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

Annia Susin Streher [1*]; Ricardo da Silva Torres[2], Leonor Patrícia Cerdeira Morellato [3]; Thiago Sanna Freire Silva [4].

[1] Universidade Estadual Paulista (Unesp), Instituto de Biociências, Rio Claro, São Paulo, Brazil.

[2] Department of ICT and Natural Sciences, NTNU - Norwegian University of Science and Technology, Ålesund, Norway.

[3] Universidade Estadual Paulista (Unesp), Instituto de Biociências, Departamento de Biodiversidade, Phenology Lab, Rio Claro, São Paulo, Brazil.

[4] 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**

**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$^2$), rarely observed in tropical or temperate forests, and demonstrated that values above 300 g/m$^2$ 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. Introduction

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

A wide set of leaf traits, including many of the LES traits, can be detected and accurately predicted using leaf spectral reflectance data (Asner et al., 2016; Cavender-Bares et al., 2017; Curran et al., 2001; Serbin et al., 2014). Still, despite its ecological relevance, the relationship between leaf-level spectral reflectance and important functional foliar traits such as LMA and LDMC remains under-explored, and is mainly focused on plants from forested ecosystems (Van Cleemput et al., 2018). There is also an apparent inconsistency with the trait names used by the remote sensing community and by ecologists (Homolová et al., 2013). In the ecology literature, LMA is the ratio of leaf dry weight (mass) per leaf area (g $m^{-2}$), while LDMC is an investment index, determined by the ratio between leaf dry and fresh weights (g/g) (Pérez-Harguindeguy et al., 2013). However, several remote sensing studies use the terms ''leaf dry matter content'' or ''dry matter content'' when actually referring to LMA (Homolová et al., 2013), and also refer to the ratio between leaf fresh and dry weights (LDMC) as quantification of "leaf water content" (Ball et al., 2015; Cheng et al., 2011). Although LDMC is mathematically related to leaf water content (LWC = 1- LDMC, Pérez-Harguindeguy et al., 2013), ecologists tend to consider LMA, LDMC, and LWC as separate traits.

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

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

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

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

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

## 2. Materials and Methods

*2.1 Study area and sampling design*

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

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

We sampled leaf traits and leaf reflectance spectra during the October 2016 – March 2017 growing season (Streher *et al.* 2017). Our study design included five sampling sites distributed along the elevation gradient, from 820 m to 1500 m, based on the natural environmental stratification of elevation and edaphic conditions (Mattos et al. 2019). Within each elevation, four transects of 250 m, distant at least 50 m from each other, were established based on expert knowledge and interpretation of high-resolution aerial images, ensuring the inclusion of all vegetation types (a proxy for edaphic conditions and resulting functional assemblages) found within each site (see Mattos et al. 2019, for detailed description of vegetations and soil). Our samples thus encompassed all types of *cerrado* and *campo rupestre* vegetation*,* and are hereafter referred to as *campo rupestre*, as this was the dominant vegetation sampled. Sampling points were established at 7 m intervals along each transect, with a 3.5 m search radius delimited around each point. Within each search radius, we identified and sampled three individual plants, applying the following selection criteria: 1) we identified the three individuals closest to the center of the search radius belonging to morphotypes not sampled before in the same transect; 2) if less than three individuals from new morphotypes were found, we sampled the closest individuals to the center of the search radius, regardless of species, to reach three samples per sampling point. This sampling strategy was designed to ensure maximal sampling of morphotypic variation and maximizing trait variability, while still reflecting the relative abundances of different mophotypes. For each individual plant, three fully-expanded sun leaves were sampled. In total, we sampled 4944 leaves from 1648 individual plants, encompassing all observed growth form and representing the majority of plant phenotypes found at Serra do Cipó.

### *2.2 Plant growth form definitions*

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

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

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

*2.3 Leaf trait measurements*

For trees and shrubs, we harvested branches of individual canopies containing sunlit and mature leaves, while for grasses we sampled the whole plant, keeping roots when possible (Pérez-Harguindeguy et al., 2013). We followed partial rehydration protocols by immediately storing the samples in moistened sealed plastic bags, under elevated $CO_2$ concentrations and saturated air humidity, stored in lightproof containers filled with ice (Garnier et al., 2001; Pérez-Harguindeguy et al., 2013). We kept the samples at ~ 4 ºC in the dark, and measurements were taken between six to eight hours after harvesting. From each branch/individual sampled, we removed three healthy leaves with no serious herbivore or pathogen damage, including petioles, blotted them dry to remove surface water, immediately weighed them to determine saturated fresh mass (Garnier *et al.*, 2001) and then measured reflectance spectra. All spectral measurements were taken within the same day (Foley et al., 2006), between six to eight hours after branch harvesting (see next section). We then determined one-sided leaf area (Pérez-Harguindeguy *et al.,* 2013) by photographing each leaf under a straight overhead (nadir) view, while gently pressing individual leaves between a glass plate and a sheet of paper including a printed distance scale, ensuring photo scale calibration and thus accurate area measurements. We then calculated leaf area using the ImageJ2 software (Schindelin et al., 2015). After photographing, we oven-dried leaf samples at 80 °C for 72 hours to determine leaf dry mass to the nearest 0.01 g. We computed LMA ($g/m^2$) as the ratio between dry mass and leaf area, and LDMC (g/g), as the ratio between leaf fresh mass and dry mass (Pérez-Harguindeguy *et al.,* 2013).

*2.4 Leaf spectral measurements*

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

Bi-directional reflectance measurements were taken for the same three replicate leaves from which LMA and LDMC were estimated, immediately after obtaining saturated fresh mass. Leaves were arranged over a large black non-reflective surface, covering the whole diameter of the contact probe (10 mm) and ensuring that no light escaped the measurement. Plants with small leaves or leaflets were arranged so that the FOV was fully covered, without any gaps or excessive overlap, using more than a single leaf or leaflet when necessary. For each leaf, ten measurements were taken at one to six different parts of the leaf adaxial surface (depending on leaf size), avoiding main veins, herbivory and pathogens damage when possible, following the protocols and standards by Asner & Martin (2009). For compound leaves, we took up to 10 measurements of different leaflets. The final leaf spectrum of each leaf was then given as the average of the 10 scans.

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

*2.5 Leaf trait predictive modeling*

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

We fit four models to predict each of the two leaf traits: a model based on all observations ("*All*"), and three models restricted by plant growth form ("*Woody*", "*Forbs*", and "*Graminoids*"), for a total of eight PLSR models. Based on the initial results, we also fitted four additional models for a subset of the original LMA dataset, comprising only values between 0 and 300 $g/m^2$. For each model, we split our data into training (70%, hereafter train set) and validation (30%, hereafter test set), using the "createDataPartition()" function from the "caret" package (Kuhn, 2008) in R, to ensure that both sets spanned the entire range of measured values for each trait. To reduce overfitting, we optimized the number of PLSR latent variables in the final models by minimizing the root mean square error (RMSE) of the prediction residual sum of squares (PRESS statistic, Chen *et al.*, 2004). For the larger datasets ("*All*", "*Woody*", and "*Graminoids*"), we calculated the PRESS statistic of successive model components using a 10-fold cross-validation scheme, while for the "*Forbs*" dataset we used a standard leave-one-out cross validation (LOOCV) analysis as recommended for datasets with fewer observations (Serbin et al., 2014). We assessed the final accuracy of each model by calculating the RMSE value between predicted and observed trait values in the test set, expressing it in the original variable units (RMSE), as percentage of the sample data range (%RMSE), and as the ratio of each model RMSE to the mean value of the trait dataset (mRMSE). Thus, we computed the coefficient of determination ($R^2$) of the observed versus predicted values of each model, to understand the percentage of variance explained by the model in the test dataset. We also report RMSECV, the RMSE obtained from the cross-validation procedure using the 10-fold or LOOCV methods, as discrepancies between RMSECV and RMSE can indicate model overfitting (Kuhn and Johnson, 2013).

Lastly, we computed the variable importance of projections (VIP, Wold (1994)) metric for each model, to identify the spectral regions that contributed the most to the prediction of each leaf trait. VIP is the weighted sum of squares of the PLSR-weights, with the weights calculated from the amount of variance from the response variable explained by each PLS component (Wold 1994).

*2.6 Spectral dissimilarities among plant growth forms*

To understand the contribution of different spectral regions to the identification of plant functional strategies, we evaluated spectral dissimilarity between plant growth forms using the Bhattacharyya distance (Bhattacharyya, 1943; Kailath, 1967) (Eqn. 1). This metric quantifies the integrated difference between two individuals of different growth forms over the full spectral range, identifying the wavelengths with the greatest discriminatory power. This metric has been successfully applied for the recognition of differences between species (Baldeck and Asner, 2014), and plants with different growing habits (Sánchez-Azofeifa et al., 2009).

$$B = \frac{1}{8}(\mu_i - \mu_j)^T \Sigma^{-1}(\mu_i - \mu_j) + \frac{1}{2}\, ln(\frac{|\Sigma|}{\sqrt{|\Sigma_i| - |\Sigma_j|}}) \qquad \textbf{Eqn 1}$$

where $\mu_i$ and $\mu_j$ are the mean values across all spectral bands for species i and j, Σi and Σj are the covariance matrices for each individual, and Σ is the pooled covariance matrix. $B$ is the Bhattacharyya distance.

We used a randomized approach to estimate the distribution of $B$ by randomly sampling 1000 pairs of spectra for each combination of growth forms ("Woody" x "Grass"; "Woody" x "Forb" and "Forb" x "Grass"), and then computing the average and spread (standard deviation) of the 1000 calculated pairwise distances for each combination.

## 3. Results

### *3.1 Leaf trait variability*

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

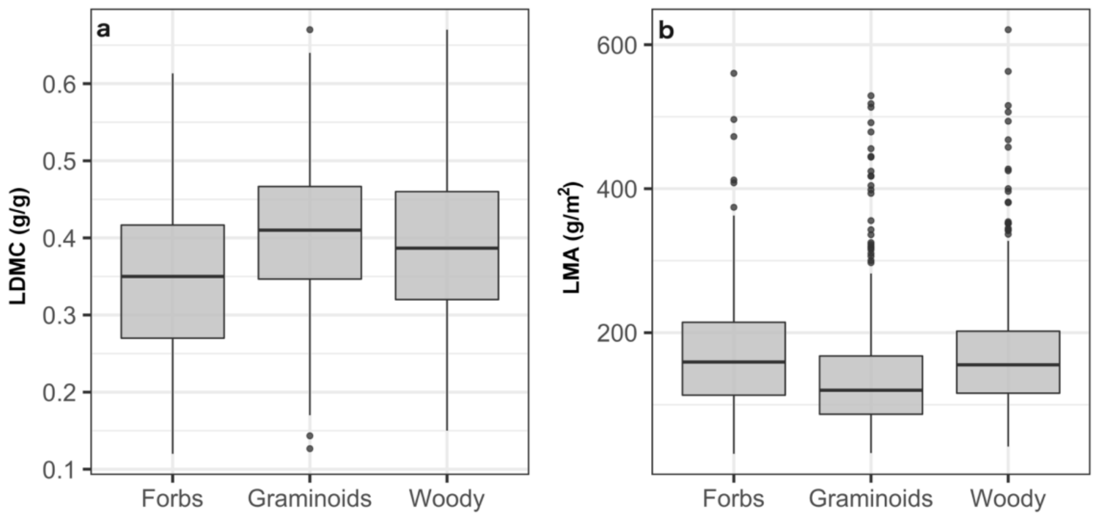


Figure 1. Variability of leaf functional traits measured for 1648 individuals of *campo rupestre* vegetation at Serra do Cipó, Southeastern Espinhaço range, Brazil, including 369 individuals of the "*Forbs"* class, 564 individuals of the "*Graminoids"* class, and 715 individuals of the "*Woody"* class.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

*3.2 PLSR modeling*

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

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

| *Growth form* | *Number of samples* | *Range of variation (min - max)* | *RMSECV* | *Final number of latent variables* | *RMSE* | *mRMSE (RMSE/ mean)* | *%RMSE (% of range)* | *$R^2$* |
|---|---|---|---|---|---|---|---|---|
| ***LDMC*** | | | | | | | | |
| *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 |
| **LMA** | | | | | | | | |
| *ALL* | 1648 | 31.80 - 620.81 (g/m$^2$) | 44.56 (g/m$^2$) | 17 | 50.58 (g/m2) | 0.32 | 8.58 % | 0.58 |
| *Graminoids* | 564 | 32.77 - 529.12 (g/m$^2$) | 44.89 (g/m$^2$) | 16 | 43.22 (g/m2) | 0.31 | 8.70 % | 0.60 |
| *Forbs* | 369 | 31.80 - 560.29 (g/m$^2$) | 53.12 (g/m$^2$) | 14 | 44.08 (g/m2) | 0.26 | 8.34 % | 0.42 |
| *Woody* | 715 | 41.89 - 620.81 (g/m$^2$) | 39.57 (g/m$^2$) | 18 | 43.33 (g/m2) | 0.26 | 7.48 % | 0.65 |
| **LMA < 300** | | | | | | | | |
| *ALL* | 1571 | 31.80 – 298.94 (g/m$^2$) | 32.00 (g/m$^2$) | 18 | 30.70 (g/m$^2$) | 0.21 | 11.49 % | 0.71 |
| *Graminoids* | 539 | 32.77 - 297.23 (g/m$^2$) | 33.56 (g/m$^2$) | 20 | 35.73 (g/m$^2$) | 0.28 | 14.45 % | 0.58 |
| *Forbs* | 337 | 31.80 - 298.94 (g/m$^2$) | 32.95 (g/m$^2$) | 19 | 35.32 (g/m$^2$) | 0.22 | 13.61 % | 0.71 |
| *Woody* | 695 | 41.89 - 298.52 (g/m$^2$) | 28.65 (g/m$^2$) | 20 | 26.23 (g/m$^2$) | 0.16 | 10.79 % | 0.78 |

Our PLSR LDMC spectral model had an overall error (RMSE) of 0.053 g/g, c.a. 9 % of the range of LDMC values of the entire dataset (Table 1 and Fig. 2). Among growth-form restricted models, accuracy was higher for *Woody* plants, with %RMSE of c.a. 10% (RMSE = 0.051 g/g). The “*Graminoids*” and “*Forbs”* models yielded similar error rates; although “*Graminoids*” models had higher overall error (RMSE = 0.059 g/g) than “*Forbs”* (RMSE = 0.055 g/g), these errors represented similar ratios of error in relation to the mean class value mRMSE = 0.15) and %RMSE considering the full range of values (“*Graminoids*” %RMSE = 11.66%; “*Forbs”* %RMSE = 11.22%).

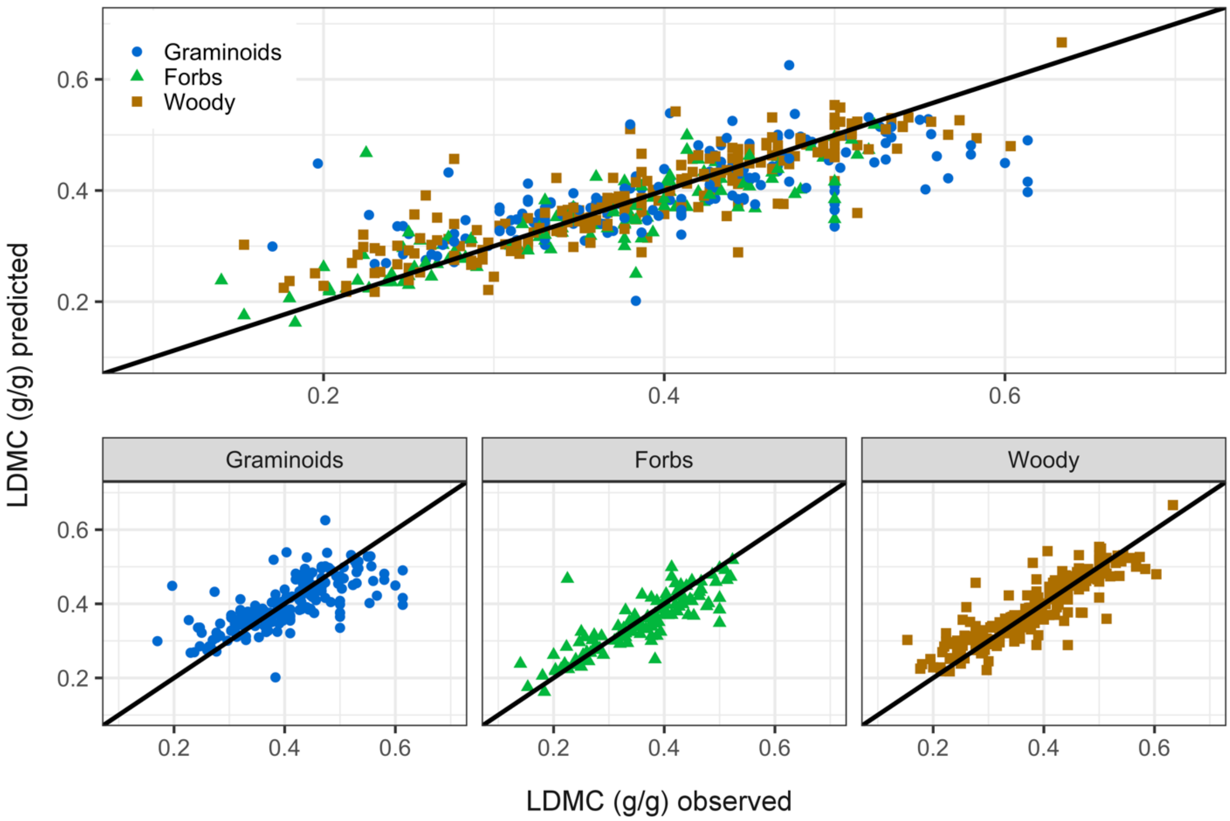


Figure 2. Leaf dry matter content (LDMC) as observed and predicted from leaf level reflectance using partial least-squares regression (PLSR) models. The upper panel shows the prediction for the total range of LDMC values (“*All*” class). The lower panels show the relationship between observed and predicted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

The PLSR model for LMA had the highest overall accuracy with a RMSE of 50.58 g/m$^2$, representing an error percentage around 8 % of the range of LMA values of the entire dataset (Table 1 and Fig. 3). The restricted models for LMA showed lower discrepancies between growth forms classes, with similar RMSE between groups. The restricted model with highest accuracy corresponded to the "*Woody"* data set, with a RMSE of 43.33 g/m$^2$ and error percentage of c.a. 7 % of the range of values within the class. While the model accuracy for the "*Graminoids*" class was similar to the "*Woody"* class (RMSE = 43.22 g/m$^2$), the error percentage of the range of values was higher (8.7%). The lowest accuracy was yielded by the "*Forbs"* restricted model, with RMSE of 44.26 g/m$^2$, ca. 8.4 % of the "*Forbs"* LMA value range.

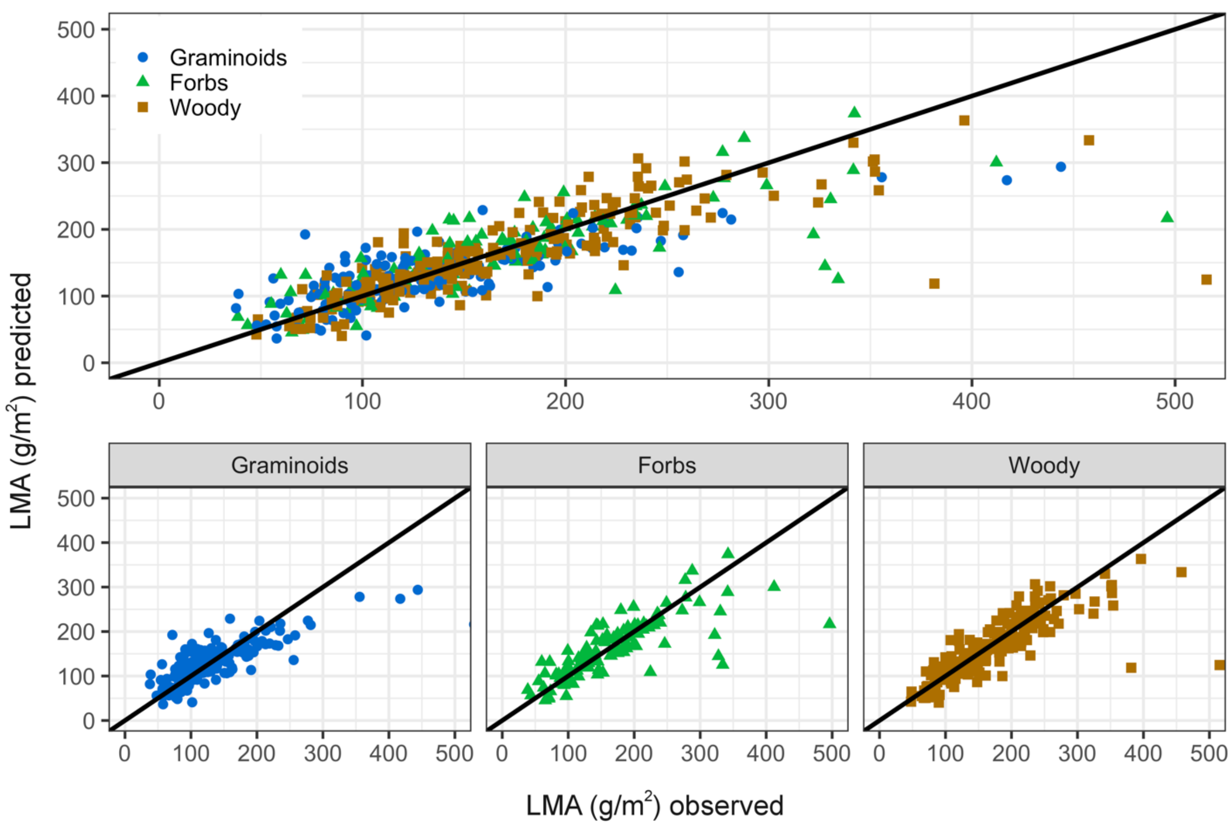


Figure 3: Partial least-squares regression (PLSR) results for observed vs. predicted leaf mass per area (LMA). The upper panel shows the prediction for the total range of LMA values ("*All*" class). The lower panels show the relationship between observed and predicted LMA values for each growth form.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

We observed a loss of predictive power for all PLSR models for high LMA values, *i.e.* above 300 g/m² (Fig. 3), while PLSR models performed only slightly worst for LDMC high values (Fig 2). To quantify the influence of this loss, we refitted the PLSR models using only LMA values between 0 and 300 g/m² (Table 1), matching the range of LMA values usually observed for tropical (Asner et al., 2011a, 2011b) and temperate (Serbin *et al.*, 2014) forested systems, which are also typically used in radiative transfer models (Féret and Asner, 2011) and most frequently reported in the literature of leaf trait spectroscopy. These restricted-range PLSR models could explained more of LMA variance ($R^2$ = 0.78) (Fig. 4), yielding an overall decrease in mRMSE of 0.21 in LMA values (Table 1 – LMA < 300 g/m²). The decrease in the overall error was also uniformly observed for models of each growth form, as so as an increase in the percentage of variance explained ($R^2$) (Table 1). The highest improvement was found for the "*Forbs"* class, with a restricted range mRMSE of 0.22, down from mRMSE= 0.31 from the full range model (Table 1 and Fig. 5). The lowest performance of the restricted model was found for "*Graminoids"* (mRMSE= 0.28), with 1-fold change improvement. Using the same approach with LDMC values above 0.05 g/g (Fig. 2), where the points start to deviate from the 1:1 line, and we found that removing these points from the analysis did not improve model accuracy and did not increase the percentage of variance explained (Fig S1 and Table S3).

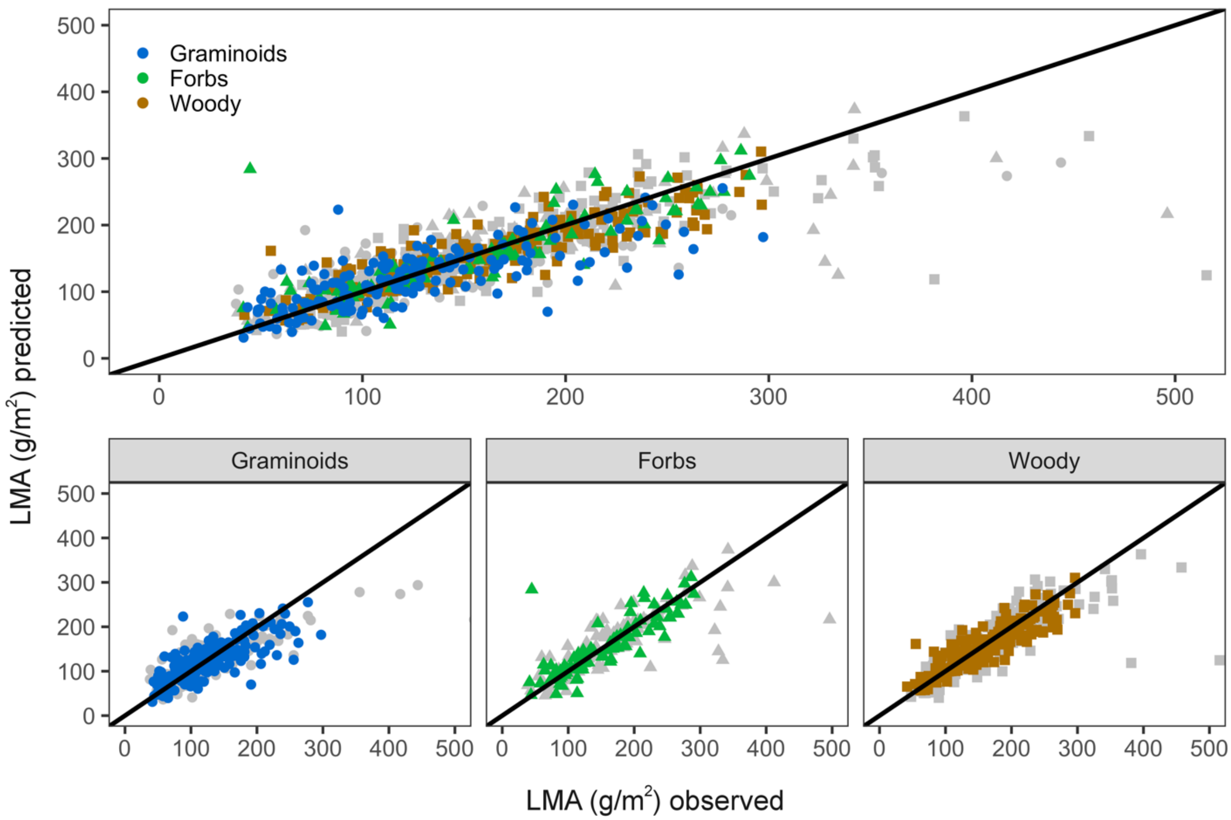


Figure 4: Partial least-squares regression (PLSR) results for observed vs. predicted leaf mass per area (LMA), with values restricted to 0 - 300 g/m². The upper panel shows the prediction for the total range of LMA values ("*All*" class). The lower panels show the relationship between observed and predicted LMA 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 LMA values above 300 g/m², which were not included in the restricted models. Black lines indicate the 1:1 relationship as reference.

Overall, VIP values had consistent patterns across the spectrum, with a few notable variations from specific wavelengths (Fig. 5). For LDMC, the wavelength region centered in 1400 nm yielded the highest VIP value, but wavelengths in the visible (VIS) (550 to 650 nm), red-edge (700-750 nm), and in the shortwave infrared (SWIR) (around 1700 and 1900 nm) were also important (Fig. 4a). The most important spectral region for LMA was the red-edge (700-750 nm), followed by the VIS region at the wavelength centered in 550 nm (Fig. 5b). The VIP metric also varied in the position of peak importance among growth forms for both traits, but specially for LMA, where a SWIR spectral region from 1900 to 2100 nm stood out for the "*Graminoids*" form (Fig. 5b). The red-edge (700-750 nm) was the spectral region with the closest agreement of VIP values among growth forms for both leaf traits.

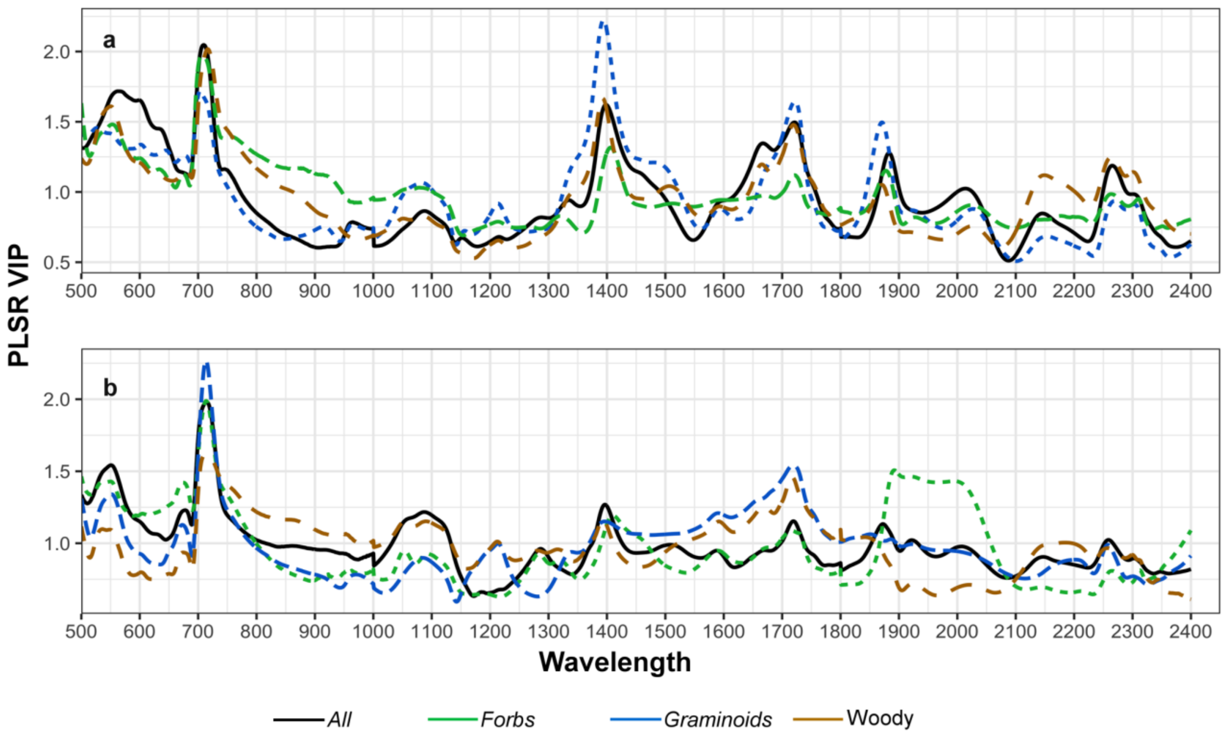


Figure 5: Partial Least Squares Regression (PLSR) variable importance of prediction (VIP) plotted by wavelength for **(a)** leaf dry matter content (LDMC), and **(b)** leaf mass per area (LMA), measured for *campo rupestre* plants at Serra do Cipó, Southern Espinhaço Range, Brazil. Colored lines represent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

*3.3 Leaf reflectance spectra dissimilarity among growth forms*

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

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

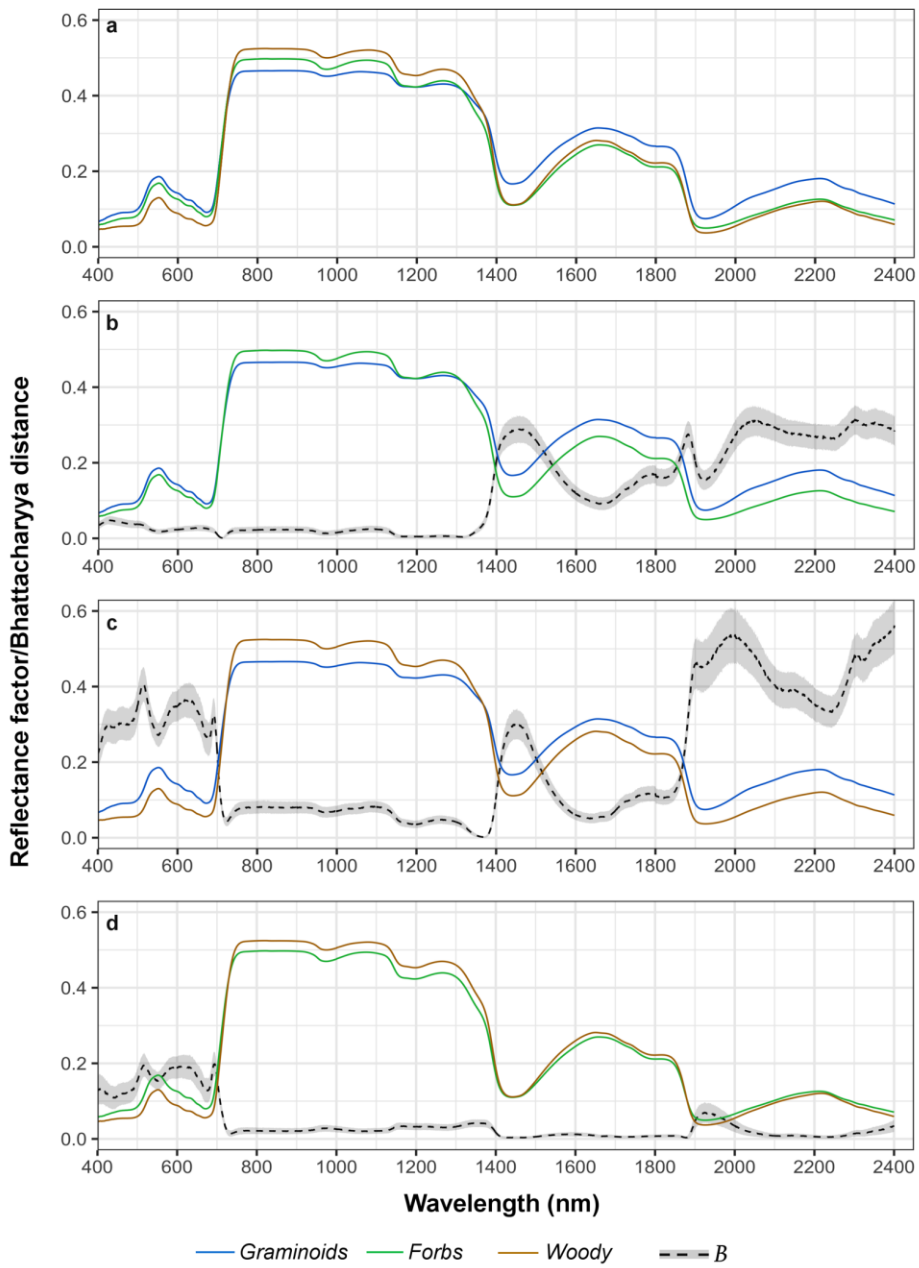


Figure 6. Comparison of leaf reflectance spectral averages per growth form **(a)**, and the spectral dissimilarity (Bhattacharyya distance) between growth forms across the full wavelength range (400 – 2400 nm): **(b)** “*Forbs”* and “*Graminoids*”, **(c)** “*Woody”* and “*Graminoids”* and **(d)** “*Woody”* and “*Forbs”*. The peaks observed on the Bhattacharyya index (*B*, dashed line and the gray shaded area represents ± 1 standard deviation) indicate the spectral bands with highest dissimilarities among growth forms.

## 4. Discussion

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

Considering the small variation in leaf traits, our results reinforce the potential of PLSR and spectroscopy to quantitatively describe structural foliar properties. Our general models were able to successfully explain variations related to leaf strategy without bias towards any growth form, going one step further towards the development of generalized global models. Still, the restricted PLSR models had overall better performances for woody plants than other growth forms for both measured traits. Our error rates for woody species (%RMSE 7% - 10%) are comparable to rates observed for tropical (Asner *et al.*, 2011b; %RMSE = 5.9%) and temperate forests (Serbin *et al.*, 2014; %RMSE= 10.1%). To the best of our knowledge, there is a small number of studies addressing PLSR-spectroscopy modelling of LMA and LDMC from "herbaceous" plants, with emphasis on grasses. Our modelling resulted in an equal predictive performance for LMA on grasses in relation to previous studies (Wang et al., 2019; %RMSE 12%), and slightly lower for LDMC (Roelofsen et al., 2014; RMSE = 0.10).

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

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

Future aerial and orbital remote sensors and missions may provide a better and urgently needed synoptic view of terrestrial ecosystem dynamics, as long as they allow for a high enough frequency of observations to capture specific phenological stages, thus yielding information on temporal leaf trait variation, a key information still mostly unexplored in trait-based ecology. Considering the spectral wavelengths identified in our analyses, multispectral sensors with multiple, high signal-to-noise spectral bands in the red-edge (700-750 nm) and SWIR (around 1700 and 1900 nm) regions would bring us to the next level in scaling-up functional diversity patterns to larger regions.

*4.1 Insights from full reflectance spectra on plant functional characterization*

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

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

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

Eudicot herbs and sub-shrubs represented an intermediate growth form between woody plants and grasses. On one hand, they were differentiated from grasses by the amount of leaf water and structural properties absorbing along the SWIR, and lower photosynthetic rates than grasses, in contrast to the subtle differences found in the VIS from woody plants. The lack of proper spectral discrimination can be due to our inclusion of herbs and sub-shrubs within the same growth form due to sample size limitations. Sub-shrubs are unique since they have leaf anatomys similar to herbs (Rossatto et al., 2015), but are functionally clustered with trees and shrubs (Rossatto and Franco, 2017). This implies that although they are on an evolutionary trajectory of ecological convergence with herbaceous plants, they are not phylogenetically independent of the tree lineages from which they have evolved (Rossatto and Franco, 2017; Simon et al., 2009).

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

## 5. Concluding remarks

We accurately predicted LMA and LDMC for seasonally dry tropical plants from spectroscopy, even though these traits had little variation among growth forms, reinforcing the ability of leaf spectroscopy to predict functional leaf traits. However, we also found an important limitation in using PLSR methods to predict high LMA values (> 300 $g/m^2$), resulting in underestimated values for LMA ranges that have been seldom addressed in the literature before. There are currently large biases in the sampling of plant traits and related spectra, favoring humid forested systems, hindering our understanding of spectroscopic relationships and limiting our ability to make reliable inferences and apply them to global biodiversity science. Further work in determining whether limitations in LMA prediction are a methodological shortcoming from PLSR and/or a biophysical limitation of spectroscopy in high LMA environments is thus imperative.

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

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

***Acknowledgments***

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

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

## References


Abrahão, A., de Britto Costa, P., Lambers, H., Andrade, S.A.L., Sawaya, A.C.H.F., Ryan, M.H., Oliveira, R.S., 2018. Soil types filter for plants with matching nutrient-acquisition and -use traits in hyperdiverse and severely nutrient-impoverished *campos rupestres* and *cerrado* in Central Brazil, Journal of Ecology. https://doi.org/10.1111/1365-2745.13111

Alcantara, S., de Mello-Silva, R., Teodoro, G.S., Drequeceler, K., Ackerly, D.D., Oliveira, R.S., 2015. Carbon assimilation and habitat segregation in resurrection plants: a comparison between desiccation- and non-desiccation-tolerant species of Neotropical Velloziaceae (Pandanales). Funct. Ecol. 29, 1499–1512. https://doi.org/10.1111/1365-2435.12462

Ali, A.M., Darvishzadeh, R., Skidmore, A.K., Duren, I. van, Heiden, U., Heurich, M., 2016. Estimating leaf functional traits by inversion of PROSPECT: Assessing leaf dry matter content and specific leaf area in mixed mountainous forest. Int. J. Appl. Earth Obs. Geoinf. 45, 66–76. https://doi.org/10.1016/j.jag.2015.11.004

Alves, R., Silva, N., Oliveira, J., Medeiros, D., 2014. Circumscribing campo rupestre – megadiverse Brazilian rocky montane savanas. Brazilian J. Biol. 74, 355–362. https://doi.org/10.1590/1519-6984.23212

Asner, G.P., Knapp, D.E., Anderson, C.B., Martin, R.E., Vaughn, N., 2016. Large-scale climatic and geophysical controls on the leaf economics spectrum. Proc. Natl. Acad. Sci. 113, E4043–E4051. https://doi.org/10.1073/pnas.1604863113

Asner, G.P., Martin, R.E., 2009. Airborne spectranomics: mapping canopy chemical and taxonomic diversity in tropical forests. Front. Ecol. Environ. 7, 269–276. https://doi.org/10.1890/070152

Asner, G.P., Martin, R.E., Carranza-Jiménez, L., Sinca, F., Tupayachi, R., Anderson, C.B., Martinez, P., 2014. Functional and biological diversity of foliar spectra in tree canopies throughout the Andes to Amazon region. New Phytol. 204, 127–139. https://doi.org/10.1111/nph.12895

Asner, G.P., Martin, R.E., Knapp, D.E., Tupayachi, R., Anderson, C., Carranza, L., Martinez, P., Houcheime, M., Sinca, F., Weiss, P., 2011a. Spectroscopy of canopy chemicals in humid tropical forests. Remote Sens. Environ. 115, 3587–3598. https://doi.org/10.1016/j.rse.2011.08.020

Asner, G.P., Martin, R.E., Tupayachi, R., Emerson, R., Martinez, P., Sinca, F., Powell, G.V.N., Wright, S.J., Lugo, A.E., 2011b. Taxonomy and remote sensing of leaf mass per area (LMA) in humid tropical forests. Ecol. Appl. 21, 85–98. https://doi.org/10.1890/09-1999.1

Baldeck, C. a., Asner, G.P., 2014. Improving remote species identification through efficient training data collection. Remote Sens. 6, 2682–2698. https://doi.org/10.3390/rs6042682

Ball, A., Sanchez-Azofeifa, A., Portillo-Quintero, C., Rivard, B., Castro-Contreras, S., Fernandes, G., 2015. Patterns of leaf biochemical and structural properties of Cerrado life forms: Implications for remote sensing. PLoS One 10, 1–15. https://doi.org/10.1371/journal.pone.0117659

Bhattacharyya, A., 1943. On a measure of divergence between two statistical populations defined by their probability distributions. Bull. Calcutta Math. Soc. 35, 99–109.

Brum, M., Teodoro, G.S., Abrahão, A., Oliveira, R.S., 2017. Coordination of rooting depth and leaf hydraulic traits defines drought-related strategies in the campos rupestres, a tropical montane biodiversity hotspot. Plant Soil 420, 467–480. https://doi.org/10.1007/s11104-017-3330-x

Castro-Esau, K.L., Sánchez-Azofeifa, G.A., Caelli, T., 2004. Discrimination of lianas and trees with leaf-level hyperspectral data. Remote Sens. Environ. 90, 353–372. https://doi.org/10.1016/j.rse.2004.01.013

Castro-Esau, K.L., Sánchez-Azofeifa, G.A., Rivard, B., Wright, S.J., Quesada, M., 2006. Variability in leaf optical properties of mesoamerican trees and the potential for species classification. Am. J. Bot. 93, 517–530. https://doi.org/10.3732/ajb.93.4.517
Cavender-Bares, J., Gamon, J.A., Hobbie, S.E., Madritch, M.D., Meireles, J.E., Schweiger, A.K., Townsend, P.A., 2017. Harnessing plant spectra to integrate the biodiversity sciences across biological and spatial scales. Am. J. Bot. 104, 966–969. https://doi.org/10.3732/ajb.1700061
Chavana-Bryant, C., Malhi, Y., Wu, J., Asner, G.P., Anastasiou, A., Enquist, B.J., Cosio Caravasi, E.G., Doughty, C.E., Saleska, S.R., Martin, R.E., Gerard, F.F., Chavana-Bryant, C., Malhi, Y., Wu, J., Asner, G.P., Anatasiou, A., Enquist, B.J., Saleska, S.R., Doughty, C., and Gerard, F., de la Riva, E.G., Olmo, M., Poorter, H., Ubera, J.L., Villar, R., Mahowald, N., Lo, F., Zheng, Y., Harrison, L., Funk, C., Lombardozzi, D., Goodale, C., Poorter, H., Niinemets, U., Poorter, L., Wright, I.J., Villar, R., Ke, Y., Im, J., Park, S., Gong, H., 2016. Leaf aging of Amazonian canopy trees revealed by spectral and physiochemical measurements. New Phytol. 11, 215. https://doi.org/10.1111/nph.13853
Chen, S., Hong, X., Harris, C.J., Sharkey, P.M., 2004. Sparse Modeling Using Orthogonal Forward Regression With PRESS Statistic and Regularization. IEEE Trans. Syst. Man Cybern. Part B 34, 898–911. https://doi.org/10.1109/TSMCB.2003.817107
Cheng, T., Rivard, B., Sánchez-Azofeifa, A., 2011. Spectroscopic determination of leaf water content using continuous wavelet analysis. Remote Sens. Environ. 115, 659–670. https://doi.org/10.1016/j.rse.2010.11.001
Cheng, T., Rivard, B., Sánchez-Azofeifa, A.G., Féret, J.-B., Jacquemoud, S., Ustin, S.L., 2014. Deriving leaf mass per area (LMA) from foliar reflectance across a variety of plant species using continuous wavelet analysis. ISPRS J. Photogramm. Remote Sens. 87, 28–38. https://doi.org/10.1016/j.isprsjprs.2013.10.009
Curran, P.J., 1989. Remote sensing of foliar chemistry. Remote Sens. Environ. 30, 271–278. https://doi.org/10.1016/0034-4257(89)90069-2
Curran, P.J., Dungan, J.L., Macler, B.A., Plummer, S.E., Peterson, D.L., 1992. Reflectance spectroscopy of fresh whole leaves for the estimation of chemical concentration. Remote Sens. Environ. 39, 153–166. https://doi.org/10.1016/0034-4257(92)90133-5
Curran, P.J., Dungan, J.L., Peterson, D.L., 2001. Estimating the foliar biochemical concentration of leaves with reflectance spectrometry. Remote Sens. Environ. 76, 349–359. https://doi.org/10.1016/S0034-4257(01)00182-1
Dayrell, R.L.C., Arruda, A.J., Pierce, S., Negreiros, D., Meyer, P.B., Lambers, H., Silveira, F.A.O., 2018. Ontogenetic shifts in plant ecological strategies. Funct. Ecol. 0–2. https://doi.org/10.1111/1365-2435.13221
Díaz, S., Kattge, J., Cornelissen, J.H.C., Wright, I.J., Lavorel, S., Dray, S., Reu, B., Kleyer, M., Wirth, C., Colin Prentice, I., Garnier, E., Bönisch, G., Westoby, M., Poorter, H., Reich, P.B., Moles, A.T., Dickie, J., Gillison, A.N., Zanne, A.E., Chave, J., Joseph Wright, S., Sheremet'ev, S.N., Jactel, H., Baraloto, C., Cerabolini, B., Pierce, S., Shipley, B., Kirkup, D., Casanoves, F., Joswig, J.S., Günther, A., Falczuk, V., Rüger, N., Mahecha, M.D., Gorné, L.D., 2016. The global spectrum of plant form and function. Nature 529, 167–171. https://doi.org/10.1038/nature16489
Doughty, C.E., Asner, G.P., Martin, R.E., 2011. Predicting tropical plant physiology from leaf and canopy spectroscopy. Oecologia 165, 289–299. https://doi.org/10.1007/s00442-010-1800-4
Doughty, C.E., Santos-Andrade, P.E., Goldsmith, G.R., Blonder, B., Shenkin, A., Bentley, L.P., Chavana-Bryant, C., Huaraca-Huasco, W., Díaz, S., Salinas, N., Enquist, B.J., Martin, R., Asner, G.P., Malhi,

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

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

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

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

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

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

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

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

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

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

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

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

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

Gitelson, A.A., Gritz †, Y., Merzlyak, M.N., 2003. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. J. Plant Physiol. 160, 271–282. https://doi.org/10.1078/0176-1617-00887

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

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

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

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

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

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

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

Kikuzawa, K., Lechowicz, M.J., 2011. Ecology of Leaf Longevity, Ecological Research Monographs, Ecological Research Monographs. Springer Tokyo, Tokyo. https://doi.org/10.1007/978-4-431-53918-6

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

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

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

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

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

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

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

Negreiros, D., Le Stradic, S., Fernandes, G.W., Rennó, H.C., 2014. CSR analysis of plant functional types in highly diverse tropical grasslands of harsh environments. Plant Ecol. 215, 379–388. https://doi.org/10.1007/s11258-014-0302-6

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Wang, Z., Townsend, P.A., Schweiger, A.K., Couture, J.J., Singh, A., Hobbie, S.E., Cavender-Bares, J., 2019. Mapping foliar functional traits and their uncertainties across three years in a grassland experiment. Remote Sens. Environ. 221, 405–416. https://doi.org/10.1016/j.rse.2018.11.016

Warming, E., 1908. Lagoa Santa: Contribuição para a geographia phytobiologica. Arq. da Real Soc. Dinamarqueza das Sci. Naturaes e Math. VI.

Wold, S., 1994. PLS for multivariate linear modeling., in: Waterbeemd, H. van de (Ed.), Chemometric Methods in Molecular Design. Verlag-Chemie, Weinheim, Germany, pp. 195–218.

Wold, S., Sjöström, M., Eriksson, L., 2001. PLS-regression: a basic tool of chemometrics. Chemom. Intell. Lab. Syst. 58, 109–130. https://doi.org/10.1016/S0169-7439(01)00155-1

Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J.H.C., Diemer, M., Flexas, J., Garnier, E., Groom, P.K., Gulias, J., Hikosaka, K., Lamont, B.B., Lee, T., Lee, W., Lusk, C., Midgley, J.J., Navas, M.-L., Niinemets, Ü., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V.I., Roumet, C., Thomas, S.C., Tjoelker, M.G., Veneklaas, E.J., Villar, R., 2004. The worldwide leaf economics spectrum. Nature 428, 821–827. https://doi.org/10.1038/nature02403

Wu, J., Chavana-Bryant, C., Prohaska, N., Serbin, S.P., Guan, K., Albert, L.P., Yang, X., van Leeuwen, W.J.D., Garnello, A.J., Martins, G., Malhi, Y., Gerard, F., Oliviera, R.C., Saleska, S.R., 2017. Convergence in relationships between leaf traits, spectra and age across diverse canopy environments and two contrasting tropical forests. New Phytol. 214, 1033–1048. https://doi.org/10.1111/nph.14051

Zappi, D., Milliken, W., Nicholas Hind, D.J., Biggs, N., Rando, J., Malcolm-Tompkins, P., Mello-Silva, R., 2014. Plantas do Setor Noroeste da Serra do Cipó, Minas Gerais - guia ilustrado. https://doi.org/10.13140/2.1.3216.9286

# Supplementary material for:

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

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

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

```
ldmc.aov <- aov(log10(LDMC) ~ growth_form, data = new_df)
lma.aov <- aov(LMA ~ growth_form, data = new_df)
```

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

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

**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 | |

## 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.

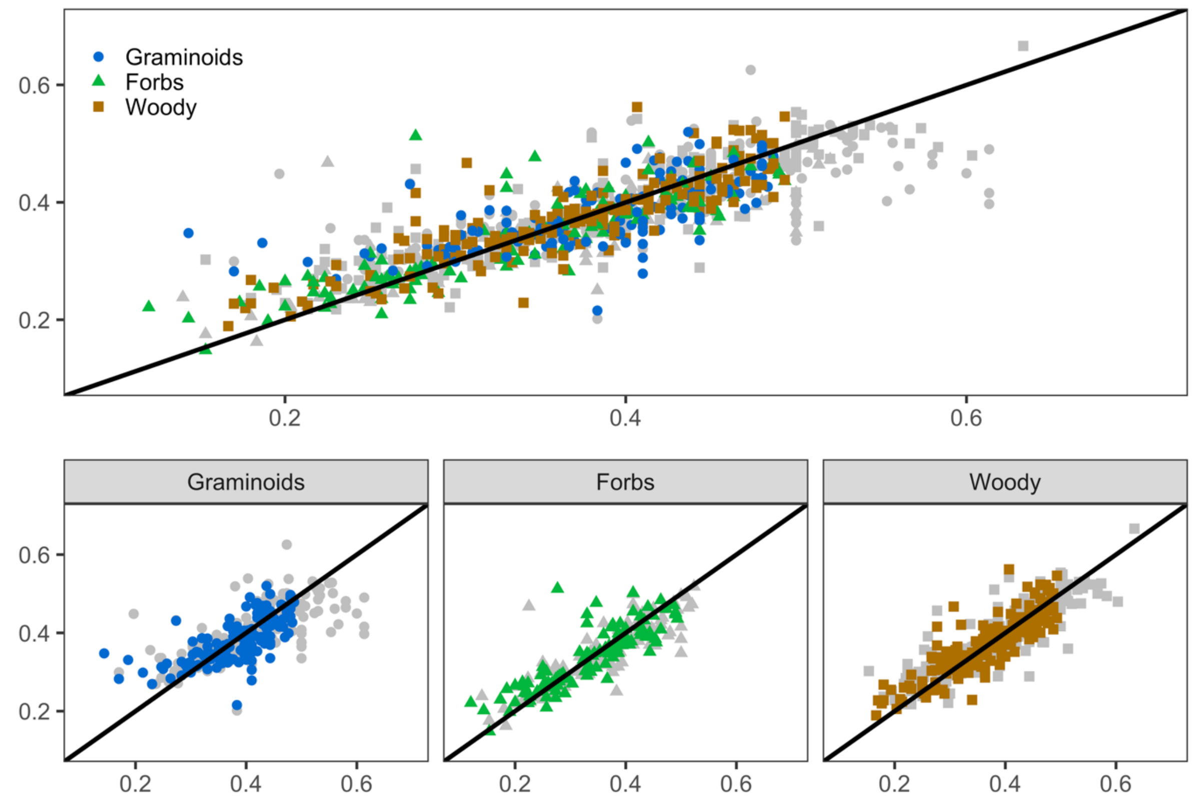


**Figure S1**: Partial least-squares regression (PLSR) results for observed vs. predicted leaf dry matter content (LDMC) with values restricted to 0 – 0.05 g/g. The upper panel shows the prediction for the total range of LDMC values ("*All*" class). The lower panels show the relationship between observed 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.

**Table S3.** Results of the partial least-squares regression (PLSR) modeling and cross-validation for each leaf trait, showing the number of samples and range of trait variation for the global data set (all) and per growth form. RMSECV is the root mean square error (RMSE) of the cross-validation procedure with train data set; RMSE is the measured error using the test data; mRMSE is the ratio of the error of each model in relation to the mean values (RMSE/mean); and the RMSE percentage (%RMSE) shows the error of each model as a percentage of the observed data range. Predicted R2 shows the predictive quality of each model. All results are presented for the entire range of LMA and LDMC values ("*All*" class) and per growth form. "*LMA < 300*" represents the data set containing only LMA values bellow 300 g/m$^2$.

| Growth form | Number of samples | Range of variation (min - max) | RMSECV | Final number of latent variables | RMSE | mRMSE (RMSE/ mean) | %RMSE (% of range) | $R^2$ |
|---|---|---|---|---|---|---|---|---|
| **LDMC** | | | | | | | | |
| 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 |
| **LDMC < 0.05** | | | | | | | | |
| 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 |