**Effects of distinct data processing in the phylogenetic signature of foliar spectra of regenerating plants in Neotropical Forest gaps**

**Écio Souza Diniza,\*, Cibele Hummel do Amarala,b, Lucas Arthur de Almeida Tellesa, and João Augusto Alves Meira-Netoa**

a Laboratório de Geoprocessamento e Sensoriamento Remoto, Universidade Federal de Viçosa, Departamento de Engenharia Florestal, Campus UFV s/n, 36570-000, Viçosa, Minas Gerais, Brazil

b Laboratório de Ecologia e Evolução de Plantas, Departamento de Biologia Vegetal, Universidade Federal de Viçosa, Campus UFV s/n, 36570-000, Viçosa, Minas Gerais, Brazil

**\*** Corresponding author: [eciodiniz@gmail.com](mailto:eciodiniz@gmail.com)

**Abstract.** Important biochemical traits may be robustly mapped by optical spectroscopy, allowing inferences about phylogenetic conservation in plant species. However, distinct types of data processing might lead to distinct patterns of phylogenetic signal in the foliar spectra. Thus, investigate the standard analytical approaches in order to understand their influence over the phylogenetic signal is essential. In this context, this study investigates how untransformed and transformed foliar-spectral data affects the phylogenetic signal of plant species located in regenerating forest gaps. Spectroscopic measurements from the adaxial surface of leaf samples were taken under standard light and temperature conditions for 53 regenerating plant species with a field spectroradiometer. Then, the average spectral signature for each specie was considered under two types of data processing: untransformed (raw reflectance spectra) and transformed (normalization by first derivative). Examined spectral regions for untransformed and transformed wavelengths were VIS (visible region: 400–700 nm), NIR (near infrared region: 701–1349 nm), SWIR-1 (short-wave infrared region part one: 1551–1849 nm) and SWIR-2 (short-wave infrared region part two: 2051–2450). Evolutionary conservation was evaluated through Blomberg (K) and Pagel (λ) metrics, which were calculated for all 1649 bands considering the average specie’s spectra. Thus, the percentage of wavelength with significant phylogenetic signal (K and λ) was quantified in both types of spectral processing. Significant phylogenetic signal was found for transformed spectra in NIR and SWIR-1 regions, along with reduced portions in SWIR-2. For untransformed spectra, there was significant signal mainly in SWIR-2. In conclusion, main results indicate that normalization by first derivative performs better in disentangling overlapping wavelengths. Thus, the transformed spectra can highlight the phylogenetic signal of plant features that are underemphasized in untransformed foliar spectra.

**Keywords**: Foliar Spectra, Data Normalization, Phylogenetic Conservation, Plant Biochemical Traits, Reflected Wavelengths, Short-Wave Infrared.

**1. Introduction**

Plant assemblages are structured according to several abiotic and biotic factors that directly influence the species’ functional traits (e.g., biochemistry, anatomy, morphology), whose interaction is capable of driving communities with high (over-dispersed) or low (clustered) phylogenetic diversity (Webb et al. 2002; Cavender-Bares et al. 2009). Evolutionary conservation of species’ traits, which is demonstrated by the presence of phylogenetic signal, correlates to assembly of phylogenetically clustered or over-dispersed communities 1,3. Many important ecophysiological plant traits (e.g. N, Cellulose, Tannin, Lignin) might be potentially conserved along evolution in phylogenetically close clades, which determine the species potential to adapt and succeed in the underlying environmental conditions (Ackerly et al. 2000; Ackerly 2009). Such biochemical/physiological traits may be robustly mapped by optical spectroscopic imaging (Curran 1989; Doughty et al. 2011).

Applications from the last decade demonstrate that optical spectroscopy is a promissory, accurate and quick tool to track plant functional characteristics and their plasticity. It allows robust measurements of foliar reflectance and absorption features, which are produced by electronic process and chemical bonds in leaf biochemistry (Madritch et al. 2014; Singh et al. 2015). Innovative studies (Cavender-Bares et al. 2016; McManus et al. 2016; Diniz et al. 2021b) show that the combination of remote and proximal sensing with community phylogenetic information, allows to investigate the evolutionary conservation of biochemical foliar traits (e.g. evolutionary conservation of N, Cellulose, Lignin). Those characteristics are related with distinct regions of the optical spectrum, what enable to identify the respective wavelengths that constitute the spectral-phylogenetic signature/signal of each feature (i.e., from visible to short-wave infrared regions – VSWIR).

The combined application of remote sensing techniques, such as imaging spectroscopy, and phylogenetic data, can be an important allied in ecological investigations, since it enables to understand evolutionary patterns of the vegetation spectral signatures at different scales, e.g. landscapes and ecosystems (Schweiger et al. 2018; Diniz et al. 2021b). Even highly diverse and complex communities, such as, Neotropical forests, can benefit from collecting vast amounts of plant spectral properties at broader scales (Amaral et al. 2018; Schweiger et al. 2018), which allows inferences about community functioning according to observed levels of phylogenetic conservation of ecophysiological features registered in spectral signatures.

Despite the potential and robustness of using foliar spectra signatures to infer about the concentrations of biochemical and physiological plant traits, the raw (non-normalized) spectra usually contains high-frequency noise along the spectrum(Rinnan et al. 2009; Beal and Eamon 2010; González-Fernández et al. 2019), which can bias and limit conclusive interpretations regarding ecological patterns. Thus, when testing for phylogenetic signal in biological traits, e.g. leaf biochemistry inferred by spectra, it is recommended to standardize the trait values with a suitable transformation technique(Pavoine and Ricotta 2013). Derivative transformations like first and higher order are commonly used to smooth spectral data and correct baseline effects, thus better discriminating overlapping and differences between spectrum bands, while keeping reflectance peaks in the same place as in the original reflectance spectra(Norris and Williams 1984; Rinnan et al. 2009; Silalahi et al. 2018). Despite of being a key factor in applying image spectroscopy for phylogenetic analysis, the impact of data processing (i.e., raw or transformed spectra) on phylogenetic signal is still not addressed. Neglect its effect may lead to biased interpretations and conclusions about evolutionary conservation across spectrum.

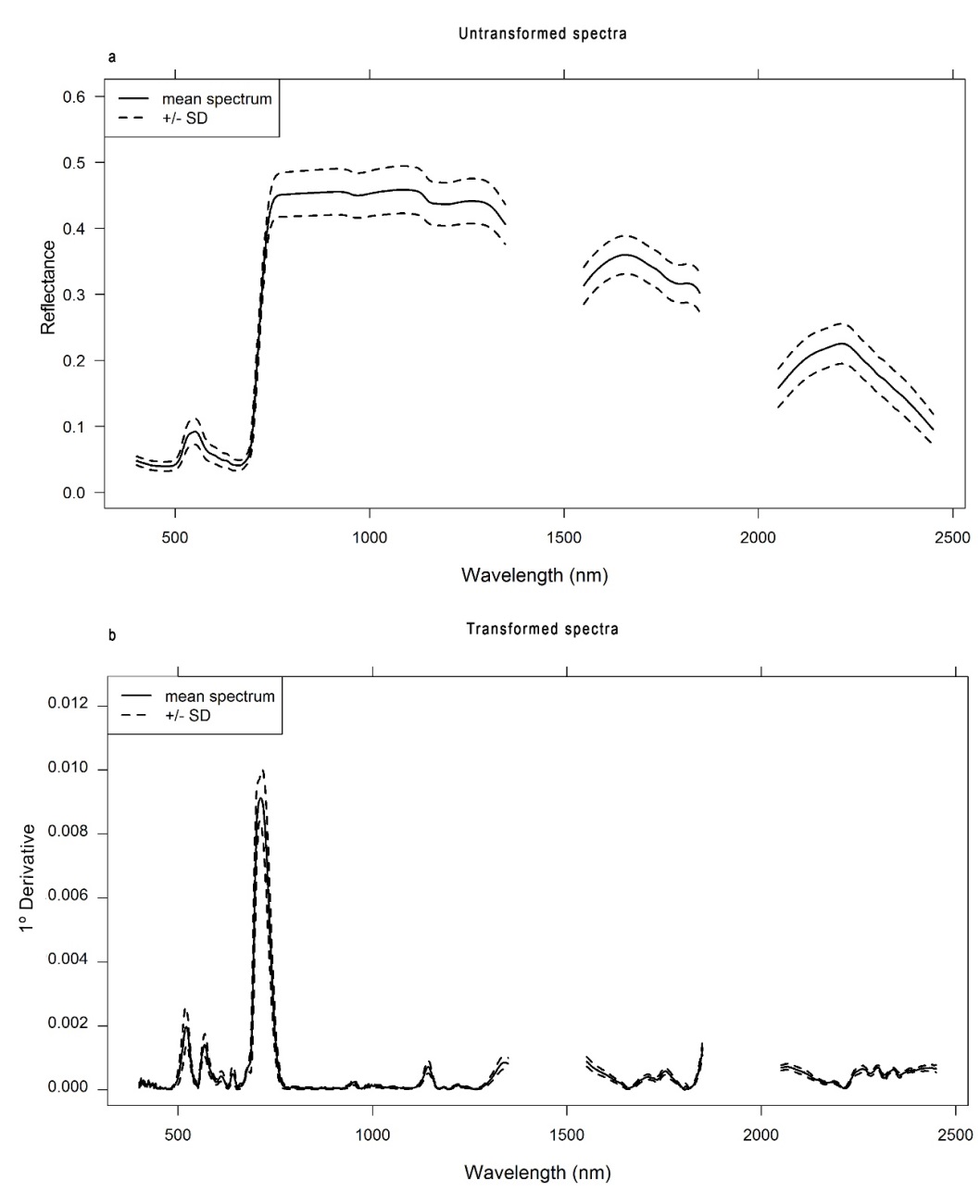
Therefore, understanding how spectral data processing may affect the outcomes of phylogenetic signals of wavelengths is of major importance, especially for plant communities under regenerating stage. In the varied luminosity conditions inherent from gap-phase, the plant photosynthetic plasticity required to succeed in such environments is expected to be markedly expressed in its spectral signature (Chazdon et al. 1996; Cavender-Bares et al. 2016). Thus, data processing may highlight or underestimate spectral differences among taxa. In this study, we aimed to investigate how raw (untransformed) and first derivative (transformed) data affects the phylogenetic signal outcomes, concerning the reflected foliar spectra from plant species located in regenerant forest gaps. The null hypothesis states that both raw and transformed spectra does not change the overall trends of phylogenetic signal of wavelengths across spectral regions.

**2 Material and methods**

*2.1 Data collection and spectral processing*

Spectroscopic measurements were carried out in the adaxial surface of leaf samples from 53 regenerating plant species existing in the ground of forest gaps from two protected seasonal semi-deciduous montane forests (Mata da Biologia - site 1 and Mata do Paraíso - site 2), both located in Viçosa county (20°45′14″S, 42°45′53″W). This region, known as “Zona da Mata Mineira”, is inserted in the Atlantic Forest domain, Minas Gerais State, Southeast Brazil. The measurements were taken under standard light and temperature conditions by using an ASD FieldSpec 4 spectroradiometer equipped with Plant Probe and Leaf Clip accessories (ASD Inc., Boulder, CO, USA). The average spectral signature from five leaves per individual was adopted as the spectral signature value for each specie. For more detail and information about study site, specific settings of sample collection and spectroscopic measurements, see Diniz, et al. (Diniz et al. 2021b).

The averaged spectra per specie was considered as untransformed reflectance, while the transformed spectra were achieved by normalization of first derivative. We used the Derivative Gap functionality from ASD Fieldspec 4, whose algorithm employees a specified gap distance to take the differences instead of adjacent data points when transforming raw spectra by the first derivative. For further details about the theoretical backbone of derivative approximation and mathematical equations that this algorithm is based on, see Morrey (Morrey 1968). It was considered the spectrum (untransformed and transformed) spanning from 400 nm to 2450 nm wavelength region, removing low signal-to-noise regions (350–399 nm and 2451–2500 nm) and wavelengths between 1350–1550 nm as well as 1850–2050 nm in order to avoid noise generated by water absorption, common in those regions. Thus, data accounted for 1649 nm-channels. Therefore, the examined spectral regions were: VIS (visible region: 400–700 nm), NIR (near infrared region: 701–1349 nm), SWIR-1 (short-wave infrared region part one: 1551–1849 nm) and SWIR-2 (short-wave infrared region part two: 2051–2450). Averaged untransformed reflectance spectra (Fig. 1a) and first derivative transformed spectra (Fig. 1b), concerning the 53 sampled species in both forest gaps/sites are presented in Fig. 1. To access the complete species data of average leaf spectra (transformed and untransformed), please see Diniz, et al. (Diniz et al. 2021b).



**Fig. 1.** Averaged untransformed (a) and transformed (b) reflected foliar spectra of all 53 regenerating plant species sampled in Semideciduous Atlantic Forest gaps located in Southeast Brazil.

*2.2 Phylogenetic tree reconstruction and signal*

The phylogenetic tree was constructed from the species pool (53 regenerating species) concerning both study sites. Thus, we pruned the mega-tree GBOTB pruning to our species list under phylogenetic hypothesis of scenario 3 by using the function phylo.maker, which is available in V.PhyloMaker package (Jin and Qian 2019) for R programing language version 3.6.1 (R Development Core Team 2019). The extraction of genus or family-level largest cluster's root and basal node information was created through the mega-tree by using build.nodes.1 function (Jin and Qian 2019). Then, all 1649 bands processed by splice correction and first derivative were tested according to Blomberg’s K (Blomberg et al. 2003) and Pagel’s Lambda (λ) metrics (Pagel 1999; Freckleton et al. 2002) for phylogenetic signal against the hypothesis of signal absence or congruence with stochasticity in character evolution (i.e. trait) under Brownian Motion Evolution (BM) model. According to BM, a trait evolves in small amount during a given evolutionary interval and in random manner 4. Thus, tests for phylogenetic signal of traits can be interpreted as indicators for homologous (similar trait values share a common ancestry, i.e. evolutionary divergence) or homoplastic (similar trait values area share independently in separate lineages, i.e. convergence) trait conservatism in evolutionary lineages (Stearns and Hoekstra 2005; Losos 2008).

K and λ were calculated using phyloSignal function from phylosignal R package with 10,000 randomizations (Keck F, Rimet F, Bouchez A 2016). Those metrics display complementary role when testing phylogenetic signal, i.e. Blomberg's K is accurate in capturing small changes in the phylogenetic trait distribution that are not related to BM strength, while Pagel's λ is robust for assessing random trait distribution under BM (Münkemüller et al. 2012). For further details about formulas and statistics behind both metrics, please see Blomberg et al. (Blomberg et al. 2003) and Freckleton et al. (Freckleton et al. 2002).

Regarding metric values, K = 0 indicates absence of phylogenetic signal; K = 1, trait distribution perfectly conforms to BM; K < 1 indicates fewer trait similarities among closely related species than what would be expected under BM, i.e. phylogenetic conservation of convergent (homoplastic) traits; K > 1 indicates greater phylogenetic signal among closely related species than what would be expected under BM, i.e. phylogenetic conservation of homologous (divergent) traits (Blomberg et al. 2003). The λ metric that maximizes the likelihood of trait data (e.g., foliar spectra) given a BM is interpreted as λ = 0, indicating no signal; whereas λ =1 suggests conformity to BM. Conversely, values between 0 and 1, i.e. λ < 1, indicate less phylogenetic signal than expected and values > 1 indicate more signal than expected under BM, although λ is not always defined for values greater than one once off-diagonal elements in the variance-covariance matrix cannot be larger than the diagonal elements, thus restricting λ’s upper limit (Pagel 1999; Freckleton et al. 2002). We accounted for the percentages of phylogenetic signal (evaluated considering the distinct values =0, =1, >1 and < 1 of K and λ) in untransformed and transformed spectra, through quantifying the wavelengths within each spectral region with significant signal. To calculate the overall percentage of significant phylogenetic signal per spectral region, we compared their significant K and λ number of wavelengths to the total 1649 wavelengths.

**3 Results**

Overall, untransformed (reflectance) and transformed (1º derivative) spectra revealed close phylogenetic signal trends (K and λ) in all analyzed spectral regions (VIS, NIR, SWIR-1 and SWIR-2; Table 1). According to significant values lower than 1, found for K and λ metrics (Table 1), the major percentage of phylogenetic signature in untransformed and transformed spectra suggested homoplasy, i.e., evolutionary convergence, indicating shared wavelengths among more distant phylogenetic relatives than what would be expected under Brownian Motion Evolution. In short, reflected wavelengths indicate an overall trend to evolutionary convergence regardless the type of spectra processing (untransformed or transformed).

While the lowest percentage of phylogenetic signal (K and λ) was found in the NIR region of untransformed spectra, both spectra (untransformed and transformed) evidenced low and high percentage for VIS (400–700 nm) and SWIR 1-2 (1551–1849 nm; 2051–2450 nm), respectively (Table 1). NIR presented intermediate portions of phylogenetic signatures in the transformed spectra, but no significance when analyzing the untransformed one (Table 1). This and the significant phylogenetic signal found in SWIR-2 for K, among close relatives (99% of significance), and λ, among distant relatives (100%) (Table 1), constituted the major distinction between phylogenetic tests conducted under splice corrected (untransformed) compared to the 1ª derivative normalized (transformed) spectral data. Wavelengths from all analyzed spectral regions that demonstrate the most significant phylogenetic signal according K and λ metrics are shown in the table 2. To assess the full phylogenetic signal outcomes from transformed and untransformed spectra for both metrics (K and λ), please see Diniz et al. (Diniz et al. 2021a).

**Table 1.** Percentage of wavelengths showing significant phylogenetic signal (calculated using Blomberg's K and Pagel's λ metrics) in each region of untransformed (values outside brackets) and transformed (1ª derivative, values between brackets) foliar spectra. The first derivative outcomes correspond to values reported by Diniz et al. (Diniz et al. 2021b).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Metric | VIS  (400-700nm) | NIR  (701-1349nm) | SWIR-1  (1551-1849nm) | SWIR-2  (2051-2450nm) |
| *% K total* | 3.33 [3.33] | 1.07 [30.92] | 29.66 [54.66] | 100 [30.42] |
| *%K>1* | 0 [0] | 0 [2.92] | 0 [7.66] | 99 [2.49] |
| *%K<1* | 3.33 [3.33] | 1.07 [28] | 29.66 [47] | 1 [27.93] |
| *% K-all spectra* | 0.6 [0.6] | 0.42 [12.18] | 5.39 [9.93] | 24.24 [7.39] |
| *% λ total* | 0 [0.33] | 0 [26.46] | 34 [40.33] | 100 [27.68] |
| *% λ>1* | 0 [0] | 0 [0.61] | 0 [0] | 0 [0.49] |
| *% λ<1* | 0 [0.33] | 0 [25.84] | 34 [40.33] | 100 [27.18] |
| *% λ -all spectra* | 0 [0.06] | 0 [10.4] | 6.17 [7.33] | 24.25 [6.71] |

% K total: total percentage of wavelengths with significant phylogenetic signal derived using Blomberg's K; %K > 1 and %K < 1 are, respectively, the percentage of wavelengths with significant K > 1 and K < 1; %K-all-spectra: percentage of wavelengths with a significant phylogenetic signal in relation to the total number of wavelengths (1649) in the spectra. % λ total: total percentage of wavelengths with significant phylogenetic signal derived using Pagel's λ. % λ > 1 and % λ < 1 are, respectively, the percentage of wavelengths with significant λ > 1 and λ < 1; % λ-all-spectra: percentage of wavelengths with significant phylogenetic signal in relation to the total number of wavelengths.

**Table 2.** Wavelengths across spectral regions showing the highest significance of phylogenetic signal according to K and λ metrics.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Metric | VIS  (nm) | NIR  (nm) | SWIR-1  (nm) | SWIR-2  (nm) |
| *Untransformed* |  |  |  |  |
| *K* | 439 | 711-719 | 1567-1849 | 2051-2450 |
|  |  |  |  |  |
| *λ* | 439 | 711-719 | 1567-1849 | 2051-2450 |
| *Transformed* |  |  |  |  |
| *K* | 551 | 931-933 | 1731-1826 | 2435-2449 |
|  |  | 940-943 |  |  |
| *λ* | 551 | 931-933 | 1731-1826 | 2435-2449 |
|  |  | 940-943 |  |  |

**4 Discussion**

Our Findings indicated that transformation by first derivative highlights and normalize some wavelengths in NIR and SWIR-1 spectral regions, through reducing some portion of wavelengths with significant phylogenetic signal in SWIR-2 (Diniz et al. 2021b). Among highlighted portions of the spectra, some small ones were associated with spectral slopes of absorption features that are associated to enhancement of several biochemical compounds in plants(Diniz et al. 2021b). This becomes more evident when analyzing the wavelengths with highest phylogenetic signals (Diniz et al. 2021a). Such findings corroborate the premise that trait data (e.g. foliar spectra) transformation should be applied to limit the impact of outliers or noise when testing phylogenetic signal in order to evidence significant signal for trait values that otherwise could be underestimated due statistical type I error (Pavoine and Ricotta 2013). In addition, overestimated phylogenetic signal can be avoided by relating high resolution phylogenies with normalized/transformed trait data (Davies et al. 2012; Münkemüller et al. 2012). The phylogeny used in this study was fully resolved, i.e., bifurcated entirely regarding branch lengths and nodes, thus reinforcing unbiased estimations of phylogenetic signal for both untransformed and transformed foliar spectra.

While the vast majority of wavelengths from SWIR-2 (i.e., about 300 bands) had shown the highest phylogenetic signal in untransformed spectra, smaller portions (~10 to 100 bands) had shown the highest significance across transformed spectra. The transformed spectra also highlighted the high phylogenetic importance of NIR and SWIR-1 spectral regions for the plant’s absorption features (Diniz et al. 2021b). These phylogenetic signals are associated with nitrogen and carbon-rich compounds, rather than spectrally remarkable pigments (e.g. chlorophyll at 680 nm) or water (e.g. 1200 nm) (Kokaly et al. 2003). Moreover, the underemphasized visible spectral region (VIS) corroborates to the importance of C and N molecules in the phylogenetic structure of studied plant communities. For VIS, only wavelengths at 551 nm from transformed spectra showed a significant phylogenetic signal, while untransformed spectra showed significant signal in green-edge region, between 439 and 551 nm. Therefore, although underemphasized, VIS demonstrate important evolutionary conservation associated with Phycoerythrin (Kumar et al. 2002), a protein-pigment complex from the light-harvesting phycobiliprotein family related to photosynthetic efficiency.

In addition to IR (Infrared) wavelengths that are associated to nitrogen (Curran 1989; Kokaly 2001) and RuBISCO (Davies and Grant 1988), it can be inferred that in forest floor with high variation of illumination conditions, phylogenetic signals from regenerating species leaves seems to be associated with N and C rich compounds, whose role is both structural (e.g. lignin, cellulose) and physiological (e.g., phycoerythrin, rubisco). Thus, analyzing the results from untransformed and transformed spectral data made possible to disentangle the biochemical meaning of phylogenetic signals associated with VSWIR (400-2450 nm) light usage under such environmental condition.

In summary, results indicate that the normalization by first derivative transformation highlights the phylogenetic signal of spectral features that are underemphasized in the untransformed spectra. This suggests that this transformation better disentangle overlapping wavelengths and, consequently, provide more reliable phylogenetic signal outcomes. Therefore, such transformed spectra should be considered when drawing main interpretations about phylogenetic signal of foliar spectra. However, it is important to notice that in both cases (untransformed and transformed spectra) the highlighted wavelengths are mainly associated to N- and C- bonds present in plant structural and physiological compounds (e.g., lignin, cellulose, and proteins). They also belong to spectral regions (specially SWIR) recognized by high portions of overlapping bends (Curran 1989; Curran et al. 2001; Kokaly et al. 2003, 2009). Thus, we cannot assure that each data processing generated absolutely contrasting results. Further investigations would benefit from being able to predict underlying traits of leaves based on foliar spectra (untransformed and transformed) through supervised models like PLS regression, which is recognized as a model with reduced prediction error when applied on spectra data normalized by first derivative (Silalahi et al. 2018). This, in turn, would enable the selection of the most important wavelengths to each foliar trait, as well as test them for phylogenetic signal considering both data, untransformed and transformed (Schweiger et al. 2018; Silalahi et al. 2018).

**5 Conclusions**

Our findings lead us to conclude that normalizing foliar spectra data by applying robust transformation techniques, such as the first derivative, is a recommendable procedure since it may highlight phylogenetic signal in underemphasized spectral features concomitantly to the standardization of overestimated and overlapped wavelengths. Thus, the employment of such normalization can assist in reaching higher reliability in the interpretations of phylogenetic conservation of reflected wavelengths across the foliar spectra. Nevertheless, further studies conducted under larger sampling, especially including actual biochemical traits, are necessary to confirm if in fact the accuracy of the normalization provided by first derivative stands out.

**Declaration of competing interests**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Aknowledgements**

We thank the Departamento de Engenharia Florestal of the Universidade Federal de Viçosa - UFV (Brazil) for supporting field campaigns, the undergraduate students for supporting data collection, the Laboratório de Ecologia e Evolução de Plantas (LEEP) of UFV for providing the spectrometer and necessary equipment to conduct fieldwork, and FAPEMIG for funding support.

**Funding**

This research was funded by FAPEMIG (Fundação de Amparo à Pesquisa do Estado de Minas Gerais), grant number APQ-03066-17.

**References**

Ackerly D (2009) Conservatism and diversification of plant functional traits: Evolutionary rates versus phylogenetic signal. Proc Natl Acad Sci U S A 106 Suppl:19699–706. https://doi.org/10.1073/pnas.0901635106

Ackerly DD, Dudley SA, Sultan SE, et al (2000) The Evolution of Plant Ecophysiological Traits: Recent Advances and Future Directions: New research addresses natural selection, genetic constraints, and the adaptive evolution of plant ecophysiological traits. BioScience 50:979–995

Amaral CH do, Almeida TIR, Filho CRS, et al (2018) Characterization of indicator tree species in neotropical environments and implications for geological mapping. Remote Sensing of Environment 216:385–400

Beal D, Eamon M (2010) Preliminary Results of Testing and a Proposal for Radiometric Error Correction Using Dynamic, Parabolic Linear Transformations of “Stepped” Data - PCORRECT.EXE. Analytical Spectral Devices

Blomberg SP, Garland T, Ives AR (2003) Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution (N Y) 57:717–745. https://doi.org/10.1111/j.0014-3820.2003.tb00285.x

Cavender-Bares J, Kozak KH, Fine PVA, Kembel SW (2009) The merging of community ecology and phylogenetic biology. Ecology Letters 12:693–715. https://doi.org/10.1111/j.1461-0248.2009.01314.x

Cavender-Bares J, Meireles JE, Couture JJ, et al (2016) Associations of Leaf Spectra with Genetic and Phylogenetic Variation in Oaks: Prospects for Remote Detection of Biodiversity. Remote Sensing 8:1–17

Chazdon RL, Pearcy RW, Lee DW, Fetcher N (1996) Photosynthetic responses of tropical forest plants to contrasting light environments. In: Mulkey SS, Chazdon AP, Smith RL (eds) Tropical forest plant ecophysiology. Chapman & Hall, New York, p p.5-55

Curran PJ (1989) Remote sensing of foliar chemistry. Remote Sensing Environment 30:271–278

Curran PJ, Dungan JL, Peterson DL (2001) Estimating the foliar biochemical concentration of leaves with reflectance spectrometry: testing the Kokaly and Clark methodologies. Remote Sensing of Environment 76:349–359. https://doi.org/10.1016/S0034-4257(01)00182-1

Davies AMC, Grant A (1988) Near infrared spectroscopy for the analysis of specific molecules in food. Royal Society of Chemistry 46–51

Davies TJ, Kraft NJ, Salamin N, Wolkovich EM (2012) Incompletely resolved phylogenetic trees inflate estimates of phylogenetic conservatism. Ecology 93:242–247

Diniz ES, Amaral CH, Meira-Neto JAA (2021a) Phylogenetic Signal for transformed and untransformed Foliar Spectra of Regenerating Plants. Mendeley Data. https://doi.org/10.17632/jc2nrty843.1

Diniz ES, Amaral CH, Sardinha ST, et al (2021b) Phylogenetic signatures in reflected foliar spectra of regenerating plants in Neotropical forest gaps. Remote Sensing Environment 253:112172. https://doi.org/10.1016/j.rse.2020.112172

Doughty CE, Asner GP, Martin RE (2011) Predicting tropical plant physiology from leaf and canopy spectroscopy. Oecologia 165:289–299. https://doi.org/10.1007/s00442-010-1800-4

Freckleton RP, Harvey PH, Pagel M (2002) Phylogenetic analysis and comparative data: a test and review of evidence. The American Naturalist 160:712–726

González-Fernández AB, Sanz-Ablanedo E, Gabella VM, et al (2019) Field Spectroscopy: A Non-Destructive Technique for Estimating Water Status in Vineyards. Agronomy 9:1–19. https://doi.org/10.3390/agronomy9080427

Jin Y, Qian H (2019) V.PhyloMaker: an R package that can generate very large phylogenies for vascular plants. Ecography 42:1353–1359. https://doi.org/10.1111/ecog.04434

Keck F, Rimet F, Bouchez A FA (2016) phylosignal: an R package to measure, test, and explore the phylogenetic signal. Ecology and Evolution 6:2774–2780

Kokaly RF (2001) Investigating a physical basis for spectroscopic estimates of leaf nitrogen concentration. Remote Sensing of Environment 75:153–161

Kokaly RF, Asner GP, Ollinger S V, et al (2009) Characterizing canopy biochemistry from imaging spectroscopy and its application to ecosystem studies. Remote Sensing of Environment 113:S78--S91. https://doi.org/10.1016/j.rse.2008.10.018

Kokaly RF, Despain DG, Clark RN, Livo KE (2003) Mapping vegetation in Yellowstone National Park using spectral feature analysis of AVIRIS data. Remote Sensing of Environment 84:437–456

Kumar L, Schmidt K, Dury S, Skidmore A (2002) Imaging spectrometry and vegetation science. In: van der Meer FD, de Jong SM (eds) Imaging spectrometry. Springer, pp 111–155

Losos JB (2008) Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. Ecology Letters 11:995–1003. https://doi.org/10.1111/j.1461-0248.2008.01229.x

Madritch MD, Kingdon CC, Singh A, et al (2014) Imaging spectroscopy links aspen genotype with below-ground processes at landscape scales. Philosophical Transactions of the Royal Society B: Biological Sciences , 20130194 369:20130194

McManus KM, Asner GP, Martin RE, et al (2016) Phylogenetic structure of foliar spectral traits in tropical forest canopies. Remote Sensing 8:1–16. https://doi.org/10.3390/rs8030196

Morrey JR (1968) On Determining Spectral Peak Positions from Composit Spectra with a Digital Computer. Analytical Chemistry 40:905–914

Münkemüller T, Lavergne S, Bzeznik B, et al (2012) How to measure and test phylogenetic signal. Methods in Ecology and Evolution 3:743–756. https://doi.org/10.1111/j.2041-210X.2012.00196.x

Norris KH, Williams PC (1984) Optimization of mathematical treatments of raw near-infrared signal in the measurement of protein in hard red spring wheat. i. influence of particle size. Cereal Chemestry 61:158–165

Pagel M (1999) Inferring the historical patterns of biological evolution. Nature 401:877–884. https://doi.org/10.1038/44766

Pavoine S, Ricotta C (2013) Testing for phylogenetic signal in biological traits: The ubiquity of cross-product statistics. Evolution (N Y) 67:828–840. https://doi.org/10.1111/j.1558-5646.2012.01823.x

R Development Core Team . (2019) R: A language and environment for statistical computing

Rinnan A, van den Berg F, Engelsen SB (2009) Review of the most common pre-processing techniques for near-infrared spectra. Trends in Analytical Chemistry 28:

Schweiger AK, Cavender-Bares J, Townsend PA, et al (2018) Plant spectral diversity integrates functional and phylogenetic components of biodiversity and predicts ecosystem function. Nature Ecology & Evolution 2:976–982. https://doi.org/10.1038/s41559-018-0551-1

Silalahi DD, Midib H, Arasanb J, et al (2018) Robust generalized multiplicative scatter correction algorithm on pre-processing of near infrared spectral data. Vib Spectrosc 97:55–65

Singh A, Serbin SP, McNeil BE, et al (2015) Imaging spectroscopy algorithms for mapping canopy foliar chemical and morphological traits and their uncertainties. Ecological Applications 25:2180–2197

Stearns SC, Hoekstra RF (2005) Evolution: an introduction, 2nd edn. Oxford University Press, Oxford

Webb CO, Ackerly DD, Mcpeek M a, Donoghue MJ (2002) Phylogenies and Community Ecology. Annual Review of Ecology and Systematics 33:475–505

**Caption List**

**Figure 1.** Averaged untransformed (a) and transformed (b) reflected foliar spectra of all 53 regenerating plant species sampled in Semideciduous Atlantic Forest gaps located in Southeast Brazil.

**Table 1.** Percentage of wavelengths showing significant phylogenetic signal (calculated using Blomberg's K and Pagel's λ metrics) in each region of untransformed (values outside brackets) and transformed (1ª derivative, values between brackets) foliar spectra. The first derivative outcomes correspond to values reported by Diniz et al. (Diniz et al. 2021b).

**Table 2.** Wavelengths across spectral regions showing the highest significance of phylogenetic signal according to K and λ metrics.