The Optimal Leaf Biochemical Selection for Mapping Species Diversity Based on Imaging Spectroscopy

Yujin Zhao, Yuan Zeng *, Dan Zhao, Bingfang Wu and Qianjun Zhao

Key Laboratory of Digital Earth Science, Institute of Remote Sensing and Digital Earth (RADI),
Chinese Academy of Science, Datun Road, Chaoyang District, Beijing 100101, China; zhaoyj@radi.ac.cn (Y.Z.);
zhaoz@radi.ac.cn (D.Z.); wubf@radi.ac.cn (B.W.); qjzhao@cashq.ac.cn (Q.Z.)

* Correspondence: zengyuan@radi.ac.cn; Tel.: +86-10-6480-6217; Fax: +86-10-6485-8721

Academic Editors: Parth Sarathi Roy and Prasad S. Thenkabail
Received: 30 October 2015; Accepted: 2 February 2016; Published: 8 March 2016

Abstract: Remote sensing provides a consistent form of observation for biodiversity monitoring across space and time. However, the regional mapping of forest species diversity is still difficult because of the complexity of species distribution and overlapping tree crowns. A new method called “spectranomics” that maps forest species richness based on leaf chemical and spectroscopic traits using imaging spectroscopy was developed by Asner and Martin. In this paper, we use this method to detect the relationships among the spectral, biochemical and taxonomic diversity of tree species, based on 20 dominant canopy species collected in a subtropical forest study site in China. Eight biochemical components (chlorophyll, carotenoid, specific leaf area, equivalent water thickness, nitrogen, phosphorus, cellulose and lignin) are quantified by spectral signatures ($R^2 = 0.57–0.85$, $p < 0.01$). We also find that the simulated maximum species number based on the eight optimal biochemical components is approximately 15, which is suitable for most 30 m $\times$ 30 m forest sites within this study area. This research may support future work on regional species diversity mapping using airborne imaging spectroscopy.

Keywords: forest biodiversity; imaging spectroscopy; biochemical components; partial least squares; Monte-Carlo; species richness

1. Introduction

Forest biodiversity is a key element in the provisioning of ecosystem services, function and stability, and therefore is critical from ecological, conservation and management standpoints [1,2]. The loss of biodiversity due to human actions and climate change is a major global issue. However, limited data at regional scales constrains our understanding of the spatial distribution patterns and temporal changes in forest biodiversity with conservation efforts [3]. Satellite and airborne remote sensing provide a consistent method of observation for biodiversity monitoring across space and time. It enables indirect detection of habitat quality and heterogeneity or essential environmental parameters as surrogates of forest biodiversity, and it directly describes species presence, species richness and diversity [4–7]. However, it is still difficult to identify the number of canopy species or particular species of interest in some regions with complex species distributions. Recently, past studies in tropical forests have demonstrated that leaf biochemical variation is often dominated by the taxonomic diversity of plant species [8]. Given the causal linkages among spectral, chemical, and taxonomic diversity, Asner and Martin [9,10] have introduced a new approach called “spectranomics” that monitors tropical forest species diversity based on the taxonomic variations in leaf chemical and spectroscopic traits. In contrast to the empirical statistical method, the advantages of this approach include a biophysical foundation and better general applicability to different study sites.
Leaf biochemical properties are the critical determinant factors of plant physiology and ecosystem function [11,12]. The biochemical differences among species can alter the distribution of absorbed photosynthetically active radiation and drive variations in biogeochemical processes [8]. Although variations in climate, phenology and soil substrate might affect biogeochemistry and ecosystem function, high-diversity plant communities often contribute to significant heterogeneity in biochemical properties. Past studies have shown that the leaf biochemical properties mainly control the absorption features caused by biochemical composition control the shape of the leaf spectral optical properties and, thus, the optical remote sensing of canopies [13]. The variations in the remotely sensed spectral properties could be used to assess plant species diversity; this is known as the Spectral Variations Hypothesis (SVH) [14–16]. According to the theory of SVH, the spectral heterogeneity of remotely sensed images is correlated with the spatial heterogeneity of the environment, particularly of the plant communities, which in turn is linked to species richness [17]. A high spectral resolution is also critical to capture the detailed biochemical and biophysical information of forest canopy species [18]. The potential linkages between the taxonomic and biochemical diversity of forest species indicate a new direction to combine high spectral resolution remote sensing to map forest species diversity.

Imaging spectroscopy, also called hyperspectral remote sensing, provides continuous narrow-band spectral information that can be associated with biochemical properties [13,19,20]. The technological advances and increasing number of imaging spectrometers have greatly expanded the availability of high-fidelity data to improve the accuracy of remotely sensed estimates of leaf and canopy biochemical components [21–29]. If biochemical-spectral signatures are species specific, imaging spectroscopy may be used for species diversity monitoring based on interspecies differences in biochemical components.

The subtropical forest ecosystem occupies a quarter of the land area of China and involves remarkably high tree species diversity. Because of the importance of mapping the species distribution of subtropical forest for conservation strategies, we acquired the airborne hyperspectral and LiDAR data as well as field-measured leaf biochemical and spectral data of the Shennongjia National Forest Natural Reserve, which is the world’s only intact subtropical forest ecosystem in the middle latitude zone [30]. We expect to use these remote sensing and field-measured data to develop an approach for forest species diversity mapping based on variations in biochemical components. Consequently, the first step is to find the potential ties among the spectral, biochemical and taxonomic diversity using the field-measured leaf biochemical and spectroscopic data, which is the focus of this paper, and this primary study will support subsequent regional species diversity mapping using airborne hyperspectral and LiDAR data in future work.

The main objective of this paper is to determine the optimal biochemical components based on the “spectranomics” [10] method using imaging spectroscopy for mapping the forest species diversity. The optimal leaf biochemical selection is based on the criterion that the leaf optimal components should be well predicted by the spectral properties and that their combinations can distinguish a sufficient number of species. The detailed research questions include the following: (1) Do forest species have unique biochemical properties? (2) Do forest species have unique spectral signatures? (3) Do the spectral signatures of forest species correspond to their biochemical properties, and which are the optimal matches? (4) How do chemical and spectral variations track species richness, and where is the saturation point?

2. Materials

2.1. Study Site

Our study area is located in the Shennongjia National Forest Natural Reserve (31°21′20″–31°36′20″N, 110°03′05″–110°33′50″E) in Xingshan county, Hubei province, China (Figure 1). This region lies on the southern slope of the Shennongjia Forest District and in the northeast of the Three Gorges region and is in the transitional zone from the mid-subtropical belt to the northern
subtropical belt. It is characterized by a species-rich and multi-layered community of trees, lianas, and other epiphytic plants. The study area ranges in elevation from 826 m to 1761 m above sea level. Because of the influence of the southeast monsoon, the average precipitation is approximately 100–150 mm per month, but in the spring–summer (April–September) season, it can be approximately 200–300 mm per month [31,32]. This study site is mainly occupied by natural subtropical evergreen broadleaved forest, deciduous broadleaved forest and coniferous forest, as well as a mixture of these forests. A total of 22 field sample plots (30 m × 30 m) were simultaneously measured based on different forest distribution patterns and topographic strata.

2.2. Field Biochemical and Spectral Data

Based on the fieldwork, we selected a total of 20 dominant tree species (Table 1). To explore the causal linkage among the biochemical and spectral properties as well as the species diversity, we collected top-of-canopy leaves for each dominant tree species within 22 sample plots to measure their biochemical and spectral properties.

![Image](https://example.com/image)

**Figure 1.** Location of study area with field-measured sample plots.

**Table 1.** Description of sixteen biochemical values of twenty species.

<table>
<thead>
<tr>
<th>Tree Species</th>
<th>EWT</th>
<th>SLA</th>
<th>Chl-a</th>
<th>Chl-b</th>
<th>Ca</th>
<th>C</th>
<th>N</th>
<th>K</th>
<th>P</th>
<th>Mg</th>
<th>Zn</th>
<th>Mn</th>
<th>B</th>
<th>Cel</th>
<th>Lig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Populus davidiana</td>
<td>0.019</td>
<td>54.66</td>
<td>14.82</td>
<td>4.57</td>
<td>6.4</td>
<td>9.28</td>
<td>0.28</td>
<td>2.01</td>
<td>0.14</td>
<td>2.32</td>
<td>0.44</td>
<td>0.0457</td>
<td>0.152</td>
<td>1.06</td>
<td>2.68</td>
</tr>
<tr>
<td>Quercus aliena</td>
<td>0.0075</td>
<td>94.71</td>
<td>7.07</td>
<td>4.43</td>
<td>2.96</td>
<td>5.24</td>
<td>0.21</td>
<td>0.77</td>
<td>0.13</td>
<td>1.29</td>
<td>0.14</td>
<td>0.0023</td>
<td>0.061</td>
<td>0.62</td>
<td>1.46</td>
</tr>
<tr>
<td>Platycladus strobilaceae</td>
<td>0.017</td>
<td>77.21</td>
<td>6.48</td>
<td>3.37</td>
<td>1.3</td>
<td>6.1</td>
<td>0.19</td>
<td>1.59</td>
<td>0.11</td>
<td>2.83</td>
<td>0.28</td>
<td>0.0015</td>
<td>0.009</td>
<td>0.94</td>
<td>1.7</td>
</tr>
<tr>
<td>Quercus glandulifera</td>
<td>0.022</td>
<td>54.41</td>
<td>6.62</td>
<td>2.57</td>
<td>3.86</td>
<td>9.12</td>
<td>0.31</td>
<td>0.95</td>
<td>0.25</td>
<td>1.85</td>
<td>0.33</td>
<td>0.0039</td>
<td>0.008</td>
<td>1.09</td>
<td>2.65</td>
</tr>
<tr>
<td>Betula luminifera</td>
<td>0.029</td>
<td>48.82</td>
<td>8.4</td>
<td>3.48</td>
<td>4.92</td>
<td>10.17</td>
<td>0.46</td>
<td>1.76</td>
<td>0.39</td>
<td>1.01</td>
<td>0.53</td>
<td>0.0071</td>
<td>0.082</td>
<td>0.52</td>
<td>2.15</td>
</tr>
<tr>
<td>Quercus variabilis</td>
<td>0.025</td>
<td>39.20</td>
<td>21.68</td>
<td>5.87</td>
<td>9.44</td>
<td>12.93</td>
<td>0.40</td>
<td>1.34</td>
<td>0.17</td>
<td>2.51</td>
<td>0.43</td>
<td>0.0064</td>
<td>0.499</td>
<td>0.66</td>
<td>3.67</td>
</tr>
<tr>
<td>Rhhus potanini</td>
<td>0.020</td>
<td>66.21</td>
<td>11.48</td>
<td>3.32</td>
<td>4.83</td>
<td>7.18</td>
<td>0.23</td>
<td>1.76</td>
<td>0.17</td>
<td>2.05</td>
<td>0.31</td>
<td>0.0022</td>
<td>0.004</td>
<td>0.63</td>
<td>0.95</td>
</tr>
<tr>
<td>Juglans cathayensis</td>
<td>0.034</td>
<td>62.17</td>
<td>9.81</td>
<td>3.70</td>
<td>7.56</td>
<td>7.41</td>
<td>0.33</td>
<td>1.28</td>
<td>0.23</td>
<td>6.31</td>
<td>0.96</td>
<td>0.0043</td>
<td>0.013</td>
<td>0.78</td>
<td>1.71</td>
</tr>
<tr>
<td>Sorbus folgneri</td>
<td>0.019</td>
<td>62.21</td>
<td>11.25</td>
<td>4.02</td>
<td>5.47</td>
<td>7.83</td>
<td>0.18</td>
<td>1.78</td>
<td>0.12</td>
<td>1.89</td>
<td>0.48</td>
<td>0.0085</td>
<td>0.206</td>
<td>3.04</td>
<td>3.32</td>
</tr>
<tr>
<td>Populus davidiana</td>
<td>0.0081</td>
<td>97.71</td>
<td>8.60</td>
<td>2.87</td>
<td>3.68</td>
<td>4.99</td>
<td>0.13</td>
<td>0.85</td>
<td>0.19</td>
<td>1.82</td>
<td>0.32</td>
<td>0.0040</td>
<td>0.009</td>
<td>0.61</td>
<td>0.78</td>
</tr>
<tr>
<td>Sorbus folgneri</td>
<td>0.0063</td>
<td>132.46</td>
<td>2.42</td>
<td>0.23</td>
<td>2.19</td>
<td>3.66</td>
<td>0.11</td>
<td>0.56</td>
<td>0.06</td>
<td>1.01</td>
<td>0.14</td>
<td>0.0027</td>
<td>0.016</td>
<td>0.79</td>
<td>1.25</td>
</tr>
<tr>
<td>Castanea seguinii</td>
<td>0.026</td>
<td>47.31</td>
<td>4.86</td>
<td>1.69</td>
<td>4.02</td>
<td>10.34</td>
<td>0.17</td>
<td>0.90</td>
<td>0.19</td>
<td>2.51</td>
<td>0.67</td>
<td>0.0069</td>
<td>0.404</td>
<td>1.45</td>
<td>3.57</td>
</tr>
<tr>
<td>Dendrobenthamia japonica</td>
<td>0.029</td>
<td>50.24</td>
<td>13.54</td>
<td>5.57</td>
<td>6.27</td>
<td>9.45</td>
<td>0.15</td>
<td>1.23</td>
<td>0.10</td>
<td>4.47</td>
<td>0.77</td>
<td>0.0026</td>
<td>0.014</td>
<td>1.53</td>
<td>2.06</td>
</tr>
<tr>
<td>Clethra carvaleri</td>
<td>0.020</td>
<td>110.2</td>
<td>7.08</td>
<td>2.27</td>
<td>2.72</td>
<td>4.21</td>
<td>0.09</td>
<td>1.75</td>
<td>0.05</td>
<td>0.87</td>
<td>0.38</td>
<td>0.0074</td>
<td>0.37</td>
<td>0.44</td>
<td>1.51</td>
</tr>
<tr>
<td>Rhus chinensis</td>
<td>0.027</td>
<td>38.97</td>
<td>24.12</td>
<td>11.55</td>
<td>9.75</td>
<td>12.14</td>
<td>0.45</td>
<td>3.48</td>
<td>0.42</td>
<td>5.77</td>
<td>1.33</td>
<td>0.0051</td>
<td>0.016</td>
<td>1.35</td>
<td>3.40</td>
</tr>
<tr>
<td>Rhus communis</td>
<td>0.039</td>
<td>48.09</td>
<td>19.34</td>
<td>6.86</td>
<td>8.11</td>
<td>10.03</td>
<td>0.23</td>
<td>2.31</td>
<td>0.12</td>
<td>4.9</td>
<td>0.59</td>
<td>0.0031</td>
<td>0.013</td>
<td>0.74</td>
<td>3.06</td>
</tr>
<tr>
<td>Cunninghamia lanceolata</td>
<td>0.024</td>
<td>54.12</td>
<td>9.79</td>
<td>3.33</td>
<td>3.88</td>
<td>9.02</td>
<td>0.15</td>
<td>1.37</td>
<td>0.15</td>
<td>2.62</td>
<td>0.36</td>
<td>0.0031</td>
<td>0.18</td>
<td>1.19</td>
<td>4.97</td>
</tr>
<tr>
<td>Pinus massoniana</td>
<td>0.030</td>
<td>44.12</td>
<td>12.47</td>
<td>3.85</td>
<td>5.89</td>
<td>11.59</td>
<td>0.25</td>
<td>1.24</td>
<td>0.19</td>
<td>2.63</td>
<td>0.65</td>
<td>0.0085</td>
<td>0.499</td>
<td>1.13</td>
<td>5.46</td>
</tr>
<tr>
<td>Fagus engleriensis</td>
<td>0.011</td>
<td>83.35</td>
<td>10.08</td>
<td>3.72</td>
<td>4.44</td>
<td>5.97</td>
<td>0.15</td>
<td>0.61</td>
<td>0.13</td>
<td>2.4</td>
<td>0.37</td>
<td>0.0033</td>
<td>0.047</td>
<td>0.39</td>
<td>1.93</td>
</tr>
<tr>
<td>Pinus armandii</td>
<td>0.023</td>
<td>78.12</td>
<td>5.50</td>
<td>1.79</td>
<td>1.79</td>
<td>6.46</td>
<td>0.15</td>
<td>0.82</td>
<td>0.15</td>
<td>0.79</td>
<td>0.18</td>
<td>0.0047</td>
<td>0.053</td>
<td>0.67</td>
<td>2.73</td>
</tr>
</tbody>
</table>

Notes: The unit of EWT is g cm⁻²; SLA, cm² g⁻¹ and other biochemical components, mg 10⁻³ cm⁻².

Biochemistry. We measured chlorophyll a and b (Chl-a, Chl-b), total carotenoids (Car), specific leaf area (SLA), equivalent water thickness (EWT), total carbon (C), nitrogen (N), phosphorus (P),...
lignin (Lig), cellulose (Cel), and some trace elements, including calcium (Ca), magnesium (Mg), zinc (Zn), manganese (Mn), and boron (B), as the major biochemical components. Leaf biochemicals can be partitioned functionally into major groups related to light capture and growth, longevity and defense, and maintenance and metabolism [33]. Chlorophyll a (Chl-a), chlorophyll b (Chl-b) and carotenoids (Car) play an important role in the photosynthetic reaction process by light capture and transfer [22,28]. Specific leaf area (SLA; cm$^2$·g$^{-1}$) is an important leaf structural property positively related to potential relative growth rate (RGR), and it tends to scale positively with the mass-based light-saturated photosynthetic rate [34]. Leaf water is a vital indicator of vegetation moisture stress [35]. Secondary metabolites such as lignin and cellulose contribute to leaf defense and longevity [36]. Leaf concentrations of N, P and trace elements are critical predictors of vegetation health and basic biogeochemical cycling [11]. Each of these leaf parameters also makes a demonstrated contribution to the spectroscopy and thus optical remote sensing of canopies [13,19].

Small branches were collected from fully sunlit portions of the uppermost canopies of each species and were stored in plastic bags on ice for transport to the laboratory. Between 25 and 150 fresh leaves were randomly selected and then stored in a −80 °C freezer until pigment analyses (Chl-a, Chl-b, and Car) were performed in the laboratory. For needle-leaved species, the entire leaves of the branch were sampled. Chl-a, Chl-b, and Car contents were determined based on multiwavelength analysis at 470, 645, 662, and 710 nm [37]. Additional samples (40–250 leaves) were scanned for leaf area determination and weighed within 2 h of collection. Foliar samples were dried at 70 °C for at least 72 h and weighed to determine leaf water concentration ((fresh mass-dry mass)/dry mass) and specific leaf area (SLA; cm$^2$·g$^{-1}$). All of the dried leaves were analyzed for C and N concentration using the elemental analyzer method, P and K using Inductively Coupled Plasma-Atomic Emission Spectrometry (ICP-AES) after microwave cooking, Cel and Lig by organic solvent extraction and the other trace elements (Ca, Mg, Zn, Mn and B) by atomic absorption spectrophotometry. For each species, all leaf chemical measurements were expressed on an area basis (i.e., µg·cm$^{-2}$) through dividing the measured chemical concentration by SLA. This made the leaf biochemical units correspond to the remotely sensed estimates of the pixel-level canopy biochemical concentration. Chlorophyll a and b (Chl-a, Chl-b) were summed for a single measurement of total chlorophyll (Chl) concentration.

Spectroscopy. Hemispherical reflectance spectra with 350–2500 nm wavelengths were measured on 10 fresh leaves of each species immediately after detachment from the branch in the field. The leaf spectral reflectance was obtained using a portable field spectrometer SVC HR-768i (Spectra Vista, Poughkeepsie, NY, USA) attached via a fiber optic cable to a Li-Cor 1800 integrating sphere (Li-Cor 1800-12S, Li-COR, Inc., Lincoln, NE, USA) with a light source for full-range spectral measurements. The SVC spectrometer acquired measurements with a 1.5-nm spectral sampling interval in 350–1000 nm, 7.6 nm in 1000–1890 and 5 nm in 1890–2500. However, some data with a wavelength less than 400 nm and more than 2400 nm were unavailable because of a large amount of noise. Hence, they were excluded. The measured leaf spectra of each species were then averaged to minimize the effect of noise on the spectral response. A moving Savitzky–Golay filter [38] was finally applied to further smooth the spectra.

2.3. Statistical Analysis

To determine the biochemical diversity between species, we first standardized the measured biochemical values of each species based on the Min-Max normalization method:

$$\Lambda_{ij} = (M_{ij} - \min(M_{ij}))/\left(\max(M_{ij}) - \min(M_{ij})\right)$$

where $\Lambda_{ij}$ is the normalized value of biochemical component $j$ of species $i$, $M_{ij}$ is the measured value of biochemical component $j$ of species $i$, and $\min(M_{ij})$, $\max(M_{ij})$ are the min and max values of $M_{ij}$, respectively. Hierarchical cluster analysis was then used to determine whether the spectral signatures of the species are unique. Through clustering, those species with similar spectral signatures
are sorted into the same group. The Ward’s minimum variance method was applied as the criterion of hierarchical clustering analysis [39]. The distance between two clusters is the ANOVA (analysis of variance) sum of squares between the two clusters, with all the variables added up:

$$D_{KL} = \frac{||\overrightarrow{X_K} - \overrightarrow{X_L}||^2}{N_K + N_L}$$ \hspace{1cm} (2)

where $D_{KL}$ is the statistical distance between cluster $K$ and cluster $L$, $X_K$ and $X_L$ are the mean vectors for cluster $K$ and cluster $L$, respectively, and $N_K$ and $N_L$ are the number of observations in cluster $K$ and cluster $L$, respectively. At each generation of cluster, the within-cluster sum of squares is minimized over all partitions obtainable by merging two clusters from the previous generation. This clustering analysis finally generated a statistically dendrogram indicating the organization of species based on the degree of their spectral association.

Following the analysis of biochemical and spectral diversity, we used constrained partial least squares (PLS-PRESS) regression to determine the relationship between the species’ spectral and biochemical signatures. Without the consideration of multiple correlations among variables, the PLS approach utilizes the continuous, full-range spectrum rather than a band-by-band analysis to calculate the relative contribution of each biochemical constituent to the spectral signatures of the species [11]. The cluster and PLS-PRESS analyses were performed using the SAS JMP 9.0 statistical software package.

To explore how biochemical and spectral diversity change with species richness, the Monte-Carlo simulation technique was used to calculate the average change in the biochemical and spectral variance with increasing species diversity [18,40]. A single species is randomly selected from the total population of 20 species, and the average biochemical and spectral values are recorded (e.g., Chl). Other species are sequentially randomly selected and combined with the previously selected species to track the change in the variance of biochemical or spectral values. These random selections of species are repeated until the entire community is populated. The simulation is carried out 1000 times as the same operation. The model can be run for a single standardized biochemical component, multiple biochemicals (e.g., Chl and EWT) or any number ($n$) of combinations ($a_i$) of leaf biochemical components ($j$) per species ($i$). However, the change ($CV_{total}$) in biochemical combinations ($a_i$) is calculated as the square root of variance ($CV$) summed over all the normalized values of leaf biochemical component ($j$):

$$CV_{total} = \sum_{j=1}^{n} CV_j$$

where $\Lambda_j$ is the normalized value of biochemical component $j$ calculated in Equation (1). $S$ is the number of species. $CV_j$ is the square root of variance of biochemical component $j$. Similar to the biochemical combinations ($a_i$), an optical equivalent ($\lambda_i$) is also provided for analyzing taxonomic variation, but with a Savitzky–Golay filtered leaf average reflectance of tree species within 400–2400 nm.

3. Results

3.1. Forest Species Have Unique Biochemical Properties

To utilize the spectral variations caused by biochemical differences to predict species diversity, the essential prerequisite is to ensure the biochemical diversity among species. In this study, we standardize the absolute value of each measured leaf biochemical constituent (Table 1) to visually determine inter-species variation in leaf biochemical contents (Figure 2). The length of each color
segment expresses absolute biochemical differences among species and the relative importance of each leaf biochemical component to the total biochemical portfolio of species. For single biochemical components, no two species are absolutely equal, but many are quite similar. However, the combination of leaf biochemical components is different for any two species. For example, the EWT contents in *Clethra cavalerei* and *Swida macrophylla* are similar, but their carotenoid contents are quite different. The biochemical diversity between species will increase when additional biochemical components are combined.

Figure 2. The 16 standardized biochemical components of 20 dominant tree species.

Here, we do not take the internal factors of the intraspecies variations in biochemical components and external drivers of the change of climate and soil fertility into account [41,42] because a past study found that these effects are small compared to interspecies variation in biochemistry when considering the relative contribution of each constituent to the integrated biochemical portfolio of species [8]. The biochemical variation among species would be able to make a powerful contribution to predicting species diversity.

3.2. Forest Species Have Distinctive Spectral Signatures

The Savitzky–Golay filtered leaf average reflectance of each species shows different field spectral signatures among species (Figure 3). The spectral differences among species change with wavelength-associated biochemical absorption signatures. The significant spectral variations are mostly in the near-infrared (NIR; 700–1300 nm) and short-wave-infrared (SWIR; 1500–2400 nm) regions. The NIR spectral signatures are mainly controlled by variation in leaf water concentration and leaf structure, related to SLA [35]. The spectral reflectance variation in SWIR are also influenced by leaf water content, but with significant contributions from protein N, cellulose and lignin contents [19]. In the visible region (VIR; 400–700 nm), the spectral reflectance among species performs obviously variable around 550 nm and 670 nm, which are closely associated with the absorption of chlorophyll, carotenoid, and xanthophyll [43–45].
Although the broadleaved tree species has less-obvious differences in spectral properties, the same family or genus still can be identified. Not all of the biochemical components could be well predicted by spectral reflectance. Among the 16 biochemical components, eight of them (Chl, Car, SLA, EWT, N, P, Cel and Lig) show a relatively greater importance in driving the spectral variation of species.

The result of hierarchical cluster analysis of 20 dominant tree species also indicates the distinctive spectral signatures of most species (Figure 4). The color within the cluster diagram quantitatively depicts the spectral properties of each species, with yellows–reds and greens–blues showing high and low reflectance, respectively. The dendrogram at the right shows the statistical similarity among species. The different color groups at the left denote the cluster group of tree species. As is clear in the cluster diagram, even small differences in color quantitatively express the differences among the spectral characteristics. This leads to our understanding that, similar to biochemical uniqueness, few species have identical spectral signatures, but many are close. It is obvious that the coniferous and broadleaved tree species could be divided into two groups (Cunninghamia lanceolata, Pinus massoniana and Pinus armandii belong to coniferous tree species). Although the broadleaved tree species has less-obvious differences in spectral properties, the same family or genus still can be identified.

![Figure 3. Savitzky–Golay filtered field spectral signatures of 20 dominant tree species.](image1)

**Figure 3.** Savitzky–Golay filtered field spectral signatures of 20 dominant tree species.

![Figure 4. Hierarchical cluster analysis of 20 dominant species based on their spectral reflectance signatures from 400 to 2400 nm.](image2)

**Figure 4.** Hierarchical cluster analysis of 20 dominant species based on their spectral reflectance signatures from 400 to 2400 nm.
3.3. Spectral Signature Response to Biochemical Properties

We have explored the biochemical diversity and spectral diversity above, but another key step for forest biodiversity mapping is to determine whether the biochemical constituents of species can be estimated by their spectral signatures. PLS regression analysis is applied, and the two first components of the model are computed with cross validation. The prediction strength ($R^2$) provides a relative degree of the importance of each biochemical component in determining the spectral reflectance of all species (Figure 5). The results show that: (1) Chl-a, Chl-b and SLA are strongly predicted ($R^2 = 0.85$); (2) Car and EWT are also significantly quantified ($R^2 = 0.75–0.78$); (3) N, P, Cel and Lig concentrations are relatively well expressed by spectral signatures ($R^2 = 0.57–0.69$); and (4) the trace elements, such as Ca, Mg, Zn, Mn, and B, have no obvious relationship to leaf spectral properties. Not all of the biochemical components could be well predicted by spectral reflectance. Among the 16 biochemical components, eight of them (Chl, Car, SLA, EWT, N, P, Cel and Lig) show a relatively greater importance in driving the spectral variation of species.

![Figure 5. Cont.](image-url)
Figure 5. PLS scatterplots show the relationship between the field-measured biochemical components and those predicted by the field spectral reflectance.

The spectral reflectance weightings of the PLS regression analysis reveal the wavelength-specific importance for predicting the fifteen leaf chemical properties and SLA (Figure 6). The wavelengths of maximum importance are those with spectral reflectance weightings that diverge from zero. Although the spectral reflectance weightings are variable for different leaf chemical components, their change tendencies are almost consistent (except Zn) but opposite those of SLA. This may be explained by our area-basis biochemical measurements. In the VIR (especially within 500–700 nm), the strong PLS reflectance weightings indicate that the most important contributions come from Chl-a, Chl-b and Car, followed by SLA, C, Mg, Lig, Ca, K, N, P, Cel and others less so. The spectral reflectance weightings in the NIR are similar to those in the VIR, but EWT is also an important contributor. Cel, C, SLA and most pigments play a significant role in controlling the SWIR reflectance among species.
The spectral reflectance weightings of the PLS regression analysis reveal the wavelength-specific importance for predicting the fifteen leaf chemical properties and SLA (Figure 6). The wavelengths of maximum importance are those with spectral reflectance weightings that diverge from zero. Although the spectral reflectance weightings are variable for different leaf chemical components, their change tendencies are almost consistent (except Zn) but opposite those of SLA. This may be explained by our area-basis biochemical measurements. In the VIR (especially within 500–700 nm), the strong PLS reflectance weightings indicate that the most important contributions come from Chl-a, Chl-b and Car, followed by SLA, C, Mg, Lig, Ca, K, N, P, Cel and others less so. The spectral reflectance weightings in the NIR are similar to those in the VIR, but EWT is also an important contributor. Cel, C, SLA and most pigments play a significant role in controlling the SWIR reflectance among species.

3.4. Relationships among Biochemical Diversity, Spectroscopic Diversity and Species Richness

Figure 7 shows changes in biochemical and spectroscopic diversity with increased species richness based on the different biochemical component combinations using Monte-Carlo simulation. As the species richness increases, the biochemical and spectroscopic diversity raise nonlinearly, finally reaching saturation. A comparison of different numbers of combinations of biochemical components indicates that the higher numbers of biochemical combinations reach saturation later. Overall, the combination of all 16 measured biochemical components ($\alpha$) could provide more richness for tracking species diversity. It is also notable that the spectroscopic diversity for tracking species richness performed better than the biochemical diversity. This may be due to more biochemical and structural information being contained in the spectral properties of species.

On the other hand, the saturation point of the biochemical diversity can be regarded as the maximum species number that could be recognized by the variation in the biochemical components. The maximum species number is only three for single Chl, then increases to approximately 9, 11 and 13 for the two combinations (Chl + EWT), the four biochemical combinations (Chl + EWT + Cel + Lig) and the six biochemical combinations (Chl + EWT + Cel + Lig + SLA + N), respectively. For the eight biochemical combinations (Chl + EWT + Cel + Lig + SLA + N + Car + P), each of which could be well predicted by spectral signatures, the maximum species number expectedly reaches approximately 15, which is suitable for most 30 m $\times$ 30 m forest sites within this study area. Thus,
the eight biochemical components (Chl, Car, SLA, EWT, N, P, Cel and Lig) are confirmed to be the optimal biochemical components among the 16 components used for species richness monitoring. It is unsurprising that the maximum number of recognized species continues to increase when considering all 16 biochemical components.

Figure 6. Partial least squares regression (PLSR) reflectance weighting factors for fifteen leaf chemical properties and SLA.

3.4. Relationships among Biochemical Diversity, Spectroscopic Diversity and Species Richness

Figure 7 shows changes in biochemical and spectroscopic diversity with increased species richness based on the different biochemical component combinations using Monte-Carlo simulation. As the species richness increases, the biochemical and spectroscopic diversity raise nonlinearly, finally reaching saturation. A comparison of different numbers of combinations of biochemical components indicates that the higher numbers of biochemical combinations reach saturation later. Overall, the combination of all 16 measured biochemical components (α) could provide more richness for tracking species diversity. It is also notable that the spectroscopic diversity for tracking species richness performed better than the biochemical diversity. This may be due to more biochemical and structural information being contained in the spectral properties of species.

Figure 7. Change in biochemical and spectroscopic diversity with increased species richness based on the different numbered biochemical component combinations using Monte-Carlo simulation. α denotes the combination of all 16 measured biochemical components. λ denotes the combination of all measured leaf spectral reflectance signatures from 400 to 2400 nm.

4. Discussion

The leaf biochemical and spectral properties of subtropical forest species are highly diverse, further supporting the “spectranomics” method for monitoring forest species diversity. The PLS regression analysis also reveals the good relationship between the leaf biochemical and spectral properties among species. Therefore, the spectral diversity can serve as a radical surrogate for biochemical diversity, contributing to our understanding of the biophysical foundation of Spectral Variation Hypothesis (SVH). However, to successfully apply the “spectranomics” method, it is necessary to ensure the adequate capture of biochemical variation among species. An airborne or space-borne imaging spectrometer needs to be introduced to translate biochemical and spectral diversity into taxonomic diversity.

Our study in this paper suggests that high taxonomic diversity creates significant biochemical variation. The between-species biochemical differences can drive the variation in biogeochemical processes by affecting the uptake, storage and transformation of C, N, P and other nutrients, and they are likely to forecast the response of forest ecosystems to environmental change [8,11]. Although it remains unclear how the variations in climate, topography and soil substrate gradients drive the biogeochemical process in subtropical forests, we contend that the species-driven biochemical variation will exceed the effects of environmental change on biochemistry. In addition, whereas the biochemical differences within species might not be significant compared to that among species, additional biochemical data across different environmental gradients for a specific species could better increase our understanding of species-driven biochemical variation in subtropical forests.

Leaf-level PLS analysis indicates that Chl, Car, SLA, EWT, N, P, Cel and Lig of sixteen leaf biochemical components are highly correlated with reflectance spectra (Figure 5), with PLS reflectance weightings revealing the wavelength-specific contributions for estimating the leaf biochemical properties (Figure 6). For the eight selected optimal biochemical components, Chl and Car are...
heavily weighted in the VIR and NIR (510–800 nm), whereas SLA, EWT, N, P, Cel and Lig are expressed in the NIR and SWIR (700–2400 nm). Previous studies have often focused on the estimation of leaf chemicals from reflectance spectra with the absorption features of individual biochemical constituents, such as: (a) Cel and Lig were strongly correlated with three absorption features, centered near 1700, 2100, and 2300 nm \[19,36\]; (b) the spectroscopic estimation of N was also associated to the chlorophyll (Chl) absorption feature near 680 nm \[46\] and protein-related absorption at 2050 and 2170 nm \[36\]; and (c) EWT could be estimated using the near-infrared (NIR) water features near 867 and 1190 nm \[25,47,48\].

The accurate retrievals of leaf biochemical properties are still difficult based on canopy reflectance derived from airborne imaging spectroscopy because of the effect of canopy structure \[29\]. New airborne remote sensing technologies, especially LiDAR (Light Detection and Ranging), and physical modeling approaches, such as 4-SCALE \[49\], ACRM (Two-Layer Canopy Reflectance Model) \[50\], FRT (Forest Reflectance and Transmittance) \[51\] and PROSPECT-DART (Discrete Anisotropic Radiative Transfer) \[52\], have been developed. LiDAR is an active remote sensing technology, and its scanners can emit a high-frequency pulse that can penetrate vegetation canopy gaps and, by recording the return time and intensity of backscatter from targets, provide detailed information of both the ground and multiple points within the forest canopy \[53–55\]. LiDAR has shown enormous potential for capturing canopy structural characteristics, especially in the retrieval of leaf area index (LAI) and tree height, as well as the isolation of individual tree crowns \[53,56–58\]. It is also possible to apply these structural parameters derived from LiDAR to scale the canopy reflectance to the leaf level based on newly developed physically based radiative transfer models, such as LIBRAT \[59\] and DART 5 \[60\].

In this paper, we emphasize the potential tie among the biochemical diversity, spectral diversity and taxonomical diversity of forest species based on the Monte-Carlo simulation technique. It is worth noting that the key point of our simulations is to track the change in leaf biochemical and spectral diversity with species richness rather than to simulate whole canopies, as might be viewed by remote sensing satellites. The variation in canopy structural properties, especially LAI, can decrease the sensitivity of canopy biochemical and spectral variations among species, as does viewing geometry and extra-foliar, epiphyll growth \[9\]. Moreover, the effects of the structural diversity on taxonomic diversity variations challenge our ability to measure the effect of biochemical diversity. Therefore, it is critical for the “spectranomics” method to separate the effects of the structural and biochemical variations on spectral variation among species.

The eight selected biochemical components are suitable for diversity mapping in subtropical forests with a similar species distribution located in the south of China. However, in extremely high species richness (more than 15) regions, the “spectranomics” method would probably be restricted. On the other hand, it is still difficult to confirm how biochemical signatures are expressed taxonomically by species, genus, or family. A coordinated acquisition of more field measurements of biochemical and environmental data in combination with hyperspectral data and optimal algorithms is needed to indicate clearly the relationships among biochemical diversity, spectral diversity and taxonomical diversity in species-rich forests.

5. Conclusions

In this study, we try to understand the intrinsic linkage among biochemical diversity, spectral diversity and taxonomical diversity of forest species in a subtropical forest study site. Our results indicate that increasing species richness corresponds to increasing spectral diversity by way of increasing biochemical diversity. However, not all leaf biochemical components can be well predicted by their spectral reflectance. Thus, we need to determine the optimal biochemical combinations while considering their relative importance in driving spectral variation among species. We finally select Chl, Car, SLA, EWT, N, P, Cel and Lig, a total of eight optimal biochemical components for mapping species diversity using imaging spectroscopy. We also find that the simulated max species number is 15 based on the eight selected biochemical components.
This study supports our future work in regional species diversity mapping using airborne imaging spectroscopy based on the interspecies variations in the optimal biochemical components. However, the accuracy of remotely sensed estimations of leaf biochemical concentrations is often affected by canopy structural variations, leaving it unclear as to whether plant species or functional type can be identified directly from spectral signatures. Although many studies with imaging spectroscopy have determined the direct species-spectral relationship that mostly allows for species identification or plant functional type classification \[61–63\], the confounding effects of biochemical and structural properties on spectral signatures hinder their scalability and expansibility to other regions. Therefore, the biochemical and structural properties should be separated from the spectral reflectance of forest canopy with scaling of the canopy reflectance to the leaf level, yet it is still difficult to resolve this problem when depending on imaging spectroscopy alone. LiDAR can capture valuable 3D canopy structural parameters and isolate individual tree crowns, thus helping solve the scaling issues with the combination of physically based radiative transfer models \[64\].

In addition, the structural diversity of species shape traits leads to another theoretical basis of distinguishing species communities \[8,65,66\]. Several studies have demonstrated that the classification accuracy of tree species could be improved when taking LiDAR-derived tree height into consideration \[67,68\]. Moreover, with the isolation of individual tree crowns, we can calculate the uniform value for each of the eight optimal biochemical components and tree heights at the individual tree scale and then cluster every tree to different species; clustering results within the 30 m × 30 m or even 50 m × 50 m window indeed reflect the species richness. Consequently, the combination of airborne imaging spectroscopy and LiDAR can be expected to provide higher accuracy for predicting regional forest species diversity with the support of the optimal biochemical components.

Acknowledgments: This work was supported by the National Natural Science Foundation of China (No. 41201351) and the Director Foundation of the Institute of Remote Sensing and Digital Earth, CAS. We thank Zongqiang Xie, Wenting Xu, and Changming Zhao from the Shengnongjia Biodiversity Research Station of CAS for their support with the fieldwork. We also thank Zhugeng Duan, Wenwen Gao and Hong Chi for their assistance in the field sample collection.

Author Contributions: Yujin Zhao collected field data, processed and analyzed the data, and wrote the manuscript. Yuan Zeng assisted in guiding the design of the experiment, discussing the results, and revising the manuscript. Dan Zhao helped design the experiment and contributed to the manuscript. Bingfang Wu and Qianjun Zhao supervised this study.

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

References
Remote Sens. 2016, 8, 216


33. Asner, G.; Martin, R.; Suhaili, A. Sources of canopy chemical and spectral diversity in lowland bornean forest. *Ecosystems* 2012, 15, 504–517. [CrossRef]


