Spectral Diversity Successfully Estimates the \( \alpha \)-Diversity of Biocrust-Forming Lichens

Javier Blanco-Sacristán 1,*, Cinzia Panigada 1, Giulia Tagliabue 1, Rodolfo Gentili 1, Roberto Colombo 1, Mónica Ladrón de Guevara 2,3, Fernando T. Maestre 4 and Micol Rossini 1

Article

Abstract: Biocrusts, topsoil communities formed by mosses, lichens, liverworts, algae, and cyanobacteria, are a key biotic component of dryland ecosystems worldwide. Experiments carried out with lichen- and moss-dominated biocrusts indicate that climate change may dramatically reduce their cover and diversity. Therefore, the development of reproducible methods to monitor changes in biocrust diversity and abundance across multiple spatio-temporal scales is key for evaluating how climate change may impact biocrust communities and the myriad of ecosystem functions and services that rely on them. In this study, we collected lichen-dominated biocrust samples from a semi-arid ecosystem in central Spain. Their \( \alpha \)-diversity was then evaluated using very high spatial resolution hyperspectral images (pixel size of 0.091 mm) measured in laboratory under controlled conditions. Support vector machines were used to map the biocrust composition. Traditional \( \alpha \)-diversity metrics (i.e., species richness, Shannon’s, Simpson’s, and Pielou’s indices) were calculated using lichen fractional cover data derived from their classifications in the hyperspectral imagery. Spectral diversity was calculated at different wavelength ranges as the coefficient of variation of different regions of the reflectance spectra of lichens and as the standard deviation of the continuum removal algorithm (SD_CR). The accuracy of the classifications of the images obtained was close to 100%. The results showed the best coefficient of determination \( (r^2 = 0.47) \) between SD_CR calculated at 680 nm and the \( \alpha \)-diversity calculated as the Simpson’s index, which includes species richness and their evenness. These findings indicate that this spectral diversity index could be used to track spatio-temporal changes in lichen-dominated biocrust communities. Thus, they are the first step to monitor \( \alpha \)-diversity of biocrust-forming lichens at the ecosystem and regional levels, a key task for any program aiming to evaluate changes in biodiversity and associated ecosystem services in drylands.

Keywords: biocrusts; biological soil crust; spectral diversity; chlorophyll; continuum removal; biodiversity; \( \alpha \)-diversity; support vector machine; remote sensing

1. Introduction

Biocrusts are communities formed by photoautotrophic (algae, lichens, cyanobacteria, liverworts, and bryophytes) and heterotrophic (bacteria, fungi, protozoa, and nematodes) organisms that live on the soil surface and cover a large part of the non-vegetated surface in drylands worldwide [1]. These
Remote Sens. 2019, 11, 2942 2 of 16

communities influence fundamental ecosystem processes in drylands, including—but not limited to—nutrient cycling, soil respiration, and runon-runoff dynamics [2], and are critical for maintaining the multiplicity of ecosystem services they provide [3]. Attributes of biocrust communities, such as their cover, composition, and diversity, largely modulate their impacts on multiple ecosystem functions simultaneously [4,5], and thus have been suggested as indicators of ecosystem functioning in drylands [6,7].

Multiple lines of evidence suggest that ongoing climate change can dramatically affect biocrust communities, reducing their cover and diversity [8]. Reference [9] estimated a global reduction of their cover by 40% globally within the next 65 years, and quick reductions in their cover have already been observed in experiments simulating climate change [10–12]. As biocrusts contribute to ecosystem multifunctionality [2,11,13,14], changes in their composition, cover and diversity could lead to a reduction of the capacity of drylands to provide essential ecosystem services such as atmospheric CO$_2$ sequestration and the maintenance of soil fertility. For this reason, finding accurate and operational methods to estimate the cover and diversity of biocrust constituents is a key goal for any ecosystem monitoring program in drylands.

Remote sensing has been highlighted several times as an important tool for biodiversity monitoring and conservation [15–18]. It is one of the most cost-effective approaches to identify biodiversity hotspots and to predict changes in species composition, potentially providing repeated measurements and making it possible to study temporal changes in biodiversity [19]. In this context, the spectral variation hypothesis (SVH) proposes that the larger the spectral heterogeneity of an environment is, the higher the number of species found here will be [20]. This hypothesis has been validated several times using $\alpha$-diversity metrics with vascular plants [21–25]. Even though measures of spectral diversity based on spectral dispersion (i.e., the coefficient of variation (CV) or the standard deviation (SD)) have shown good correlations with the richness, diversity, evenness, and composition of vascular plants [25–27], the SVH has never been tested using lichens.

The use of optical remote sensing of biocrusts so far has mainly focused on mapping and monitoring their distribution [28–33] and only one study [34] investigated lichens’ $\alpha$-diversity at the landscape level through correlation of high-resolution data with field samples. Since biocrusts are spectrally characterized by narrow absorption features in specific spectral regions [35], hyperspectral data have shown potential when discriminating them from vegetation and other soil features (e.g., [35–37]). However, confounding factors such as water content and biocrust tridimensional structure might hinder their spectral characterization. For this reason, several authors [31,33,35,38–40] proposed the use of the continuum removal (CR) algorithm [41] to ensure the comparability of biocrusts’ absorption features. The CR quantifies the absorption features at specific wavelengths, normalizing the reflectance spectra to a common baseline. This is achieved by approximating the continuum between local spectral maxima through straight-line segments: a value of 1 is assigned to the local maxima, and a value between 0 and 1 is obtained in correspondence of the absorption features. This approach might be particularly useful to enhance the discrimination of lichens, which are characterized by subtle spectral differences [42].

In this framework, the objective of this study was to apply the SVH to capture the $\alpha$-diversity of lichens to support dryland monitoring by (i) evaluating the potential of hyperspectral high spatial resolution images to identify biocrust-forming lichens (at the genus level); (ii) exploring the relationships between spectral diversity metrics and the $\alpha$-diversity of lichens within biocrust communities.

2. Materials and Methods

2.1. Study Area and Sampling

The study area is located in Aranjuez, central Spain (40°01'53.9"N 3°32'50.8"W; Figure 1). The climate is semi-arid Mediterranean, with mean annual temperature and rainfall of 15 °C and 349 mm, respectively. The plant cover is below 40%, and is dominated by Macrochloa tenacissima and other
small shrubs, such as *Helianthemum squamatum* and *Gypsophila struthium*. A well-developed biocrust community, a mixture of lichens and mosses, is present between vegetation, covering up to more than 50% of the plant interspaces. In lichen-dominated biocrusts, *Acarospora* spp., *Buellia* spp., *Diploschistes* spp., *Fulgensia* spp., *Psora* spp., and *Squamarina* spp. are the most abundant genera, while in moss-dominated biocrusts, *Pleurochaete squarrosa*, *Tortula revolvens*, and *Didymodon acutus* are the most abundant mosses. Since this study focused on lichens, we decided to cluster the mosses identified (under the term *Moss*). Finally, in patches of soil dominated by cyanobacteria, genera as *Microcoleus* spp., *Tolypothrix* spp., and *Nostoc* spp. are present [43]. See [44] for a species list of the biocrust-forming lichens and mosses in our study area.

Following the sampling protocol described in [30], we collected 54 biocrust samples using Petri dishes of 8.5 cm of diameter from 18 plots (three samples per plot randomly chosen). These plots were distributed following two altitudinal transects in two separated slopes of the study area with different exposures (north- vs. south-facing). Each transect was divided in three levels depending on the height of the transect respect to its base, placing three plots on each level.

2.2. Hyperspectral Imagery Acquisition

Hyperspectral images of the biocrust samples were acquired with a hyperspectral scanning imaging system [45] measuring spectral reflectance in 840 bands in the visible (VIS) and near-infrared (NIR) spectral region (i.e., 380–1000 nm), with a spectral resolution of 2–3 nm (calculated at full width at half maximum; Figure 2). The system consists of a custom high-precision linear stage that embeds a hyperspectral imaging spectrometer (Hyperspec® VNIR, HeadWall Photonics, USA) and a dedicated halogen stable light source (i.e., 600 or 1000 W, LOT Quantum Design). Powered by an electrical engine, both the spectrometer and the halogen light source are able to move back and forth at a defined speed. The system is a push-broom scanner and it measures lines of 1004 pixels while moving forward by means of the linear stage until an image of the whole studied object is collected. We used a water-filter tube between the lamp and the mirror reflecting the light to the samples to avoid sample overheating.
Two sets of images were collected: one with the dry samples (dry set) and another 10 min after evenly watering the samples with 30 mL of distilled water (wet set). A total of 18 images (six samples on each set; Figure 3) with a spatial resolution of 0.091 mm were taken. A calibrated white Lambertian Spectralon® panel (Labsphere, North Sutton, USA) was placed close to the samples to calculate the reflectance as the ratio between the radiance reflected by the biocrust samples and the panel. The instrument’s dark current signal was measured by manually closing the imaging spectrometer aperture prior to capturing each image and subtracted from the measured radiance.

Figure 2. Picture of the system used to acquire the hyperspectral imagery. The system consists of: (a) samples holder and motor driver; (b) HeadWall VNIR camera thermostated by heaters; (c) stable halogen lamp; (d) calibrated Lambertian Spectralon panel; (e) PC connected to the spectrometer by a dedicated interface.

Two sets of images were collected: one with the dry samples (dry set) and another 10 min after evenly watering the samples with 30 mL of distilled water (wet set). A total of 18 images (six samples on each set; Figure 3) with a spatial resolution of 0.091 mm were taken. A calibrated white Lambertian Spectralon® panel (Labsphere, North Sutton, USA) was placed close to the samples to calculate the reflectance as the ratio between the radiance reflected by the biocrust samples and the panel. The instrument’s dark current signal was measured by manually closing the imaging spectrometer aperture prior to capturing each image and subtracted from the measured radiance.

Figure 3. True color composite samples (upper images) from one of the hyperspectral images used and the classification derived (lower images).

2.3. Images Processing and Classification

Both sets of images were pre-processed and classified to characterize the composition of each sample. First, the parts in the images not belonging to the samples were masked and the reflectance was computed between 420 and 900 nm. To reduce the noise in the measured signal, a Savitzky-Golay smoothing filter [46] was applied using a 20 band-window width. A minimum noise fraction (MNF) transformation was applied to the smoothed reflectances to synthesize the main information in a reduced number of MNF components as well as to remove the residual noise in the data. The first 15 MNF components of each image transformation were used as input of the classification algorithm.

The training set for the classification was created by selecting pure endmembers of each classes by visual-identification on the images of the samples. The following classes were considered in the classification process: Acarospora (Acarospora spp.), Buellia (Buellia spp.), Diplochistes (Diploschistes spp.),
Fulgensia (*Fulgensia* spp.), *Psora* (*Psora* spp.), *Squamarina* (*Squamarina* spp.), *Moss* and *Bare Soil*. Biocrust classification was performed with a supervised machine learning method, using a total of around 1000 pixels per class identified in each image to train the support vector machine (SVM) algorithm.

The SVM is a supervised classification method based on statistical learning theory [47,48]. Using training samples from the classes of interest, it separates them by a decision surface, called hyperplane, that maximizes the margin between them. The closest training samples to the hyperplane are the ones used by the algorithm, called support vectors. SVM is a binary classifier in its simplest form, but can also act as a multiclass classifier by creating a binary classifier for each possible pair of classes. SVM uses a penalty parameter that allows a certain degree of misclassification. Fine tuning is important to avoid overfitting of the model because increasing its value increases the cost of misclassifying points and forces the creation of a more accurate model, which might not be generalizable. We used a radial basis kernel function, which usually performs well in remote sensing applications [49]. This kernel is controlled by two parameters that determine the final classification accuracy, the penalty parameter (C) and the width of the Gaussian kernel (c). A large C reduces the training dataset error, but may result in model overfitting to the training data, reducing their generalizability. We used the SVM in a pairwise classification way to classify the images and evaluate the best parameters to use for each classification using the LIBVSM library [31,50].

2.4. Validation of Classifications

The validation set for each classification was created selecting 300 pixels per class of pure endmembers by visual-identification. We assessed the accuracy of the classifications using the receiver operating characteristics (ROC) technique [51] and the area under the curve (AUC) [52,53] derived from it. A ROC graph is a two-dimensional depiction of a classifier’s performance [54] and is constructed calculating the sensitivity and specificity of the resulting classification for each possible classification threshold, where

\[
\text{sensitivity} = \frac{a}{a + c}
\]

\[
\text{specificity} = \frac{b}{b + d},
\]

with \(a\) and \(d\) as the true positives and the true negatives for a certain classification, respectively, and \(b\) and \(c\) as the corresponding false positives and false negatives. The ‘sensitivity’ is the probability that a pixel of a particular class is correctly classified, while the ‘specificity’ is the probability that a pixel not belonging to that class is correctly classified. In this way, the best performing classification would be that with the highest possible value of both sensitivity and specificity. We evaluated the probability of detection (calculated as the sensitivity) against the false positive rate (calculated as \(1 – \text{specificity}\)). A ROC graph was calculated for each of the classifications produced (i.e., 18 classifications, six samples each), averaging the curves obtained for each class.

From each ROC graph, we calculated its corresponding area under the curve (AUC), that is an overall quantitative performance score of the classification that allows to reduce the ROC performance to a single scalar value independent of a single prediction threshold [52,54]. The AUC represents the probability that a randomly chosen positive sample is correctly classified with greater suspicion than a randomly chosen negative one [53]. This value might range from 0.5 (a random assignment to the class of interest) to 1 (a perfect classification). The AUC values were calculated by joining the points of the ROC through a composite trapezoid rule using the AUC function from the R package DescTools [55]. To assess classification accuracy, we also evaluated the average accuracy creating a confusion matrix with the average accuracies from the classifications of each set of images (i.e., dry and wet), extracting the overall accuracy and Cohen’s kappa coefficient [56]. The points used to create these matrices were the same used to create the ROC curves.
2.5. Computation of the Spectral Diversity

Using the classifications previously obtained, we calculated, for each biocrust sample, the average spectral reflectance of each lichen genera from the reflectance images. We then calculated the pigment absorption features by the continuum removal method, normalizing the spectra to a common baseline. Several metrics were then calculated as indicators of spectral diversity for each sample: the coefficient of variation (CV; i.e., ratio of the standard deviation to the mean), calculated between 420–900 nm (\(CV_{420-900}\)), between 550–750 nm (\(CV_{550-750}\)), and at 680 nm (\(CV_{680}\)), and the standard deviation of the continuum removal (SD_CR), calculated between 550–750 nm (SD_CR\(550-750\); i.e., absorption feature related to the presence of chlorophyll) and at 680 nm (SD_CR\(680\); i.e., maximum peak of this absorption feature).

2.6. Biodiversity Metrics

The fractional cover of each lichen class (classified at the genus level) was used to calculate the following \(\alpha\)-diversity metrics for each sample: species richness (\(S\)), Shannon’s index (\(H'\); [57]), Simpson’s index (\(D\); [58]) and Pielou’s index (\(J'\); [59]). The details of the calculation are reported in Table 1. These metrics were calculated using the ‘vegan’ R package, version 2.4–5 [60].

<table>
<thead>
<tr>
<th>(\alpha)-Diversity Metric</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Species richness ((S))</td>
<td>(S = \text{Number of classes})</td>
</tr>
<tr>
<td>Shannon’s index ((H'))</td>
<td>(H' = -\sum p_i \ln (p_i))</td>
</tr>
<tr>
<td>Reciprocal of Simpson’s index ((D))</td>
<td>(D = 1 / \sum p_i^2)</td>
</tr>
<tr>
<td>Pielou’s index ((J'))</td>
<td>(J' = H' / \ln (S))</td>
</tr>
</tbody>
</table>

2.7. Statistical Analysis

To evaluate the spectral diversity metrics that better capture the \(\alpha\)-diversity of biocrust-forming lichens, we evaluated through linear regression models the relationships between spectral diversity (calculated as the average CV and CR of the three samples from each plot) and \(\alpha\)-diversity metrics (species richness, Shannon’s index, Simpson’s index, and Pielou’s index), calculated using the fractional cover of each plot (the average of three samples). Due to the high heterogeneity of four plots, their samples were not averaged, and the values of the single samples were used in the analysis. Species richness was calculated as the total number of genera observed in the three samples from each plot. Due to the small sample size (\(n = 26\)) in this analysis, the cross-validated statistics obtained with the leave-one-out cross-validation procedure were also computed to compare performances of different spectral diversity metrics in predicting \(\alpha\)-diversity.

3. Results

3.1. Classifications and Accuracy Evaluation

The classification (Figure 3) of both dry and wet images was highly accurate. The high values of the AUC derived from the ROC curves showed that the SVMs used were successfully trained to classify biocrust-forming lichens (Figure 4).
The classification of the wet dataset had an average AUC of 0.95 (Figure 4), a kappa coefficient of 0.97 and an overall accuracy of 97.83% (Table 2). The lowest accuracies were achieved classifying Moss and Fulgensia, with AUCs of 0.91 and 0.92, respectively. The most accurate results were obtained for Buellia and Psora, with AUCs values of 0.99 and 0.98, respectively. Bare soil was the most confused class, mostly with moss, even though it was better classified than in the dry set of images, where it presented an AUC of 0.82. The dry set was less accurate, presenting an average AUC of 0.93, an average kappa coefficient of 0.95 and an overall accuracy of 95.69%. All the lichens were accurately classified, with values of AUC over 0.95, similar to those obtained from the wet set. Moss presented the lowest accuracy, with an AUC of 0.75, and was the most misclassified, being confused mostly with bare soil.

Table 2. Average confusion matrix obtained crossing the ground truth (columns) with the results of the classification (lines) performed on wet samples. 300 pixels per class in each image were randomly selected as validation set to create this matrix. A total of 20,400 random selected pixels were used. The ground truth (%) shows the class distribution in percent for each ground truth class.

<table>
<thead>
<tr>
<th>Ground truth (%)</th>
<th>Acarospora</th>
<th>Bare Soil</th>
<th>Buellia</th>
<th>Diploschistes</th>
<th>Fulgensia</th>
<th>Moss</th>
<th>Psora</th>
<th>Squamarina</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acarospora</td>
<td>99.96</td>
<td>0</td>
<td>0</td>
<td>0.04</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Bare soil</td>
<td>0.04</td>
<td>97.55</td>
<td>0</td>
<td>0.72</td>
<td>0.4</td>
<td>0.32</td>
<td>1.06</td>
<td>0</td>
</tr>
<tr>
<td>Buellia</td>
<td>0</td>
<td>0</td>
<td>93.45</td>
<td>4.26</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.98</td>
</tr>
<tr>
<td>Diploschistes</td>
<td>0</td>
<td>1.63</td>
<td>6.49</td>
<td>94.89</td>
<td>0.24</td>
<td>0</td>
<td>1.3</td>
<td>0.23</td>
</tr>
<tr>
<td>Fulgensia</td>
<td>0</td>
<td>0.04</td>
<td>0</td>
<td>0</td>
<td>99.31</td>
<td>0</td>
<td>0.05</td>
<td>0</td>
</tr>
<tr>
<td>Moss</td>
<td>0</td>
<td>0.76</td>
<td>0</td>
<td>0</td>
<td>0.05</td>
<td>99.65</td>
<td>0.3</td>
<td>0</td>
</tr>
<tr>
<td>Psora</td>
<td>0</td>
<td>0.02</td>
<td>0</td>
<td>0</td>
<td>0.05</td>
<td>0.03</td>
<td>97.29</td>
<td>0</td>
</tr>
<tr>
<td>Squamarina</td>
<td>0</td>
<td>0.06</td>
<td>0.09</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>98.79</td>
<td>0</td>
</tr>
</tbody>
</table>

3.2. Spectral Characterization of Biocrusts

Dry biocrusts had increasing reflectances from the blue region until 700 nm. In the visible region the reflectances of various biocrust classes differ because of different pigment content and composition (Figure 5) differences in the NIR region are mainly related to the biocrust tridimensional structure. Lichens presented higher reflectances than bare soil and moss across the whole spectrum. Their higher reflectance in the NIR region is related to their more developed structure, which causes multi-scattering of light in this region. Acarospora, Squamarina, Buellia, and Diploschistes are light-colored lichens that showed higher reflectances in the visible region; Psora presented the lowest reflectance among lichens. Mosses had the lowest reflectance due to their darker color and less developed structure, resulting in a reduction of light scattering. The application of the continuum removal algorithm in the spectral range 450–900 nm highlighted the absorption features caused by different pigments. The chlorophyll absorption feature at 680 nm was present in all the classes. Bare soil showed a weak absorption feature.
at 680 nm, evidencing the presence of cyanobacteria colonizing the soils interspace within lichens and moss patches. *Psora* showed an absorption feature around 550 nm due to phycoerythrin, which is absent in the other lichens. An absorption feature at 500 nm related to the presence of carotenoids was observed in mosses. *Fulgensia* presented an absorption peak around 490 nm due to the presence of carotenoids or phycoerythrin.

**Figure 5.** Top: RGB composites of some of the samples used in the study, showing the change of color that takes place from dry (first row) to wet (second row) state of biocrusts. Down: Mean reflectance spectra and mean continuum removal absorption spectra between 450–900 nm of the biocrust classes and bare soil studied in this work. The red dashed line marks the absorption feature at 680 nm related to the presence of chlorophyll. (a) Dry samples; (b) Wet samples.

Watering had a strong effect on biocrust optical properties (Figure 5b), causing a decrease of reflectance across the whole spectrum. The change was particularly evident in the absorption feature caused by chlorophyll at ~680 nm, which became deeper, evidencing the abrupt change in the reflectance from red to NIR, typical of vegetation (i.e., red edge region). The continuum removal algorithm allowed to enhance this absorption feature at ~680 nm related to the activation of chlorophyll after watering, increasing its variation (Figure 6).
3.3. Fractional Cover of Biocrusts and Diversity Metrics

The most abundant classes were moss and bare soil, covering more than 76% and 46% of the surface in some samples, respectively (Table 3). The most abundant lichen was *Diploschistes*, while the least were *Psora* and *Buellia*, which appeared scarcely. Species richness ranged from 2 to 6. Shannon’s index ranged from 0.18 to 1.53, with a mean value of 1.03, and Simpson’s index from 0.08 to 0.77, with a mean of 0.56. Pielou’s index ranged from 0.14 to 0.95, with a mean value of 0.66.

Table 3. Fractional cover values of the classes evaluated in this work. Samples: number of samples where the class was identified. Plots: number of plots where the class was identified. Mean Fc, Max Fc, Min Fc, SD Fc: mean, maximum, minimum and standard deviation of the fractional cover observed for each class. Classes: Acarospora: *Acarospora* spp.; Buellia: *Buellia* spp., Bare Soil: bare soil; Diploschistes: *Diploschistes* spp., Fulgensia: *Fulgensia* spp.; Moss: mosses; Psora: *Psora* spp.; Squamarina: *Squamarina* spp.

<table>
<thead>
<tr>
<th>Class</th>
<th>Samples</th>
<th>Plots</th>
<th>Mean Fc (%)</th>
<th>Max Fc (%)</th>
<th>Min Fc (%)</th>
<th>SD Fc (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acarospora</td>
<td>37</td>
<td>17</td>
<td>3.8</td>
<td>32.8</td>
<td>0.3</td>
<td>5.8</td>
</tr>
<tr>
<td>Bare Soil</td>
<td>54</td>
<td>18</td>
<td>21.6</td>
<td>46.1</td>
<td>5.1</td>
<td>8.1</td>
</tr>
<tr>
<td>Buellia</td>
<td>33</td>
<td>13</td>
<td>2.4</td>
<td>13.8</td>
<td>0.5</td>
<td>3</td>
</tr>
<tr>
<td>Diploschistes</td>
<td>54</td>
<td>18</td>
<td>14.6</td>
<td>53.1</td>
<td>0.1</td>
<td>11</td>
</tr>
<tr>
<td>Fulgensia</td>
<td>53</td>
<td>18</td>
<td>12</td>
<td>25.4</td>
<td>0.7</td>
<td>7.3</td>
</tr>
<tr>
<td>Moss</td>
<td>54</td>
<td>8</td>
<td>43.5</td>
<td>76.9</td>
<td>4.9</td>
<td>17</td>
</tr>
<tr>
<td>Psora</td>
<td>41</td>
<td>16</td>
<td>1.9</td>
<td>12.6</td>
<td>0.1</td>
<td>2.6</td>
</tr>
<tr>
<td>Squamarina</td>
<td>27</td>
<td>13</td>
<td>4.3</td>
<td>4.3</td>
<td>0.3</td>
<td>4.5</td>
</tr>
</tbody>
</table>

3.4. Relationships between Biodiversity and Spectral Diversity

The correlation analysis between spectral diversity (CV and SD_CR) and $\alpha$-diversity metrics (i.e., species richness, Shannon’s Index, Simpson’s Index, and Pielou’s Index) showed that there were no significant correlations with the dry samples (results not shown). Conversely, positive and statistically significant relationships were found when the samples were wet (Table 4).
with values to the original analysis, showing the stability of the predictions made by the linear regression models calculated between spectral diversity metrics (standard deviation of the continuum removal calculated between 550–750 nm (SD_CR550–750) and at 680 nm (SD_CR680); coefficient of variation calculated between 420–900 nm (CV420–900)), between 550–750 nm (CV550–750), and at 680 nm (CV680) and α-diversity metrics (Species richness (S), Shannon’s index (H’), Simpson’s index (D), and Pielou’s Index (J’)) when the samples were wet. The statistics in cross-validation (Table 5) of the linear regressions presented similar values to the linear regression models calculated between spectral diversity metrics: (α-diversity metrics: (Shannon’s index, Simpson’s index, and Pielou’s index. Neither the SD_CR nor the CV captured the species richness. The Simpson’s index was positively correlated to the CV at all the spectral ranges analyzed (420–900, 550–750, and 680 nm), especially at 680 nm. While Simpson’s and Pielou’s indices were the metrics that correlated best with SD_CR at 680 nm (SD_CR680), especially at 680 nm. When the spectral diversity was calculated as SD_CR550–750 and SD_CR680, the relationships with α-diversity were positive and strongly significant for all the diversity metrics (Table 4). The standard deviation of the CR at 680 nm (SD_CR680), which is related to the difference in chlorophyll content, was the spectral diversity metric better related to α-diversity metrics. Neither the SD_CR nor the CV captured the species richness. The Simpson’s index was positively correlated to the CV at all the spectral ranges analyzed (420–900, 550–750, and 680 nm), especially at 680 nm. While Simpson’s and Pielou’s indices were the metrics that correlated best with SD_CR680 (r² = 0.47 and r² = 0.42, respectively; Table 4, Figure 7), they did not show any significant correlation with CV420–900 or CV550–750. The statistics in cross-validation (Table 5) of the linear regressions presented similar values to the original analysis, showing the stability of the predictions made by the linear regression models between spectral diversity and α-diversity of lichens.

Table 4. Slopes, coefficients of determination (r²) and p-values (between brackets) of the linear regression models calculated between spectral diversity metrics (standard deviation of the continuum removal calculated between 550–750 nm (SD_CR550–750) and at 680 nm (SD_CR680); coefficient of variation calculated between 420–900 nm (CV420–900)), between 550–750 nm (CV550–750), and at 680 nm (CV680) and α-diversity metrics (Species richness (S), Shannon’s index (H’), Simpson’s index (D), and Pielou’s Index (J’)) when the samples were wet.

<table>
<thead>
<tr>
<th>α-Diversity Metric</th>
<th>SD_CR550–750</th>
<th>SD_CR680</th>
<th>CV420–900</th>
<th>CV550–750</th>
<th>CV680</th>
</tr>
</thead>
<tbody>
<tr>
<td>Species richness (S)</td>
<td>0.001</td>
<td>0.003</td>
<td>0.005</td>
<td>0.013</td>
<td>0.08</td>
</tr>
<tr>
<td>Shannon’s Index (H’)</td>
<td>0.012</td>
<td>-</td>
<td>0.056</td>
<td>0.063</td>
<td>0.071</td>
</tr>
<tr>
<td>Simpson’s Index (D)</td>
<td>0.33(0.001)</td>
<td>0.41(0.0004)</td>
<td>-</td>
<td>-</td>
<td>0.16(0.03)</td>
</tr>
<tr>
<td>Pielou’s Index (J’)</td>
<td>0.39(0.0004)</td>
<td>0.47(0.0001)</td>
<td>-</td>
<td>-</td>
<td>0.26(0.007)</td>
</tr>
</tbody>
</table>

When the spectral diversity was calculated as SD_CR550–750 and SD_CR680, the relationships with α-diversity were positive and strongly significant for all the diversity metrics (Table 4). The standard deviation of the CR at 680 nm (SD_CR680), which is related to the difference in chlorophyll content, was the spectral diversity metric better related to α-diversity metrics. Neither the SD_CR nor the CV captured the species richness. The Simpson’s index was positively correlated to the CV at all the spectral ranges analyzed (420–900, 550–750, and 680 nm), especially at 680 nm. While Simpson’s and Pielou’s indices were the metrics that correlated best with SD_CR680 (r² = 0.47 and r² = 0.42, respectively; Table 4, Figure 7), they did not show any significant correlation with CV420–900 or CV550–750. The statistics in cross-validation (Table 5) of the linear regressions presented similar values to the original analysis, showing the stability of the predictions made by the linear regression models between spectral diversity and α-diversity of lichens.

Figure 7. Linear regression between the spectral diversity measured as the Standard Deviation of the Continuum removal at 680 nm and three α-diversity metrics: (a) Shannon’s index, (b) Simpson’s index, and (c) Pielou’s index. Shaded areas represent ±95% symmetrical confidence interval.

Table 5. Summary of statistics in fitting (r² and RMSE) and cross-validation (r²cv and RMSEcv) of the linear regression models between the spectral diversity measured as the standard deviation of the continuum removal at 680 nm and the α-diversity metrics (Shannon’s, Simpson’s, and Pielou’s indices). RMSE: root mean square error; r²: coefficient of determination.

<table>
<thead>
<tr>
<th>α-Diversity Metric</th>
<th>r²</th>
<th>RMSE</th>
<th>rCV²</th>
<th>RMSEcv</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shannon’s index (H’)</td>
<td>0.41</td>
<td>0.01</td>
<td>0.32</td>
<td>0.01</td>
</tr>
<tr>
<td>Simpson’s index (D)</td>
<td>0.47</td>
<td>0.009</td>
<td>0.39</td>
<td>0.01</td>
</tr>
<tr>
<td>Pielou’s index (J’)</td>
<td>0.42</td>
<td>0.009</td>
<td>0.35</td>
<td>0.01</td>
</tr>
</tbody>
</table>
4. Discussion

The high accuracies obtained using SVM to classify hyperspectral imagery reinforce their use to differentiate biocrusts [31]. Differences in the reflectance of biocrust constituents are subtle (Figure 5), but SVMs are capable of differentiating spectrally similar classes when the inputs are spectral signatures [61–63]. Even though the accuracy metrics from both sets of images were similar, the classification improved when biocrusts were wet. This result is explained by the enhanced differences that appear in the reflectance spectra when biocrusts are metabolically activated after irrigation [35]. Several studies have characterized the optical properties of biocrust communities in the same optical range used in this study [42]. However, the spatial scale of these studies only allowed characterization of biocrust communities. Conversely, the fine spatial resolution of the imaging system used in this work allowed to characterize for the first time the pure spectral signature of six lichen genera, to compare their characteristics and to capture the spectral diversity among them. In fact, we used a pixel size lower than the size of each object of interest (i.e., thallus of lichens), as suggested by [64,65].

Lichens present different structural and biochemical traits, which create wavelength-dependent variations that can be integrated by spectral diversity, as shown in vascular plants by [66]. Different genera of lichens are characterized by particular biochemical traits that promote spectral variations, so their spectral diversity can be exploited to infer their α-diversity. However, it is important to understand which are the best metrics to represent the relationship between lichen spectral diversity and α-diversity. Since the absorption peak around 680 nm has been widely investigated as a spectral feature of biocrusts and is present in all lichens (e.g., [35,40,67–69]), we tested if a spectral diversity metric focused on this feature may be suitable for monitoring changes in their composition. We found that using a spectral diversity metric based on this absorption feature (i.e., the SD_CR680) increases the spectral variability determined by the chlorophyll content of different lichens, while removing the confounding influence of other factors such as the structure of lichens. In fact, some lichen genera are characterized by a strong tridimensional component (e.g., Diploschistes spp., Squamarina spp.) relative to others, which can determine a higher intra-genera variability compared to the inter-genera variability. The normalization of the reflectance spectra to a common baseline through the CR algorithm minimized these structural effects, allowing to capture the spectral variability of lichens determined by the chlorophyll content. Conversely, the CV is influenced by both variations in the content of pigments and the structure of biocrusts, which may hide inter-genera variability. This might be the reason why the CV did not perform as well as expected in previous studies conducted with vascular plants (e.g., [23–25]), where the structural component constitutes were the main source of spectral variability. Conversely, the results obtained in this study highlight the dominant role of the chlorophyll content to determine the differences between lichen genera and the importance of using the absorption feature centered at ~680 nm to capture their spectral diversity. Nevertheless, using absorption features that are found in the shortwave infrared region of the spectra of lichens [35] might help to improve the results obtained in this study, that investigated only the visible and near infrared spectral domain.

Among the α-diversity metrics tested, the Simpson’s index correlated the most with the spectral diversity of lichens. The better performance of the Simpson’s index compared to the Shannon’s index is in agreement with recent studies conducted in tropical forests [21] and in a prairie grassland [24]. The latter found similar and weaker relationships between spectral diversity and evenness (calculated as Pielou’s index) as we did here, although they correlated these metrics with the CV instead of the SD_CR. This might be due to the fact that Shannon’s index assumes that all the species are present and randomly sampled [70], and the Simpson’s index is more sensitive to dominant or common species, as noted by [24], making it more suitable when this is the case. The similar relationships obtained in cross-validation show that these results are reliable despite their sample size.

The imagery used in this study captured with high detail the lichen genera present in our samples, allowing to identify positive and significant relationships between the spectral diversity and the α-diversity of lichens. The methodology proposed in this study should be in principle applied to monitor α-diversity of lichens at wider scales, a key task to understand the shifts in the composition
that these communities are undergoing in the actual context of global change [9–12]. However, the upsampling of this methodology may be not trivial due to the decreased spatial and/or spectral resolution when working on wider scales.

Many ecological processes maintain scale-dependent relations [24,71–73] and a sampling scale bigger than the object studied might cause a loss of information that is provided at finer spatial resolutions [26]. In order to understand if the relations found in this work hold with decreasing spatial resolution, the spatial sensitivity of the spectral diversity–biodiversity relationship should be investigated. A previous work in a prairie grassland [24] identified a strong scale dependence of the spectral diversity–biodiversity relationships and suggested that the optimal pixel size for distinguishing $\alpha$-diversity in prairie plots was similar to the size of an individual herbaceous plant (1 mm to 10 cm). This might hamper the monitoring of the $\alpha$-diversity of lichen-dominated biocrusts at landscape scale with sensors installed on satellite or airborne platforms, which most likely would have a bigger pixel size than the one used in this work.

Conversely, the use of unmanned aerial vehicles (UAVs), which can carry light-weight imaging sensors of different spatial (reaching 2–4 cm/pixel in many cases) and spectral resolutions [74] may allow replicating studies like ours at landscape level [75]. Many of these sensors do not have very high spectral resolution, but have at least one band at ~680 nm [74], which would allow to calculate the SD_CR$_{680}$ (the spectral diversity metric that presented the highest predictability of lichens’ $\alpha$-diversity in our study) to monitor $\alpha$-diversity of lichens at larger spatial scales. Achieving the results shown in this work at landscape scale might also be hampered by the spectral properties that other components (e.g., vascular plants, mosses, bare soil) have on the spectral reflectance measured from remote sensors, which might hinder the separation of the pure spectral component of lichens and to estimate their $\alpha$-diversity. This issue could be however solved using spectral mixture analysis [76], a technique that models a mixed spectrum as a combination of its spectral components weighted by the correspondent subpixel fractional covers [77] and has already proved to be successful for mapping biocrusts [31].

5. Conclusions

The very high accuracies obtained classifying the hyperspectral images using SVMs showed the reliability of this methodology to identify different biocrust-forming lichens. Therefore, we were able to extract pure spectral signatures of different biocrust constituents and to evaluate the relationships between $\alpha$-diversity and spectral diversity of lichens. We found that the SD_CR$_{680}$ nm was the spectral diversity metric that predicted the best the $\alpha$-diversity metrics that include richness and evenness in their calculations (i.e., Pielou’s and Simpson’s indices). As such, we suggest that this index could help to track spatio-temporal changes in lichen-dominated biocrust communities. In this context, the results of this study will help to improve future works upsampling the methodology here shown to coarser scales in drylands, a key task in any monitoring program aiming to assess the impacts of ongoing climate change and desertification processes in these environments.

**Author Contributions:** J.B.-S., C.P., G.T., R.G., M.L.d.G., F.T.M., and M.R. designed the field sampling; J.B.-S., C.P., G.T., R.G., M.L.d.G., conducted the field sampling; J.B.-S. and C.P. conducted imaging measurements; J.B.-S. analyzed the data; J.B.-S., C.P., G.T., and M.R. interpreted the results. All authors contributed to the discussion of the results and the final version of the manuscript.

**Funding:** The research has received funding from the European Union’s Horizon 2020 research and innovation 514 program under the Marie Skłodowska-Curie grant agreement no. 721995. F.T.M. acknowledges support from the European Research Council grant agreement no. 647038 (BIODESERT).

**Acknowledgments:** We thank Roberto Garzonio and Khelvi Biriukova for their help during the hyperspectral images acquisition, and Biagio Di Mauro for his comments during the early stages of this manuscript. We are also grateful to the three reviewers for their comments and suggestions, which improved this manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.
References


37. Weksler, S.; Rozenstein, O.; Ben-Dor, E. Mapping Surface Quartz Content in Sand Dunes Covered by Biological Soil Crusts Using Airborne Hyperspectral Images in the Longwave Infrared Region. *Minerals* 2018, 8, 318. [CrossRef]


51. Hanley, J.A.; McNeil, B.J. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982, 143, 29–36. [CrossRef]

52. Bradley, A.P. The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recognit. 1997, 30, 1145–1159. [CrossRef]


