

## Accuracy and limitations for spectroscopic prediction of leaf traits in seasonally dry tropical environments

Annia Susin Streher<sup>1\*</sup>; Ricardo da Silva Torres<sup>2</sup>, Leonor Patrícia Cerdeira Morellato<sup>3</sup>; Thiago Sanna Freire Silva<sup>4</sup>.

<sup>1</sup> Universidade Estadual Paulista (Unesp), Instituto de Biociências, Rio Claro, São Paulo, Brazil.

<sup>2</sup> Department of ICT and Natural Sciences, NTNU - Norwegian University of Science and Technology, Ålesund, Norway.

<sup>3</sup> Universidade Estadual Paulista (Unesp), Instituto de Biociências, Departamento de Biodiversidade, Phenology Lab, Rio Claro, São Paulo, Brazil.

<sup>4</sup> Biological and Environmental Sciences, Faculty of Natural Resources, University of Stirling. Stirling, UK, FK9 4LA.

\* **Corresponding author** currently at Remote Sensing Division, National Institute for Space Research (INPE), Brazil. e-mail: annia.streher@gmail.com

Accepted refereed manuscript of:

Streher AS, Torres RdS, Morellato LPC & Silva TSF (2020) Accuracy and limitations for spectroscopic prediction of leaf traits in seasonally dry tropical environments. *Remote Sensing of Environment*, 244, Art. No.: 111828. DOI: <https://doi.org/10.1016/j.rse.2020.111828>

© 2020, Elsevier. Licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International <http://creativecommons.org/licenses/by-nc-nd/4.0/>

## Abstract

Generalized assessments of the accuracy of spectroscopic estimates of ecologically important leaf traits, such as leaf mass per area (LMA) and leaf dry matter content (LDMC), are still lacking for most ecosystems and particularly for non-forested and/or seasonally dry tropical vegetation. Here, we tested the ability of using leaf reflectance spectra to estimate LMA and LDMC and classify plant growth forms within the *cerrado* and *campo rupestre* vegetation, a seasonally dry non-forest vegetation types of Southeastern Brazil, filling an existing gap in published assessments of leaf optical properties and plant traits in such environments. We measured leaf reflectance spectra from 1648 individual plants comprising grasses, herbs, shrubs, and trees, developed partial least squares regression (PLSR) models linking LMA and LDMC to leaf spectra (400–2500 nm), and identified the spectral regions with the greatest discriminatory power among growth forms using Bhattacharyya distances. We accurately predicted leaf functional traits and identified different growth forms. LMA was overall more accurately predicted (RMSE = 8.58%) than LDMC (RMSE = 9.75%). Our model including all sampled plants was not biased towards any particular growth form, but growth-form specific models yielded higher accuracies and showed that leaf traits from woody plants can be more accurately estimated than for grasses and forbs, independently of the trait measured. We observed a large range of LMA values (31.80 - 620.81 g/m<sup>2</sup>), rarely observed in tropical or temperate forests, and demonstrated that values above 300 g/m<sup>2</sup> cannot be accurately estimated. Our results suggest that spectroscopy may have an intrinsic saturation point, and/or that PLSR, the current approach of choice for estimating traits from plant spectra, is not able to model the entire range of LMA values. This finding has very important implications to our ability to use field, airborne, and orbital spectroscopic methods to derive generalizable functional information. We thus highlight the need for increasing spectroscopic sampling and research efforts in drier non-forested environments, where environmental pressures lead to leaf adaptations and allocation strategies that are very different from forested ecosystems, producing thicker leaves. Our findings also confirm that leaf reflectance spectra can provide important information regarding differences in leaf metabolism, structure, and chemical composition. Such

information enabled us to accurately discriminate plant growth forms in these environments regardless the lack of variation in leaf economics traits, encouraging further adoption of remote sensing methods by ecologists and allowing a more comprehensive assessment of plant functional diversity.

**Keywords:** *leaf spectroscopy; LMA; LDMC; partial least squares regression (PLSR); plant functional traits, campo rupestre; cerrado.*

1 **1. Introduction**

2 Trade-offs in acquisition and allocation of resources to support growth, survival, and reproduction can lead  
3 to a variety of plant functional strategies, which have been the main focus of so-called “trait-based ecology”  
4 (Violle et al., 2007). In this context, leaf structural properties or ‘traits’ are essential variables - they are  
5 relatively easy to measure and indicate fundamental trade-offs in plant survival strategies (Díaz et al., 2016;  
6 Wright et al., 2004). Two very important functional leaf traits are leaf mass per area (LMA), a key trait  
7 related to plant growth and representing the trade-off between the energetic cost of leaf construction and  
8 the achieved light intercepting area (Poorter et al., 2009), and leaf dry matter content (LDMC), which  
9 captures the investment trade-off between structural versus liquid-phase processes (Hodgson et al., 2011;  
10 Kikuzawa and Lechowicz, 2011). Both traits have been extensively studied since they are key components  
11 of the “leaf economics spectrum” (LES) (Wright et al., 2004), an important functional dimension  
12 representing a continuum of carbon and nutrient investment strategies and leaf persistence. In the LES  
13 context, low LMA and LDMC values suggest rapid production of biomass, lower physical strength, and  
14 shorter leaf lifespan, while high values suggest efficient conservation of nutrients, slow growth rates, and  
15 long-lived leaves (Garnier et al., 2001).

16 A wide set of leaf traits, including many of the LES traits, can be detected and accurately predicted using  
17 leaf spectral reflectance data (Asner et al., 2016; Cavender-Bares et al., 2017; Curran et al., 2001; Serbin et  
18 al., 2014). Still, despite its ecological relevance, the relationship between leaf-level spectral reflectance and  
19 important functional foliar traits such as LMA and LDMC remains under-explored, and is mainly focused  
20 on plants from forested ecosystems (Van Cleemput et al., 2018). There is also an apparent inconsistency  
21 with the trait names used by the remote sensing community and by ecologists (Homolová et al., 2013). In  
22 the ecology literature, LMA is the ratio of leaf dry weight (mass) per leaf area ( $\text{g m}^{-2}$ ), while LDMC is an  
23 investment index, determined by the ratio between leaf dry and fresh weights (g/g) (Pérez-Harguindeguy  
24 et al., 2013). However, several remote sensing studies use the terms “leaf dry matter content” or “dry  
25 matter content” when actually referring to LMA (Homolová et al., 2013), and also refer to the ratio between

26 leaf fresh and dry weights (LDMC) as quantification of “leaf water content” (Ball et al., 2015; Cheng et al.,  
27 2011). Although LDMC is mathematically related to leaf water content ( $LWC = 1 - LDMC$ , Pérez-  
28 Harguindeguy et al., 2013), ecologists tend to consider LMA, LDMC, and LWC as separate traits.

29 Despite this misunderstanding among scientific fields, leaf spectral reflectance data has proven very  
30 successful for the estimation of LMA (Asner et al., 2011b; Chavana-Bryant et al., 2016; Doughty et al.,  
31 2017, 2011; Feilhauer et al., 2015; Féret et al., 2018; Serbin et al., 2014), and LDMC (Ali et al., 2016;  
32 Roelofsen et al., 2014), but the functional breadth of these studies remains limited (Homolová et al., 2013).  
33 Mixed performance results have been reported before, suggesting that LMA can be retrieved with low to  
34 moderately good accuracy (average RMSE 45%-30%, see Homolová et al., 2013 for a review), but with  
35 little agreement among physically based and empirical methods on the best spectral wavelengths for LMA  
36 estimation (Féret et al., 2018). Furthermore, most studies to date have been focused on forested systems  
37 (Van Cleemput et al., 2018).

38 There is a sufficient and well-established theoretical basis linking the spectral, chemical, and taxonomic  
39 diversity of tree species (Asner et al., 2014; Ball et al., 2015; Castro-Esau et al., 2006; Cavender-Bares et  
40 al., 2017; Curran et al., 1992; Ferreira et al., 2013; Sánchez-Azofeifa et al., 2009; Schweiger et al., 2018;  
41 Serbin et al., 2014; Sims and Gamon, 2002; Ustin and Gamon, 2010), but there are remarkable functional  
42 differences between leaves from forest plants in relation to plants from open-canopy environments. Trees  
43 reaching the top of the forest canopy have been successful in competing for light, and have consequently  
44 developed trait combinations that maximize growth rates in these environments (Falster and Westoby,  
45 2005), with more similar sun-exposed leaves in respect to growth strategy and nutrient stoichiometry  
46 (Niinemets, 2010). This is not generalizable to other vegetation types, such as savannas, due to differences  
47 in biomass allocation; savanna plants tend to allocate less biomass to leaves and stems than forest  
48 individuals (Hoffmann and Franco, 2003), as competition shifts from light towards water and other limiting  
49 resources, as well as being influenced by adaptations to fire, resulting in much greater plasticity of leaf  
50 structural traits (Hoffmann and Franco, 2003).

51 Diversification of leaf functional strategies is also conditioned by the integration of multiple traits at the  
52 plant level, underlined by the overall growth form of the plant (Rossato et al., 2015). The larger phenotypic  
53 plasticity of leaves and growth forms in savannas may thus affect the consistency of leaf trait-reflectance  
54 relationships, and potentially limits the utility of empirical trait–spectra relations usually applied in forested  
55 systems. A recent meta-analysis has shown that, from a structural perspective, only leaf area index has been  
56 extensively addressed by grassland and shrubland spectroscopy studies (Van Cleemput et al., 2018) and the  
57 number of studies predicting LMA and/or LDMC is very limited in these systems (Ball et al., 2015;  
58 Roelofsen et al., 2014; Wang et al., 2019).

59 In order to achieve a truly global remote sensing framework for assessing plant functional diversity, more  
60 effort is needed in sampling grassland and shrubland ecosystems on arid and tropical regions, in terms of  
61 both plant traits and spectroscopic measurements (Jetz et al., 2016; Martin et al., 2012; Schimel et al., 2015;  
62 Van Cleemput et al., 2018). This shortfall sets a fundamental limit to our knowledge regarding the  
63 generality of correlations between optical and structural traits (Van Cleemput et al., 2018) from plants with  
64 different growth forms, life histories, and deciduousness strategies, and is crucial for further adoption of  
65 spectroscopic approaches by ecologists, given the increasing availability and affordability of data generated  
66 by hyperspectral sensors.

67 Here, we measured LMA and LDMC, two ecologically-relevant functional leaf traits (Violle et al., 2007;  
68 Díaz et al., 2016; Feilhauer et al., 2018; Shipley et al., 2006) together with leaf-level spectral reflectance,  
69 discriminating among dominant growth forms found in *cerrado* and *campo rupestre* vegetation occurring  
70 along a seasonally dry tropical landscape. We then assessed the potential of spectroscopy to predict  
71 structural traits in such tropical and seasonally-dry environments, by addressing the following questions:  
72 (i) does the relationship between leaf spectra and leaf traits as we know it from forests hold on a grass-  
73 shrubby-dominated and water limited environment? and given that variations in leaf reflectance should  
74 come from variations in leaf chemistry and structure, (ii) do spectral reflectance provides more evidence of  
75 plant functional strategies than usually measured functional traits in seasonally dry environments?

76 **2. Materials and Methods**

77 *2.1 Study area and sampling design*

78 The Espinhaço Mountain Range, in Southeastern Brazil, is among the most ancient landscapes on Earth,  
79 having remarkably high levels of diversity and endemism with more than 5000 described plant species  
80 (Fernandes, 2016; Fernandes et al., 2018; Silveira et al., 2016). Located at the southern portion of the  
81 Espinhaço Range, the Serra do Cipó subregion (19°23'29.8" S, 43°32'00.7" W) is also known for its  
82 megadiverse vegetation, with more than 1800 species recorded within a 200 km<sup>2</sup> area (Alves et al., 2014;  
83 Giuliatti et al., 1987). The climate of Serra do Cipó is marked by strong seasonality with two  
84 distinguishable seasons: a warm rainy season from October to April (average temperatures between 18 °C  
85 and 28 °C; monthly precipitation > 60 mm) and a cold dry season from May to September (average  
86 temperatures between 13 °C and 25 °C; monthly precipitation <40 mm) (Fernandes et al., 2016; ANA  
87 2017).

88 The rugged topography of Serra do Cipó provides a complex combination of topographic and edaphic  
89 conditions, which can lead to frequent and abrupt changes in vegetation structure and composition, where  
90 a large variety of plant growth forms and phenotypes assemble (Schaefer et al., 2016; Silveira et al.,  
91 2016). At lower elevations, a gradient of *cerrado* vegetation types differing from each other in structure,  
92 composition, and deciduousness can be found, while above 1000 m, natural areas of *campo rupestre*  
93 *sensu stricto* (Silveira et al., 2016) growing on shallow soils dominate the landscape. *Campo rupestre* has  
94 been described as a montane, fire-prone grassland vegetation growing on sandy, stony, or waterlogged  
95 soils, interspersed with rock outcrops dominated by evergreen shrubs, forbs and a few herbs (Morellato &  
96 Silveira 2018).

97 We sampled leaf traits and leaf reflectance spectra during the October 2016 – March 2017 growing season  
98 (Streher et al. 2017). Our study design included five sampling sites distributed along the elevation  
99 gradient, from 820 m to 1500 m, based on the natural environmental stratification of elevation and  
100 edaphic conditions (Mattos et al. 2019). Within each elevation, four transects of 250 m, distant at least 50

101 m from each other, were established based on expert knowledge and interpretation of high-resolution  
102 aerial images, ensuring the inclusion of all vegetation types (a proxy for edaphic conditions and resulting  
103 functional assemblages) found within each site (see Mattos et al. 2019, for detailed description of  
104 vegetations and soil). Our samples thus encompassed all types of *cerrado* and *campo rupestre* vegetation,  
105 and are hereafter referred to as *campo rupestre*, as this was the dominant vegetation sampled.  
106 Sampling points were established at 7 m intervals along each transect, with a 3.5 m search radius  
107 delimited around each point. Within each search radius, we identified and sampled three individual plants,  
108 applying the following selection criteria: 1) we identified the three individuals closest to the center of the  
109 search radius belonging to morphotypes not sampled before in the same transect; 2) if less than three  
110 individuals from new morphotypes were found, we sampled the closest individuals to the center of the  
111 search radius, regardless of species, to reach three samples per sampling point. This sampling strategy  
112 was designed to ensure maximal sampling of morphotypic variation and maximizing trait variability,  
113 while still reflecting the relative abundances of different morphotypes. For each individual plant, three  
114 fully-expanded sun leaves were sampled. In total, we sampled 4944 leaves from 1648 individual plants,  
115 encompassing all observed growth form and representing the majority of plant phenotypes found at Serra  
116 do Cipó.

117

## 118 *2.2 Plant growth form definitions*

119 We followed the ‘growth form’ classification system proposed by Dansereau (1951), which relies on the  
120 forms (morphological aspects and height) shown by plants in their aboveground structure, and has already  
121 been applied to *cerrado* plants by Rossatto & Franco (2017). The plants at Serra do Cipó encompass an  
122 array of woody and herbaceous growth forms, comprising trees, shrubs, sub-shrubs, herbs, and grasses  
123 (Zappi et al., 2014, Mattos et al. 2019). Based on the proposed classification system and field  
124 observations, we classified all the growth forms encountered into three dominant classes found in *cerrado*  
125 (Warming, 1908):



- 126 ● “*Woody*”: taller plants with secondary vascular growth, such as trees (woody plants with a  
127 defined stem, taller than 2m) and shrubs (height between 2 and 3 m, without a dominant stem and  
128 having lignified branches and stems);
- 129 ● “*Forbs*”: plants with herbaceous and/or partially lignified stems, but with herbaceous branches,  
130 such as herbs (small eudicots from 0.1– 0.6m height, with herbaceous stems and branches) and  
131 sub-shrubs (plants with 0.5 – 1m height, generally with a thickened, partially lignified stem, and  
132 with aerial parts growing annually from an underground woody xylopodium);
- 133 ● “*Graminoids*”: monocot plants, including grasses and sedges from the *Poaceae*, *Xyridaceae*, and  
134 *Cyperaceae* family.

135

136 From the 1648 sampled individuals, 369 (22%) were classified as “*Forbs*”, 564 (34%) as “*Graminoids*”  
137 and 715 (54%) as “*Woody*”. We randomly subset 300 samples of each growth form group and then  
138 performed a One-Way ANOVA to compare if trait data is significantly different between growth  
139 forms. We tested for homoscedasticity and the normality distribution of residuals using standardized  
140 residuals versus fitted values scatter plots and Shapiro–Wilk test. When normality could not be accessed,  
141 log-transformed response variables were used. Post hoc Tuckey tests were applied in order to test for  
142 differences among groups of plant forms.

143

### 144 2.3 Leaf trait measurements

145 For trees and shrubs, we harvested branches of individual canopies containing sunlit and mature leaves,  
146 while for grasses we sampled the whole plant, keeping roots when possible (Pérez-Harguindeguy et al.,  
147 2013). We followed partial rehydration protocols by immediately storing the samples in moistened sealed  
148 plastic bags, under elevated CO<sub>2</sub> concentrations and saturated air humidity, stored in lightproof containers  
149 filled with ice (Garnier et al., 2001; Pérez-Harguindeguy et al., 2013). We kept the samples at ~ 4 °C in  
150 the dark, and measurements were taken between six to eight hours after harvesting. From each  
151 branch/individual sampled, we removed three healthy leaves with no serious herbivore or pathogen

152 damage, including petioles, blotted them dry to remove surface water, immediately weighed them to  
153 determine saturated fresh mass (Garnier *et al.*, 2001) and then measured reflectance spectra. All spectral  
154 measurements were taken within the same day (Foley *et al.*, 2006), between six to eight hours after branch  
155 harvesting (see next section). We then determined one-sided leaf area (Pérez-Harguindeguy *et al.*, 2013)  
156 by photographing each leaf under a straight overhead (nadir) view, while gently pressing individual  
157 leaves between a glass plate and a sheet of paper including a printed distance scale, ensuring photo scale  
158 calibration and thus accurate area measurements. We then calculated leaf area using the ImageJ2 software  
159 (Schindelin *et al.*, 2015). After photographing, we oven-dried leaf samples at 80 °C for 72 hours to  
160 determine leaf dry mass to the nearest 0.01 g. We computed LMA (g/m<sup>2</sup>) as the ratio between dry mass  
161 and leaf area, and LDMC (g/g), as the ratio between leaf fresh mass and dry mass (Pérez-Harguindeguy *et*  
162 *al.*, 2013).

163

#### 164 *2.4 Leaf spectral measurements*

165 We acquired leaf spectra using a full-range (350–2500 nm) ASD FieldSpec 4 Standard spectroradiometer  
166 (Analytical Spectral Devices, ASD, Malvern, Worcestershire, UK), with a spectral resolution of 3 nm in  
167 the VNIR and 10 nm in the SWIR, and wavelength accuracy of 0.5 nm. We used the ASD leaf probe  
168 accessory, which measures the spectral reflectance at close range from the leaf. The probe contains its  
169 own calibrated light source and the measuring end of a bare fiber-optic cable (25° field-of-view (FOV))  
170 mounted at 42° perpendicular to the contact surface (Serbin *et al.*, 2014), minimizing measurement errors  
171 produced by variations in illumination geometry.

172 Bi-directional reflectance measurements were taken for the same three replicate leaves from which LMA  
173 and LDMC were estimated, immediately after obtaining saturated fresh mass. Leaves were arranged over  
174 a large black non-reflective surface, covering the whole diameter of the contact probe (10 mm) and  
175 ensuring that no light escaped the measurement. Plants with small leaves or leaflets were arranged so that  
176 the FOV was fully covered, without any gaps or excessive overlap, using more than a single leaf or leaflet

177 when necessary. For each leaf, ten measurements were taken at one to six different parts of the leaf  
178 adaxial surface (depending on leaf size), avoiding main veins, herbivory and pathogens damage when  
179 possible, following the protocols and standards by Asner & Martin (2009). For compound leaves, we took  
180 up to 10 measurements of different leaflets. The final leaf spectrum of each leaf was then given as the  
181 average of the 10 scans.

182 To ensure measurement quality and improve signal-to-noise ratio (SNR), we re-calibrated the  
183 spectrometer for dark current and stray light between each set of leaf replicates, using a white reflectance  
184 reference (Spectralon; Labsphere Inc., Durham, NH, USA). Recorded spectra were read using the  
185 “FieldSpectra” package (Serbin *et al.*, 2014) of the R statistical language, version 3.4.0 (R Development  
186 Core Team 2007), and underwent quality assurance by visual assessment. Finally, we averaged the  
187 triplicate measurements of all leaf traits and leaf reflectance to the individual level, and trimmed the full-  
188 range leaf spectra at the far edges (450 to 2400 nm), to remove data with low SNR.

189

### 190 *2.5 Leaf trait predictive modeling*

191 We used partial least squares regression (PLSR) models (Geladi and Kowalski, 1986; Wold *et al.*, 2001),  
192 adapting the approach from Serbin *et al.* (2014), to predict LMA and LDMC from leaf spectral properties.  
193 PLSR is the most employed method for relating leaf spectroscopy and leaf traits, due to its capacity to  
194 compensate for multicollinearity and reduce a large predictor matrix down to a relatively low number of  
195 predictors, the non-correlated latent components (Feilhauer *et al.*, 2015; Serbin *et al.*, 2014; Wu *et al.*,  
196 2017).

197 We fit four models to predict each of the two leaf traits: a model based on all observations (“*All*”), and  
198 three models restricted by plant growth form (“*Woody*”, “*Forbs*”, and “*Graminoids*”), for a total of eight  
199 PLSR models. Based on the initial results, we also fitted four additional models for a subset of the  
200 original LMA dataset, comprising only values between 0 and 300 g/m<sup>2</sup>. For each model, we split our data  
201 into training (70%, hereafter train set) and validation (30%, hereafter test set), using the  
202 “createDataPartition()” function from the “caret” package (Kuhn, 2008) in R, to ensure that both sets

203 spanned the entire range of measured values for each trait. To reduce overfitting, we optimized the  
204 number of PLSR latent variables in the final models by minimizing the root mean square error (RMSE) of  
205 the prediction residual sum of squares (PRESS statistic, Chen *et al.*, 2004). For the larger datasets (“*All*”,  
206 “*Woody*”, and “*Graminoids*”), we calculated the PRESS statistic of successive model components using a  
207 10-fold cross-validation scheme, while for the “*Forbs*” dataset we used a standard leave-one-out cross  
208 validation (LOOCV) analysis as recommended for datasets with fewer observations (Serbin *et al.*, 2014).  
209 We assessed the final accuracy of each model by calculating the RMSE value between predicted and  
210 observed trait values in the test set, expressing it in the original variable units (RMSE), as percentage of  
211 the sample data range (%RMSE), and as the ratio of each model RMSE to the mean value of the trait  
212 dataset (mRMSE). Thus, we computed the coefficient of determination ( $R^2$ ) of the observed versus  
213 predicted values of each model, to understand the percentage of variance explained by the model in the  
214 test dataset. We also report RMSECV, the RMSE obtained from the cross-validation procedure using the  
215 10-fold or LOOCV methods, as discrepancies between RMSECV and RMSE can indicate model  
216 overfitting (Kuhn and Johnson, 2013).  
217 Lastly, we computed the variable importance of projections (VIP, Wold (1994)) metric for each model, to  
218 identify the spectral regions that contributed the most to the prediction of each leaf trait. VIP is the  
219 weighted sum of squares of the PLSR-weights, with the weights calculated from the amount of variance  
220 from the response variable explained by each PLS component (Wold 1994).

221

## 222 *2.6 Spectral dissimilarities among plant growth forms*

223 To understand the contribution of different spectral regions to the identification of plant functional  
224 strategies, we evaluated spectral dissimilarity between plant growth forms using the Bhattacharyya  
225 distance (Bhattacharyya, 1943; Kailath, 1967) (Eqn. 1). This metric quantifies the integrated difference  
226 between two individuals of different growth forms over the full spectral range, identifying the  
227 wavelengths with the greatest discriminatory power. This metric has been successfully applied for the

228 recognition of differences between species (Baldeck and Asner, 2014), and plants with different growing  
229 habits (Sánchez-Azofeifa et al., 2009).

230

$$231 \quad B = \frac{1}{8} (\mu_i - \mu_j)^T \Sigma^{-1} (\mu_i - \mu_j) + \frac{1}{2} \ln\left(\frac{|\Sigma|}{\sqrt{|\Sigma_i| |\Sigma_j|}}\right) \quad \text{Eqn 1}$$

232

233

234 where  $\mu_i$  and  $\mu_j$  are the mean values across all spectral bands for species  $i$  and  $j$ ,  $\Sigma_i$  and  $\Sigma_j$  are the

235 covariance matrices for each individual, and  $\Sigma$  is the pooled covariance matrix.  $B$  is the Bhattacharyya

236 distance.

237

238 We used a randomized approach to estimate the distribution of  $B$  by randomly sampling 1000 pairs of

239 spectra for each combination of growth forms (“Woody” x “Grass”; “Woody” x “Forb” and “Forb” x

240 “Grass”), and then computing the average and spread (standard deviation) of the 1000 calculated pairwise

241 distances for each combination.

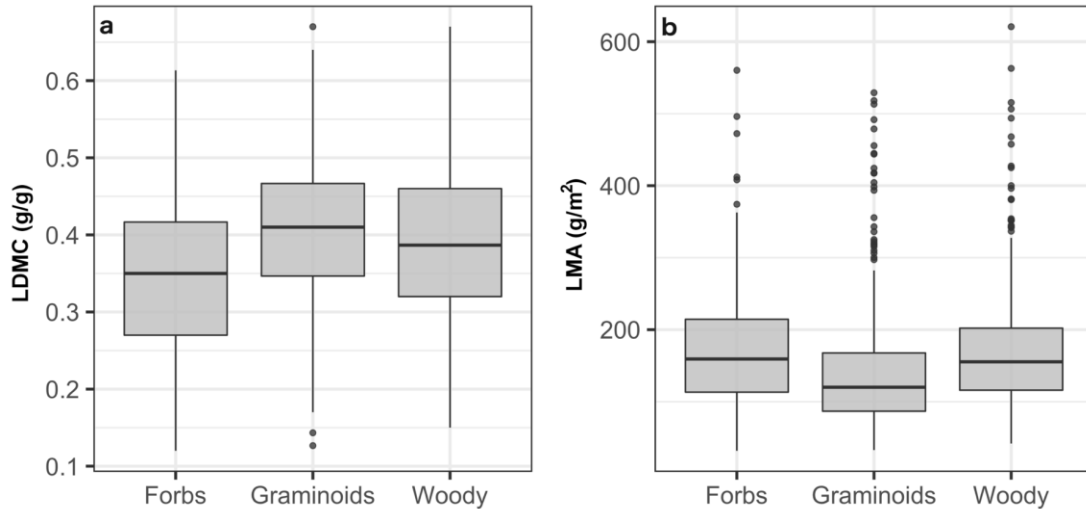
242

243 **3. Results**

244

245 *3.1 Leaf trait variability*

246 Differences in LDMC and LMA were subtle among growth forms (LDMC : $F_{2,897} = 24.44$  ,  $p < 0.001$ ;  
247 LMA:  $F_{2,897} = 16.21$ ,  $p < 0.001$ ) (Fig. 1, and Supplementary material S1). Overall LDMC values varied  
248 between 0.12 and 0.67 g/g, with a similar range of variation between growth forms (Fig. 1 and Table 1),  
249 with the largest LDMC range observed for “*Graminoids*” (0.12 – 0.67 g/g) and the smallest for “*Forbs*”  
250 (0.12 – 0.61 g/g). Average LDMC values per growth form were lowest for “*Forbs*” (mean = 0.34;  
251 standard error of the mean (se) =  $\pm 0.004$  g/g), followed by “*Woody*” ( $0.38 \pm 0.003$  g/g) and  
252 “*Graminoids*” ( $0.41 \pm 0.003$  g/g) (Fig. 1). Post hoc comparisons using Tukey test showed that there was a  
253 significant difference between the mean LDMC of “*Forbs*” and other growth forms, with woody plants  
254 showing an average of LDMC 0.05 g/g higher than “*Forbs*”, while “*Graminoids*” had an average LDMC  
255 value of -0.06 g/g lower than “*Forbs*” (Table S2). The total measured range of LMA values was 31.8 to  
256 621 g/m<sup>2</sup>. Average LMA values by growth form were lowest for “*Graminoids*” ( $137.9 \pm 3.31$  g/m<sup>2</sup>), and  
257 similar for the other two growth forms, with “*Woody*” having lower standard error among all growth  
258 forms ( $168.7 \pm 4.05$  g/m<sup>2</sup> for “*Forbs*”,  $167.9 \pm 2.76$  g/m<sup>2</sup> for “*Woody*”) (Fig. 1). “*Graminoids*” had the  
259 smallest LMA range (32.8 – 529 g/m<sup>2</sup>), and woody plants the largest LMA range (41.9 – 621 g/m<sup>2</sup>). The  
260 mean LMA values of “*Graminoids*” differ from the other growth forms, with LMA mean values lower  
261 than “*Woody*” and “*Forbs*” (30.93 g/m<sup>2</sup>, 28.35 respectively) (Table S2).



262

263 Figure 1. Variability of leaf functional traits measured for 1648 individuals of *campo rupestre* vegetation  
 264 at Serra do Cipó, Southeastern Espinhaço range, Brazil, including 369 individuals of the “*Forbs*” class,  
 265 564 individuals of the “*Graminoids*” class, and 715 individuals of the “*Woody*” class. **(a)** Leaf dry matter  
 266 content (LDMC); **(b)** leaf mass per area (LMA). Differences in LDMC and LMA were subtle among  
 267 growth forms, but statistically significant (LDMC:  $F_{2,897} = 24.44$ ,  $p < 0.001$ ; LMA:  $F_{2,897} = 16.21$ ,  $p <$   
 268  $0.001$ ) (Table S1).

269

### 270 3.2 PLSR modeling

271 Both leaf traits were predicted with high accuracy from reflectance measurements of fresh leaf material,  
 272 and no models showed signs of overfitting (Table 1). Overall, LMA was estimated from leaf reflectance  
 273 with higher accuracy (%RMSE = 8.58 %) than LDMC (%RMSE = 9.75 %), however the predicted values  
 274 from the LDMC PLSR model explained more (68%) of the variance of the predicted values than the  
 275 LMA PLSR model (58%) (Table 1). In general, “*Graminoids*” were the growth form with the worst  
 276 modelling performance for both traits, while “*Woody*” was the most accurate estimated growth form  
 277 (Table 1).

278 **Table 1.** Results of the partial least-squares regression (PLSR) modeling and cross-validation for each  
 279 leaf trait, showing the number of samples and range of trait variation for the global data set (all) and per  
 280 growth form. RMSECV is the root mean square error (RMSE) of the cross-validation procedure with  
 281 train data set; RMSE is the measured error using the test data; mRMSE is the ratio of the error of each  
 282 model in relation to the mean values (RMSE/mean); and the RMSE percentage (%RMSE) shows the error  
 283 of each model as a percentage of the observed data range. R<sup>2</sup> shows the goodness-of-fit between the  
 284 observations and the predicted values of each model. All results are presented for the entire range of LMA  
 285 and LDMC values (“All” class) and per growth form. “LMA < 300” represents the data set containing  
 286 only LMA values below 300 g/m<sup>2</sup>.

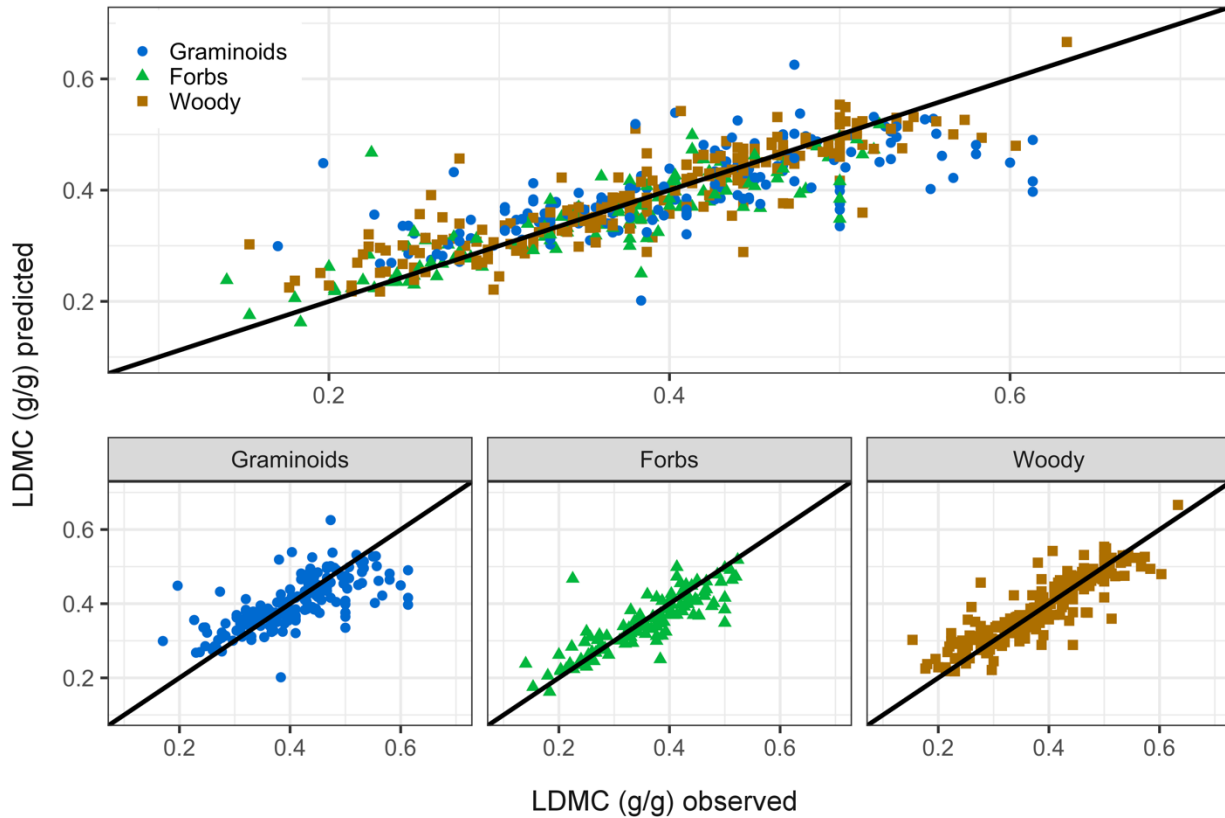
<i>Growth form</i>	<i>Number of samples</i>	<i>Range of variation (min - max)</i>	<i>RMSECV</i>	<i>Final number of latent variables</i>	<i>RMSE</i>	<i>mRMSE (RMSE/mean)</i>	<i>%RMSE (% of range)</i>	<i>R<sup>2</sup></i>
<b>LDMC</b>								
<i>ALL</i>	1648	0.12-0.67 (g/g)	0.052 (g/g)	20	0.053 (g/g)	0.13	9.75 %	0.68
<i>Graminoids</i>	564	0.12-0.67 (g/g)	0.063 (g/g)	17	0.059 (g/g)	0.15	11.66 %	0.48
<i>Forbs</i>	369	0.12-0.61 (g/g)	0.046 (g/g)	13	0.055 (g/g)	0.15	11.22 %	0.73
<i>Woody</i>	715	0.15-0.67 (g/g)	0.043 (g/g)	18	0.051 (g/g)	0.13	9.98%	0.78
<b>LMA</b>								
<i>ALL</i>	1648	31.80 - 620.81 (g/m <sup>2</sup> )	44.56 (g/m <sup>2</sup> )	17	50.58 (g/m <sup>2</sup> )	0.32	8.58 %	0.58
<i>Graminoids</i>	564	32.77 - 529.12 (g/m <sup>2</sup> )	44.89 (g/m <sup>2</sup> )	16	43.22 (g/m <sup>2</sup> )	0.31	8.70 %	0.60
<i>Forbs</i>	369	31.80 - 560.29 (g/m <sup>2</sup> )	53.12 (g/m <sup>2</sup> )	14	44.08 (g/m <sup>2</sup> )	0.26	8.34 %	0.42
<i>Woody</i>	715	41.89 - 620.81 (g/m <sup>2</sup> )	39.57 (g/m <sup>2</sup> )	18	43.33 (g/m <sup>2</sup> )	0.26	7.48 %	0.65
<b>LMA &lt; 300</b>								
<i>ALL</i>	1571	31.80 – 298.94 (g/m <sup>2</sup> )	32.00 (g/m <sup>2</sup> )	18	30.70 (g/m <sup>2</sup> )	0.21	11.49 %	0.71
<i>Graminoids</i>	539	32.77 - 297.23 (g/m <sup>2</sup> )	33.56 (g/m <sup>2</sup> )	20	35.73 (g/m <sup>2</sup> )	0.28	14.45 %	0.58
<i>Forbs</i>	337	31.80 - 298.94 (g/m <sup>2</sup> )	32.95 (g/m <sup>2</sup> )	19	35.32 (g/m <sup>2</sup> )	0.22	13.61 %	0.71
<i>Woody</i>	695	41.89 - 298.52 (g/m <sup>2</sup> )	28.65 (g/m <sup>2</sup> )	20	26.23 (g/m <sup>2</sup> )	0.16	10.79 %	0.78

287

288 Our PLSR LDMC spectral model had an overall error (RMSE) of 0.053 g/g, c.a. 9 % of the range of  
 289 LDMC values of the entire dataset (Table 1 and Fig. 2). Among growth-form restricted models, accuracy



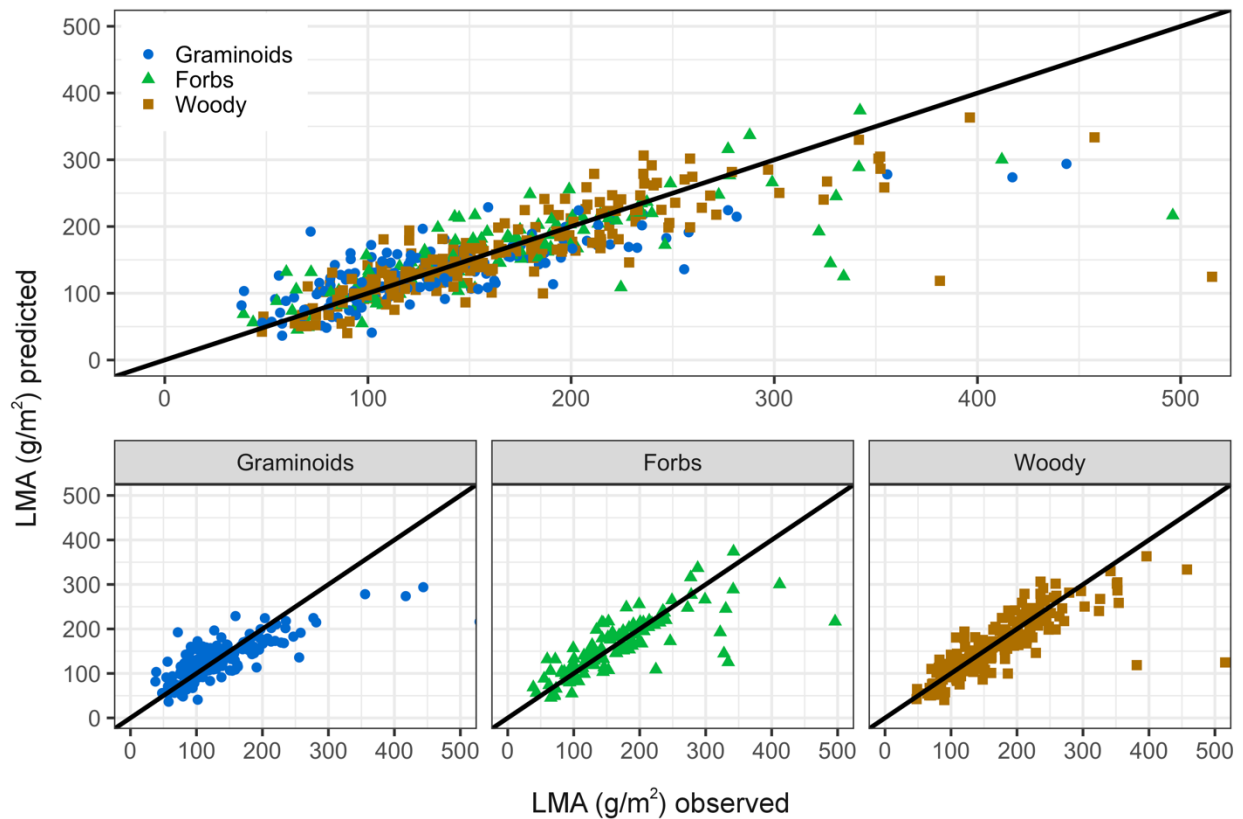
290 was higher for *Woody* plants, with %RMSE of c.a. 10% (RMSE = 0.051 g/g). The “*Graminoids*” and  
 291 “*Forbs*” models yielded similar error rates; although “*Graminoids*” models had higher overall error  
 292 (RMSE = 0.059 g/g) than “*Forbs*” (RMSE = 0.055 g/g), these errors represented similar ratios of error in  
 293 relation to the mean class value mRMSE = 0.15) and %RMSE considering the full range of values  
 294 (“*Graminoids*” %RMSE = 11.66%; “*Forbs*” %RMSE = 11.22%).



295  
 296 Figure 2. Leaf dry matter content (LDMC) as observed and predicted from leaf level reflectance using  
 297 partial least-squares regression (PLSR) models. The upper panel shows the prediction for the total range  
 298 of LDMC values (“*All*” class). The lower panels show the relationship between observed and predicted  
 299 LDMC values for each growth form. Symbols and colors indicate the growth form of each individual  
 300 plant: blue dots as “*Graminoids*”; green triangles as “*Forbs*”, and brown squares as “*Woody*”. Black lines  
 301 indicate the 1:1 relationship as reference.

302 The PLSR model for LMA had the highest overall accuracy with a RMSE of 50.58 g/m<sup>2</sup>, representing an  
 303 error percentage around 8 % of the range of LMA values of the entire dataset (Table 1 and Fig. 3). The

304 restricted models for LMA showed lower discrepancies between growth forms classes, with similar  
 305 RMSE between groups. The restricted model with highest accuracy corresponded to the “*Woody*” data  
 306 set, with a RMSE of 43.33 g/m<sup>2</sup> and error percentage of c.a. 7 % of the range of values within the class.  
 307 While the model accuracy for the “*Graminoids*” class was similar to the “*Woody*” class (RMSE = 43.22  
 308 g/m<sup>2</sup>), the error percentage of the range of values was higher (8.7%). The lowest accuracy was yielded by  
 309 the “*Forbs*” restricted model, with RMSE of 44.26 g/m<sup>2</sup>, ca. 8.4 % of the “*Forbs*” LMA value range.

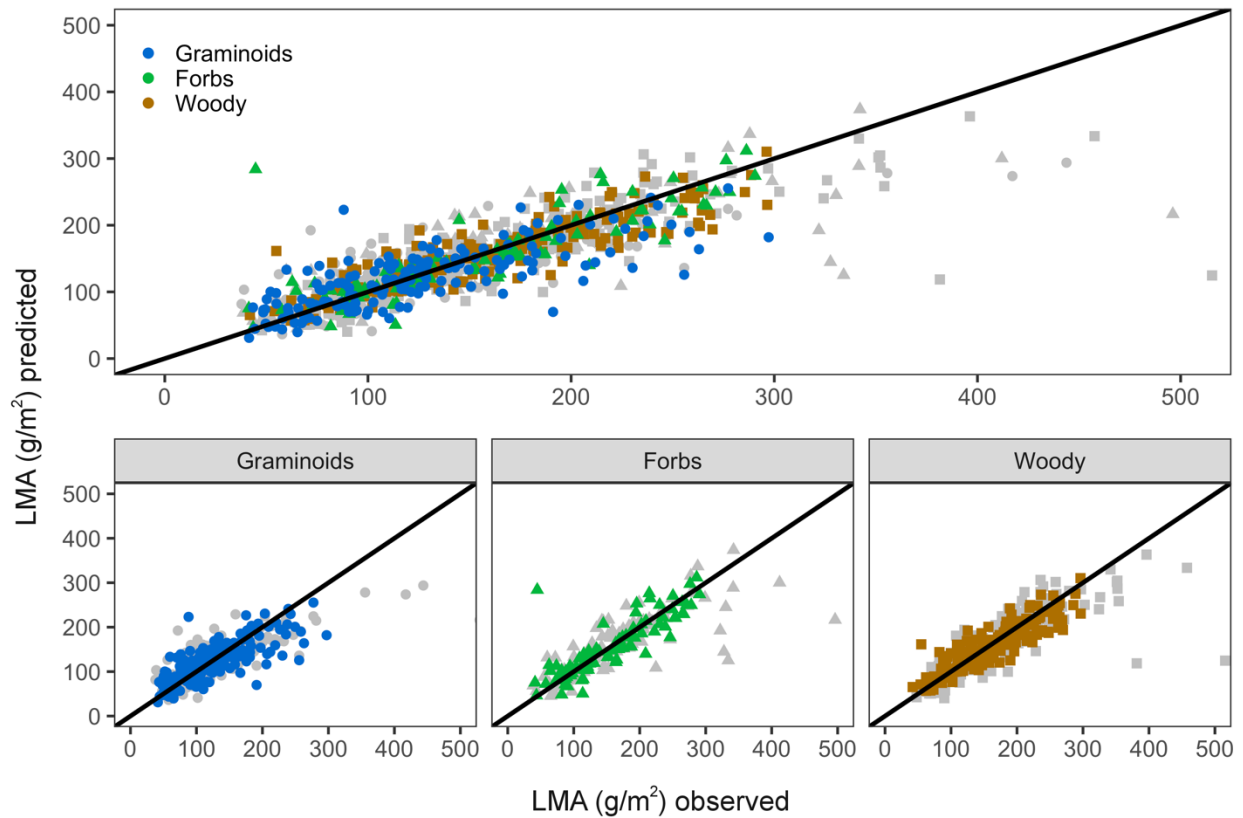


310  
 311 Figure 3: Partial least-squares regression (PLSR) results for observed vs. predicted leaf mass per area  
 312 (LMA). The upper panel shows the prediction for the total range of LMA values (“*All*” class). The lower  
 313 panels show the relationship between observed and predicted LMA values for each growth form. Symbols  
 314 and colors indicate the growth form of each individual plant: blue dots as “*Graminoids*”; green triangles  
 315 as “*Forbs*”, and brown squares as “*Woody*”. Black lines indicate the 1:1 relationship as reference.

316

317 We observed a loss of predictive power for all PLSR models for high LMA values, *i.e.* above 300 g/m<sup>2</sup>  
318 (Fig. 3), while PLSR models performed only slightly worst for LDMC high values (Fig 2). To quantify  
319 the influence of this loss, we refitted the PLSR models using only LMA values between 0 and 300 g/m<sup>2</sup>  
320 (Table 1), matching the range of LMA values usually observed for tropical (Asner et al., 2011a, 2011b)  
321 and temperate (Serbin *et al.*, 2014) forested systems, which are also typically used in radiative transfer  
322 models (Féret and Asner, 2011) and most frequently reported in the literature of leaf trait spectroscopy.  
323 These restricted-range PLSR models could explained more of LMA variance ( $R^2 = 0.78$ ) (Fig. 4),  
324 yielding an overall decrease in mRMSE of 0.21 in LMA values (Table 1 – LMA < 300 g/m<sup>2</sup>). The  
325 decrease in the overall error was also uniformly observed for models of each growth form, as so as an  
326 increase in the percentage of variance explained ( $R^2$ ) (Table 1). The highest improvement was found for  
327 the “*Forbs*” class, with a restricted range mRMSE of 0.22, down from mRMSE= 0.31 from the full range  
328 model (Table 1 and Fig. 5). The lowest performance of the restricted model was found for “*Graminoids*”  
329 (mRMSE= 0.28), with 1-fold change improvement. Using the same approach with LDMC values above  
330 0.05 g/g (Fig. 2), where the points start to deviate from the 1:1 line, and we found that removing these  
331 points from the analysis did not improve model accuracy and did not increase the percentage of variance  
332 explained (Fig S1 and Table S3).

333



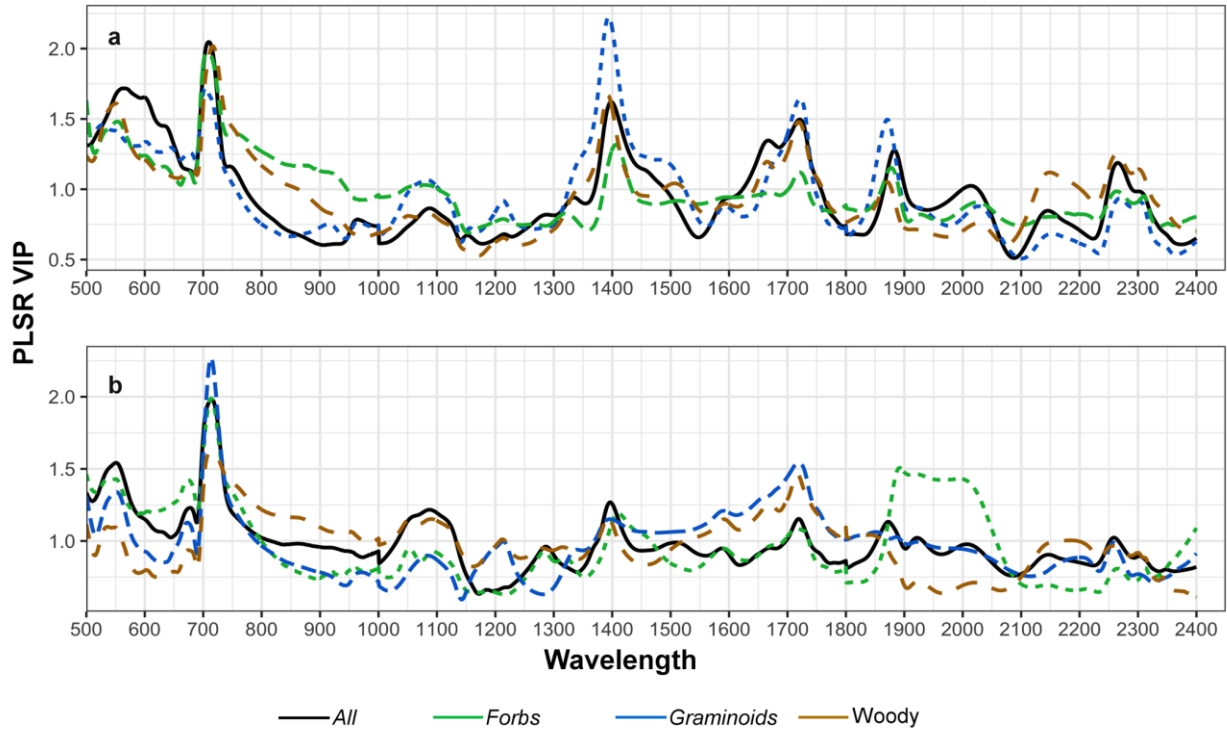
334

335 Figure 4: Partial least-squares regression (PLSR) results for observed vs. predicted leaf mass per area  
 336 (LMA), with values restricted to 0 - 300 g/m<sup>2</sup>. The upper panel shows the prediction for the total range of  
 337 LMA values (“All” class). The lower panels show the relationship between observed and predicted LMA  
 338 values for each growth form class. Symbols and colors indicate the growth form of each individual plant:  
 339 blue dots as “Graminoids”; green triangles as “Forbs”, and brown squares as “Woody”. Gray squares  
 340 comprise original LMA values above 300 g/m<sup>2</sup>, which were not included in the restricted models. Black  
 341 lines indicate the 1:1 relationship as reference.

342

343 Overall, VIP values had consistent patterns across the spectrum, with a few notable variations from  
 344 specific wavelengths (Fig. 5). For LDMC, the wavelength region centered in 1400 nm yielded the highest  
 345 VIP value, but wavelengths in the visible (VIS) (550 to 650 nm), red-edge (700-750 nm), and in the  
 346 shortwave infrared (SWIR) (around 1700 and 1900 nm) were also important (Fig. 4a). The most

347 important spectral region for LMA was the red-edge (700-750 nm), followed by the VIS region at the  
 348 wavelength centered in 550 nm (Fig. 5b). The VIP metric also varied in the position of peak importance  
 349 among growth forms for both traits, but specially for LMA, where a SWIR spectral region from 1900 to  
 350 2100 nm stood out for the “*Graminoids*” form (Fig. 5b). The red-edge (700-750 nm) was the spectral  
 351 region with the closest agreement of VIP values among growth forms for both leaf traits.



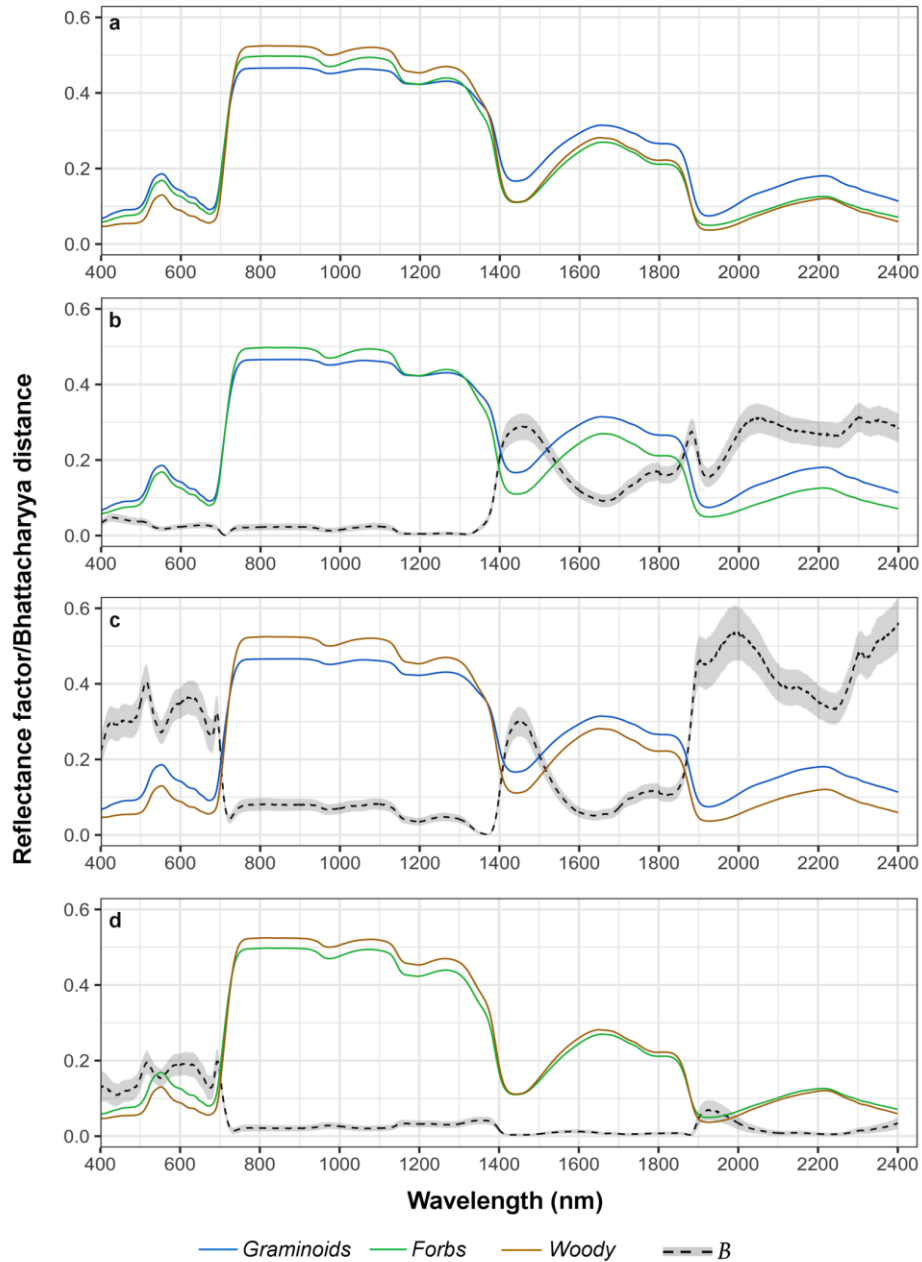
352  
 353 Figure 5: Partial Least Squares Regression (PLSR) variable importance of prediction (VIP) plotted by  
 354 wavelength for (a) leaf dry matter content (LDMC), and (b) leaf mass per area (LMA), measured for  
 355 *campo rupestre* plants at Serra do Cipó, Southern Espinhaço Range, Brazil. Colored lines represent the  
 356 three growth forms investigated in this study with the green dashed line representing “*Forbs*”, the blue  
 357 dashed line representing “*Graminoids*” and the brown dashed line representing “*Woody*”. The black solid  
 358 line represents “*All*” growth forms combined.

359  
 360

361 3.3 Leaf reflectance spectra dissimilarity among growth forms

362 Overall, full leaf reflectance spectra were able to track the expected ecophysiological changes in leaves  
363 from different growth forms (Fig. 6a). Reflectance measurements showed a reduction in reflectance along  
364 VIS wavelengths and a steep red-edge transition around 700 nm, where variance in reflectance of all  
365 plants was very low. Minor water absorption features were visible around 1000 and 1200 nm, while major  
366 absorption features stood out around 1400 and 1900 nm for all the three growth forms. Comparisons  
367 among growth forms showed that “*Woody*” plants had the lowest reflectance on the VIS range and the  
368 highest reflectance on the NIR region (Fig. 6a). The average reflectance spectra of “*Graminoids*” plants  
369 had the opposite pattern, with the highest reflectance in the VIS and SWIR, and lowest in the NIR regions  
370 (Fig. 6a). “*Forbs*” had intermediate reflectance values, with a spectral profile closer to “*Graminoids*” in  
371 the VIS region, while more similar to “*Woody*” in the SWIR (Fig. 6a).

372 Bhattacharyya distances ( $B$ ) indicated a greater degree of dissimilarity between the leaf reflectance spectra  
373 of “*Woody*” and “*Graminoids*” plants at the VIS (400 – 700 nm), around 1500 nm, and highest at the  
374 edge of the SWIR ( $\geq 1900$  nm) (Fig. 6c), in comparison to other pairwise interactions (Fig. 6b; 6d). As  
375 “*Forbs*” is an intermediate group between “*Graminoids*” and “*Woody*” plants, the dissimilarity between  
376 these pairs of interactions was subtler. The 1450 nm wavelength feature and the SWIR region yielded the  
377 highest degree of separability between “*Forbs*” and “*Graminoids*” (Fig. 6b), while “*Forbs*” and “*Woody*”  
378 were the most spectrally similar growth forms, as indicated by the smallest values of  $B$ , with the VIS  
379 region having the highest degree of separability (Fig. 6d).



380

381 Figure 6. Comparison of leaf reflectance spectral averages per growth form (a), and the spectral

382 dissimilarity (Bhattacharyya distance) between growth forms across the full wavelength range (400 –

383 2400 nm): (b) “Forbs” and “Graminoids”, (c) “Woody” and “Graminoids” and (d) “Woody” and

384 “Forbs”. The peaks observed on the Bhattacharyya index (B, dashed line and the gray shaded area

385 represents  $\pm 1$  standard deviation) indicate the spectral bands with highest dissimilarities among growth

386 forms.

#### 387 4. Discussion

388 Modern spectroscopy theory states that leaf reflectance spectra are quantitatively linked to leaf functional  
389 traits, particularly to LMA (Ustin & Gamon, 2010; Asner et al., 2011b; Serbin et al., 2019). Conversely,  
390 our results show that the high LMA values observed in our water limited, grassland-shrubland dominated  
391 system were partially correlated to leaf reflectance, saturating above 300 g/m<sup>2</sup>, differing from the  
392 expectations based mostly on LMA values observed for moist, forested systems. An important result from  
393 our study is that more efforts are needed to fully understand the relative influence of possible  
394 methodological shortcomings versus the biophysical limitations for predicting high LMA values from  
395 spectroscopy, which is paramount for developing models that will help to expand trait databases in order  
396 to address the known bias in geographical observational datasets and large-scale assessment of functional  
397 diversity (Schimel et al., 2015; Jetz et al., 2016; Van Cleemput et al., 2018). Our results support that  
398 spectroscopy is able to discriminate among woody, herbaceous, and graminoid growth-forms, as also  
399 shown by other studies (Knapp and Carter, 1998; Sánchez-Azofeifa et al., 2009), however we show that  
400 differences between growth forms in *campo rupestre* plants likely arise mainly from chemical leaf  
401 variation that are not captured by leaf structural trait variation. This illustrates the utility of the spectral  
402 approach in providing rapid, relatively low-cost and nondestructive measurements of key plant traits,  
403 highlighting that full-spectrum leaf profiles carry more ecological information than individual LES traits  
404 *per se*.

405 Considering the small variation in leaf traits, our results reinforce the potential of PLSR and spectroscopy  
406 to quantitatively describe structural foliar properties. Our general models were able to successfully  
407 explain variations related to leaf strategy without bias towards any growth form, going one step further  
408 towards the development of generalized global models. Still, the restricted PLSR models had overall  
409 better performances for woody plants than other growth forms for both measured traits. Our error rates for  
410 woody species (%RMSE 7% - 10%) are comparable to rates observed for tropical (Asner *et al.*, 2011b;  
411 %RMSE = 5.9%) and temperate forests (Serbin *et al.*, 2014; %RMSE= 10.1%). To the best of our



412 knowledge, there is a small number of studies addressing PLSR-spectroscopy modelling of LMA and  
413 LDMC from “herbaceous” plants, with emphasis on grasses. Our modelling resulted in an equal  
414 predictive performance for LMA on grasses in relation to previous studies (Wang et al., 2019; %RMSE  
415 12%), and slightly lower for LDMC (Roelofsen et al., 2014; RMSE = 0.10).

416 Although our empirical models provided good estimates of both leaf traits, it underestimated LMA values  
417 above 300 g/m<sup>2</sup>. Trees usually have LMA values up to ~350 g/m<sup>2</sup>, and most of the literature on empirical  
418 and radiative transfer models has tested the ability of spectroscopy to quantify LMA up to this value  
419 (Asner et al., 2011b; Cheng et al., 2014; Doughty et al., 2017; Feilhauer et al., 2015; Féret et al., 2018;  
420 Serbin et al., 2014). The global range of LMA variation spans two orders of magnitude (14 -1515 g/m<sup>2</sup>;  
421 Glopnet data – Wright *et al.*, 2004), and most studies of forest systems capture only c.a. 20% of this  
422 range. Our dataset covers c.a. 39% of the LMA worldwide variance. When we refitted our PLSR models  
423 constraining LMA values up to 300 g/m<sup>2</sup>, our predictive power improved considerably for all models  
424 (Table 1 and Fig. 5), particularly for eudicot herbs and sub-shrubs. Two key implications emerge from  
425 this result: 1) the PLSR method may not be able to predict large LMA variations; and/or 2) spectroscopy  
426 may not be sensitive to variations of high LMA values (*i.e.*, it has a saturation point). Multivariate linear  
427 non-parametric approaches like PLSR are considered state-of-the- art for operational mapping  
428 applications (Verrelst et al., 2015), and have been shown to perform comparably and equally well to other  
429 non-linear non-parametric methods like Random Forest, Support Vector Machine and Gaussian Processes  
430 Regression (Feilhauer et al., 2015; Van Cleemput et al., 2018; Wang et al., 2019). Our results set an  
431 important direction for future studies, showing the need to increase efforts in sampling leaf spectra for  
432 seasonally dry and dry vegetation sites, open, high light environments (*i.e.*, high LMA), and plants with  
433 contrasting resource use strategies. That is essential if we expect to fully understand and characterize the  
434 sensitivity of leaf spectroscopy and the feasibility of developing general, globally applicable methods for  
435 spectral LMA quantification.

436 The spectral regions selected for predicting LDMC were conservative among growth forms, and were  
437 associated with the red-edge inflection position centered at 740 nm, and a water absorption feature found  
438 at 1400 nm. The red edge is an inflection point where a steep increase in reflectance from the VIS (where  
439 chlorophyll absorbs light in the red region for photosynthesis), towards the NIR wavelengths occurs,  
440 where the intensification of the NIR reflectance is correlated with the increase of leaf thickness (Horler et  
441 al., 1983; Sims and Gamon, 2002). The relationship between spectra and LDMC is fundamentally the  
442 relationship of leaf water content, and leaf structure (carbon), reflecting the ecological significance of  
443 LDMC, which is an investment ratio in cell structure (red edge) versus fluid cell content (water  
444 absorption band) (Kikuzawa and Lechowicz, 2011; Shipley et al., 2006). The red edge was also the most  
445 important spectral region to predict LMA for all growth forms assessed, despite the SWIR being usually  
446 reported as the most important region of the spectrum for this trait in forest systems (Asner et al., 2011b).  
447 Nonetheless, Roelofsen *et al.* (2014) and Wang et al. (2019) have also found the VIS and NIR regions to  
448 be important for predicting the LMA of grasses. The red edge region is known for being strongly related  
449 to chlorophyll content (Curran et al., 2001), but this relationship is affected by variation in leaf thickness  
450 (Gitelson et al., 2003; Sims and Gamon, 2002). This is also consistent with the link between LMA and  
451 plant investment in chemical compounds distributed throughout the leaf mesophyll, which strongly affect  
452 leaf thickness and mass (Asner et al., 2011b; Poorter et al., 2009). Therefore, although unexpected, we do  
453 not consider the importance of red edge in predicting LMA a spurious correlation, and this interrelation  
454 can indicate structural limitations to photosynthesis as a result of increased LMA (Niinemets, 1999).

455 Future aerial and orbital remote sensors and missions may provide a better and urgently needed synoptic  
456 view of terrestrial ecosystem dynamics, as long as they allow for a high enough frequency of observations  
457 to capture specific phenological stages, thus yielding information on temporal leaf trait variation, a key  
458 information still mostly unexplored in trait-based ecology. Considering the spectral wavelengths  
459 identified in our analyses, multispectral sensors with multiple, high signal-to-noise spectral bands in the

460 red-edge (700-750 nm) and SWIR (around 1700 and 1900 nm) regions would bring us to the next level in  
461 scaling-up functional diversity patterns to larger regions.

#### 462 4.1 Insights from full reflectance spectra on plant functional characterization

463 Contrary to expectations, at Serra do Cipó LMA and LDMC values were very similar between growth  
464 forms, and the values found for grasses, eudicots herbs, and sub-shrubs are comparable to those found for  
465 woody plants. Usually, plants from the *cerrado* ground-layer are described as having thin, mesomorphic  
466 leaves (*i.e.*, low LMA and LDMC), since this stratum is completely destroyed during the passage of fire,  
467 while woody plants have thick and rigid sclerophyllous leaves, with large amounts of mechanical tissue,  
468 palisade parenchyma, and a well-developed vascular system (Rossatto et al., 2015; Rossatto & Franco,  
469 2017). The overall leaf structural similarity found among growth forms at Serra do Cipó can be linked to  
470 leaf persistence during drought conditions (Brum et al., 2017; Negreiros et al., 2014), with plants from  
471 abundant families (e.g., *Velloziaceae*, here classified as Forbs, and *Cyperaceae*, here classified as  
472 Graminoids), having species with desiccation-tolerant strategies and dormancy during the dry season  
473 (resurrection plants) (Alcantara et al., 2015; Oliveira et al., 2005). The high average values of LDMC  
474 found among growth forms can also be associated with the ability of species to endure very low water  
475 potentials and persist under dry conditions (Brum et al., 2017; Markesteijn et al., 2011; Oliveira et al.,  
476 2016).

477 Despite sharing very similar functional trait values, *campo rupestre* growth forms could be well  
478 distinguished based solely on leaf reflectance spectra. Our findings indicate that there are significant  
479 differences in pigment composition, and leaf anatomy, and consequently optical properties between  
480 growth forms that the two key LES traits did not capture. Over commonly measured traits, leaf spectra  
481 have the advantage of incorporating more of the total variation associated with leaf chemistry, anatomy  
482 and morphology into a single easy measurement, including variations that are difficult to measure or may  
483 be of unrecognized importance (Schweiger et al., 2018).

484 The potential of using leaf reflectance to discriminate growth forms is not new *per se* (Asner et al., 2011a;  
485 Ball et al., 2015; Castro-Esau et al., 2004; Knapp and Carter, 1998; Sánchez-Azofeifa et al., 2009). But  
486 our results are unique in the sense that the use of full reflectance spectrum allowed us to draw insights on  
487 leaf growth/allocation strategies, in a case where LMA and LDMC, two widely used functional traits, did  
488 not translate into the expected dissimilarities between growth forms. All growth forms had a substantial  
489 amount of mesophyll tissue, indicated by the high reflectance values along the NIR, but the mesophyll of  
490 trees and shrubs were generally thicker in comparison to other growth forms. This can be grasped from  
491 the fact that reflectance will increase when the amount of scattering structures per unit thickness increases  
492 (Knapp and Carter, 1998; Ustin and Gamon, 2010). The fact that NIR reflectance values from grasses  
493 were consistently lower than other growth forms indicates that lack of LMA variation is not a  
494 consequence of leaf thickness, which is highly correlated with NIR wavelengths (Knapp and Carter,  
495 1998), but most likely related to variations in leaf area (Streher et al, unpublished results from the same  
496 dataset). Woody plants and grasses had reflectance spectra with the largest differences in magnitude, and  
497 spectroscopy was able to capture the expected patterns: grasses had the highest VIS and lowest NIR  
498 reflectance, while woody plants had the opposite profile. The predominance of C4 grasses in *campo*  
499 *rupestre* suggests that grasses should have higher photosynthetic rates per unit of leaf area in comparison  
500 with other growth forms (Rossatto et al., 2015). The SWIR was the most important region to discriminate  
501 woody plants from grasses, suggesting differences in structural components, water content and water-use  
502 strategies (Curran, 1989) between these two growth forms, not captured by LMA and LDMC.

503 Eudicot herbs and sub-shrubs represented an intermediate growth form between woody plants and  
504 grasses. On one hand, they were differentiated from grasses by the amount of leaf water and structural  
505 properties absorbing along the SWIR, and lower photosynthetic rates than grasses, in contrast to the  
506 subtle differences found in the VIS from woody plants. The lack of proper spectral discrimination can be  
507 due to our inclusion of herbs and sub-shrubs within the same growth form due to sample size limitations.  
508 Sub-shrubs are unique since they have leaf anatomys similar to herbs (Rossatto et al., 2015), but are

509 functionally clustered with trees and shrubs (Rossatto and Franco, 2017). This implies that although they  
510 are on an evolutionary trajectory of ecological convergence with herbaceous plants, they are not  
511 phylogenetically independent of the tree lineages from which they have evolved (Rossatto and Franco,  
512 2017; Simon et al., 2009).

513 Leaf anatomy has been shown to diverge among growth forms, as plant form (Santiago and Wright, 2007)  
514 is related to leaf structure in environments characterized by frequent fire and highly seasonal rainfall  
515 (Rossatto et al., 2015). In our study site, the severely P-impoverished and shallow soils with low moisture  
516 retention impose a strong environmental filter (Abrahão et al., 2018), leading to a general convergence in  
517 ecological strategies, not reflecting the expected functional differences between leaf growth forms. The  
518 very high LMA and LDMC of scleromorphous leaves from different growth forms from *campo rupestre*  
519 places them in the stress-tolerant corner of Grime's C-S-R scheme (Dayrell et al., 2018; Negreiros et al.,  
520 2014). At a first glance, the use of soft leaf structural traits to distinguish growth forms in Serra do Cipó  
521 would restrict the use of "growth forms" as functional groups. Nevertheless, leaf spectral profiles shows  
522 that plant growth forms are still distinguishable within the multivariate trait space, particularly for traits  
523 related to photosynthetic activity, water-use strategies and lignin content, emphasized by the selection of  
524 VIS and SWIR regions to discriminate the growth forms assessed here.

525

## 526 **5. Concluding remarks**

527 We accurately predicted LMA and LDMC for seasonally dry tropical plants from spectroscopy, even  
528 though these traits had little variation among growth forms, reinforcing the ability of leaf spectroscopy to  
529 predict functional leaf traits. However, we also found an important limitation in using PLSR methods to  
530 predict high LMA values ( $> 300 \text{ g/m}^2$ ), resulting in underestimated values for LMA ranges that have been  
531 seldom addressed in the literature before. There are currently large biases in the sampling of plant traits  
532 and related spectra, favoring humid forested systems, hindering our understanding of spectroscopic  
533 relationships and limiting our ability to make reliable inferences and apply them to global biodiversity

534 science. Further work in determining whether limitations in LMA prediction are a methodological  
535 shortcoming from PLSR and/or a biophysical limitation of spectroscopy in high LMA environments is  
536 thus imperative.

537 A second key contribution from our study is showing that leaf reflectance carries more ecological  
538 information than commonly-used individual LES traits, at least when characterizing plant functional  
539 diversity in a seasonally dry, tropical area. By using full spectrum data, we revealed an idiosyncrasy of  
540 *campo rupestre* vegetation, showing that plant growth forms differ more in biochemical leaf traits than in  
541 the expected structural leaf aspects. The integrative depiction of foliar chemistry and morphology yielded  
542 by spectroscopy is thus essential to understand the response and resilience of vegetation to continued  
543 global change. Spectroscopy provides rapid, standardized, cost-effective, and easily replicated  
544 measurements that add more information about life-history strategies than measuring individual traits  
545 (Cavender-Bares et al., 2017; Schweiger et al., 2018), better enabling us to describe variability of leaf  
546 functional traits across different spatial and temporal scales (Serbin et al., 2014; Wang et al., 2018, 2019).

547 We thus recommend two directions for further work on plant spectroscopic modeling. First, although  
548 spectroscopy offers a powerful tool for acquiring trait data across scales, to fully understand the  
549 sensitivity and potential of leaf reflectance for plant ecology researchers should focus on sampling  
550 vegetations with contrasting life-history strategies and leaf longevities, from forests to grasslands and  
551 across wider seasonality gradients, producing reliable and standardized data and methods that can support  
552 global models relating foliar traits to leaf spectroscopy. Second, to enable a global understanding of trait-  
553 spectra relationship we stress the importance of reporting proper statistical information (*e.g.* goodness-of-  
554 fit-statistics, sample sizes, etc.), and standardization in trait nomenclature following known protocols, to  
555 simplify future comparisons between geographical locations and vegetation types. Advancing on these  
556 fronts will enable us to better understand plant trait variability and reduce uncertainties in functional  
557 spectroscopic ecology.

558

559 *Acknowledgments*

560 The authors thank two anonymous reviewers and Ricardo Dallagnol for comments on previous versions  
561 that improved the quality of this manuscript. Our research was supported by São Paulo Research  
562 Foundation (FAPESP) (grants: FAPESP-Microsoft Research Institute #2013/50155-0 and #2009/54208-  
563 6) and by the National Council for Scientific and Technological Development (CNPq) (grant: CNPq-PVE  
564 #400717/2013-1). ASS received a FAPESP scholarships (grants: #2015/17534-3 and BEPE #2016/00757-  
565 2 and #2017/ 01912-4). LPCM, RST and TSFS received research productivity grants from CNPq  
566 (#310761/2014-0, #311820/2018-2, #307560/2016-3, and #310144/2015-9, respectively). We thank  
567 ICMBio for granting the permits needed to work at Serra do Cipo National Park (PNSC) and its buffer  
568 zone. We also thank the Reserva Vellozia, Pousada Pouso do Elefante and Cedro Company for allowing  
569 access to private areas around the PNSC, and PELD-CRSC for the infrastructure and support. The authors  
570 thank Soizig Le Stradic, and MGG Camargo for helping with the setting of the sampling transects, Luis  
571 Fernando Campanha, Renata Martins, João Sobreiro, and Julio Alves for helping with field work. We are  
572 very thankful to our colleagues from the Phenology Lab, and Ecodyn Lab for their helpful insights and  
573 discussions.

574

575 **Authors Contribution:** Conceived and designed the study: ASS, TSFS, LPCM; collected data: ASS;  
576 analyzed data: ASS, TSFS and RST; wrote and revised the manuscript: ASS, RST, LPCM, and TSFS.

577

578 **References**

- 579
- 580 Abrahão, A., de Britto Costa, P., Lambers, H., Andrade, S.A.L., Sawaya, A.C.H.F., Ryan, M.H., Oliveira,  
581 R.S., 2018. Soil types filter for plants with matching nutrient-acquisition and -use traits in  
582 hyperdiverse and severely nutrient-impooverished *campos rupestres* and *cerrado* in Central Brazil,  
583 *Journal of Ecology*. <https://doi.org/10.1111/1365-2745.13111>
- 584 Alcantara, S., de Mello-Silva, R., Teodoro, G.S., Drequeceler, K., Ackerly, D.D., Oliveira, R.S., 2015.  
585 Carbon assimilation and habitat segregation in resurrection plants: a comparison between  
586 desiccation- and non-desiccation-tolerant species of Neotropical Velloziaceae (Pandanales). *Funct.*  
587 *Ecol.* 29, 1499–1512. <https://doi.org/10.1111/1365-2435.12462>
- 588 Ali, A.M., Darvishzadeh, R., Skidmore, A.K., Duren, I. van, Heiden, U., Heurich, M., 2016. Estimating  
589 leaf functional traits by inversion of PROSPECT: Assessing leaf dry matter content and specific leaf  
590 area in mixed mountainous forest. *Int. J. Appl. Earth Obs. Geoinf.* 45, 66–76.  
591 <https://doi.org/10.1016/j.jag.2015.11.004>
- 592 Alves, R., Silva, N., Oliveira, J., Medeiros, D., 2014. Circumscribing campo rupestre – megadiverse  
593 Brazilian rocky montane savanas. *Brazilian J. Biol.* 74, 355–362. [https://doi.org/10.1590/1519-](https://doi.org/10.1590/1519-6984.23212)  
594 [6984.23212](https://doi.org/10.1590/1519-6984.23212)
- 595 Asner, G.P., Knapp, D.E., Anderson, C.B., Martin, R.E., Vaughn, N., 2016. Large-scale climatic and  
596 geophysical controls on the leaf economics spectrum. *Proc. Natl. Acad. Sci.* 113, E4043–E4051.  
597 <https://doi.org/10.1073/pnas.1604863113>
- 598 Asner, G.P., Martin, R.E., 2009. Airborne spectranomics: mapping canopy chemical and taxonomic  
599 diversity in tropical forests. *Front. Ecol. Environ.* 7, 269–276. <https://doi.org/10.1890/070152>
- 600 Asner, G.P., Martin, R.E., Carranza-Jiménez, L., Sinca, F., Tupayachi, R., Anderson, C.B., Martinez, P.,  
601 2014. Functional and biological diversity of foliar spectra in tree canopies throughout the Andes to  
602 Amazon region. *New Phytol.* 204, 127–139. <https://doi.org/10.1111/nph.12895>
- 603 Asner, G.P., Martin, R.E., Knapp, D.E., Tupayachi, R., Anderson, C., Carranza, L., Martinez, P.,  
604 Houcheime, M., Sinca, F., Weiss, P., 2011a. Spectroscopy of canopy chemicals in humid tropical  
605 forests. *Remote Sens. Environ.* 115, 3587–3598. <https://doi.org/10.1016/j.rse.2011.08.020>
- 606 Asner, G.P., Martin, R.E., Tupayachi, R., Emerson, R., Martinez, P., Sinca, F., Powell, G.V.N., Wright,  
607 S.J., Lugo, A.E., 2011b. Taxonomy and remote sensing of leaf mass per area (LMA) in humid  
608 tropical forests. *Ecol. Appl.* 21, 85–98. <https://doi.org/10.1890/09-1999.1>
- 609 Baldeck, C. a., Asner, G.P., 2014. Improving remote species identification through efficient training data  
610 collection. *Remote Sens.* 6, 2682–2698. <https://doi.org/10.3390/rs6042682>
- 611 Ball, A., Sanchez-Azofeifa, A., Portillo-Quintero, C., Rivard, B., Castro-Contreras, S., Fernandes, G.,  
612 2015. Patterns of leaf biochemical and structural properties of Cerrado life forms: Implications for  
613 remote sensing. *PLoS One* 10, 1–15. <https://doi.org/10.1371/journal.pone.0117659>
- 614 Bhattacharyya, A., 1943. On a measure of divergence between two statistical populations defined by their  
615 probability distributions. *Bull. Calcutta Math. Soc.* 35, 99–109.
- 616 Brum, M., Teodoro, G.S., Abrahão, A., Oliveira, R.S., 2017. Coordination of rooting depth and leaf  
617 hydraulic traits defines drought-related strategies in the campos rupestres, a tropical montane  
618 biodiversity hotspot. *Plant Soil* 420, 467–480. <https://doi.org/10.1007/s11104-017-3330-x>
- 619 Castro-Esau, K.L., Sánchez-Azofeifa, G.A., Caelli, T., 2004. Discrimination of lianas and trees with leaf-  
620 level hyperspectral data. *Remote Sens. Environ.* 90, 353–372.  
621 <https://doi.org/10.1016/j.rse.2004.01.013>



622 Castro-Esau, K.L., Sánchez-Azofeifa, G.A., Rivard, B., Wright, S.J., Quesada, M., 2006. Variability in  
623 leaf optical properties of mesoamerican trees and the potential for species classification. *Am. J. Bot.*  
624 93, 517–530. <https://doi.org/10.3732/ajb.93.4.517>

625 Cavender-Bares, J., Gamon, J.A., Hobbie, S.E., Madritch, M.D., Meireles, J.E., Schweiger, A.K.,  
626 Townsend, P.A., 2017. Harnessing plant spectra to integrate the biodiversity sciences across  
627 biological and spatial scales. *Am. J. Bot.* 104, 966–969. <https://doi.org/10.3732/ajb.1700061>

628 Chavana-Bryant, C., Malhi, Y., Wu, J., Asner, G.P., Anastasiou, A., Enquist, B.J., Cosio Caravasi, E.G.,  
629 Doughty, C.E., Saleska, S.R., Martin, R.E., Gerard, F.F., Chavana-Bryant, C., Malhi, Y., Wu, J.,  
630 Asner, G.P., Anastasiou, A., Enquist, B.J., Saleska, S.R., Doughty, C., and Gerard, F., de la Riva,  
631 E.G., Olmo, M., Poorter, H., Ubersa, J.L., Villar, R., Mahowald, N., Lo, F., Zheng, Y., Harrison, L.,  
632 Funk, C., Lombardozzi, D., Goodale, C., Poorter, H., Niinemets, U., Poorter, L., Wright, I.J., Villar,  
633 R., Ke, Y., Im, J., Park, S., Gong, H., 2016. Leaf aging of Amazonian canopy trees revealed by  
634 spectral and physiochemical measurements. *New Phytol.* 11, 215. <https://doi.org/10.1111/nph.13853>

635 Chen, S., Hong, X., Harris, C.J., Sharkey, P.M., 2004. Sparse Modeling Using Orthogonal Forward  
636 Regression With PRESS Statistic and Regularization. *IEEE Trans. Syst. Man Cybern. Part B* 34,  
637 898–911. <https://doi.org/10.1109/TSMCB.2003.817107>

638 Cheng, T., Rivard, B., Sánchez-Azofeifa, A., 2011. Spectroscopic determination of leaf water content  
639 using continuous wavelet analysis. *Remote Sens. Environ.* 115, 659–670.  
640 <https://doi.org/10.1016/j.rse.2010.11.001>

641 Cheng, T., Rivard, B., Sánchez-Azofeifa, A.G., Féret, J.-B., Jacquemoud, S., Ustin, S.L., 2014. Deriving  
642 leaf mass per area (LMA) from foliar reflectance across a variety of plant species using continuous  
643 wavelet analysis. *ISPRS J. Photogramm. Remote Sens.* 87, 28–38.  
644 <https://doi.org/10.1016/j.isprsjprs.2013.10.009>

645 Curran, P.J., 1989. Remote sensing of foliar chemistry. *Remote Sens. Environ.* 30, 271–278.  
646 [https://doi.org/10.1016/0034-4257\(89\)90069-2](https://doi.org/10.1016/0034-4257(89)90069-2)

647 Curran, P.J., Dungan, J.L., Macler, B.A., Plummer, S.E., Peterson, D.L., 1992. Reflectance spectroscopy  
648 of fresh whole leaves for the estimation of chemical concentration. *Remote Sens. Environ.* 39, 153–  
649 166. [https://doi.org/10.1016/0034-4257\(92\)90133-5](https://doi.org/10.1016/0034-4257(92)90133-5)

650 Curran, P.J., Dungan, J.L., Peterson, D.L., 2001. Estimating the foliar biochemical concentration of leaves  
651 with reflectance spectrometry. *Remote Sens. Environ.* 76, 349–359. [https://doi.org/10.1016/S0034-4257\(01\)00182-1](https://doi.org/10.1016/S0034-4257(01)00182-1)

652

653 Dayrell, R.L.C., Arruda, A.J., Pierce, S., Negreiros, D., Meyer, P.B., Lambers, H., Silveira, F.A.O., 2018.  
654 Ontogenetic shifts in plant ecological strategies. *Funct. Ecol.* 0–2. <https://doi.org/10.1111/1365-2435.13221>

655

656 Díaz, S., Kattge, J., Cornelissen, J.H.C., Wright, I.J., Lavorel, S., Dray, S., Reu, B., Kleyer, M., Wirth, C.,  
657 Colin Prentice, I., Garnier, E., Bönsch, G., Westoby, M., Poorter, H., Reich, P.B., Moles, A.T.,  
658 Dickie, J., Gillison, A.N., Zanne, A.E., Chave, J., Joseph Wright, S., Sheremet'ev, S.N., Jactel, H.,  
659 Baraloto, C., Cerabolini, B., Pierce, S., Shipley, B., Kirkup, D., Casanoves, F., Joswig, J.S.,  
660 Günther, A., Falczuk, V., Rüger, N., Mahecha, M.D., Gorné, L.D., 2016. The global spectrum of  
661 plant form and function. *Nature* 529, 167–171. <https://doi.org/10.1038/nature16489>

662 Doughty, C.E., Asner, G.P., Martin, R.E., 2011. Predicting tropical plant physiology from leaf and  
663 canopy spectroscopy. *Oecologia* 165, 289–299. <https://doi.org/10.1007/s00442-010-1800-4>

664 Doughty, C.E., Santos-Andrade, P.E., Goldsmith, G.R., Blonder, B., Shenkin, A., Bentley, L.P., Chavana-  
665 Bryant, C., Huaraca-Huasco, W., Díaz, S., Salinas, N., Enquist, B.J., Martin, R., Asner, G.P., Malhi,

666 Y., 2017. Can Leaf Spectroscopy Predict Leaf and Forest Traits Along a Peruvian Tropical Forest  
667 Elevation Gradient? *J. Geophys. Res. Biogeosciences* 122, 2952–2965.  
668 <https://doi.org/10.1002/2017JG003883>

669 Falster, D.S., Westoby, M., 2005. Alternative height strategies among 45 dicot rain forest species from  
670 tropical Queensland, Australia. *J. Ecol.* 93, 521–535. [https://doi.org/10.1111/j.0022-](https://doi.org/10.1111/j.0022-0477.2005.00992.x)  
671 [0477.2005.00992.x](https://doi.org/10.1111/j.0022-0477.2005.00992.x)

672 Feilhauer, H., Asner, G.P., Martin, R.E., 2015. Multi-method ensemble selection of spectral bands related  
673 to leaf biochemistry. *Remote Sens. Environ.* 164, 57–65. <https://doi.org/10.1016/j.rse.2015.03.033>

674 Feilhauer, H., Schmid, T., Faude, U., Sánchez-Carrillo, S., Cirujano, S., 2018. Are remotely sensed traits  
675 suitable for ecological analysis? A case study of long-term drought effects on leaf mass per area of  
676 wetland vegetation. *Ecol. Indic.* 88, 232–240. <https://doi.org/10.1016/j.ecolind.2018.01.012>

677 Féret, J.-B., Asner, G.P., 2011. Spectroscopic classification of tropical forest species using radiative  
678 transfer modeling. *Remote Sens. Environ.* 115, 2415–2422.  
679 <https://doi.org/10.1016/j.rse.2011.05.004>

680 Féret, J.B., le Maire, G., Jay, S., Berveiller, D., Bendoula, R., Hmimina, G., Cheraiet, A., Oliveira, J.C.,  
681 Ponzoni, F.J., Solanki, T., de Boissieu, F., Chave, J., Nouvellon, Y., Porcar-Castell, A., Proisy, C.,  
682 Soudani, K., Gastellu-Etchegorry, J.P., Lefèvre-Fonollosa, M.J., 2018. Estimating leaf mass per area  
683 and equivalent water thickness based on leaf optical properties: Potential and limitations of physical  
684 modeling and machine learning. *Remote Sens. Environ.* 1–14.  
685 <https://doi.org/10.1016/j.rse.2018.11.002>

686 Fernandes, G., Almeida, H.A., Nunes, C.A., Xavier, J.H.A., Beirão, N.S.C., Carneiro, M.A.A.,  
687 Cornelissen, T., Neves, F.S., Ribeiro, S.P., Nunes, Y.R.F., Pires, A.C. V., Beirão, M. V., 2016.  
688 Cerrado to Rupestrian Grasslands: Patterns of Species Distribution and the Forces Shaping Them  
689 Along an Altitudinal Gradient, in: Fernandes, G.W. (Ed.), *Ecology and Conservation of*  
690 *Mountaintop Grasslands in Brazil*. Springer International Publishing, pp. 345–371.  
691 <https://doi.org/10.1007/978-3-319-2980>

692 Fernandes, G.W., 2016. *Ecology and Conservation of Mountaintop grasslands in Brazil*, 1<sup>o</sup>. ed. Springer  
693 International Publishing, Cham. <https://doi.org/10.1007/978-3-319-29808-5>

694 Fernandes, G.W., Barbosa, N.P.U., Alberton, B., Barbieri, A., Dirzo, R., Goulart, F., Guerra, T.J.,  
695 Morellato, L.P.C., Solar, R.R.C., 2018. The deadly route to collapse and the uncertain fate of  
696 Brazilian rupestrian grasslands. *Biodivers. Conserv.* 27, 2587–2603. [https://doi.org/10.1007/s10531-](https://doi.org/10.1007/s10531-018-1556-4)  
697 [018-1556-4](https://doi.org/10.1007/s10531-018-1556-4)

698 Ferreira, M.P., Grondona, A.E.B., Rolim, S.B.A., Shimabukuro, Y.E., 2013. Analyzing the spectral  
699 variability of tropical tree species using hyperspectral feature selection and leaf optical modeling. *J.*  
700 *Appl. Remote Sens.* 7, 073502. <https://doi.org/10.1117/1.JRS.7.073502>

701 Foley, S., Rivard, B., Sanchez-Azofeifa, G.A., Calvo, J., 2006. Foliar spectral properties following leaf  
702 clipping and implications for handling techniques. *Remote Sens. Environ.* 103, 265–275.  
703 <https://doi.org/10.1016/j.rse.2005.06.014>

704 Garnier, E., Shipley, B., Roumet, C., Laurent, G., 2001. A standardized protocol for the determination of  
705 specific leaf area and leaf dry matter content. *Funct. Ecol.* 15, 688–695.  
706 <https://doi.org/10.1046/j.0269-8463.2001.00563.x>

707 Geladi, P., Kowalski, B.R., 1986. Partial least-squares regression: a tutorial. *Anal. Chim. Acta* 185, 1–17.  
708 [https://doi.org/10.1016/0003-2670\(86\)80028-9](https://doi.org/10.1016/0003-2670(86)80028-9)

709 Gitelson, A.A., Gritz †, Y., Merzlyak, M.N., 2003. Relationships between leaf chlorophyll content and

710 spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant  
711 leaves. *J. Plant Physiol.* 160, 271–282. <https://doi.org/10.1078/0176-1617-00887>

712 Giuliatti, A.M., Menezes, N.L., Pirani, J.R., Meguro, M., Wanderley, M.G.L., 1987. Flora da Serra do  
713 Cipó, Minas Gerais: caracterização e lista das espécies. *Bol. Botânica da Univ. São Paulo.*

714 Hodgson, J.G., Montserrat-Martí, G., Charles, M., Jones, G., Wilson, P., Shipley, B., Sharafi, M.,  
715 Cerabolini, B.E.L., Cornelissen, J.H.C., Band, S.R., Bogard, A., Castro-Díez, P., Guerrero-Campo,  
716 J., Palmer, C., Pérez-Rontomé, M.C., Carter, G., Hynd, A., Romo-Díez, A., de Torres Espuny, L.,  
717 Royo Pla, F., 2011. Is leaf dry matter content a better predictor of soil fertility than specific leaf  
718 area? *Ann. Bot.* 108, 1337–1345. <https://doi.org/10.1093/aob/mcr225>

719 Hoffmann, W.A., Franco, A.C., 2003. Comparative growth analysis of tropical forest and savanna woody  
720 plants using phylogenetically independent contrasts. *J. Ecol.* 91, 475–484.  
721 <https://doi.org/10.1046/j.1365-2745.2003.00777.x>

722 Homolová, L., Malenovský, Z., Clevers, J.G.P.W., García-Santos, G., Schaepman, M.E., 2013. Review of  
723 optical-based remote sensing for plant trait mapping. *Ecol. Complex.* 15, 1–16.  
724 <https://doi.org/10.1016/j.ecocom.2013.06.003>

725 Horler, D.N.H., Dockray, M., BARBER, J., 1983. The red edge of plant leaf reflectance. *Int. J. Remote*  
726 *Sens.* 4, 273–288. <https://doi.org/10.1080/01431168308948546>

727 Jetz, W., Cavender-Bares, J., Pavlick, R., Schimel, D., Davis, F.W., Asner, G.P., Guralnick, R., Kattge, J.,  
728 Latimer, A.M., Moorcroft, P., Schaepman, M.E., Schildhauer, M.P., Schneider, F.D., Schrodtt, F.,  
729 Stahl, U., Ustin, S.L., 2016. Monitoring plant functional diversity from space. *Nat. Plants* 2, 16024.  
730 <https://doi.org/10.1038/nplants.2016.24>

731 Kailath, T., 1967. The Divergence and Bhattacharyya Distance Measures in Signal Selection. *IEEE Trans.*  
732 *Commun. Technol.* 15, 52–60. <https://doi.org/10.1109/TCOM.1967.1089532>

733 Kikuzawa, K., Lechowicz, M.J., 2011. Ecology of Leaf Longevity, *Ecological Research Monographs*,  
734 *Ecological Research Monographs*. Springer Tokyo, Tokyo. [https://doi.org/10.1007/978-4-431-](https://doi.org/10.1007/978-4-431-53918-6)  
735 [53918-6](https://doi.org/10.1007/978-4-431-53918-6)

736 Knapp, A.K., Carter, G.A., 1998. Variability in Leaf Optical Properties Among 26 Species from a Broad  
737 Range of Habitats. *Am. J. Bot.* 85, 940. <https://doi.org/10.2307/2446360>

738 Kuhn, M., 2008. Building Predictive Models in R Using the caret Package. *J. Stat. Softw.* 28.  
739 <https://doi.org/10.18637/jss.v028.i05>

740 Kuhn, M., Johnson, K., 2013. Applied Predictive Modeling. Springer New York, New York, NY.  
741 <https://doi.org/10.1007/978-1-4614-6849-3>

742 Markesteijn, L., Poorter, L., Bongers, F., Paz, H., Sack, L., 2011. Hydraulics and life history of tropical  
743 dry forest tree species: coordination of species' drought and shade tolerance. *New Phytol.* 191, 480–  
744 495. <https://doi.org/10.1111/j.1469-8137.2011.03708.x>

745 Martin, L.J., Blossey, B., Ellis, E., 2012. Mapping where ecologists work: Biases in the global  
746 distribution of terrestrial ecological observations. *Front. Ecol. Environ.* 10, 195–201.  
747 <https://doi.org/10.1890/110154>

748 Mattos, J. S., Camargo, M. G. G., Morellato, L. P. C., Batalha, M. A. 2019. Plant phylogenetic diversity  
749 of tropical mountaintop rocky grasslands: local and regional constraints. *Plant Ecology* 220(12):  
750 1119-1129.

751 Morellato, L. P. C., & Silveira, F. A. 2018. Plant life in campo rupestre: New lessons from an ancient  
752 biodiversity hotspot. *Flora*, 238, 1-10. <https://doi.org/10.1016/j.flora.2017.12.001>

753 Negreiros, D., Le Stradic, S., Fernandes, G.W., Rennó, H.C., 2014. CSR analysis of plant functional types

754 in highly diverse tropical grasslands of harsh environments. *Plant Ecol.* 215, 379–388.  
755 <https://doi.org/10.1007/s11258-014-0302-6>

756 Niinemets, Ü., 2010. Responses of forest trees to single and multiple environmental stresses from  
757 seedlings to mature plants: Past stress history, stress interactions, tolerance and acclimation. *For.*  
758 *Ecol. Manage.* 260, 1623–1639. <https://doi.org/10.1016/j.foreco.2010.07.054>

759 Niinemets, Ü., 1999. Research review. Components of leaf dry mass per area - thickness and density -  
760 alter leaf photosynthetic capacity in reverse directions in woody plants. *New Phytol.* 144, 35–47.  
761 <https://doi.org/10.1046/j.1469-8137.1999.00466.x>

762 Oliveira, R.S., Abrahão, A., Pereira, C., Teodoro, G.S., Mauro Brum, S., Alcantara, U., Lambers, H.,  
763 2016. Ecophysiology of Campos Rupestres Plants, in: Fernandes, G.W. (Ed.), *Ecology and*  
764 *Conservation of Mountain Top Grasslands in Brazil*. Springer International Publishing, pp. 228–262.  
765 <https://doi.org/10.1007/978-3-319-29808-5>

766 Oliveira, R.S., Dawson, T.E., Burgess, S.S.O., 2005. Evidence for direct water absorption by the shoot of  
767 the desiccation-tolerant plant *Vellozia flavicans* in the savannas of central Brazil. *J. Trop. Ecol.* 21,  
768 585–588. <https://doi.org/10.1017/S0266467405002658>

769 Pérez-Harguindeguy, N., Díaz, S., Garnier, E., Lavorel, S., Poorter, H., Jaureguiberry, P., Bret-Harte,  
770 M.S., Cornwell, W.K., Craine, J.M., Gurvich, D.E., Urcelay, C., Veneklaas, E.J., Reich, P.B.,  
771 Poorter, L., Wright, I.J., Ray, P., Enrico, L., Pausas, J.G., de Vos, A.C., Buchmann, N., Funes, G.,  
772 Quétier, F., Hodgson, J.G., Thompson, K., Morgan, H.D., ter Steege, H., Sack, L., Blonder, B.,  
773 Poschlod, P., Vaieretti, M. V., Conti, G., Staver, A.C., Aquino, S., Cornelissen, J.H.C., 2013. New  
774 handbook for standardised measurement of plant functional traits worldwide. *Aust. J. Bot.* 61, 167.  
775 <https://doi.org/10.1071/BT12225>

776 Poorter, H., Niinemets, U., Poorter, L., Wright, I.J., Villar, R., 2009. Causes and consequences of  
777 variation in leaf mass per area (LMA): a meta-analysis. *New Phytol.* 182, 565–588.  
778 <https://doi.org/10.1111/j.1469-8137.2009.02830.x>

779 Roelofsen, H.D., van Bodegom, P.M., Kooistra, L., Witte, J.M., 2014. Predicting leaf traits of herbaceous  
780 species from their spectral characteristics. *Ecol. Evol.* 4, 706–719. <https://doi.org/10.1002/ece3.932>

781 Rossatto, D.R., Franco, A.C., 2017. Expanding our understanding of leaf functional syndromes in  
782 savanna systems: the role of plant growth form. *Oecologia* 183, 953–962.  
783 <https://doi.org/10.1007/s00442-017-3815-6>

784 Rossatto, D.R., Kolb, R.M., Franco, A.C., 2015. Leaf anatomy is associated with the type of growth form  
785 in Neotropical savanna plants. *Botany* 93, 507–518. <https://doi.org/10.1139/cjb-2015-0001>

786 Sánchez-Azofeifa, G.A., Castro, K., Wright, S.J., Gamon, J., Kalacska, M., Rivard, B., Schnitzer, S.A.,  
787 Feng, J.L., 2009. Differences in leaf traits, leaf internal structure, and spectral reflectance between  
788 two communities of lianas and trees: Implications for remote sensing in tropical environments.  
789 *Remote Sens. Environ.* 113, 2076–2088. <https://doi.org/10.1016/j.rse.2009.05.013>

790 Santiago L.S., Wright, S.J., 2007. Leaf functional traits of tropical forest plants in relation to growth form.  
791 *Funct. Ecol.* 21, 19–27. <https://doi.org/10.1111/j.1365-2435.2006.01218.x>

792 Schaefer, C.E.G.R., Corrêa, G.R., Candido, H.G., Arruda, D.M., Nunes, J.A., Araujo, R.W., Rodrigues,  
793 P.M.S., Filho, E.I.F., Pereira, A.F.S., Brandão, P.C., NeriCarlos, A. V., 2016. The physical  
794 environment of Rupestrian Grasslands (Campos Rupestres) in Brazil: geological, geomorphological  
795 and pedological characteristics, and interplays, in: Fernandes, G.W. (Ed.), *Ecology and*  
796 *Conservation of Mountain Top Grasslands in Brazil*. Springer International Publishing, pp. 15–53.  
797 [https://doi.org/10.1007/978-3-319-29808-5\\_2](https://doi.org/10.1007/978-3-319-29808-5_2)

798 Schimel, D., Pavlick, R., Fisher, J.B., Asner, G.P., Saatchi, S., Townsend, P., Miller, C., Frankenberg, C.,  
799 Hibbard, K., Cox, P., 2015. Observing terrestrial ecosystems and the carbon cycle from space. *Glob.*  
800 *Chang. Biol.* 21, 1762–1776. <https://doi.org/10.1111/gcb.12822>

801 Schindelin, J., Rueden, C.T., Hiner, M.C., Eliceiri, K.W., 2015. The ImageJ ecosystem: an open platform  
802 for biomedical image analysis. *Mol. Reprod. Dev.* 82, 518–529.

803 Schweiger, A.K., Cavender-Bares, J., Townsend, P.A., Hobbie, S.E., Madritch, M.D., Wang, R., Tilman,  
804 D., Gamon, J.A., 2018. Plant spectral diversity integrates functional and phylogenetic components  
805 of biodiversity and predicts ecosystem function. *Nat. Ecol. Evol.* 2, 976–982.  
806 <https://doi.org/10.1038/s41559-018-0551-1>

807 Serbin, S.P., Singh, A., McNeil, B.E., Kingdon, C.C., Townsend, P.A., 2014. Spectroscopic  
808 determination of leaf morphological and biochemical traits for northern temperate and boreal tree  
809 species. *Ecol. Appl.* 24, 1651–1669. <https://doi.org/10.1890/13-2110.1>

810 Serbin, S. P., Wu, J., Ely, K. S., Kruger, E. L., Townsend, P. A., Meng, R., Wolfe, Brett T., Chlus, A.,  
811 Wang, Z., & Rogers, A. (2019). From the Arctic to the tropics: multibiome prediction of leaf mass  
812 per area using leaf reflectance. *New Phytol.* 224(4), 1557-1568.

813 Shipley, B., Lechowicz, M.J., Wright, I., Reich, P.B., 2006. Fundamental trade-offs generating the  
814 worldwide leaf economics spectrum. *Ecology* 87, 535–541. <https://doi.org/10.1890/05-1051>

815 Silveira, F.A.O., Negreiros, D., Barbosa, N.P.U., Buisson, E., Carmo, F.F., Carstensen, D.W., Conceição,  
816 A. a., Cornelissen, T.G., Echternacht, L., Fernandes, G.W., Garcia, Q.S., Guerra, T.J., Jacobi, C.M.,  
817 Lemos-Filho, J.P., Le Stradic, S., Morellato, L.P.C., Neves, F.S., Oliveira, R.S., Schaefer, C.E.,  
818 Viana, P.L., Lambers, H., 2016. Ecology and evolution of plant diversity in the endangered campo  
819 rupestre: a neglected conservation priority. *Plant Soil* 403, 129–152. [https://doi.org/10.1007/s11104-](https://doi.org/10.1007/s11104-015-2637-8)  
820 [015-2637-8](https://doi.org/10.1007/s11104-015-2637-8)

821 Simon, M.F., Grether, R., de Queiroz, L.P., Skema, C., Pennington, R.T., Hughes, C.E., 2009. Recent  
822 assembly of the Cerrado, a neotropical plant diversity hotspot, by in situ evolution of adaptations to  
823 fire. *Proc. Natl. Acad. Sci.* 106, 20359–20364. <https://doi.org/10.1073/pnas.0903410106>

824 Sims, D.A., Gamon, J.A., 2002. Relationships between leaf pigment content and spectral reflectance  
825 across a wide range of species, leaf structures and developmental stages. *Remote Sens. Environ.* 81,  
826 337–354. [https://doi.org/10.1016/S0034-4257\(02\)00010-X](https://doi.org/10.1016/S0034-4257(02)00010-X)

827 Streher, A.S., Sobreiro, J.F.F., Morellato, L.P.C., Silva, T.S.F., 2017. Land Surface Phenology in the  
828 Tropics: The Role of Climate and Topography in a Snow-Free Mountain. *Ecosystems* 20, 1436–  
829 1453. <https://doi.org/10.1007/s10021-017-0123-2>

830 Ustin, S.L., Gamon, J.A., 2010. Remote sensing of plant functional types. *New Phytol.* 186, 795–816.  
831 <https://doi.org/10.1111/j.1469-8137.2010.03284.x>

832 Van Cleemput, E., Vanierschot, L., Fernández-Castilla, B., Honnay, O., Somers, B., 2018. The functional  
833 characterization of grass- and shrubland ecosystems using hyperspectral remote sensing: trends,  
834 accuracy and moderating variables. *Remote Sens. Environ.* 209, 747–763.  
835 <https://doi.org/10.1016/j.rse.2018.02.030>

836 Violle, C., Navas, M. L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I., & Garnier, E., 2007. Let the  
837 concept of trait be functional!. *Oikos*, 116 (5), 882-892. [https://doi.org/10.1111/j.0030-](https://doi.org/10.1111/j.0030-1299.2007.15559.x)  
838 [1299.2007.15559.x](https://doi.org/10.1111/j.0030-1299.2007.15559.x)

839 Wang, R., Gamon, J.A., Cavender-Bares, J., Townsend, P.A., Zyguelbaum, A.I., 2018. The spatial  
840 sensitivity of the spectral diversity-biodiversity relationship: An experimental test in a prairie  
841 grassland. *Ecol. Appl.* 28, 541–556. <https://doi.org/10.1002/eap.1669>

842 Wang, Z., Townsend, P.A., Schweiger, A.K., Couture, J.J., Singh, A., Hobbie, S.E., Cavender-Bares, J.,  
843 2019. Mapping foliar functional traits and their uncertainties across three years in a grassland  
844 experiment. *Remote Sens. Environ.* 221, 405–416. <https://doi.org/10.1016/j.rse.2018.11.016>  
845 Warming, E., 1908. Lagoa Santa: Contribuição para a geographia phytobiologica. *Arq. da Real Soc.*  
846 *Dinamarqueza das Sci. Naturaes e Math.* VI.  
847 Wold, S., 1994. PLS for multivariate linear modeling., in: Waterbeemd, H. van de (Ed.), *Chemometric*  
848 *Methods in Molecular Design.* Verlag-Chemie, Weinheim, Germany, pp. 195–218.  
849 Wold, S., Sjöström, M., Eriksson, L., 2001. PLS-regression: a basic tool of chemometrics. *Chemom.*  
850 *Intell. Lab. Syst.* 58, 109–130. [https://doi.org/10.1016/S0169-7439\(01\)00155-1](https://doi.org/10.1016/S0169-7439(01)00155-1)  
851 Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., Cavender-Bares, J.,  
852 Chapin, T., Cornelissen, J.H.C., Diemer, M., Flexas, J., Garnier, E., Groom, P.K., Gulias, J.,  
853 Hikosaka, K., Lamont, B.B., Lee, T., Lee, W., Lusk, C., Midgley, J.J., Navas, M.-L., Niinemets, Ü.,  
854 Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V.I., Roumet, C., Thomas, S.C.,  
855 Tjoelker, M.G., Veneklaas, E.J., Villar, R., 2004. The worldwide leaf economics spectrum. *Nature*  
856 428, 821–827. <https://doi.org/10.1038/nature02403>  
857 Wu, J., Chavana-Bryant, C., Prohaska, N., Serbin, S.P., Guan, K., Albert, L.P., Yang, X., van Leeuwen,  
858 W.J.D., Garnello, A.J., Martins, G., Malhi, Y., Gerard, F., Oliviera, R.C., Saleska, S.R., 2017.  
859 Convergence in relationships between leaf traits, spectra and age across diverse canopy  
860 environments and two contrasting tropical forests. *New Phytol.* 214, 1033–1048.  
861 <https://doi.org/10.1111/nph.14051>  
862 Zappi, D., Milliken, W., Nicholas Hind, D.J., Biggs, N., Rando, J., Malcolm-Tompkins, P., Mello-Silva,  
863 R., 2014. *Plantas do Setor Noroeste da Serra do Cipó, Minas Gerais - guia ilustrado.*  
864 <https://doi.org/10.13140/2.1.3216.9286>  
865  
866

867  
868  
869  
870

**Supplementary material for:**

**Accuracy and limitations for spectroscopic prediction of leaf traits in seasonally dry tropical environments**

871 Annia Susin Streher; Ricardo da Silva Torres, Leonor Patrícia Cerdeira Morellato; Thiago Sanna Freire Silva  
872  
873

874 **S1: ANOVA results comparing trait variation (LDMC and LMA) between growth forms**

875

876

877 *ldmc.aov <- aov(log10(LDMC) ~ growth\_form, data = new\_df)*

878 *lma.aov <- aov(LMA ~ growth\_form, data = new\_df)*

879

880

881 **Table S1.** Anova table comparing the means of LDMC and LMA values among growth forms.

882

LDMC					
	DF	SUM of Squares	Mean Square	F-value	PR(>F)
Growth form	2	0.626	0.31296	36.54	5.55 e <sup>-16</sup>
Residuals	897	7.683	0.00856		
LMA					
	DF	SUM of Squares	Mean Square	F-value	PR(>F)
Growth form	2	242218	121109	21.15	1.05e <sup>-09</sup> ***
Residuals	897	5135423	5725		

883

884

885 **Table S2.** Multiple comparison Tuckey test comparing growth forms.

LDMC					
	Estimate	Std. Error	t value	Pr(> t )	
forbs - graminoids	-0.078234	0.009436	-8.291	<1e-04 ***	
woody - graminoids	-0.019328	0.009436	-2.048	0.101	
woody - forbs	0.058906	0.009436	6.243	<1e-04 ***	
LMA					
	Estimate	Std. Error	t value	Pr(> t )	
forbs - graminoids	33.121	6.178	5.361	<1e-05 ***	
woody - graminoids	36.267	6.178	5.870	<1e-05 ***	
woody - forbs	3.146	6.178	0.509	0.867	

886

887

888

889

890

891

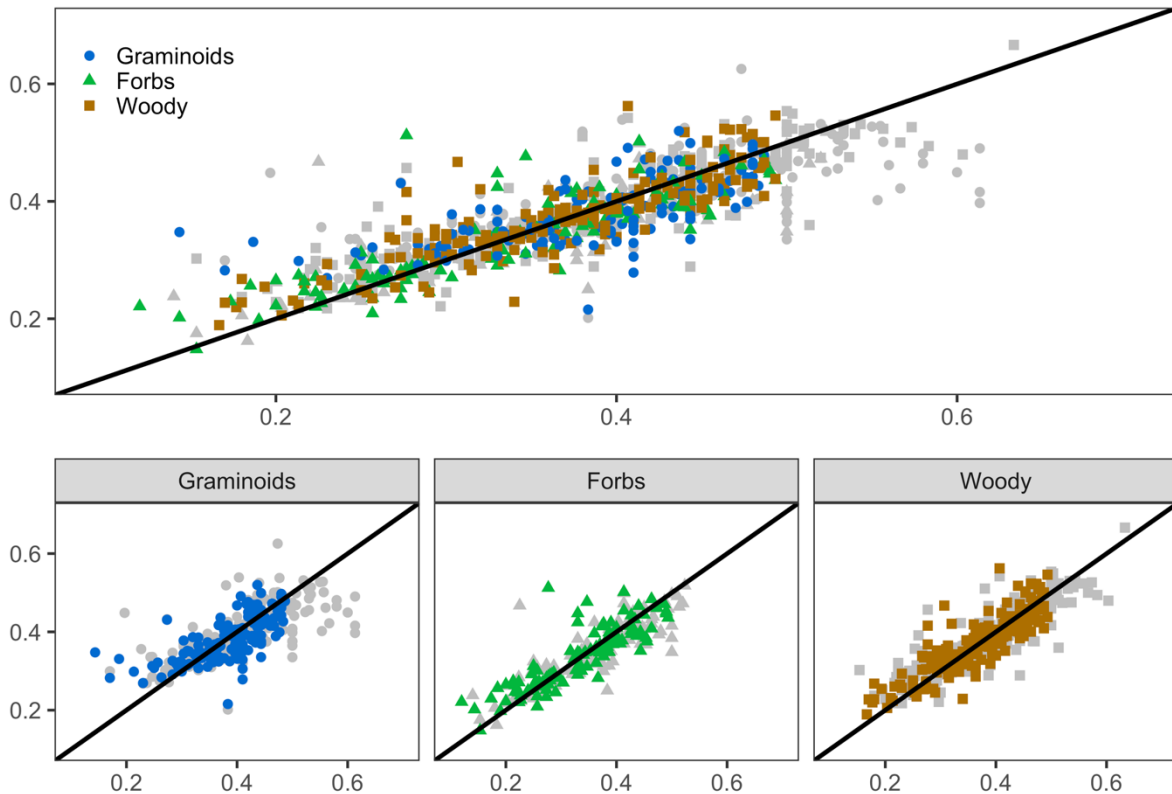
892



893  
894  
895  
896  
897  
898  
899  
900

## S2: LDMC spectroscopy saturation analyses

Looking to figure 2 of the main text is possible observe that approximately near 0.5 g/g the model does not capture properly the data variability. We perform the same approach as we did for LMA, and run PLSR restricting values up to 0.5 g/g, and then assessed the new model with the same metrics (Table 1, main text). Contrary to LMA, this new modelling did not show any improvement in comparison to the full LDMC model.



901  
902  
903  
904  
905  
906  
907  
908  
909  
910

**Figure S1:** Partial least-squares regression (PLSR) results for observed vs. predicted leaf dry matter content (LDMC) with values restricted to 0 – 0.05 g/g. The upper panel shows the prediction for the total range of LDMC values (“All” class). The lower panels show the relationship between observed and predicted LDMC values for each growth form class. Symbols and colors indicate the growth form of each individual plant: blue dots as “Graminoids”; green triangles as “Forbs”, and brown squares as “Woody”. Gray squares comprise original LDMC values above 0.05 g/g, which were not included in the restricted models. Black lines indicate the 1:1 relationship as reference.

911 **Table S3.** Results of the partial least-squares regression (PLSR) modeling and cross-validation for each  
 912 leaf trait, showing the number of samples and range of trait variation for the global data set (all) and per  
 913 growth form. RMSECV is the root mean square error (RMSE) of the cross-validation procedure with  
 914 train data set; RMSE is the measured error using the test data; mRMSE is the ratio of the error of each  
 915 model in relation to the mean values (RMSE/mean); and the RMSE percentage (%RMSE) shows the error  
 916 of each model as a percentage of the observed data range. Predicted R<sup>2</sup> shows the predictive quality of  
 917 each model. All results are presented for the entire range of LMA and LDMC values (“All” class) and per  
 918 growth form. “LMA < 300” represents the data set containing only LMA values bellow 300 g/m<sup>2</sup>.

919

Growth form	Number of samples	Range of variation (min - max)	RMSECV	Final number of latent variables	RMSE	mRMSE (RMSE/mean)	%RMSE (% of range)	R <sup>2</sup>
<b>LDMC</b>								
ALL	1648	0.12-0.67 (g/g)	0.052 (g/g)	20	0.053 (g/g)	0.13	9.75 %	0.68
Graminoids	564	0.12-0.67 (g/g)	0.063 (g/g)	17	0.059 (g/g)	0.15	11.66 %	0.48
Forbs	369	0.12-0.61 (g/g)	0.046 (g/g)	13	0.055 (g/g)	0.15	11.22 %	0.73
Woody	715	0.15-0.67 (g/g)	0.043 (g/g)	18	0.051 (g/g)	0.13	9.98%	0.78
<b>LDMC &lt; 0.05</b>								
ALL	1441	0.12-0.49 (g/g)	0.045 (g/g)	20	0.04 (g/g)	0.12	12.20 %	0.68
Graminoids	470	0.12-0.49 (g/g)	0.055(g/g)	12	0.04 (g/g)	0.12	12.95 %	0.45
Forbs	350	0.12-0.49 (g/g)	0.048 (g/g)	12	0.05 (g/g)	0.14	13.9 %	0.72
Woody	621	0.15-0.49(g/g)	0.048 (g/g)	7	0.03 (g/g)	0.10	11.11 %	0.72

920

921

922