



## “Flower power”: How flowering affects spectral diversity metrics and their relationship with plant diversity

Michela Perrone<sup>a,\*</sup>, Luisa Conti<sup>a</sup>, Thomas Galland<sup>b</sup>, Jan Komárek<sup>a</sup>, Ondřej Lagner<sup>a</sup>, Michele Torresani<sup>c</sup>, Christian Rossi<sup>d</sup>, Carlos P. Carmona<sup>e</sup>, Francesco de Bello<sup>b,f</sup>, Duccio Rocchini<sup>a,g</sup>, Vítězslav Moudrý<sup>a</sup>, Petra Šímová<sup>a</sup>, Simonetta Bagella<sup>h</sup>, Marco Malavasi<sup>h</sup>

<sup>a</sup> Department of Spatial Sciences, Czech University of Life Sciences Prague, Kamýcká 129, 165 00 Praha-Suchbát, Czech Republic

<sup>b</sup> Department of Botany, University of South Bohemia, Na Zlaté stoce 1, 370 05 České Budějovice, Czech Republic

<sup>c</sup> Free University of Bolzano-Bozen, Faculty of Agricultural, Environmental and Food Sciences, Universitätsplatz 5 - piazza Università, 5, 39100 Bozen-Bolzano, Italy

<sup>d</sup> Department of Geoinformation, Swiss National Park, Runatsch 124, Chasè Planta-Wildenberg, 7530, Zernez, Switzerland

<sup>e</sup> Department of Botany, University of Tartu, J. Liivi 2, 50409 Tartu linn, Estonia

<sup>f</sup> Centro de Investigaciones sobre Desertificación, Universidad de Valencia, Crta. Moncada-Náquera Km 4,5, 46113 Moncada (Valencia), Spain

<sup>g</sup> BIOME Lab, Department of Biological, Geological and Environmental Sciences, Alma Mater Studiorum University of Bologna, Via Irnerio, 42, 40126 Bologna, Italy

<sup>h</sup> Department of Chemistry, Physics, Mathematics and Natural Sciences, University of Sassari, Via Vienna 2, 07100 Sassari, Italy

### ARTICLE INFO

#### Keywords:

Grasslands  
Remote sensing  
Phenology  
UAV  
Spectral variation hypothesis

### ABSTRACT

Biodiversity monitoring is constrained by cost- and labour-intensive field sampling methods. Increasing evidence suggests that remotely sensed spectral diversity (SD) is linked to plant diversity, holding promise for monitoring applications. However, studies testing such a relationship reported conflicting findings, especially in challenging ecosystems such as grasslands, due to their variety and high temporal dynamism. It follows that a thorough investigation of the key factors influencing these relationships, such as the metrics applied (i.e., continuous, categorical) and phenology (e.g., flowering), is necessary. The present study aims to assess the effect of flowering on the applicability of six different SD metrics for plant diversity monitoring at the local scale and to investigate how spatial resolution affects the results. Taxonomic diversity was calculated based on data collected in 159 plots of 1.5 m × 1.5 m with experimental mesic grassland communities. Spectral information was collected using a UAV-borne sensor measuring reflectance across six bands in the visible and near-infrared range at ~2 cm spatial resolution. Our results showed that, in the presence of flowering, the relationship between SD and plant diversity is significant and positive only when SD is calculated using categorical metrics. Despite the observed significance, the variance explained by the models was very low, with no evident differences when resampling spectral data to coarser pixel sizes. Such findings suggest that new insights into the possible confounding effects on the SD-plant diversity relationship in grassland communities are needed to use SD for monitoring purposes.

### 1. Introduction

The current global biodiversity change rate makes species conservation one of the most pressing priorities of our times (Wilting et al., 2017). While growing attention has been paid to species conservation, biodiversity evaluation and monitoring are still limited and lack standardized methods for quick and scalable data collection (Palmer et al., 2002; Skidmore et al., 2015; Wang and Gamon, 2019). In turn, remote sensing represents a continuous source of data which potentially concerns various aspects of biodiversity, collected across diverse scales in a

uniform, borderless, and repeatable way (Turner, 2014). Among the available approaches for the remote monitoring of biodiversity, increasing evidence suggests that remotely sensed spectral diversity (SD), i.e., the variability in the electromagnetic radiation reflected by (vegetated) surfaces measured in the visible, near-infrared and short-wave infrared regions (400–2400 nm), is linked to plant diversity (Rocchini et al., 2015; Schweiger et al., 2018). The SD concept was originally framed within the spectral variation hypothesis (SVH; Palmer et al., 2002), which assumes an indirect relationship between spectral and plant diversity through environmental ‘surrogacy’, i.e., the

\* Corresponding author.

E-mail address: [perrone@fzp.czu.cz](mailto:perrone@fzp.czu.cz) (M. Perrone).

<https://doi.org/10.1016/j.ecoinf.2024.102589>

Received 21 December 2023; Received in revised form 5 April 2024; Accepted 5 April 2024

Available online 9 April 2024

1574-9541/© 2024 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

landscape heterogeneity captured by remote sensing (Hauser et al., 2021; Torresani et al., 2019). When relying on medium/coarse spatial resolution data, the link between spectral and plant diversity can only be indirect, as pixel size exceeds the size of individual plants. Conversely, multispectral data with fine spatial resolution provided by uncrewed aerial vehicles (UAVs), i.e., in the order of magnitude of centimeters, hold promise for capturing the direct link between the spectral and plant diversity at the leaf and canopy level in herbaceous communities (Rossi et al., 2021a; Thornley et al., 2023). However, when the pixel size corresponds to the size of individual plants (or even the sub-individual level), the intra-specific (or intra-individual) variability in the plant optical traits can play a role in the SD-biodiversity relationship.

Studies testing the relationship between SD and plant diversity at fine scales have reported conflicting findings, especially in challenging ecosystems such as grasslands, the biodiversity of which is threatened by their degradation and fragmentation due to anthropogenic factors (Thornley et al., 2023). Their vast global distribution (Gibson, 2009) and the broad range of ecosystem services they provide make grassland conservation a priority (Hein et al., 2006; Suding, 2011), yet their extensive monitoring is prevented by logistical constraints. Due to its efficiency and cost-effectiveness, SD-based plant diversity monitoring offers an appealing alternative to traditional sampling techniques in grasslands.

Nonetheless, the path towards the routine implementation of SD into grassland monitoring is hampered by the high temporal and spatial variability of such ecosystems, which often exhibit a complex community structure, particularly in natural or semi-natural grasslands (Wilson et al., 2012). Recent studies have highlighted the impact of plant phenology (e.g., flowering, leaf emergence and senescence; Rossi et al., 2021b; Thornley et al., 2022) when addressing the relationship between SD and plant diversity in grasslands. Despite the dynamism of grassland communities, acquiring repeated intra-annual spectral data is time- and resource-consuming. It follows that spectral data are usually acquired to ensure a temporal match with vegetation surveys when the presence of dead biomass and exposed soil is supposed to be the lowest (Asner, 1998). However, this also implies that the species present would potentially be flowering, causing unwanted additional sources of spectral variability. The presence of flowering is likely to have an impact on the SD of a community, as it leads to a high variation in important optical traits and, thus, in the spectral signatures of individual species (Conti et al., 2021; Fassnacht et al., 2022; Gholizadeh et al., 2019; Heumann et al., 2015; Thornley et al., 2022). Being non-photosynthetic organs with distinct spectral features, especially in the visible domain (Schiefer et al., 2021), flowers can affect SD, and the extent of this impact is likely to depend on the characteristics of the SD metrics applied.

The metrics used to quantify SD can be either continuous (i.e., based on variation in traditional vegetation reflectance indices or on the full spectral information) or categorical (i.e., relying on the categorization of the spectral space; Féret and Asner, 2014; Wang and Gamon, 2019). Although continuous SD metrics are the most represented and widely tested, they show weakness in capturing plant diversity by measuring the degree of contrast in the reflectance, which is more related to the specific identity of the plant species rather than their quantity. Indeed, in cases where a few spectrally distinct species are present in a community, the spectral heterogeneity can be very high even if the plant diversity is low (Fassnacht et al., 2022). Following the same reasoning, continuous SD metrics are likely to be influenced by flowering, soil presence, or the amount of biomass. In contrast, categorical SD metrics are considered less sensitive to extreme reflectance values (e.g., from background material) as extreme values would represent distinct categories alongside other equally significant ones (Rossi et al., 2021a). Even though some studies have tested and compared different SD metrics (Frye et al., 2021; Gholizadeh et al., 2018, 2020; Perrone et al., 2023; Rossi et al., 2021a; Schmidlein and Fassnacht, 2017; Warren et al., 2014), no consensus has been reached on which metrics would be the

best proxy for plant diversity during flowering.

The aim of this study is to assess the applicability of SD for plant species richness monitoring at the local scale. Here we test six different SD metrics (detailed in Table 1) while considering flowering as a potential factor affecting the spectral-plant diversity relationship using ground observations and multispectral UAV data from mesic grassland communities in South Bohemia (Czech Republic). Additionally, we investigate if changing the spatial resolution (from 2 cm to 5 cm) influences the observed relationship. Building upon Conti et al. (2021) findings, this study also considers the influence of bare soil and shadows on the relationship between spectral and plant diversity.

## 2. Material and methods

### 2.1. Field data

Our study is based on vegetation data collected in a permanent grassland experiment run by the University of South Bohemia and located in a mesic meadow in the Vysočina region (South Bohemia, Czech Republic, 49.331 N, 15.003E). As Galland et al. (2019) described, the sowing experiment consisted of 40 mesic grassland communities covering independent gradients of plant functional and phylogenetic diversity. These experimental communities derived from a specific treatment, i.e., either the sowing of a combination of six species drawn from a pool of 19 mesic meadow species naturally present in the area or

**Table 1**  
Description of the spectral diversity metrics used in this study.

Spectral diversity metric	Description	Reference	Dedicated package
Standard deviation of NDVI (sdNDVI)	The square root of the variance in the NDVI values	(Wang et al., 2016)	
Rao's Q entropy index (RaoQ)	This index quantifies the difference in reflectance values between two pixels drawn randomly with replacement from a set of pixels by considering their abundance and their pairwise distance	(Rocchini et al., 2021a)	rasterdiv R package, v. 0.3.2 (Rocchini et al., 2021c)
Mean pairwise distance (MPD)	Mean pairwise Euclidean distance between pixels in the space defined by the first two principal components of the spectral data	(Rocchini et al., 2004)	
Convex hull volume (CHV)	The smallest possible convex volume encompassing all pixels using the first three principal components of the spectral data	(Dahlin, 2016)	
Coefficient of variation of reflectance (CVrefl)	Average coefficient of variation over all available wavelengths	(Wang et al., 2016)	
Spectral species richness (SpSpR)	The number of spectral species within each plot. Spectral species are identified based on k-means clustering of a random subset of the first three components of the UAV image and then applied to the whole PCA image to assign a cluster to each pixel.	(Féret and Asner, 2014)	biodivMapR R package, v.1.9.4 (Féret and de Boissieu, 2020, 2022)

from monoculture treatments. Each six-species community was sown in two randomly situated  $1.5\text{ m} \times 1.5\text{ m}$  plots (one fertilized and one unfertilized), while monoculture plots were sown in three replicates each. Thus, the study site encompassed a total of 196 plots, with a buffer zone of 0.5 m between them (Fig. 1). Due to the subsequent processing of UAV data (i.e., excluding corrupted areas of the orthomosaic), only 159 out of the 196 plots were considered in this study. Plant species composition within each plot was measured in June 2019. The actual species richness (SR) of the considered experimental communities varied between 12 and 36 species due to the later plot invasion by spontaneously colonising species.

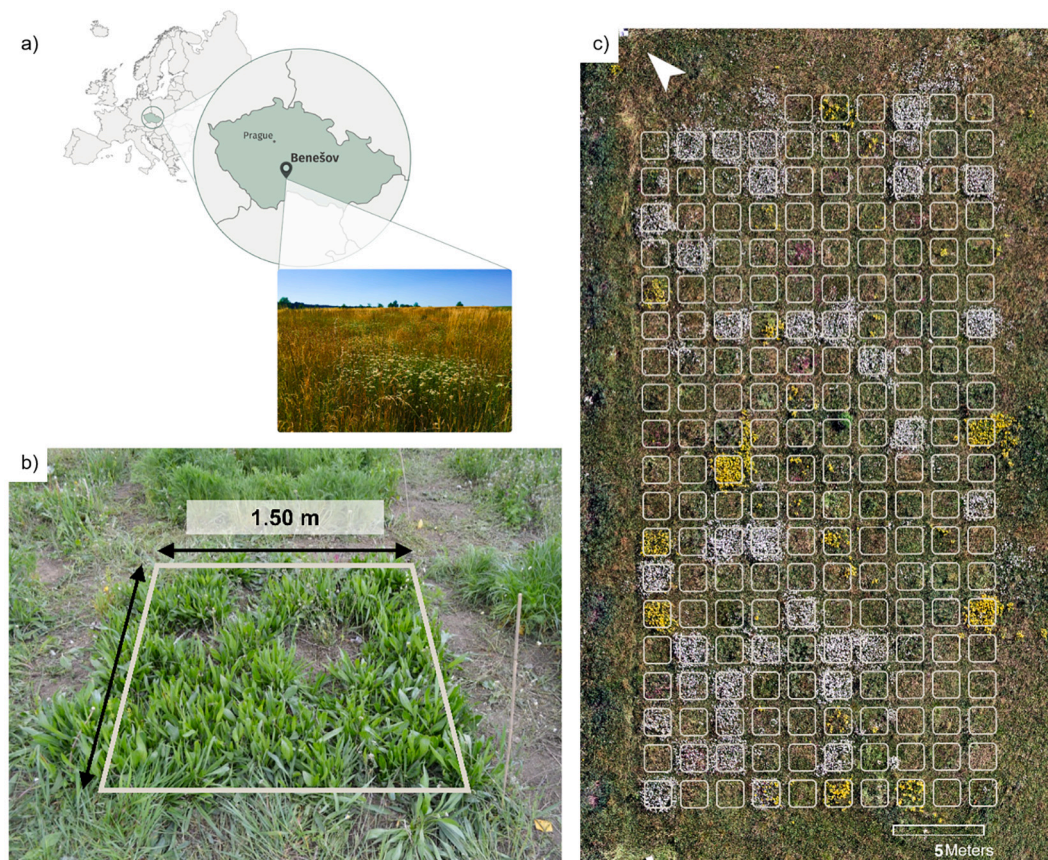
## 2.2. Spectral data

We used multispectral UAV imagery with  $\sim 2\text{ cm}$  spatial resolution (80% frontal and 70% side tile overlapping, 30 m flight altitude) acquired shortly after vegetation sampling (July 2019) using a Kingfisher multicopter (Robodrone Industries, Brno, Czech Republic) equipped with a Micro-MCA6 camera (Tetracam Inc., Chatsworth, CA, United States) measuring reflectance across six bands in the 490–900 nm range of the electromagnetic spectrum with a 10–20 nm-width. The obtained spectral data were processed using a Structure from Motion and Multi-View Stereo algorithms in the Agisoft Metashape image-matching software (Agisoft LLC, St. Petersburg, Russia) to generate an ortho-mosaic. Georeferencing was performed with spatial error of 2 cm using ground control points surveyed through a Leica GPS1200 GNSS receiver (Leica Geosystems AG, Heerbrugg, Switzerland) in RTK mode. The ortho-mosaic was radiometrically calibrated using white and grey calibration targets for which the spectral properties were known from manufacturer-provided spectrometer measurements. The ortho-mosaic portions exhibiting processing artefacts were excluded from further analyses, resulting in 37 out of 196 study plots being left out.

Previous studies have reported that the presence of bare soil, due to its spectrum being substantially different from that of the vegetation, proved to significantly deteriorate SD performance in quantifying plant diversity (Gholizadeh et al., 2018; Hauser et al., 2021). Moreover, the spectral signal has proven to be sensitive to the vertical structure of the plant community, since a complex vertical structure of the community may generate an “occlusion effect” that would lead to the obscuration of shorter species and the presence of shadowed pixels (Conti et al., 2021). Here, we masked shadowed areas and non-vegetated pixels from the multiband ortho-mosaic to mitigate their influence on SD. Bare soil pixels were masked using an NDVI threshold of 0.3, while shadowed pixels were identified and removed through an Expectation Maximization (EM) unsupervised classification. Moreover, the presence of flowering plants within the plots was assessed visually from an RGB composite and attributed to each plot (i.e., a plot was classified as flowering given  $>10\%$  flower cover). Finally, we rescaled the ortho-mosaic to 5-cm spatial resolution (using bilinear resampling) to test if coarsening the spatial resolution can mitigate the effect of intraspecific (and intra-individual) variability of optical traits.

## 2.3. Spectral diversity metrics

As there is no consensus on the best-performing method to quantify SD, we applied six different SD metrics within the  $1\text{ m} \times 1\text{ m}$  core area of each  $1.5\text{ m} \times 1.5\text{ m}$  experimental plot. Specifically, we computed six continuous SD metrics, namely: the standard deviation of NDVI (sdNDVI), the Rao’s Q entropy index (RaoQ), the mean pairwise distance of reflectance values (MPD), the coefficient of variation of the full spectral range reflectance (CVrefl), and the convex hull volume (CHV). We also computed one categorical metric, spectral species richness (SpSpR) (see Table 1 for further details on the metrics used). For the spectral species mapping we used the `biodivMapR` R package (for a



**Fig. 1.** (a) Location of the study area. (b) Example of a  $1.5\text{ m} \times 1.5\text{ m}$  surveyed plot. (c) Natural-colour mosaic of the whole study area.

complete description of the method, please see [Féret and de Boissieu, 2020](#)). To perform the mapping, the number of  $k$  clusters was set to 50 (a realistic number of the spectral species present based on ground data and as a trade-off between performance and computational efficiency). The unsupervised clustering of a random subset of pixels was repeated eight times and the mean SpSpR value calculated from these repetitions was assigned to each respective plot.

### 2.4. Statistical analyses

To test the relationship between SR and each of the metrics tested, we modelled the variation in SD through Generalized Linear Mixed Effects Models ([McCulloch and Neuhaus, 2014](#)) (Gamma family, log link) using the `lme4` R package *v.1.1-33* ([Bates et al., 2015](#)). In our models, SR and the binary variable of flowering (presence/absence) were used as fixed-effect predictors. At the same time, the type of grassland community (i.e., monoculture, low/high functional diversity, low/high phylogenetic diversity, fertilized/unfertilised) was included as a random effect to account for the interdependence among communities in our dataset. To assess the goodness of fit for our models, we calculated the marginal ( $R_m^2$ ) and conditional ( $R_c^2$ ) pseudo-R squared ([Nakagawa et al., 2017](#)) for each model. These metrics quantify the proportion of variance explained by fixed effects and the complete model (i.e., both fixed and random effects), respectively. We calculated  $R_m^2$  and  $R_c^2$  using the `r.squaredGLMM()` function of the `MuMIn` R package, *v.1.43.17*. Data and scripts are provided at [https://github.com/MichelaPerrone/SVH\\_Bene\\_sov.git](https://github.com/MichelaPerrone/SVH_Bene_sov.git) under CC-BY license.

### 3. Results

When comparing the spectra of two sample plots (i.e., one with and one without flowering), significant differences were observed in the distribution of reflectance values in all six spectral bands (Kolmogorov-Smirnov Test  $p$ -values  $< 0.01$ ) at both spatial resolutions tested. [Fig. 2](#) shows the spectra and pixel values distributions for the 2-cm ortho-

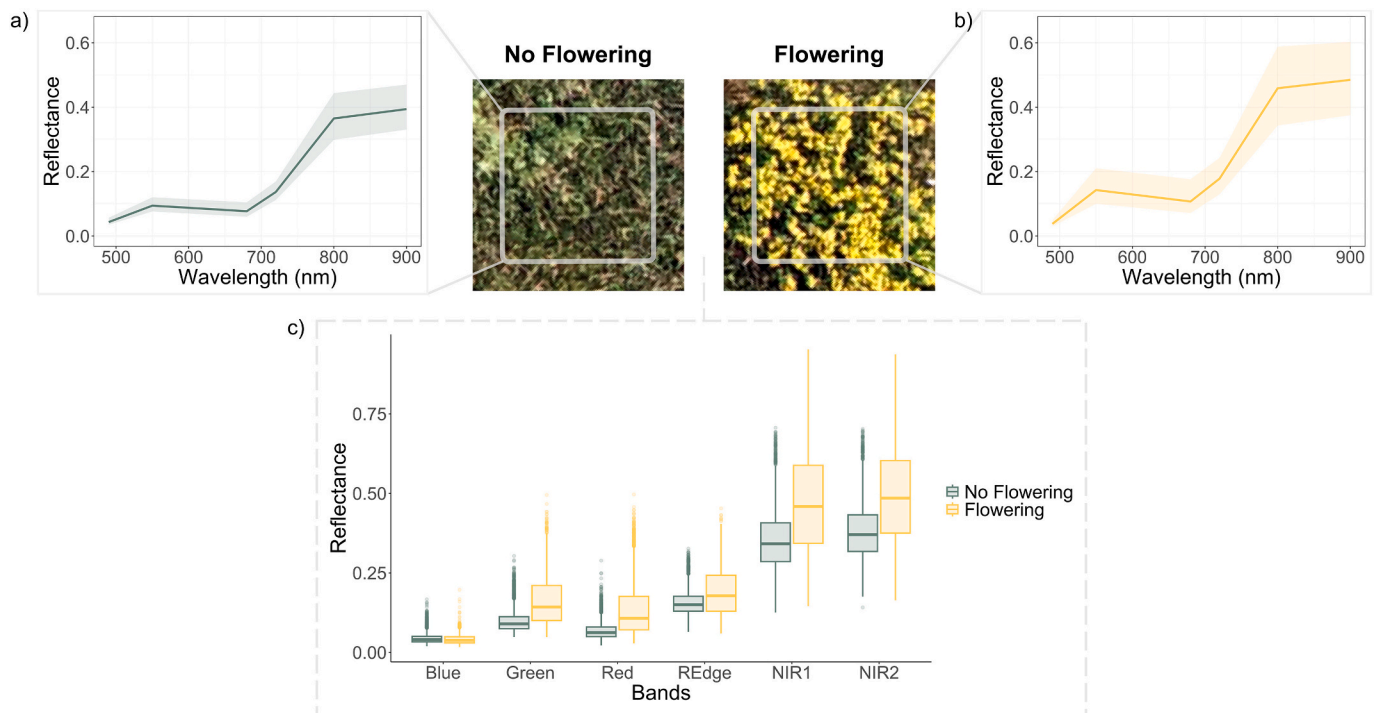
mosaic; the same information for the 5-cm ortho-mosaics can be found in [Appendix A](#). The strongest differences (i.e., larger effect sizes, or higher Kolmogorov's  $D$  values) in the distribution of values between the experimental plots were observed for the Green, Red, and NIR bands ([Table 2](#)), indicating that flowering affects both the visible and NIR regions of the spectrum. Besides, we observed that coarsening the spatial resolution from 2 to 5 cm did not impact the results above.

At both the tested spatial resolutions, the models showed a significant and positive relationship between SR and SD only when SD was measured using categorical metrics (i.e., spectral species richness) ([Table 3](#); [Fig. 3](#)). Conversely, when the SD response variable was calculated using continuous metrics, flowering was the only variable significantly (positively) associated with SD ([Table 3](#); [Fig. 3](#)). In such models, we observed relatively high  $R_m^2$  values, ranging between 0.24 (sdNDVI, 2-cm spatial resolution) and 0.51 (CVrefl, 2-cm spatial resolution). In contrast, models based on spectral species richness had very low  $R_m^2$  values ([Table 3](#); [Fig. 4](#)), with no evident differences observed between the two spatial resolutions. Despite the significant relationship between SD and SR, the low variance explained by the fixed effects in the spectral species richness models suggests that other factors play a more prominent role at this ecological scale and within these settings. Moreover, we observed that very fine-spatial resolution data, such as data

**Table 2**

D statistic values of the Kolmogorov-Smirnov tests evaluating the differences between the reflectance values in the two experimental plots shown in [Fig. 2](#). Significance codes: \*\*\* (p-value  $< 0.001$ ), \*\* (p-value  $< 0.01$ ), \* (p-value  $< 0.05$ ).

	Blue	Green	Red	Red Edge	NIR1	NIR2
Spatial resolution: 2 cm	0.12***	0.43***	0.43***	0.28***	0.37***	0.39***
Spatial resolution: 5 cm	0.12***	0.48***	0.53***	0.32***	0.40***	0.41***

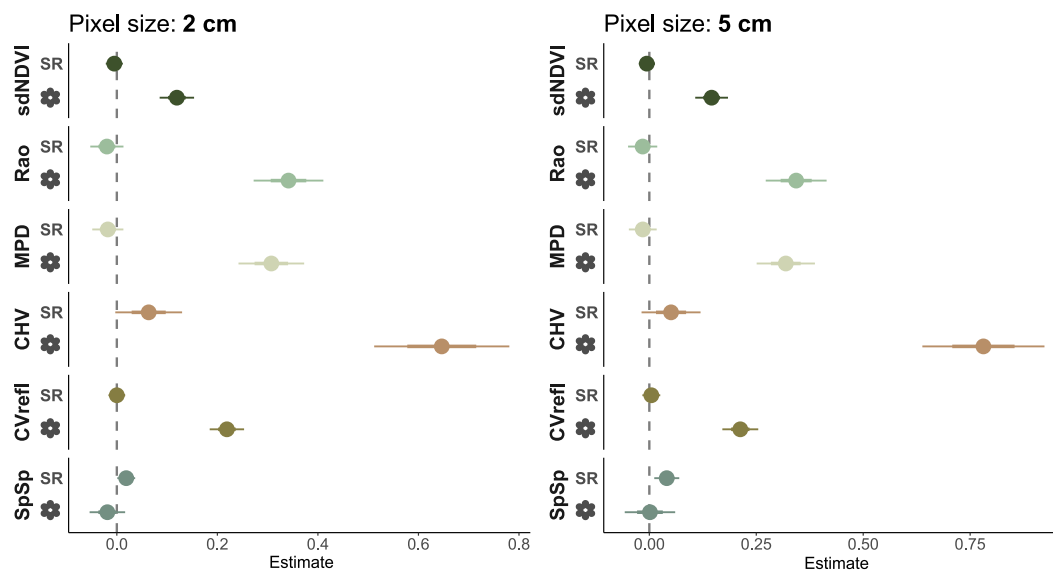


**Fig. 2.** Sample spectra of experimental plots without (a) and with (b) flowering from the 2-cm resolution ortho-mosaic after masking. Bold lines represent the median spectra within the  $1\text{ m} \times 1\text{ m}$  core area of each plot (after masking). The shadowed area represents the interval of the pixel values between the 1st and 3rd quartiles for each band. (c) Box plot of the reflectance values within the two sample experimental plots selected.

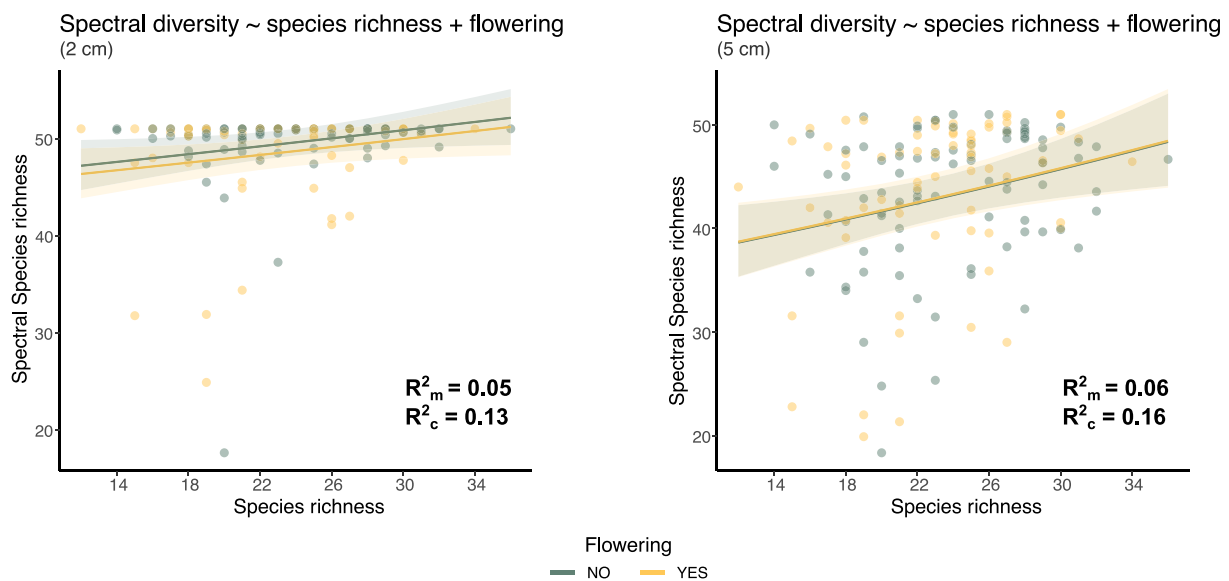
**Table 3**

Summary of the models' results. The columns refer to the SD metric used as the response variable; rows refer to the coefficient estimates (coeff. est.) of the fixed-effect explanatory variables (log-transformed values) and the models' marginal and conditional  $R^2$  values. Significance codes: \*\*\* (p-value <0.001), \*\* (p-value <0.01), \* (p-value <0.05).

Spatial resolution: 2 cm							
	sdNDVI	RaoQ	MPD	CHV	CVrefl	SpSpR	
SR coeff. est.	$-4.99e^{-3}$	-0.02	-0.02	0.06	$-1.89e^{-5}$	0.02*	
Flowering coeff. est.	0.12***	0.34***	0.30***	0.65***	0.22***	-0.02	
$R_m^2$	0.24	0.39	0.37	0.35	0.51	0.05	
$R_c^2$	0.29	0.50	0.48	0.44	0.58	0.13	
Spatial resolution: 5 cm							
	sdNDVI	RaoQ	MPD	CHV	CVrefl	SpSpR	
SR coeff. est.	$-5.94e^{-3}$	-0.02	-0.02	0.05	$4.66e^{-3}$	0.04**	
Flowering coeff. est.	0.15***	0.34***	0.32***	0.78***	0.21***	$1.46e^{-3}$	
$R_m^2$	0.28	0.37	0.36	0.41	0.38	0.06	
$R_c^2$	0.29	0.49	0.47	0.48	0.50	0.16	



**Fig. 3.** Coefficient plots of the mixed models illustrating the estimated (log-transformed) values for each explanatory variable, along with their corresponding 50% (inner error bars) and 95% (outer error bars) confidence intervals. Colours refer to the individual SD metrics used as the response variable.



**Fig. 4.** Predicted relationships between SD (calculated as spectral species richness) and species richness derived from the mixed-effect models for the two spatial resolutions tested based on the presence of flowering (see legend).

with a 2-cm resolution, are not beneficial when using categorical metrics. Indeed, with such a fine spatial resolution, we observed that a high number of spectral species are detected in most of the plots, regardless of the actual ground species richness (Fig. 4).

## 4. Discussion

### 4.1. SD metrics and flowering

The community's phenological pattern impacts the effective application of SD for biodiversity monitoring as plant species exhibit distinct physiological and structural characteristics at different phenological stages, causing changes in their spectral signatures. The presence of blooms (that are markedly different in their spectral characteristics from the plant green organs) impacts the use of continuous SD metrics as such metrics are highly sensitive to extreme reflectance values. On the contrary, our results confirm that by using categorical SD metrics, extreme pixel values are less likely to have a disproportionate influence on the estimated SD, since they would represent distinct categories alongside others of equal importance (Wang and Gamon, 2019). Notably, increasing the range of spectral information considered, i.e., from bands at specific wavelengths (sdNDVI) to all available spectral bands (RaoQ, MPD, CHV, CVrefl), did not substantially improve the performance of the continuous metrics, thus indicating that using only a subset of the available spectral information is not what limits the successful application of SD. To use any of the tested continuous SD metrics, the spectral data acquisition timing turned out to be relevant. Indeed, the asynchrony in phenology among species present in the observed community presented an additional challenge to the applicability of the SVH for estimating biodiversity, as the number of co-present phenological stages (both at the leaf- and flower-level) in a plot confounds the observed SD (Fassnacht et al., 2022; Thornley et al., 2022). Thus, the timing of data acquisition is paramount in identifying possible associations between SD and species diversity. For this reason, categorical metrics are highly advantageous in this context, as they ideally enable overcoming the issue of seasonal change in the optical properties of plant species due to their phenology. Indeed, this type of metric could allow the identification of spatial units with consistent spectral homogeneity through time (Perrone et al., 2023).

Despite their advantages, the successful use of categorical metrics for estimating spectral diversity can be challenged by the necessity of selecting appropriate settings by the user, in particular, the optimal number of clusters into which the spectral space should be partitioned. While it is possible to use virtual dimensionality to get the number of clusters in the case of hyperspectral data, the situation is different in the case of multispectral data; there, it requires going through a time-consuming trial-and-error procedure that should consider the specific plant communities observed, the reliability of the results, and the computation intensity (Féret and de Boissieu, 2020; Perrone et al., 2023; Rocchini et al., 2021b).

The critical role played by the spatial resolution of the data used should be noted as well. Previous studies on SD highlighted how pixel size affects the relationship between spectral and plant diversity, which becomes considerably weaker with progressively coarser resolutions (Gholizadeh et al., 2018, 2019; Wang et al., 2016). However, the dependence of the relationship on pixel size is contingent on the SD metrics used (Gholizadeh et al., 2018) and sampling plot size, as well as on the specifics of the study site (Gholizadeh et al., 2022). Thus, we argue that when relying on very fine spatial resolutions that allow capturing individual objects, the risk is to introduce noise (e.g., due to shadows) into the data (Conti et al., 2021), which could lead to exceeding the number of clusters (i.e., of spectral species) identified within each area unit (e.g., plot) (Fig. 3). Therefore, users must strike a balance between spatial resolution and the desired level of spectral discrimination to avoid this issue and obtain meaningful results.

### 4.2. Confounding factors

The complexity of the herbaceous communities' vertical structure can lead to a negative relationship between SD and plant diversity due to the "occlusion effect" caused by taller species (Conti et al., 2021). Given such previous knowledge, we applied a shadow and bare soil masking on the original ortho-mosaics to improve the correlation and focus on the specific issues we wanted to test (i.e., the influence of flowering on SD). Indeed, as shown by Gholizadeh et al. (2018), filtering out pixels that capture information from non-vegetation sources improves the performance of SD metrics, allowing a (partial) correction for such confounding factors. In principle, such correction is possible when relying on data with fine spatial resolution (e.g., proximal or UAV-borne sensors) since the spectral signature of individual pixels is more likely to belong to a single object type. Moving to coarser spatial resolutions, such as with airborne and spaceborne sensors, spectral unmixing would be required to correctly identify mixed pixels and correct for bare soil presence, which allows estimating the per-pixel percentage of constituent spectra, or endmembers (Asner and Heidebrecht, 2002; Gholizadeh et al., 2018; Rossi and Gholizadeh, 2023).

In light of the low explanatory power of the models based on spectral species richness, additional confounding factors may have played a role in the computed SD. As highlighted by Rossi et al. (2021a), the presence and ratio of live and dead biomass, together with the total biomass (Villoslada et al., 2020), may lead to higher spectral variability. The impact varies depending on ecosystem type (Rossi et al., 2021a) and specific plant community characteristics, particularly at fine spatial resolutions (Villoslada et al., 2020), emphasizing the contextual dependence of the spectral-plant diversity relationship (Fassnacht et al., 2022). Variables such as species composition, life forms (e.g., graminoids, forbs, legumes), and their prevalence within the community may also affect SD metrics (Gholizadeh et al., 2019; Wang et al., 2018b; Rossi et al., 2021a). As an example, Imran et al. (2021) observed a stronger SD-plant diversity link in artificial, species-poor ecosystems compared to species-rich natural grasslands. Despite being an experimental site, our study area has undergone partial re-naturalisation due to the spontaneous colonisation by local species, increasing species richness within each plot. Such a decrease in the strength of the SD-plant diversity link in high species richness habitats further hinders the operational applicability of SD for biodiversity monitoring purposes in natural and semi-natural habitats.

### 4.3. Limitations

In the present study, we were interested in assessing if the simple presence of flowering has a confounding effect on the relationship between plant and spectral diversity measured through different metric types. Therefore, flowering was included in the analyses as a binary (presence/absence) variable. We did not attempt a qualitative or quantitative estimation of the flowering presence, as it would have diverted from our original aim. While we acknowledge that gathering information on the specific flower colour, size, and spatial cover would help characterize the presence of flowering within each plot, such information could not be obtained given the data available (both in terms of ground and spectral data). Nevertheless, characterizing and assessing the presence of flowering based on specific flower types is a promising topic for future research, which could include combining hyperspectral data with species-level ground data on flowering plants. This could help in acquiring further insight into the confounding effects of flowering on SD.

Moreover, the experimental setup on which our study is based would have determined the impossibility of further testing the influence of spatial resolution. Indeed, due to the small core plot size (1 m × 1 m) and pixel masking to reduce the impact of bare soils and shadowed areas, further rescaling (i.e., over a 5-cm spatial resolution) would have resulted in few pixels available for SD computation and, thus, would

have introduced a major source of error (i.e., small sample size). Besides, our aim was to assess the direct link between SD and plant diversity, which implies the matching between pixel and individual size. Thus, coarsening the spatial resolution further would have been in contrast with our assumptions. Nevertheless, according to what has been previously shown by Fassinacht et al. (2022), we suppose that, at resolutions coarser than the ones tested here, the smoothing effect on extreme values would have resulted in lower SD values. However, we argue that (given the sensitivity of continuous SD metrics to extreme values) coarsening the spatial resolution would reduce, but not resolve, the overestimation of ground plant diversity when using such metrics in the presence of flowering. Additional analyses testing the impact of further coarsening the pixel size on the ability to map plant diversity of flowering communities using SD could be beneficial; however, given the small size of individual plots in our setting, this was not possible to perform within the frame of this study.

Additionally, we did not consider abundance-based indices (e.g., Shannon's H) for both plant and spectral diversity due to the possible mismatch between the actual field-sampled species abundance and the retained spectral information after pixel masking. However, while species richness is the most widespread species diversity metric in SD studies, we acknowledge that abundance-based metrics have proven to be more strongly related to SD on several occasions (Oldeland et al., 2010; Torresani et al., 2019; Wang et al., 2018).

Finally, due to the unavailability of total and relative (live, dead) biomass data, as well as data on life forms abundance, we were unable to take this type of confounding factors into account in our study. To gain a deeper understanding of how all the factors identified so far (i.e., bare soil, phenological features, biomass, vertical complexity, community composition) and their potential interactions affect the plant-spectral diversity relationship, future experiments should be designed to consider them simultaneously and assess their impact on different types of grasslands.

## 5. Conclusions

The reliability of using SD to monitor plant diversity is a matter of controversy and may need more consistency in specific settings, especially in dynamic ecosystems such as grasslands. In this study, we investigated the SD-biodiversity relationship in mesic grassland communities by testing how flowering may influence the estimation of plant species richness using different SD metrics. According to our results, the presence of flowering proved to impair the ability of continuous SD metrics to reflect plant diversity, with flowering likely being the main source of spectral variance within plots. On the contrary, categorical SD metrics appear less influenced by flowering, confirming the better suitability of this type of metrics observed in previous studies (Rossi et al., 2021a). Nevertheless, when calculating SD using categorical metrics, species richness only explains a small portion of the variability in spectral heterogeneity at both spatial resolutions tested. We hypothesise that such a low explanatory power should be ascribed to the presence of additional confounding factors (e.g., dead biomass,

community composition) that have previously proven to interfere with the estimation of grassland diversity (Gholizadeh et al., 2019; Rossi et al., 2021a; Schweiger et al., 2015; Villoslada et al., 2020). Thus, we encourage future investigations to systematically consider all possible confounding factors when testing the spectral diversity-biodiversity relationship in different types of grasslands. Finally, to define the actual possibilities and technical constraints of the relationship, future research should aim at identifying the optimal trade-off between the spatial and spectral resolutions of the remote sensing data used to assess plant diversity in grasslands while incorporating the temporal variations in the spectral signal. In this framework, comparing spectral signatures over an entire growing season (spatio-temporal spectral diversity) could be critical in estimating plant diversity.

## CRedit authorship contribution statement

**Michela Perrone:** Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation. **Luisa Conti:** Validation, Methodology, Formal analysis. **Thomas Galland:** Data curation. **Jan Komárek:** Writing – original draft, Software, Data curation. **Ondřej Lagner:** Software, Data curation. **Michele Torresani:** Writing – original draft, Software, Methodology. **Christian Rossi:** Writing – original draft, Software, Methodology. **Carlos P. Carmona:** Data curation. **Francesco de Bello:** Data curation. **Duccio Rocchini:** Writing – original draft, Validation, Supervision, Software. **Vítězslav Moudrý:** Writing – original draft, Supervision. **Petra Šimová:** Supervision, Resources. **Simonetta Bagella:** Supervision. **Marco Malavasi:** Writing – original draft, Supervision, Methodology, Conceptualization.

## Declaration of competing interest

None

## Data availability

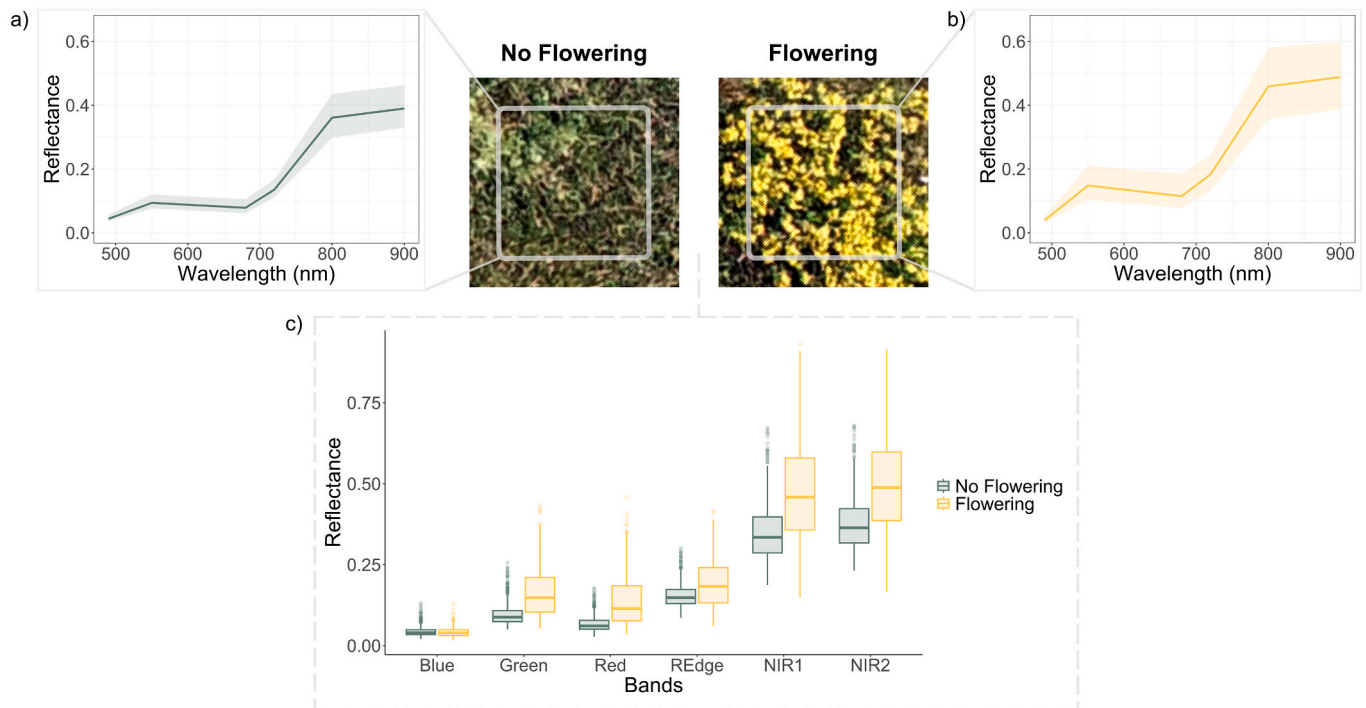
I have shared the link to the data and scripts in the manuscript.

## Acknowledgements

We thank Riccardo Testolin and Piero Zannini for useful insights and suggestions that greatly served this study. We thank Elisa Thouverai for technical support in the use of the rasterdiv R package.

MP was supported by IGA grant from the Faculty of Environmental Sciences at the Czech University of Life Sciences Prague (project n. 2022B0003). MT and DR were partially supported by the H2020 project SHOWCASE (Grant agreement No 862480). This study was funded within the Horizon Europe project EarthBridge (ID: 101079310). Funded by the European Union. Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

## Appendix A



**Fig. A1.** Sample spectra of experimental plots without (a) and with (b) flowering from the 5-cm resolution ortho-mosaic after masking. Bold lines represent the median spectra within the 1 m × 1 m core area of each plot (after masking). The shadowed area represents the interval of the pixel values between the 1st and 3rd quartiles for each band. (c) Box plot of the reflectance values within the two sample experimental plots selected

## References

- Asner, G.P., 1998. Biophysical and biochemical sources of variability in canopy reflectance. *Remote Sens. Environ.* 64, 234–253. [https://doi.org/10.1016/S0034-4257\(98\)00014-5](https://doi.org/10.1016/S0034-4257(98)00014-5).
- Asner, G.P., Heidebrecht, K.B., 2002. Spectral unmixing of vegetation, soil and dry carbon cover in arid regions: comparing multispectral and hyperspectral observations. *Int. J. Remote Sens.* 23, 3939–3958. <https://doi.org/10.1080/01431160110115960>.
- Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. <https://doi.org/10.18637/jss.v067.i01>.
- Conti, L., Malavasi, M., Galland, T., Komárek, J., Lagner, O., Carmona, C.P., de Bello, F., Rocchini, D., Šímová, P., 2021. The relationship between species and spectral diversity in grassland communities is mediated by their vertical complexity. *Appl. Veg. Sci.* 24. <https://doi.org/10.1111/avsc.12600>.
- Dahlin, K.M., 2016. Spectral diversity area relationships for assessing biodiversity in a wildland-agriculture matrix. *Ecol. Appl.* 26, 2756–2766. <https://doi.org/10.1002/eap.1390>.
- Fassnacht, F.E., Müllerová, J., Conti, L., Malavasi, M., Schmidlein, S., 2022. About the link between biodiversity and spectral variation. *Appl. Veg. Sci.* 25. <https://doi.org/10.1111/avsc.12643>.
- Féret, J.B., Asner, G.P., 2014. Mapping tropical forest canopy diversity using high-fidelity imaging spectroscopy. *Ecol. Appl.* 24 (6), 1289–1296.
- Féret, J.B., de Boissieu, F., 2020. biodivMapR: an R package for  $\alpha$ - and  $\beta$ -diversity mapping using remotely sensed images. *Methods Ecol. Evol.* 11, 64–70. <https://doi.org/10.1111/2041-210X.13310>.
- Féret, J.-B., de Boissieu, F., 2022. biodivMapR: BiodivMapR: an R package for  $\alpha$ - and  $\beta$ -diversity mapping using remotely-sensed images.
- Frye, H.A., Aiello-Lammens, M.E., Euston-Brown, D., Jones, C.S., Kilroy Mollmann, H., Merow, C., Slingsby, J.A., van der Merwe, H., Wilson, A.M., Silander, J.A., 2021. Plant spectral diversity as a surrogate for species, functional and phylogenetic diversity across a hyper-diverse biogeographic region. *Glob. Ecol. Biogeogr.* 30, 1403–1417. <https://doi.org/10.1111/geb.13306>.
- Galland, T., Adeux, G., Dvořáková, H., E-Vojtkó, A., Orbán, I., Lussu, M., Puy, J., Blažek, P., Lanta, V., Lepš, J., de Bello, F., Pérez Carmona, C., Valencia, E., Götzenberger, L., 2019. Colonization resistance and establishment success along gradients of functional and phylogenetic diversity in experimental plant communities. *J. Ecol.* 107, 2090–2104. <https://doi.org/10.1111/1365-2745.13246>.
- Gholizadeh, H., Gamon, J.A., Zyguelbaum, A.I., Wang, R., Schweiger, A.K., Cavender-Bares, J., 2018. Remote sensing of biodiversity: soil correction and data dimension reduction methods improve assessment of  $\alpha$ -diversity (species richness) in prairie ecosystems. *Remote Sens. Environ.* 206, 240–253. <https://doi.org/10.1016/j.rse.2017.12.014>.
- Gholizadeh, H., Gamon, J.A., Townsend, P.A., Zyguelbaum, A.I., Helzer, C.J., Hmimina, G.Y., Yu, R., Moore, R.M., Schweiger, A.K., Cavender-Bares, J., 2019. Detecting prairie biodiversity with airborne remote sensing. *Remote Sens. Environ.* 221, 38–49. <https://doi.org/10.1016/j.rse.2018.10.037>.
- Gholizadeh, H., Gamon, J.A., Helzer, C.J., Cavender-Bares, J., 2020. Multi-temporal assessment of grassland  $\alpha$ - and  $\beta$ -diversity using hyperspectral imaging. *Ecol. Appl.* 30. <https://doi.org/10.1002/eap.2145>.
- Gholizadeh, H., Dixon, A.P., Pan, K.H., McMillan, N.A., Hamilton, R.G., Fuhlendorf, S.D., Cavender-Bares, J., Gamon, J.A., 2022. Using airborne and DESIS imaging spectroscopy to map plant diversity across the largest contiguous tract of tallgrass prairie on earth. *Remote Sens. Environ.* 281. <https://doi.org/10.1016/j.rse.2022.113254>.
- Gibson, D.J., 2009. *Grasses and Grassland Ecology*. Oxford University Press.
- Hauser, L.T., Timmermans, J., van der Windt, N., Sil, Á.F., César de Sá, N., Soudzilovskaia, N.A., van Bodegom, P.M., 2021. Explaining discrepancies between spectral and in-situ plant diversity in multispectral satellite earth observation. *Remote Sens. Environ.* 265. <https://doi.org/10.1016/j.rse.2021.112684>.
- Hein, L., van Koppen, K., de Groot, R.S., van Ierland, E.C., 2006. Spatial scales, stakeholders and the valuation of ecosystem services. *Ecol. Econ.* 57, 209–228. <https://doi.org/10.1016/j.ecolecon.2005.04.005>.
- Heumann, B.W., Hackett, R.A., Monfils, A.K., 2015. Testing the spectral diversity hypothesis using spectroscopy data in a simulated wetland community. *Ecol. Inform.* 25, 29–34. <https://doi.org/10.1016/j.ecoinf.2014.10.005>.
- Imran, H.A., Gianelle, D., Scotton, M., Rocchini, D., Dalponte, M., Macolino, S., Vescovo, L., 2021. Potential and limitations of grasslands  $\alpha$ -diversity prediction using fine-scale hyperspectral imagery. *Remote Sens.* 13 (14), 2649.
- Mcculloch, C.E., Neuhaus, J.M., 2014. Generalized linear mixed models. In: Wiley StatsRef: Statistics Reference Online. <https://doi.org/10.1002/9781118445112.stat07540>.
- Nakagawa, S., Johnson, P.C.D., Schielzeth, H., 2017. The coefficient of determination R<sup>2</sup> and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *J. R. Soc. Interface* 14, 20170213. <https://doi.org/10.1098/rsif.2017.0213>.
- Oldeland, J., Wesulus, D., Rocchini, D., Schmidt, M., Jürgens, N., 2010. Does using species abundance data improve estimates of species diversity from remotely sensed spectral heterogeneity? *Ecol. Indic.* 10, 390–396. <https://doi.org/10.1016/j.ecolind.2009.07.012>.
- Palmer, M.W., Earls, P.G., Hoagland, B.W., White, P.S., Wohlgenuth, T., 2002. Quantitative tools for perfecting species lists. *Environmetrics* 13 (2), 121–137. <https://doi.org/10.1002/env.516>.



- Perrone, M., Di Febbraro, M., Conti, L., Divíšek, J., Chytrý, M., Keil, P., Carranza, M.L., Rocchini, D., Torresani, M., Moudrý, V., Šímová, P., Prajzlerová, D., Müllerová, J., Wild, J., Malavasi, M., 2023. The relationship between spectral and plant diversity: disentangling the influence of metrics and habitat types at the landscape scale. *Remote Sens. Environ.* 293, 113591 <https://doi.org/10.1016/j.rse.2023.113591>.
- Rocchini, D., Chiarucci, A., Loisel, S.A., 2004. Testing the spectral variation hypothesis by using satellite multispectral images. *Acta Oecol.* 26, 117–120. <https://doi.org/10.1016/j.actao.2004.03.008>.
- Rocchini, D., Hernández-Stefanoni, J.L., He, K.S., 2015. Advancing species diversity estimate by remotely sensed proxies: a conceptual review. *Ecol Inform* 25, 22–28. <https://doi.org/10.1016/j.ecoinf.2014.10.006>.
- Rocchini, D., Marcantonio, M., Da Re, D., Bacaro, G., Feoli, E., Foody, G.M., Furrer, R., Harrigan, R.J., Kleijn, D., Iannacito, M., Lenoir, J., Lin, M., Malavasi, M., Marchetto, E., Meyer, R.S., Moudry, V., Schneider, F.D., Šímová, P., Thornhill, A.H., Thouverai, E., Vicario, S., Wayne, R.K., Ricotta, C., 2021a. From zero to infinity: minimum to maximum diversity of the planet by spatio-parametric Rao's quadratic entropy. *Glob. Ecol. Biogeogr.* 30, 1153–1162. <https://doi.org/10.1111/geb.13270>.
- Rocchini, D., Salvatori, N., Beierkuhnlein, C., Chiarucci, A., de Boissieu, F., Förster, M., Garzon-Lopez, C.X., Gillespie, T.W., Haufler, H.C., He, K.S., Kleinschmit, B., Lenoir, J., Malavasi, M., Moudrý, V., Nagendra, H., Payne, D., Šímová, P., Torresani, M., Wegmann, M., Féret, J.B., 2021b. From local spectral species to global spectral communities: a benchmark for ecosystem diversity estimate by remote sensing. *Ecol Inform* 61. <https://doi.org/10.1016/j.ecoinf.2020.101195>.
- Rocchini, D., Thouverai, E., Marcantonio, M., Iannacito, M., Da Re, D., Torresani, M., Bacaro, G., Bazzichetto, M., Bernardi, A., Foody, G.M., Furrer, R., Kleijn, D., Larsen, S., Lenoir, J., Malavasi, M., Marchetto, E., Messori, F., Montagni, A., Moudrý, V., Naimi, B., Ricotta, C., Rossini, M., Santi, F., Santos, M.J., Schaeppman, M. E., Schneider, F.D., Schuh, L., Silvestri, S., Šímová, P., Skidmore, A.K., Tattoni, C., Tordini, E., Vicario, S., Zannini, P., Wegmann, M., 2021c. Rasterdiv—an information theory tailored R package for measuring ecosystem heterogeneity from space: to the origin and back. *Methods Ecol. Evol.* 12, 1093–1102. <https://doi.org/10.1111/2041-210X.13583>.
- Rossi, C., Gholizadeh, H., 2023. Uncovering the hidden: leveraging sub-pixel spectral diversity to estimate plant diversity from space. *Remote Sens. Environ.* 296, 113734 <https://doi.org/10.1016/j.rse.2023.113734>.
- Rossi, C., Kneubühler, M., Schütz, M., Schaeppman, M.E., Haller, R.M., Risch, A.C., 2021a. Spatial resolution, spectral metrics and biomass are key aspects in estimating plant species richness from spectral diversity in species-rich grasslands. *Remote Sens Ecol Conserv.* <https://doi.org/10.1002/rse2.244>.
- Rossi, C., Kneubühler, M., Schütz, M., Schaeppman, M.E., Haller, R.M., Risch, A.C., 2021b. Remote sensing of spectral diversity: a new methodological approach to account for spatio-temporal dissimilarities between plant communities. *Ecol. Indic.* 130 <https://doi.org/10.1016/j.ecolind.2021.108106>.
- Schiefer, F., Schmidtlein, S., Kattenborn, T., 2021. The retrieval of plant functional traits from canopy spectra through RTM-inversions and statistical models are both critically affected by plant phenology. *Ecol. Indic.* 121 <https://doi.org/10.1016/j.ecolind.2020.107062>.
- Schmidtlein, S., Fassnacht, F.E., 2017. The spectral variability hypothesis does not hold across landscapes. *Remote Sens. Environ.* 192, 114–125. <https://doi.org/10.1016/j.rse.2017.01.036>.
- Schweiger, A.K., Risch, A.C., Damm, A., Kneubühler, M., Haller, R., Schaeppman, M.E., Schütz, M., 2015. Using imaging spectroscopy to predict above-ground plant biomass in alpine grasslands grazed by large ungulates. *J. Veg. Sci.* 26, 175–190. <https://doi.org/10.1111/jvs.12214>.
- Schweiger, A.K., Cavender-Bares, J., Townsend, P.A., Hobbie, S.E., Madritch, M.D., Wang, R., Tilman, D., Gamon, J.A., 2018. Plant spectral diversity integrates functional and phylogenetic components of biodiversity and predicts ecosystem function. *Nat Ecol Evol* 2, 976–982. <https://doi.org/10.1038/s41559-018-0551-1>.
- Skidmore, A.K., Pettorelli, N., Coops, N.C., Geller, G.N., Hansen, M., Lucas, R., Wegmann, M., 2015. Environmental science: Agree on biodiversity metrics to track from space. *Nature* 523 (7561), 403–405.
- Suding, K.N., 2011. Toward an era of restoration in ecology: successes, failures, and opportunities ahead. *Annu. Rev. Ecol. Syst.* 42, 465–487. <https://doi.org/10.1146/annurev-ecolsys-102710-145115>.
- Thornley, R., Gerard, F.F., White, K., Verhoef, A., 2022. Intra-annual taxonomic and phenological drivers of spectral variance in grasslands. *Remote Sens. Environ.* 271 <https://doi.org/10.1016/j.rse.2022.112908>.
- Thornley, R.H., Gerard, F.F., White, K., Verhoef, A., 2023. Prediction of grassland biodiversity using measures of spectral variance: a meta-analytical review. *Remote Sens.* <https://doi.org/10.3390/rs15030668>.
- Torresani, M., Rocchini, D., Sonnenschein, R., Zebisch, M., Marcantonio, M., Ricotta, C., Tonon, G., 2019. Estimating tree species diversity from space in an alpine conifer forest: the Rao's Q diversity index meets the spectral variation hypothesis. *Ecol Inform* 52, 26–34. <https://doi.org/10.1016/j.ecoinf.2019.04.001>.
- Turner, W., 2014. Sensing biodiversity. *Science* (1979) 346, 301–302. <https://doi.org/10.1126/science.1256014>.
- Villoslada, M., Bergamo, T.F., Ward, R.D., Burnside, N.G., Joyce, C.B., Bunce, R.G.H., Sepp, K., 2020. Fine scale plant community assessment in coastal meadows using UAV based multispectral data. *Ecol. Indic.* 111 <https://doi.org/10.1016/j.ecolind.2019.105979>.
- Wang, R., Gamon, J.A., 2019. Remote sensing of terrestrial plant biodiversity. *Remote Sens. Environ.* 231 <https://doi.org/10.1016/j.rse.2019.111218>.
- Wang, R., Gamon, J.A., Emmerton, C.A., Li, H., Nestola, E., Pastorello, G.Z., Menzer, O., 2016. Integrated analysis of productivity and biodiversity in a southern Alberta prairie. *Remote Sens.* 8 <https://doi.org/10.3390/rs8030214>.
- Wang, R., Gamon, J.A., Schweiger, A.K., Cavender-Bares, J., Townsend, P.A., Zygielbaum, A.I., Kothari, S., 2018. Influence of species richness, evenness, and composition on optical diversity: a simulation study. *Remote Sens. Environ.* 211, 218–228. <https://doi.org/10.1016/j.rse.2018.04.010>.
- Warren, S.D., Alt, M., Olson, K.D., Irl, S.D.H., Steinbauer, M.J., Jentsch, A., 2014. The relationship between the spectral diversity of satellite imagery, habitat heterogeneity, and plant species richness. *Ecol Inform* 24, 160–168. <https://doi.org/10.1016/j.ecoinf.2014.08.006>.
- Wilson, J.B., Peet, R.K., Dengler, J., Pärtel, M., 2012. Plant species richness: the world records. *J. Veg. Sci.* 23, 796–802. <https://doi.org/10.1111/j.1654-1103.2012.01400.x>.
- Wiltling, H.C., Schipper, A.M., Bakkenes, M., Meijer, J.R., Huijbregts, M.A.J., 2017. Quantifying biodiversity losses due to human consumption: a global-scale footprint analysis. *Environ. Sci. Technol.* 51, 3298–3306. <https://doi.org/10.1021/acs.est.6b05296>.