., Chapman, M., Duran, S. M. et al. (2021)
749 Expanding NEON biodiversity surveys with new instrumentation and ma-
750 chine learning approaches. Ecosphere, 12, e03795.

751 Kokaly, R.F., Asner, G.P., Ollinger, S.V., Martin, M.E. & Wessman, C.A.
752 (2009) Characterizing canopy biochemistry from imaging spectroscopy and
753 its application to ecosystem studies. Remote Sensing of Environment, 113,
754 78-91.

755 Kreft, H. & Jetz, W. (2007) Global patterns and determinants of vascular
756 plant diversity. Proceedings of the National Academy of Sciences, 104,
757 5925-5930.

758 Kun, J., Li, Q., Tian, Y.-C. & Wu, B.-F. (2011) A review of classification
759 method of remote sensing imagery. Spectroscopy and Spectral Analysis,
760 31, 2618-1623.

21

⁷⁶¹ Kwok, R. (2018). Ecology's remote-sensing revolution. Nature, 556, 137-138.

⁷⁶² Jacquemoud, S. & Ustin, S.L. (2019) Leaf Optical Traits. Cambridge Uni-
⁷⁶³   versity Press, Cambridge, UK.

⁷⁶⁴ Jacquemoud, S., Ustin, S.L., Verdebout, J., Schmuck, G., Andreoli, G. &
⁷⁶⁵   Hosgood, B. (1996) Estimating leaf biochemistry using the PROSPECT
⁷⁶⁶   leaf optical properties model. Remote Sensing of Environment, 56, 194-202.

⁷⁶⁷ Joppa, L.N., Roberts, D.L., Myers, N., Pimm, S.L. (2011) Biodiversity
⁷⁶⁸   hotspots house most undiscovered plant species. Proceedings of the Na-
⁷⁶⁹   tional Academy of Sciences of the United State of America, 108, 13171-
⁷⁷⁰   13176.

⁷⁷¹ Lenoir, J., Gril, E., Durrieu, S., Horen, H., Laslier, M., Lembrechts, J. et al.
⁷⁷²   (in press). Unveil the unseen: Using LiDAR to capture time-lag dynamics
⁷⁷³   in the herbaceous layer of European temperate forests. Journal of Ecology.
⁷⁷⁴   doi: https://doi.org/10.1111/1365-2745.13837

⁷⁷⁵ Lu, D. & Weng, Q. (2005) A survey of image classification methods and
⁷⁷⁶   techniques for improving classification performance. International Journal
⁷⁷⁷   of Remote Sensing, 28, 823-870.

⁷⁷⁸ Luque, S., Pettorelli, N., Vihervaara, P. & Wegmann, M. (2018). Improving
⁷⁷⁹   biodiversity monitoring using satellite remote sensing to provide solutions
⁷⁸⁰   towards the 2020 conservation targets. Methods in Ecology and Evolution,
⁷⁸¹   9, 1784-1786.

⁷⁸² Madritch, M.D., Kingdon, C.D., Singh, A., Mock, K.E., Lindroth, R.L.,
⁷⁸³   Townsend, P.A. (2014) Imaging spectroscopy links aspen genotype with
⁷⁸⁴   below-ground processes at landscape scales. Philosophical Transactions of
⁷⁸⁵   the Royal Society B., 369, 20130194.

⁷⁸⁶ Matson, N.W.H., Mouillot, D., Lee, W.G. & Wilson, J.B. (2005) Functional
⁷⁸⁷   richness, functional evenness and functional divergence: the primary com-
⁷⁸⁸   ponents of functional diversity. Oikos, 111, 112-118.

⁷⁸⁹ Medail, F. & Quezel, P. (1997). Hot-spots analysis for conservation of plant
⁷⁹⁰   biodiversity in the Mediterranean Basin. Annals of the Missouri Botanical
⁷⁹¹   Garden, 84, 112-127.

⁷⁹² Myers, N., Mittermeier, R.A., Mittermeier, C.G., de Fonseca, G.A.B. & Kent,
⁷⁹³   J. (2000) Biodiversity hotspots for conservation priorities. Nature, 403,
⁷⁹⁴   853-858.

22

Mulatu, K.A., Decuyper, M., Brede, B., Kooistra, L., Reiche, J., Mora, B. & Herold, M. (2019) Linking Terrestrial LiDAR Scanner and Conventional Forest Structure Measurements with Multi-Modal Satellite Data. Forests, 10, 291.

Ollinger, S.V. (2011) Sources of variability in canopy reflectance and the convergent properties of plants. New Phytologist, 189, 375-394.

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

Petchey, O.L. Gaston, K.J. (2002) Functional diversity (FD), species richness and community composition. Ecology Letters, 5, 402-411.

Plaza, A., Benediktsson, J.A., Boardman, J.W., Brazile, J. et al. (2009). Recent advances in techniques for hyperspectral image processing. Remote Sensing of Environment, 113, 110-122.

Pollock, L.J., O'Connor, L.M.J., Mokany, K., Rosauer, D.F., Talluto, M.V. & Thuiller, M. (2020) Protecting biodiversity (in all its complexity): new models and methods. Trends in Ecology & Evolution, 35, 1119-1128.

Poyry, J., Bottcher, K., Fronzek, S., Gobron, N., Leinonen, R., Metsamaki, S. & Virkkala, R. (2018) Predictive power of remote sensing versus temperature-derived variables in modelling phenology of herbivorous insects. Remote Sensing in Ecology and Conservation, 4: 113-126.

Randin, C.F., M.B Ashcroft, J. Boliger, J. Cavender-Bares, N.C. Coops, et al. (2020) Modeling biodiversity in the Anthropocene using remote sensing in species distribution models. Remote Sensing of Environment, 239, 111626.

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.

Roberts, D., Gardner, M., Church, R., Ustin, S., Scheer, G. & Green, R. (1998) Mapping chaparral in the Santa Monica mountains using multiple endmember spectral mixture models. Remote Sensing of Environment, 65, 267-279.

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. & Neteler, M. (2010). Remotely sensed

23

⁸²⁹   spectral heterogeneity as a proxy of species diversity: recent advances and
⁸³⁰   open challenges. Ecological Informatics, 5, 318-329.

⁸³¹ Rocchini, D., Marcantonio, M., Da Re, D., Bacaro, G., Feoli, E., Foody,
⁸³²   G.M., Furrer, R., Harrigan, R.J., Kleijn, D., Iannacito, M., Lenoir, J.,
⁸³³   Lin, M., Malavasi, M., Marchetto, E., Meyer, R.S., Moudry, V., Schneider,
⁸³⁴   F.D., Simova, P., Thornhill, A.H, Thouverai, E., Vicario, S., Wayne, R.K.
⁸³⁵   & Ricotta, C. (2021). From zero to infinity: minimum to maximum di-
⁸³⁶   versity of the planet by spatio-parametric Rao's quadratic entropy. Global
⁸³⁷   Ecology and Biogeography, 30, 1153-1162.

⁸³⁸ Rocchini, D., Marcantonio, M., Ricotta, C. (2017). Measuring Rao's Q di-
⁸³⁹   versity index from remote sensing: an open source solution. Ecological
⁸⁴⁰   Indicators, 72, 234-238.

⁸⁴¹ Rocchini, D., Luque, S., Pettorelli, N., Bastin, L., Doktor, D., Faedi, N.,
⁸⁴²   Feilhauer, H., Féret, J.-B., Foody, G.M., Gavish, Y., Godinho, S., Kunin,
⁸⁴³   W.E., Lausch, A., Leitão, P.J., Marcantonio, M., Neteler, M., Ricotta,
⁸⁴⁴   C., Schmidtlein, S., Vihervaara, P., Wegmann, M., Nagendra, H. (2018).
⁸⁴⁵   Measuring β-diversity by remote sensing: a challenge for biodiversity mon-
⁸⁴⁶   itoring. Methods in Ecology and Evolution, 9, 1787-1798.

⁸⁴⁷ Rocchini, D., Salvatori, N., Beierkuhnlein, C., de Boissieu, F., Förster, M.,
⁸⁴⁸   Garzon-Lopez, C.X., Gillespie, T.W., Hauffe, H.C., He, K.S., Kleinschmit,
⁸⁴⁹   B., Lenoir, J., Nagendra, H., Malavasi, M., Moudrý, V., Payne, D., Šímová,
⁸⁵⁰   P., Torresani, M., Wegmann, M., Féret, J.B. (2021). From local spectral
⁸⁵¹   species to global spectral communities: a benchmark for ecosystem diver-
⁸⁵²   sity estimate by remote sensing. Ecological Informatics, 61, 101195.

⁸⁵³ Rogers, A., Medlyn, B.E., Dukes, J.S., Bonan, G., von Caemmerer, S. et al.
⁸⁵⁴   (2017) A roadmap for improving the representation of photosynthesis in
⁸⁵⁵   Earth system models. New Phytologist, 213, 22-42.

⁸⁵⁶ Schaaf, A.N., Dennison, P.E., Fryer, G.K., Roth, K.L., Roberts, D.A. (2013)
⁸⁵⁷   Mapping plant functional types at multiple spatial resolutions using imag-
⁸⁵⁸   ing spectrometry data. GIScience & Remote Sensing, 48, 324-344.

⁸⁵⁹ Schmidtlein, S. & Fassnacht, F.E. (2017) The spectral variability hypothesis
⁸⁶⁰   does not hold across landscapes. Remote Sensing of Environment, 192,
⁸⁶¹   114-125.

⁸⁶² Schmidtlein, S., Zimmermann, P., Schupferling, R. & Weiss, C. (2007) Map-
⁸⁶³   ping the floristic continuum: Ordination space position estimated from
⁸⁶⁴   imaging spectroscopy. Journal of Vegetation Science, 18, 131-140.

24

Schweiger, A.K., Laliberté, E. (2022). Plant beta-diversity across biomes captured by imaging spectroscopy. Nature Communications, 13, 2767.

Schweiger, A.K., Schultz, M., Risch, A.C., Kneubuhler, M., Haller, R., Schaepman, M.E. (2017). How to predict plant functional types using imaging spectroscopy: linking vegetation community traits, plant functional types and spectral response. Methods in Ecology and Evolution, 8, 86-95.

Serbin, S.P., Dillaway, D.N., Kruger, E.L., Townsend, P.A. (2012) Leaf optical properties reflect variation in photosynthetic metabolism and its sensitivity to temperature. Journal of Experimental Botany, 63, 489-502.

Serbin, S.P., Singh, A., McNeil, B.E., Kingdon, C.C. & Townsend, P.A. (2014) Spectroscopic determination of leaf morphological and biochemical traits for northern temperate and boreal tree species. Ecological Applications, 24, 1651-1669.

Skidmore, A.K., Pettorelli, N., Coops, N.C., Geller, G.N., Hansen, M., Lucas, R., Mücher, C.A., O'Connor, B., Paganini, M., Pereira, H.M., Schaepman, M.E., Turner, W., Wang, T., Wegmann, M. (2015) Agree on biodiversity metrics to track from space. Nature, 523, 403-405.

Skidmore, A.K., Coops, N.C., Neinavaz, E., Ali1, A., et al. (2021) Priority list of biodiversity metrics to observe from space. Nature Ecology & Evolution, 5, 896-906.

Skowronek, S., Ewald, M., Isermann, M., Van De Kerchove, R., Lenoir, J., et al. (2017) Mapping an invasive bryophyte species using hyperspectral remote sensing data. Biological Invasions, 19, 239-254.

Simonson, W. D., Allen, H. D. & Coomes, D. A. (2012) Use of an airborne lidar system to model plant species composition and diversity of Mediterranean oak forests. Conservation Biology, 26, 840-850.

Singh, A., Serbin, S.P., McNeil, B.E., Kingdon, C.C., Townsend, P.A. (2015) Imaging spectroscopy algorithms for mapping canopy foliar chemical and morphological traits and their uncertainties. Ecological Applications, 35, 2180-2197.

Somers, B. & Asner, G.P. (2014) Tree species mapping in tropical forests using multi-temporal imaging spectroscopy: wavelength adaptive spectral mixture analysis. International Journal of Applied Earth Observation and Geoinformation, 31, 57-66.

25

Torresani, M., Feilhauer, H., Rocchini, D., Féret, J.-B., Zebisch, M., Tonon, G. (2021). Which optical traits enable an estimation of tree species diversity based on the Spectral Variation Hypothesis? Applied Vegetation Science, 24: e12567.

Thorp, K.R., French, A.N. & Rango, A. (2013) Effect of image spatial and spectral characteristics on mapping semi-arid rangeland vegetation using multiple endmember spectral mixture analysis (MESMA). Remote Sensing of Environment, 132, 120-130.

Underwood, E.C., Ustin, S.L. & Ramirez, C.M. (2007) A comparison of spatial and spectral image resolution for mapping invasive plants in coastal California. Environmental Management, 39, 63-83.

Ustin, S.L. & Middleton, E.M. (2021) Current and near-term advances in Earth observation for ecological applications. Ecological Processes, 1.

Ustin, S.L. & Gamon, J.A. (2010) Remote sensing of plant functional types. New Phytologist, 186, 795-816.

Woodcock, C.E. & Strahler, A.H. (1987). The factor of scale in remote sensing. Remote sensing of Environment, 21, 311-332.

Wu, J., Rogers, A., Albert, L.P., Ely, K., Prohaska, N. et al. (2019) Leaf reflectance spectroscopy captures variation in carboxylation capacity across species, canopy environment and leaf age in lowland moist tropical forests. New Phytologist, 224, 663-674.

Zhao, Y., Zeng, Y., Zheng, Z., Dong, W., Zhao, D., Wu, B. Zhao, Q. (2018) Forest species diversity mapping using airborne LiDAR and hyperspectral data in a subtropical forest in China. Remote Sensing of Environment, 213, 104-114.

26
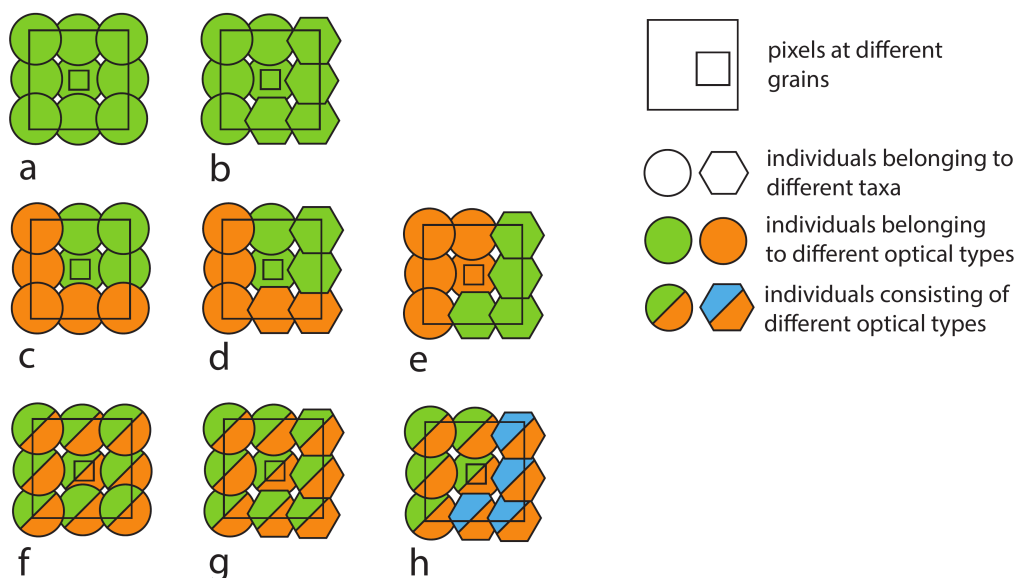
925 **Figures and boxes**

926

27

Figure 1: Some potential scenarios that can happen in optical remote sensing of vegetation canopies. The graph shows sources of variation in the relationship between species and optical types. Plants of different species might belong to different optical types, but many other situations can also be found. Optical types can be related to information of interest (e.g., species or plant traits) or to irrelevant pattern (e.g., shadows, depending on the research question). Scenario (a) represents a stand with individuals of only one single species, with a similar reflectance. In scenario (b) individuals of two species have a similar reflectance; hence they would be grouped in the same spectral species. This is further complicated once mixing individuals belonging to the same taxon but to different optical types (c) or individuals of multiple species belonging to different optical types that do not follow the species boundaries (d). What many would hope for is that plants of different species belong to different optical types, which may happen (e). Finally, the same plant individual can consist of different optical types showing different spectral properties in e.g. young versus old leaves, shadow and light, or differences in health conditions. This intra-individual mixing property will be related to all of the previous cases (f-h). Note that a stand or individual can pass through several of these scenarios in time (intra and interannual variability).

Box 1 - The spectral species algorithm at work

The spectral species concept is grounded in an algorithm which is now ready available under a free and open source R package in GitHub (`https://github.com/jbferet/biodivMapR`) named `biodivMapR` (Féret and de Boissieu, 2020), which is able to produce α- and β-diversity maps starting from the detection of spectral species based on the optical properties of vegetation in the field.

Starting from a multi- or hyper-spectral image, a spectral transformation like a Principal Component Analysis (PCA) is performed to reduce the dimensionality for further calculations (Figure 2). Based on those principal components explaining most of the variance, k-means clustering is applied to a random set of pixels in order to detect group of pixel with the same spectral reflectance, possibly related to single groups of individuals/canopies in the field that are sharing similar traits and thus that are likely to be phylogenetically related (spectral species concept). The detection of spectral species is applied backward to the whole map defining a membership probability of each pixel to a certain spectral species, as based on its spectral euclidean distance from the centroid of the previously defined clusters. A single final spectral species map can be attained by using the maximum probability for each pixel to attain to a certain spectral species. The spectral species map is further divided into elementary spatial units with a higher pixel dimension (hereafter simply units), in which calculation of α- and β-diversity can be performed, leading to crucial information on both local diversity and turnover. In order to attain α-diversity, Shannon's H' is calculated for each unit, while for β-diversity calculation Bray-Curtis dissimilarity is considered among all the possible pairs of units. The α-diversity map can be shown by directly taking into account Shannon's H' for each unit and reporting it in a final resampled map. Rather, the β-diversity map needs a further step to pass from a distance matrix to a 2D spatial representation, namely the application of Non metric Multidimensional Scaling (NMDS), by deciding the final number of reduced dimensions (in this case three in order to compose an RGB image). The final result will be a unitless β-diversity map in which different colours represent differences among species communities in space.
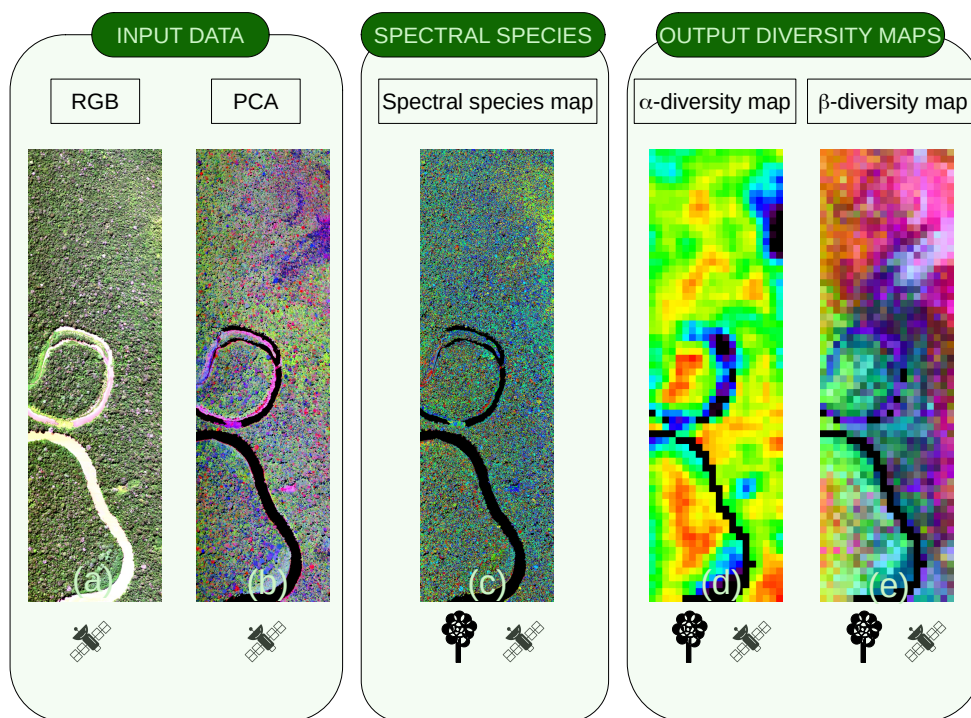
29

Figure 2: **Box 1 Figure** - The spectral species algorithm phases. The original image was acquired with the CAO AToMS imaging spectrometer during an airborne campaign over the CICRA experimental site (Amazonian Peru) (`https://www.amazonconservation.org/about/mission-vision/cicra-station/`). (a) corresponds to the RGB representation of an imaging spectroscopy subset. A standardized PCA is applied on (a) and a reduced set of components is selected (b), to maximize signal corresponding to biological patterns on forested areas and discard noisy components. Spectral species are defined for each pixel by applying an unsupervised k-means clustering on the spectral space defined by selected components (c). In this phase, a field survey recognition based on in-situ data is crucial to define the number of singular spectral signatures (spectral species) expected. The spectral species map is divided into elementary spatial units and the spectral species inventory is performed for each spatial unit, by further calculating Shannon's H and Bray-Curtis metrics to derive alpha- ((d), ranging here from minima to maxima from black to blue, green and red) and beta-diversity ((e), in which colours represent differences among spectral species) maps, respectively.

30