Plant spectra as integrative measures of plant phenotypes

Shan Kothari[1,†], Anna K. Schweiger[2,†,*]

[1]Institut de recherche en biologie végétale and Département de sciences biologiques, Université de Montréal, Montréal, Québec, Canada

[2]Remote Sensing Laboratories, Department of Geography, University of Zurich, Zurich, Switzerland

[†] Both authors contributed equally to this article

[*] Corresponding author: anna.k.schweiger@gmail.com

Abstract

1. Spectroscopy at the leaf or canopy scales is becoming one of the core tools of plant functional ecology. Remotely sensed reflectance spectra can allow ecologists to infer plant traits and strategies—and the community- or ecosystem-level processes they correlate with—continuously over unprecedented spatial scales.
2. Because of the complex entanglement of structural and chemical factors that generate spectra, it can be tricky to understand exactly what phenotypic information they contain. We discuss common approaches to estimating plant traits from spectra—radiative transfer models and multivariate empirical models—and elaborate on their strengths and limitations in terms of the causal influences of various traits on the spectrum. Many chemical traits have broad, shallow, and overlapping absorption features, and we suggest that covariance among traits may have an important role in giving empirical models the flexibility to estimate such traits.
3. While trait estimates from reflectance spectra have been used to test ecological hypotheses over the past 20 years, we review a growing body of research that uses spectra directly, without estimating specific traits. By treating positions of species in multidimensional spectral space as analogous to trait space, researchers can infer processes that structure plant communities using the information content of the full spectrum, which may be greater than any standard set of traits. We illustrate this power by

showing that co-occurring grassland species are more separable in spectral space than in trait space and that the intrinsic dimensionality of spectral data is comparable to fairly comprehensive trait data sets. Nevertheless, using spectra this way may make it harder to interpret patterns in terms of specific biological processes.

4. *Synthesis.* Plant spectra integrate many aspects of plant function. The information in the spectrum can be distilled into estimates of specific traits, or the spectrum can be used in its own right. These two approaches may be complementary—the former being most useful when specific traits of interest are known in advance and reliable models exist to estimate them, and the latter being most useful in taking advantage of the information in the full spectrum under uncertainty about which aspects of function matter most.

Keywords: dimensionality, functional ecology, leaf-level, plant traits, remote sensing, spectroscopy,

1. Introduction

The way plants interact with light is one of the key determinants of their success in particular environments and one of most important processes structuring plant communities (Canham et al. 1994; Pacala & Tilman 1994; Williams et al. 2020). The balance between striving for and avoiding excess light (Kothari et al. 2018; Kothari et al. 2021) influences plant growth and architecture (Jucker et al. 2015; Williams et al. 2017), water and resource use (Ellsworth & Reich 1993), creating complex environments that shape plant communities. The processes surrounding light use of plants give rise to a series of chemical, morphological and anatomical adaptations, ranging from trade-offs between fast and slow return on investment traits (Wright et al. 2004), to cuticular structures reducing water-loss through evaporation, and leaf thickness altering heat dissipation, many of which influence optical properties. Spectra of plants can thus be viewed as manifestations of the interactions between plants and light, providing a window into plant-environment relationships.

When light interacts with plants, or any substance for that matter, three things can happen: light can get (1) absorbed by plant tissues; (2) scattered in the forward direction, which is called transmittance; or (3) scattered in the backward direction, which is called reflectance. The partitioning among these three processes depends on the chemical and structural characteristics of plants, including leaf chemistry, anatomy and morphology (Ustin and Jacquemoud 2020)—and, when measured from a distance, canopy architecture (Serbin and Townsend 2020). Spectroscopic methods measure reflectance, absorptance or transmittance of light in many narrow contiguous bands. Measurements typically include the visible (wavelengths between 400–700 nm, VIS), near-infrared (700–1000 nm, NIR) and often also the shortwave infrared (1000–3000 nm, SWIR) ranges of the electromagnetic spectrum. Together, these ranges account for >96% of solar radiation reaching Earth's surface (American Society for Testing and Materials, 2006), so plants' interactions with light in these ranges can be particularly revealing about their adaptations to the environment.

The spectral properties of plants have been studied for several decades (Billings & Morris 1951; Gates et al. 1965; Knipling 1970; Shull 1929). Shull (1929) for instance described how variation in chlorophyll content throughout leaf ontogeny changes spectral reflectance, as well as spectral differences among plant species, between healthy and diseased plants, green and albino plants, and upper and lower surfaces of leaves. Gates et al. (1965) describe changes in spectra with leaf development and the use of spectra to differentiate plants and lichens. Starting around the 1970s more and more studies started using spectra to predict chemical and structural plant traits, and the development of first airborne sensors set in motion plant trait mapping at the canopy level (Knipling 1970). Wessman et al. (1988) used predictions from NASA's AIS (Airborne Imaging Spectrometer) to publish the first maps of canopy nitrogen (N) and lignin content.

While the first spectrometers were laboratory instruments predominantly used for leaf level studies, today spectroscopy is used across multiple scales, from the leaf level, to the proximal remote sensing level using spectrometers in a fixed (e.g., in growth chambers, as scanners in a conveyor belt setting, etc.) or moving fashion (e.g. fixed on bicycles, on unoccupied aerial vehicles (UAVs), etc.), to airborne and spaceborne levels. This makes it possible to scale between different levels of observations, although this may not be trivial (Gamon et al. 2021).

Reflectance spectroscopy, which is the most common way to measure the interactions of light with plants, is increasingly becoming a core tool in plant ecology. Applications range from modeling and predicting leaf (Asner et al. 2014; Serbin et al. 2014) and canopy traits (Asner et al. 2017, Singh et al. 2015), to detecting plant stress (Asner et al. 2016) and natural enemies (Pontius et al 2005, Sapes et al. 2022), to differentiating species and broader taxonomic clades (Féret & Asner 2012) from spectra. Indeed, maps of plant (Wang et al. 2019) and plant community traits (Cavender-Bares et al. 2022), species (Roth et al. 2016) and functional group composition (Schmidtlein et al. 2012; Schweiger et al. 2017) are highly valuable for investigating a plethora of ecological questions beyond the scale of individual research plots.

Most research in plant ecology occurs in mid-latitudinal ecosystems of the global North (Meyer et al. 2016) around the peak of the growing season. These geographic, taxonomic and temporal biases in global plant data (Jetz et al. 2016) cannot be overcome with traditional field surveys alone. Using remotely sensed spectral data in plant ecology allows consistent and repeated measurements of plant traits and plant community characteristics across large spatial extents (Schimel et al. 2013, Turner 2014), which is particularly relevant in our time of rapid global change (Cavender-Bares et al. 2021).

In addition to trait and species mapping, plant spectroscopy over the past decade has also seen the growing use of spectra as integrated measures of plant phenotypes (Cavender-Bares et al. 2017, Ustin & Gamon 2010), including in biodiversity-ecosystem function research (Schweiger et al. 2018, 2020; Williams et al. 2021) and as measures of plant diversity (Draper et al. 2019, Féret & Asner 2014, Frye et al. 2021, Rocchini et al. 2010, Schweiger et al. 2018, Wang & Gamon 2019). Instead of mapping traits, these studies use spectra of plants directly as a means to understand how plants interact with their environment. This review addresses both trait mapping and emerging approaches that use spectra directly (Figure 1). We ask: What can spectroscopy contribute to plant ecology beyond providing more estimates of things (plant traits and taxonomic classes) that we can already measure in other ways? Our question leads to a conundrum: How can spectroscopy of plants contribute to plant ecology in a meaningful way when much of the phenotypic variation that spectra measure is not well understood? Indeed, only a handful of plant traits have distinct enough absorption features that they can be predicted with physical models from spectra (Section 2). Solving this conundrum requires understanding, or at least acknowledging, the complex entanglement of structural and chemical factors that generate spectra.

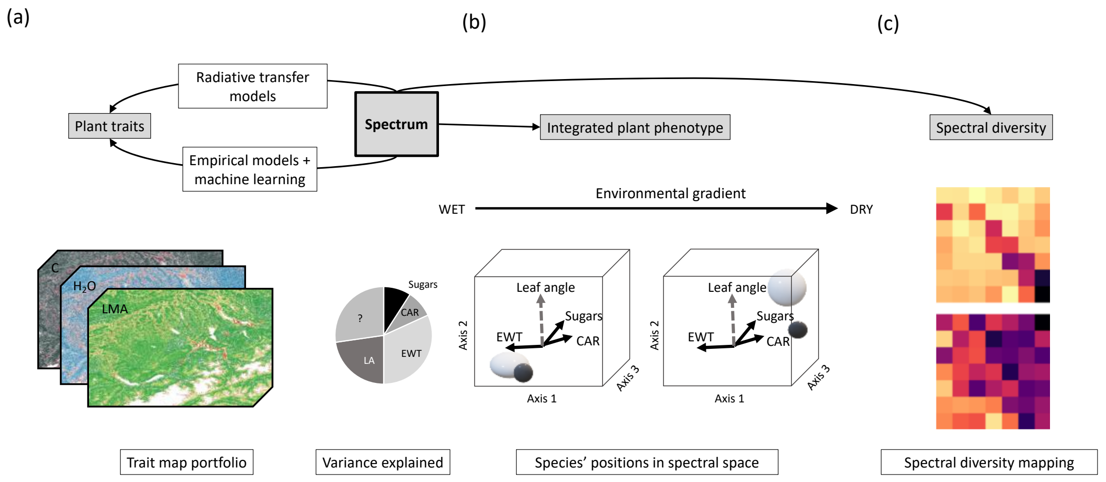

Figure 1. A conceptual figure illustrating what spectroscopy can do for plant functional ecology. (a) Plant traits (H_2O = water content, C = carbon content, N = nitrogen content, LMA = leaf mass per area) can be mapped from spectra using radiative transfer models or empirical machine learning approaches. (b) Spectra can also be used as integrated measures of plant phenotypes, for example to assess species hypervolume size and position in trait space along environmental gradients. Here, we depict a hypothetical gradient from wet to dry ecosystems. We show the position of species in a simplified multidimensional space of canopy reflectance spectra. Four traits are designated by arrows pointing in their direction of maximum correlation with the spectral axes: soluble sugars, equivalent water thickness (EWT), carotenoid content (CAR), and leaf angle distribution (LA). The arrow for LA is dashed to denote that it is a trait that is seldom measured directly, but can strongly influence canopy spectra. In our hypothetical example, drought causes the species to shift along the three spectral axes. Under drought, the species separate along Axis 2 (corresponding to LA), a difference in drought responses that may be hard to describe using traits alone. The pie chart depicts the percentage of spectral variation that might be statistically explained by these four traits, with some variation left unexplained. (c) Spectral data can also be used directly to map spectral diversity, which often corresponds with functional diversity.

In this review, we illustrate both the use of spectroscopy to derive plant traits and its potential to go beyond commonly measured traits by using spectra as integrated measures of plant phenotypes and the challenges of this approach. We will tackle the following questions:

- What is the role of trait covariance in the detection of plant traits that do not cause clearly identifiable absorption features (including non-leaf traits)?
- What can we learn by working directly with spectra compared to spectrally-derived plant traits? What are the challenges?
- How can we leverage the full information content of spectra?
- What is the intrinsic dimensionality of spectra and how should the major dimensions of spectral variation be interpreted?

Finally, we will present three case studies that illustrate the richness of information in the spectrum. We will from time to time refer to leaf level studies given that they have provided or can be assumed to provide the basis for applications at the remote sensing level. We will, however, not cover multispectral sensors, such as Landsat and Sentinel or multiband cameras, because they lack the spectral range and resolution needed to infer many traits (Shiklomanov et al. 2016) or to think about spectra in the continuous, integrative way that we seek to highlight.

2. Plant traits from spectra

2.1 Physical and empirical approaches to estimating plant traits from spectra

Estimating traits from spectra can save considerable amounts of time and resources in the lab and field. For example, a hand-held spectrometer with a leaf clip allows users to derive traits from dozens of leaves within a few hours, without destructive sampling, material transportation to the lab or chemical analyses. However, estimating traits from spectral information requires a model, typically either: (1) a physics-based radiative transfer model (Féret et al. 2021), or (2) an empirical model built using multivariate techniques, among which the most commonly used for hyperspectral data is partial least-squares regression (PLSR; Burnett et al. 2021).

Radiative transfer models (RTMs) like PROSPECT (Jacquemoud & Baret 1990) are based on physical principles which makes them generally valid for the broad class of leaves that come close to meeting their assumptions, and the approach can be extended to the canopy scale (e.g., PROSAIL; Jacquemoud et al. 2009). In PROSPECT, leaves are represented as one or several absorbing plates with rough surfaces giving rise to isotropic scattering. The model uses two classes of input variables: a leaf structure parameter representing the average number of air/cell wall interfaces within the mesophyll, and the contents of leaf biochemicals. However, only a handful of leaf traits have well-defined absorption features that influence spectral reflectance in a strong, direct, and easily characterized way. Prominent examples are pigments, dry matter content, and water content, which are incorporated in PROSPECT and can be directly estimated. In addition, the newest generation of PROSPECT splits dry matter into protein and carbon-based constituents (Féret et al. 2021).

It is highly unlikely, though, that the number of leaf traits that can be estimated from physical models will increase dramatically in the future. This is because for most biochemical traits, absorption features in the 400–2500 nm range measured by classic spectrometers are actually harmonics or overtones of fundamental features in the UV (10–400 nm) and middle-infrared (2500–6000 nm) ranges (Ustin & Jacquemoud 2020). In addition, as countless leaf characteristics, including both chemical and structural, influence reflectance in some way and their effects overlap, it is still impossible to tease apart the specific effects of individual constituents with more subtle absorption features.

This limitation of RTMs is a major motivation behind the use of empirical machine learning techniques like PLSR. PLSR was specifically designed to reduce the severity of assumptions of multivariate linear regression (Martens 2001; Wold et al. 1983). PLSR's main purpose is prediction, it does not depend on a clear, mechanistic understanding of the relationships among the dependent variables (Y) and the independent variables (X), and it allows the use of many, highly collinear predictors. These qualities make it well suited for spectral data, which are inherently multicollinear. From the original X (i.e. the sample × wavelength matrix), PLSR calculates component vectors oriented to maximize the total variance explained in X and Y and the covariance between X and Y. Models created using PLSR or similar algorithms have long been used to quantify chemical components in the pharmaceutical and agricultural sectors, and have in many cases become routine algorithms readily implemented in proprietary instrument software (Marten et al. 1989). In vegetation spectroscopy, PLSR has been successfully applied to predict leaf mass per area (LMA), the contents of N, carbon (C), carbon fractions (lignin, cellulose, hemicellulose, non-structural carbohydrates (NSCs)), phosphorus (P), potassium (K), calcium (Ca), magnesium (Mg) and pigment composition (chlorophylls and carotenoids) at the leaf level (Serbin et al. 2014, Schweiger et al. 2018) and scaled up to the canopy level (Asner et al. 2014, Singh et al. 2015, Wang et al. 2020).

2.2 Trait covariance and trait estimation

Trait mapping using PLSR models on imaging spectroscopy data is seeing growing use as a method to infer community or ecosystem processes over large scales (Asner & Vitousek 2005; Chadwick & Asner 2018; Jucker et al. 2018)—so it is important to know when trait models return accurate enough estimates, and when they instead break down. At times, PLSR and other empirical approaches to predicting traits from spectra may seem unreasonably effective. It seems apt enough that we can accurately predict chemical or structural traits that have strong absorption features, like LMA or water content, at the leaf and canopy scales (Asner et al. 2011; Serbin et al. 2019; Wang et al. 2020). But leaf-level PLSR models can let us predict (with varying accuracy) nutrients or isotopes (e.g., K, Ca, Mg; $\delta^{15}N$) that have little direct, measurable impact on leaf absorption within the VIS to SWIR range (Asner et al. 2011; Nunes et al. 2017; Kothari et al. 2021; Wang et al. 2020). Similarly, PLSR models applied to remotely sensed imagery can yield estimates of non-leaf traits like wood density (Jucker et al. 2018), or even make reasonably accurate predictions of forest plots' dominant mycorrhizal associations (Fisher et al. 2016; Sousa et al. 2021). The apparent success of these models, however modest, creates a puzzle: how could these models be used to predict traits that can't possibly have a direct effect on the spectrum?

A likely solution to this puzzle is that model-building algorithms leverage the covariance of these unobservable traits with traits that do directly influence spectral properties. For example, wood density tends to correlate with leaf or whole-canopy traits like LMA or total leaf area (Chave et al. 2009; Mencuccini et al. 2019), which influence reflectance spectra more directly. Drawing on Chadwick & Asner (2016), Nunes et al. (2017) coined the term "constellation effect" to describe how empirical models may predict a target trait by leveraging its covariance with a constellation of other traits, a phenomenon that can be confirmed using synthetic data (Figure 2). This sort of effect may be most conspicuous among traits that have no direct influence on the spectrum. However, it may also affect models for traits that do have absorption features of their own, as long as they also covary with other traits that have strong absorption features.

Ecologists often expect traits to covary with each other in ways that emerge from physical principles and evolutionary constraints (Cavender-Bares et al. 2020). Particular suites of correlated traits are given names like the leaf economics spectrum (Díaz et al. 2016; Wright et al. 2004), the wood economics spectrum (Chave et al. 2009) and Corner's rules (Corner 1949). Reich (2014) provides a broad overview of trait covariance within and across plant organs. Many of the foundational papers on trait coordination are based on analyses of global trait databases, and have validated the reliability of these correlations at global scales (Chave et al. 2009; Díaz et al. 2016; Joswig et al. 2022; Wright et al. 2004). However, these patterns need not hold at smaller phylogenetic or spatial scales—and indeed, trait correlations often become more variable or even reverse signs at finer taxonomic scales (Anderegg et al. 2018; Osnas et al. 2018; Zhou et al. 2022). Biogeographic origins can also alter the relationships between traits (Heberling et al. 2012). In addition, although the leaf economics spectrum and other such one-dimensional axes have received great attention, real networks of trait covariance can be more complex and diffuse, including traits beyond those most commonly measured (Wei et al. 2017).

The potential ubiquity and scale-dependence of constellation effects raises the question of what attitude ecologists should take towards models that may rely on them. This problem has analogues in other domains: for example, statistical models of disease risk from genomic data that work well within a population often transfer poorly among populations. Mathieson (2021) argues that this phenomenon emerges because the models leverage not just ‘core’ genes with a direct causal effect, but also ‘peripheral’ genes whose influence is contingently mediated by the core genes. We might likewise think of traits as lying along a continuum from ‘core’ to ‘peripheral,’ where core traits are those that have a strong, direct influence on the reflectance spectrum—which often allows them to be included in RTMs—as well as strong correlations with other traits. Associations between the spectrum and peripheral traits are mediated (at least in part) by correlations between core and peripheral traits (Figure 2). Even when the relationship between peripheral and core traits seems empirically reliable, it may be hard to quell some lingering doubt about using trait estimates from spectral models that rely on constellation effects.

The range of possible attitudes may be illustrated by a case study: Ollinger et al. (2008) reported strong, positive correlations among stand-level foliar N concentration, carbon assimilation, and NIR reflectance across a range of temperate and boreal forests. Re-examining the relationship, Knyazikhin et al. (2013) argued that a positive correlation between foliar N and NIR reflectance made little sense in terms of the physics of leaf-level radiative transfer. Instead, it arose as an artifact of the disparate canopy structures and leaf surface characteristics of N-poor conifer and N-rich broadleaf trees, which varied in relative abundance across the stands. The question arises: If the strong correlation between foliar N and NIR reflectance is induced by canopy structure, should the models still be used to predict foliar N or carbon assimilation?

Knyazikhin et al. (2013) proposed that attempts to retrieve leaf chemistry using remote sensing might generally be confounded by canopy structure and leaf-level albedo, and expressed skepticism at empirical approaches that are not heavily reinforced by radiative transfer modeling. In a comment, Townsend et al. (2013) argued two points. First, increased spectral range and resolution could improve direct empirical estimation of biochemical traits that, for many of the reasons we discussed earlier, are not amenable to radiative transfer modeling. Second, even when a relationship is indirect and mediated by trait covariance, it need not be considered spurious. In this case, they argued that correlations between leaf biochemistry, leaf structure, and canopy structure arise from well-described and ostensibly reliable physical and evolutionary constraints that researchers can exploit for trait mapping. One might argue that trying to avoid taking advantage of these correlations can result in worse predictions.

2.3 When are empirical models good enough?

Regardless of the merits of different modeling approaches, we can ask *when* empirical models that leverage trait covariance are sound enough to use, given that relationships may shift across regions or taxonomic scales. One might suspect that the relationship between foliar N and NIR reflectance that holds across North America’s temperate deciduous and boreal forests might do a poor job of predicting N within deciduous forests alone, or in predicting N 'out-of-sample' in other biomes—although in this case, the relationship tends to be quite robust (Hollinger et al. 2010; Wicklein et al. 2012). In general, dramatic differences between the range of environmental conditions of the training data and those of the testing data carry some risk of inaccurate or biased estimation (Schweiger et al. 2020). Best practices for predictive modeling include covering the range of values of the component(s) of interest, as well as the species, phenological stages, growth forms and light environments for which predictions shall be made (Burnett et al. 2021; Schweiger et al. 2020). Notably, although the transferability of empirical models has generally been considered limited, recent success in calibrating accurate continental scale PLSR models for a range of traits (Serbin et al. 2019; Wang et al. 2020) suggests that these models can be applied across wide geographic ranges given that the training data covers these ranges adequately. At the same time, it seems reasonable to expect that models that are specifically calibrated for local sites might outperform global models, but this still needs to be assessed. In any event, collecting external validation data to assess on-site performance of empirical models is good practice (Burnett et al. 2021).

Besides asking when empirical models are valid, we can also ask if there are a few ‘core’ traits whose clear influence on optical properties and coordination with other ‘peripheral’ traits makes it possible to estimate those other traits—and if so, what are the core traits? Determining what directly drives a multivariate empirical model, such as in PLSR, requires careful interpretation, and hints can come from multiple sources. Interpretive tools like plots of model coefficients or the variable influence in projection (VIP) metric for PLSR (Wold et al. 2001) offer heuristics for gauging which bands help predict a given trait. Bands of high importance can be checked against known absorption features of specific constituents (e.g., Curran 1989) for interpretation. For example, across studies, many traits show high VIP across the green hump and the red edge (Ely et al. 2019; Kothari et al. 2021; Yan et al. 2021), which suggests that their estimation may be aided by their covariance with pigment contents or leaf structure. Nevertheless, assigning unambiguous interpretations to these patterns is often challenging because many components’ features are shallow and overlapping.

We can also ask, when is the relationship between core and peripheral traits robust enough to estimate peripheral traits reliably? Here, looking at empirical patterns of trait covariance and examining model performance across a range of ecosystems and species may be useful. If, for instance, an empirical model shows good performance across a wide range of species but poor performance within subgroups (e.g., specific ecosystems, functional groups or species), it may suggest that the model relies on global trait covariance patterns that weaken at finer scales (Figure 2c). Poor performance on independent validation datasets may have similar implications. One interesting case study is Meacham-Hensold et al. (2019), who used transgenic lines to break the standard positive correlation between leaf N and photosynthetic parameters, but showed that PLSR models could still predict the latter. In addition, simulations coupling radiative transfer models with synthetic trait datasets may elucidate when or how well trait covariance networks propagate trait-spectral relationships to more 'peripheral' traits (Figure 2).

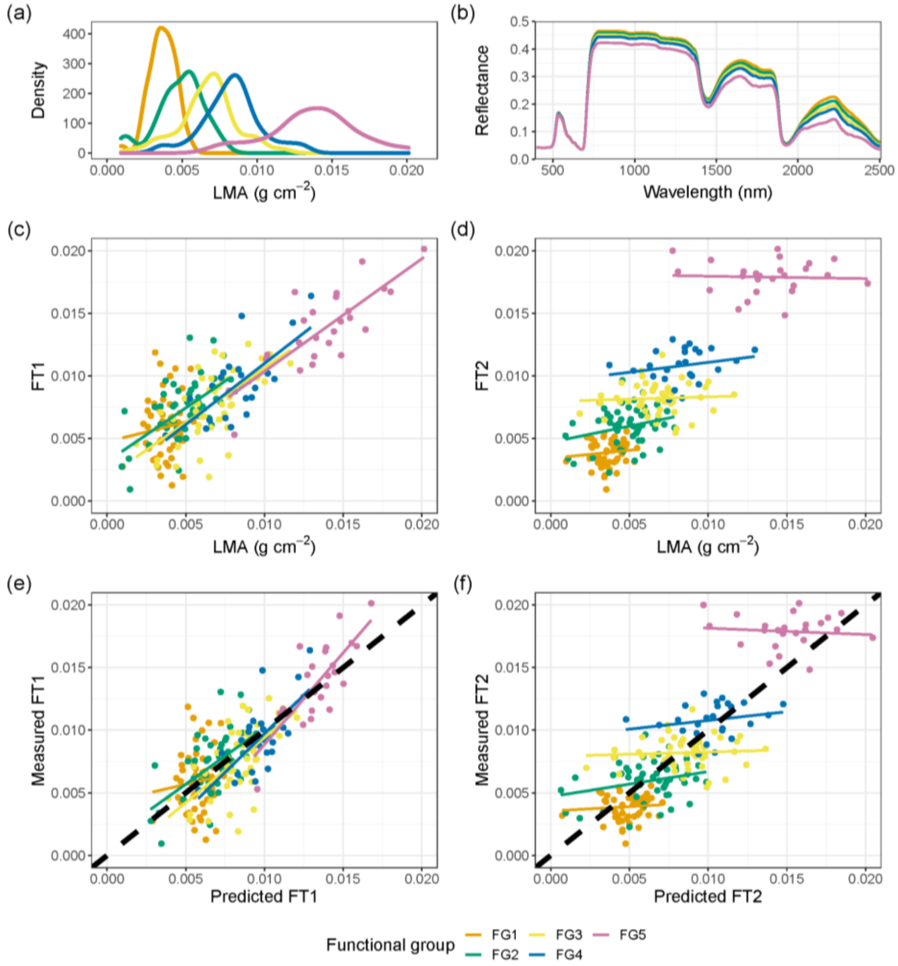

Figure 2. We can illustrate the role of covariance in trait estimation using a simple synthetic dataset. (a) We designated five functional groups (FG1-5) with different but overlapping distributions of LMA. (b) We simulated 200 spectra using PROSPECT-D (Féret et al. 2017), holding all parameters constant except LMA (equivalent to PROSPECT parameter C_m). We then simulated two other functional traits: (c) FT1 was generated by adding random noise to LMA; (d) FT2 was generated such that it correlates with LMA across but not within functional groups. While both traits are positively correlated with LMA (FT1 R^2 = 0.628; FT2 R^2 = 0.716), neither

has any direct influence on the spectrum. (e and f) Training a PLSR model on the simulated spectra with 10-fold cross-validation, we can estimate LMA (R^2 = 1; not shown), FT1 (R^2 = 0.609), and FT2 (R^2 = 0.690) across all functional groups. But while estimates of FT1 are correlated with their true values *within* functional groups, estimates of FT2 are not, because LMA and FT2 are uncorrelated within functional groups. The thick black dashed line is the 1:1 line.

These considerations about constellation effects illustrate that empirical models may often take advantage of trait covariance, especially for estimating traits that only have subtle influences on the spectrum. More generally, they also underscore the complexity of the relationships between the spectrum and traits that have varying degrees of influence on it. Traits each influence reflectance at many wavelength bands, and each band is also influenced by the complex physical interactions among many traits. These phenomena give rise to the covariance among bands (which is likely further strengthened by covariance among traits) that produces the distinctive shape of spectra from green leaves (Figure 2b). The very complexity of these interactions can make it challenging to disentangle all the information in the spectrum into a discrete set of traits. The integrative nature of reflectance spectra motivates uses that are not dependent on estimating standard traits, but use spectra in their own right.

3. Beyond traits: Treating spectra as their own entity to capture plant phenotypes

3.1 Optical types: More than the sums of plant traits?

In most cases, the spectrum of a plant is treated as an epiphenomenon of some set of underlying traits, which are the true determinants of fitness. However, there are some circumstances under which the spectrum itself may be treated as a trait subject to ecological or evolutionary selection. Naturally, absorptance in the VIS range corresponds to light harvest by photosynthetic pigments. Absorptance is also part of the leaf energy balance—a consideration that motivated some of the first studies of leaf optical properties (Billings & Morris 1951; Shull 1929). This line of research was continued by later researchers, often with a focus on avoidance of high leaf temperature (Blonder et al. 2020; Ehleringer et al. 1981; Mooney et al. 1977)—and even extended to flowers (Roddy 2019). Another context in which spectral properties themselves may or may not be adaptive is among plant-animal interactions, where both the reflectance of plant tissues and the spectral sensitivity of the animals' eyes matter. For example, the appearance of a plant to potential herbivores may influence susceptibility to herbivory (Kemp & Ellis 2019; Strauss & Cacho 2013). Competition or mutualism related to pollinator attraction may also influence the co-occurrence of flowers with different optical properties (van der Kooi et al. 2016). In such cases, the change in average spectral properties or spectral diversity through time or space might indicate which environmental pressures shape those properties.

Even when the spectrum itself does not directly influence fitness, another reason to use spectra directly, rather than estimating traits, to make ecological inferences, is that they integrate many aspects of plant form and function into one single measurement (Cavender-Bares et al. 2017). But what does this mean? If plants are more than the sum of their traits, how can we get a handle on the contributions to plant function of traits that are commonly measured vs more "fuzzy" traits that might not even yet have a name? This brings us to questions frequently asked in functional ecology: What are the "correct" traits for investigating specific ecological phenomena? What is the importance of traits that are not associated with a major, well-described axis of trait variation? And are we even able to define and measure all traits that matter for plant life? While spectroscopy will not be able to answer these questions directly, it can reduce barriers to investigating them further. For example, a large share of effort in plant functional ecology goes toward measuring a few traits (Kattge et al. 2020) which are relatively cheap and straightforward to measure and appear to be of general importance across plant communities worldwide (Westoby 1998). However, recent work suggests that the effect of functional traits on fitness is often multidimensional (Kraft et al. 2015) and dependent on local context (Blonder et al. 2018); in many cases, aspects of function that are less well-studied or more challenging to measure may be important for understanding the role of plants in their environment. Spectral measurements, which generally take less than a minute to complete, allow lots of samples to be measured in little time, which lowers the barrier to incorporating intraspecific variation, an important yet often neglected aspect in community ecology (Violle 2012). In addition, spectra can help us incorporate a broader set of plant functions than we might otherwise be able to, especially seldom-studied or difficult-to-measure traits. These may include a vast variety of defense compounds (Couture et al. 2016), cuticular waxes, leaf hairs (Ehleringer et al. 1981), and anatomical traits like mesophyll structure (Karabourniotis et al. 2021), and when measured remotely leaf inclination angles and branching structure. Many of these traits are not part of the standard repertoire of trait-based ecology (Pérez-Harguindeguy et al. 2013) and may be hard to characterize in a consistent way across the plant tree of life.

The optical type concept (Ustin & Gamon 2010) is central to using spectra of plants as their own entities. It posits that, since plants partition resources, including light, nutrients and water which all influence the spectral response, optical types can be delineated along resource use axes that can be measured spectrally and capture plant functional identity along a continuum. This concept gave rise to a series of studies using spectral diversity or the dissimilarity among plant or image spectra as means to capture plant diversity (Draper et al. 2019; Féret & Asner 2014; Gholizadeh et al. 2018; Rocchini et al. 2010; Wang & Gamon 2019). While the main axes of spectral variation, which might include resource use axes, have not yet been fully described (Section 4.3) it has been shown that spectral dissimilarity among plant species is associated with their functional dissimilarity and evolutionary divergence time (Frye et al. 2021; Schweiger et al. 2018). The association between plant spectral, functional and phylogenetic distance is currently best understood at the leaf level (Frye et al. 2021; Meireles et al. 2020; Schweiger et al. 2018). Although leaf spectra do not capture all traits that shape how plants interact with their environment, they do integrate many plant traits that are important for resource capture and stress tolerance, including the contents of pigments, water and leaf structure. Through trait covariance, they may also indirectly capture other leaf traits like macro- and micronutrients, as well as traits of other organs like flowers, stems, roots, and seeds. One advantage of using spectra as their own entity is that they incorporate more of the total variation in function associated with leaf chemistry, anatomy and morphology than what is commonly captured by standard traits, including variation that is difficult to measure or may be of unrecognized importance (Schweiger et al. 2018).

At the remote sensing level, variation in spectra captures differences in plant growth form, leaf orientation and plant architecture as well as leaf traits. These structural canopy characteristics are difficult to quantify spectrally and active remote sensing techniques, including radar (radio detection and ranging) and lidar (light detection and ranging) are much better suited to derive indicators for vegetation structure (Antonarakis et al. 2011; Bergen et al. 2009; Lenoir et al. 2022). Nevertheless, plant growth form, leaf orientation and canopy architecture create illumination patterns that are influenced by the spatial distribution of light and shade, and proportions of leaf tissue and bark (Gower et al. 1999; Kuusinen et al. 2021). Potentially, these illumination patterns influence the spectral signal in a way that contributes to spectral dissimilarities among species, broader taxonomic or functional groups. Studies partitioning the contributions of leaf traits and canopy structure to spectral diversity or species differentiation are needed to clarify the degree to which effects of canopy structure on spectra benefit or hamper plant diversity assessments using remotely sensed spectra.

3.2 The spatial organization of plant spectral diversity

In most studies using imaging spectroscopy, image pixels exceed the size of individual plants, which means that individuals from other plant species and non-vegetated areas can add to the spectral signal. In these circumstances it makes sense to consider the spatial organization of diversity and the sources of spectral information contained in a pixel. The pixel size of imaging spectroscopy data depends on multiple factors, including flying altitude and speed. While the pixel size of airborne sensors (m-scale) might allow capturing individual trees in a mature forest, the same pixel size captures plant communities in grassland ecosystems. Therefore, while it might make sense to calculate spectral alpha-diversity from image pixels acquired with airborne imaging spectroscopy as a means to estimate plant alpha-diversity in forests, the same metric might not make sense in grassland ecosystems, and one might use spectral beta-diversity to capture differences in plant community composition instead (Schweiger & Laliberté 2022). Spectral diversity can be separated into its spatial components, spectral alpha-, beta- and gamma-diversity (Laliberte et al. 2020), for example through variance partitioning (Legendre & De Cáceres 2013). This makes it possible to calculate the local contributions of individual image pixels (or communities) and the contributions of individual spectral features (or bands) to the spectral diversity, which can help illuminate the spatial and spectral sources of diversity patterns. If, especially in the context of beta-diversity assessments, the spectral contributions of non-vegetated areas should be viewed as adding to the distinctiveness of plant communities or as "noise" is, however, not yet clear.

The spatial arrangement of spectral variation can be useful for investigating community assembly or ecosystem function. Ecologists are often keen to test hypotheses about processes from patterns of functional traits or functional diversity through time or space; such hypotheses could likewise be examined through the lens of spectral identity (i.e., position in spectral space) or diversity. For example, Smith et al. (2013) argued that functional diversity-area (FAR) relationships could be used to make inferences about trait-based assembly mechanisms; this method was later extended continuously over landscapes using trait estimates derived from imaging spectroscopy (Durán et al. 2019; Schneider et al. 2017). Inspired by the FAR approach, Dahlin (2016) constructed spectral diversity-area relationships using imaging spectroscopy to consider the roles of environmental filtering and stochastic drift in a mixed agricultural-forest landscape. Draper et al. (2019) and Bongalov et al. (2019) have since addressed similar questions by comparing the roles of the environment and geographic distance in explaining spectral beta-diversity within highly diverse tropical rainforests. These studies reveal the potential for spectral data to shed light on core issues in community assembly over large scales—and they are only possible because spectral dissimilarity or diversity can act as a surrogate for functional dissimilarity or diversity.

3.3 Spectra as integrated measures of plant phenotypes

Studies that have used spectra as their own entity also include investigations into resource partitioning. Schweiger et al. (2018; 2021) have found that individual plant growth as well as community productivity can be predicted from the spectral space an individual plant or a plant community occupies. This means that individuals and communities that show more spectral variation grow more than individuals and communities that show less spectral variation. The authors propose that differences in resource use strategies, in particular in light use, might be responsible for the large hypervolumes occupied by productive individuals, species and communities. Including additional measurements of traits such as pigment composition, LMA and N content, would provide an opportunity to test this assumption.

Using plant spectra in community ecology can provide an opportunity to differentiate plant characteristics that matter under particular circumstances into those that are captured by specific sets of traits and those that remain “unmeasured” but are captured spectrally. One strategy could be partitioning the total variance of the ecological phenomenon of interest into two fractions, one that is explained by measured traits and one that is explained by spectra alone. Examining the wavelengths that contribute most to the explanatory power of the spectral component could allow drawing inferences about unmeasured traits that helped explain the ecological phenomenon of interest. Another strategy could be to investigate the main axes of spectral variation together with the traits that are associated with them. Through the successive inclusion of additional sets of traits, it might become possible to decipher which characteristics of plants are captured spectrally. It is likely though that spectra capture characteristics of plants yet to be named.

While this approach of treating spectra as a standalone entity can be powerful, it does come with challenges. A change in mean spectral identity or diversity is not as immediately interpretable in biological terms as functional identity or diversity. Using spectra on their own also makes it harder to control the weights assigned to various aspects of plant function, as one might aim to do when calculating functional diversity (Laliberté & Legendre 2010). At the leaf level, in most cases, the greatest share of spectral variation is due to overall NIR reflectance and the depth of water absorption bands in the SWIR range, which are controlled mainly by factors like leaf surface characteristics, mesophyll structure, LMA, and water content. At the canopy level and after accounting for illumination differences, leaf area and leaf angle distribution might be the dominant factors explaining spectral variation, followed by the foliar traits mentioned above. But these traits may not always be commensurately important for fitness; in some cases, the spectral features associated with aspects of plant function that matter most may account for a relatively small share of total spectral variation. In other words, although using spectra on their own may allow ecologists to take advantage of more phenotypic information than the most common plant traits, it may also not allow as much control over what information exactly is included. Methods to bring out the most salient information from spectra could be a useful topic for future research.

4. Dimensionality of spectra

As we have emphasized, plant reflectance spectra are powerful tools because they integrate many aspects of plant function into a single simple measurement. Much of this review is concerned with the question, "*what* and *how much* information about plant function is contained in the reflectance spectrum?" This question is important because the power of a trait dataset to explain or predict patterns in community ecology depends on its *intrinsic dimensionality*—the number of parameters needed to account for the dataset's properties (Laughlin 2014). The logic of this claim is that each axis of phenotypic variation (or cluster of correlated traits) can be the subject of selection by different environmental filters, so including more independent axes means that predictions can take into account the effects of more filtering variables. For efficiency's sake, a researcher might want to design a trait measurement campaign to include as many independent dimensions with as few measurements as possible (Laughlin 2014). Spectra might serve this goal—but only if the phenotypic information they contain is relevant to the environmental filters that may be acting in a given community, which may not be known in advance.

Spectral datasets are made up of a much larger number of variables (wavelength bands) than the vast majority of trait datasets, but these bands are often highly correlated—both because any given trait influences multiple bands and because traits covary. Radiative transfer models like PROSPECT can generate much of the variation found in real spectra with just a few variables (eight in PROSPECT-PRO; Féret et al. 2021), but are by necessity simplifications; the intrinsic dimensionality of spectral data has seldom if ever been explored empirically. Here, we use three example analyses to take a look at this question, with the hope that they will inspire more comprehensive investigations into the intrinsic dimensionality of spectra in the future.

4.1 Species differentiation in spectral vs trait space

To illustrate the degree of species differentiation in spectral and trait space, we use leaf spectra and trait data of 902 individuals from 14 grassland–savanna perennials sampled in 35 plots of the Cedar Creek Biodiversity (BioDIV) experiment (see Supplementary Methods). The traits included are foliar C, N, NSC, hemicellulose, cellulose, and lignin concentration (%), and the content of chlorophyll a + b (µmol m^{-2}) and ratios relative to chlorophyll content of beta-carotene, lutein, and xanthophyll (violaxanthin + antheraxanthin + zeaxanthin) pigments. We normalized each trait to mean 0 and standard deviation 1 across the dataset. We used linear discriminant analysis (LDA) to illustrate species dissimilarity in spectral and trait space, and we tested the degree to which plant species can be correctly identified based on spectra and traits using partial least squares discriminant analysis (PLSDA).

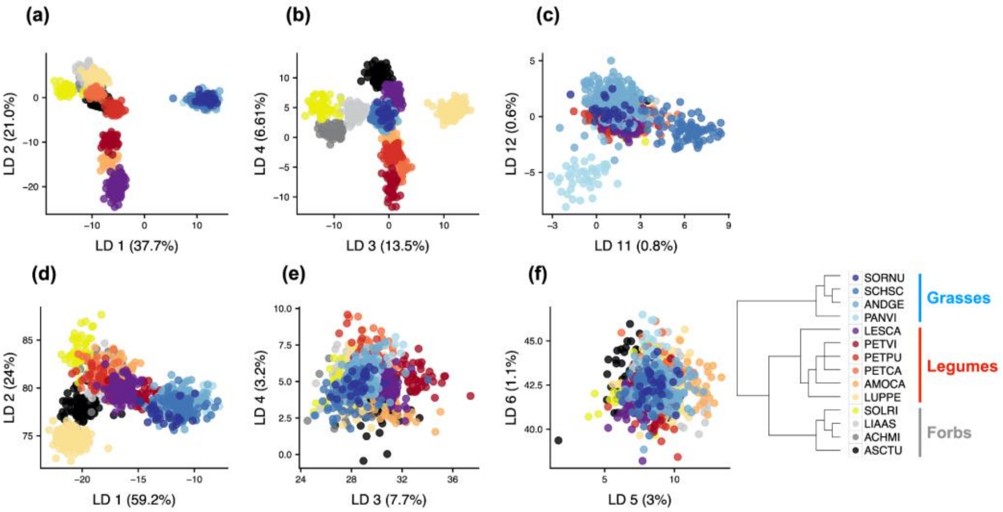

Figure 3. Species clustering along linear discriminant axes (LDs) maximizes the differences among species based on (a-c) spectra and (d-f) traits. The amount of the total variation explained by each LD is shown in parentheses; for species abbreviations and number of individuals per species see Supplementary methods. Species phylogenetic relationships and major functional groups are shown on the right.

Visual inspection of LDA results revealed that species were more distinct in spectral than in trait space (Figure 3). In spectral space, all non-graminoid species clearly separated along the first four LDs (Figure 3a, b), and LDs 11 and 12 separated the graminoids (Figure 3c). In trait space, however, only a few species formed distinct clusters, and we found no combination of LDs that separated the four graminoids species from each other (Figure 3d-f). Likewise, species identification models based on spectra (accuracy = 93 %, Kappa = 0.81) consistently outperformed species identification models based on traits (overall accuracy = 66 %, Kappa = 0.63; Figure 4).

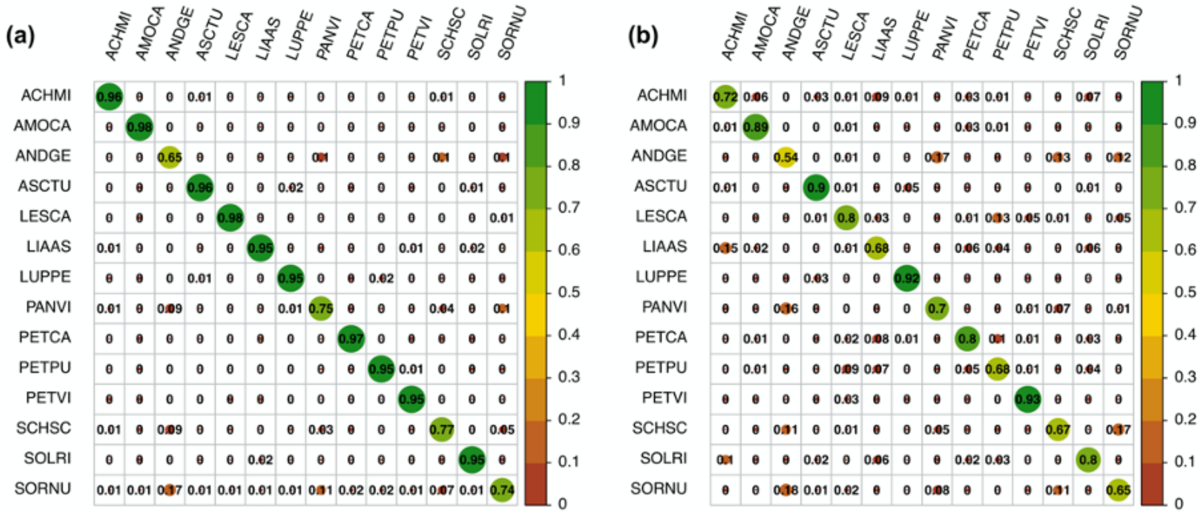

Figure 4. Species identification models. Confusion tables for PLSDA (partial least squares discriminant analysis) models showing the proportion of correctly identified (diagonal) and misidentified (off-diagonal) species based on (a) leaf spectra and (b) traits. For species abbreviations and number of individuals per species see Supplementary Methods.

To some extent, better separability of species based on spectra could be due to redundancy in the traits we measured. In our case, light gradients are probably the dominant source of environmental variation, and all leaf traits measured in our study are to some degree influenced by variation in light. For instance, the ratio of chlorophyll and carotenoid pigment levels reflects biochemical acclimation to stress under different light environments (Gamon & Berry 2012; Kothari et al. 2018). Likewise, the contents of different carbon fractions are tied to morphological adaptations, such as leaf thickness and SLA, to light gradients (Niinemets 2007). In this way, what we think of as multiple traits can also be thought of as different proxies for the same or overlapping traits (Gamon et al. 2019). Although leaf spectra do not capture all traits—in particular the traits of organs besides leaves (Schweiger et al. 2018)—they seemed to capture differences in leaf chemistry, structure and morphology among species more completely than the traits we measured. Interestingly, LDs 11 and 12 which separated the graminoid species each contributed less than 1% to the total variation in leaf spectra, while LDs 1-4 which separated the other species accounted for more than 78%. This highlights that minor axes of spectral variation can bring out important but subtle differences in species' foliar characteristics. Hypotheses regarding which foliar traits might be contributing most to LDs 11 and 12, and thus to leaf-level differences among the four graminoid species in our study, would be interesting to test. However, this would require additional trait measurements which in this case we do not have at hand. The spectral bands contributing most to species' separability along the first LD axes aligned with regions in the spectrum indicative of leaf chlorophyll, carotenoid, lignin and protein content (Figure S1a). These foliar traits also contributed most to species separability in functional trait space (Figure S1b), and all except chlorophyll content showed evidence of phylogenetic signal (Table S1), indicating that these traits might contribute to species identification across ecosystems.

4.2. Dimensionality analysis

Next, we attempted to quantify the dimensionality of spectra. We used data from the Canadian Airborne Biodiversity Observatory (CABO) comprising samples from a diverse array of ecosystems across temperate and boreal Canada, as well as one site in Australia. The traits and fresh leaf-level reflectance spectra of these samples were measured (the latter with an integrating sphere) using consistent protocols. We took all CABO data collected until 2019 (n = 1971) and subset them to include at most ten samples per species, chosen at random (leaving n = 905; species = 106; see Supplementary Methods).

We considered the following nine foliar traits: equivalent water thickness (EWT), LMA, C, N, hemicellulose, cellulose, lignin, total chlorophyll, and total carotenoids. We converted all traits besides EWT and LMA to a normalization-independent basis (*sensu* Osnas et al. 2018) to remove their statistical dependence on LMA. We also normalized the reflectance at each wavelength to mean 0 and standard deviation 1 to avoid placing emphasis on bands that have more absolute variation in reflectance. To remove highly correlated wavelengths, we subset the wavelength bands to every 20 nm from 400 to 780 nm, every 50 nm from 800 to 1350 nm, and every 25 nm from 1400 to 2400 nm, leaving 73 bands.

Researchers have often found that leaf chemical trait estimates from ground-leaf spectra are better than those from fresh-leaf spectra (Couture et al. 2016; Kothari et al. 2021; Wang et al. 2020). On a subset of CABO samples (*n* = 619), we had also measured spectra after the samples were pressed like herbarium specimens, and again after they were dried and ground (Kothari et al. 2021). We again subset these to no more than ten samples per species (*n* = 228; species = 66) and normalized and subsampled bands. Lastly, as a benchmark for our estimation methods, we used PROSPECT-D (Féret et al. 2008) to generate a synthetic dataset of 1000 spectra with a known dimensionality of 4 by independently varying leaf structure, chlorophyll, water, and dry matter content. We normalized and subsampled bands as in the real data.

We sought to describe and compare the intrinsic dimensionality of the spectral and trait datasets. Inspired by Laughlin (2014), we tried a number of techniques for estimating dimensionality (Table 1). Rather than attempt to reach a single, precise estimate of the ‘true’ dimensionality, we aimed to see what patterns emerged from the ensemble of techniques. Most of these techniques are non-linear, meaning that they can account for non-linear relationships among input variables like bands or traits.

The PROSPECT-generated dataset with a known dimensionality of 4 had estimated dimensionality between 3 and 4, which suggests that the methods we chose can capture the dimensionality of spectral data, perhaps with some tendency towards underestimation. For both the measured spectral and trait datasets, most methods returned a dimensionality of about 4-5. Laughlin (2014) showed that large whole-plant trait databases tended to return a dimensionality of 4-6 using many of the same methods. The fact that leaf-level spectra come close to (and sometimes exceed) the dimensionality of large trait databases suggests that they have great promise for explaining and predicting community assembly.

On the subset from which fresh-, pressed-, and ground-leaf spectra were all measured, fresh- and pressed-leaf spectra often had similar dimensionality (mainly 3.5-4.5), but ground-leaf spectra were considerably higher on average (mainly 4-6). We suspect this finding results from the tendency for water and leaf structure to obscure or overwhelm small absorption features of specific dry matter constituents in the SWIR range (Elvidge 1990; Peterson 1988). But given that water and leaf or canopy structure are unavoidable in remote sensing—and indeed, important for plant function—this finding may have limited relevance for inferring ecological processes over large scales.

Table 1. Estimated intrinsic dimensionality using a variety of techniques on data from the Canadian Airborne Biodiversity Observatory (CABO).

	n = 905		*n* = 228				
	Spectra	Traits	Fresh	Pressed	Ground	PROSPECT	Citation
Correlation integral	4.7	4.2	4.1	4.4	5.9	3.5	Grassberger & Procaccia (1983)
Maximum likelihood	4.4	4.5	3.8	4	4.9	3.6	Levina & Bickel (2004)
Manifold-adaptive nearest neighbor	4.3	5.0	4.5	4.4	5.8	3.4	Farahmand et al. (2007)
Non-iterative nearest neighbor	3.9	4.3	3.3	2.6	4.1	3.3	Pettis et al. (1979)
Minimum neighbor distance–maximum likelihood	4.7	4.7	4.2	4.4	5.2	3.7	Rozza et al. (2012)
PCA scree test	5	5	5	4	6	4	Cattell (1966)
Isomap scree test	3	3	3	4	3	3	Tenenbaum et al. (2000)

4.3 Leading axes of spectral variation

We emphasize above that even when using spectra directly to address ecological questions, we should ideally ground our conclusions in the underlying biology. Given an estimate of spectral dimensionality, it seems natural to wonder what aspects of plant function these dimensions correspond to. We took the CABO dataset from above (*n* = 905), with spectral bands z-standardized and downsampled, and used the manifold learning technique Isomap (Tenenbaum et al. 2000) to visualize the major dimensions of leaf spectral variation and their relationship with traits. Isomap uses a nearest neighbor-based algorithm to approximate the geodesic distances among points on a curved manifold, then applies classical multidimensional scaling (MDS) on the distance matrix. Because the relationship between spectral bands is often non-linear, Isomap may be better suited than linear methods to discovering the underlying structure of spectral data.

Much like MDS, the orientation of Isomap output is arbitrary. Here, we rotated the Isomap coordinates using Procrustes analysis to maximize its similarity to the trait dataset. We caution that, while the Procrustes analysis helps us interpret spectral variation in terms of the traits in our dataset, it may obscure dimensions of spectral variation that correspond to unmeasured traits. We focus on the first four dimensions after rotation, following the estimates in Table 1, but further dimensions may still have functional importance. We performed all Isomap-based analyses with *R* package *vegan v. 2.5.5* (Oksanen et al. 2019).

The Isomap analysis helps to visualize the spectral differences among functional groups (Figure 5). Conifers occupy much of the quadrant high on both Axes 1 and 2, while many of the samples high on Axis 1 but not 2 are forbs or wetland graminoids. Shrubs and broadleaf trees tend to be low on both axes. EWT correlates strongly with Axis 1 and LMA with Axis 2; the conifers with high values on both axes have high EWT and LMA, while many forbs and wetland graminoids have high EWT but not LMA. Functional groups separate less along Axes 3 and 4, which correlate with a greater variety of traits but less strongly. Axis 3 is associated with high C and lignin and low EWT, while Axis 4 correlates primarily with N and pigments. Low values along Axis 3 are dominated by graminoids and forbs. These plots can give us a basic sense of the way spectral variation is distributed among functional groups and the traits that underpin it.

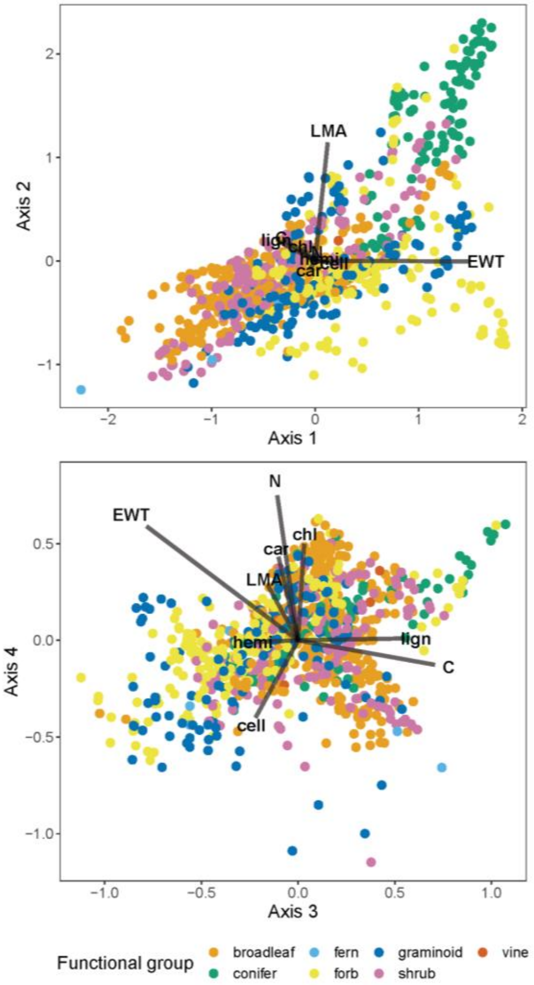

Figure 5. A visualization of CABO spectral data along four Isomap axes. Within each panel, the line segment for each trait is oriented in the direction along which it shows the greatest change within the two-dimensional space, and its relative length is proportional to the Pearson's correlation coefficient (*r*) between the trait and that direction. All traits besides LMA and EWT

were made normalization-independent (*sensu* Osnas et al. 2018). Abbreviations: hemi = hemicellulose; cell = cellulose; lign = lignin; chl = total chlorophylls; car = total carotenoids.

While the analyses in sections 4.1 and 4.2 are based on a fairly large dataset, they still represent a limited sampling of the world's floral diversity—missing, for example, all tropical biomes, drylands, and tundra. We hope these preliminary analyses provide the inspiration for future exploration with larger spectral and trait datasets. It may even be possible to carry out similar analyses using remotely sensed imagery, mapping plant traits or ecosystem functions onto the space of canopy spectral variation (Feilhauer et al. 2011). New algorithms (e.g., UMAP; McInnes et al. 2018) could allow researchers to check how their incoming spectral data fit among existing data within a better understood low-dimensional embedding, much as they can for traits (Segrestin et al. 2021). Given the right data and computational tools, ecologists could eventually describe the major axes of spectral variation, and their implications for plant strategies, just as they have described the major axes in trait variation (Díaz et al. 2016). This advance would make it easier to discern the meaning of changes in spectral identity or diversity across environmental gradients.

5. Conclusions

At the leaf and canopy scales, vegetation reflectance spectra contain an extraordinary wealth of information about the ways that plants function. For much of the early history of vegetation spectroscopy, the main way to distill this information into an interpretable form was to calculate indices or estimate traits related to specific aspects of plant function. Mainly within the last decade, researchers have begun to use the full, multidimensional spectrum itself as a tool to investigate ecological processes at the community or ecosystem scales.

In this review, we have sought to describe both the use of spectroscopy to derive plant traits as well as the potential to use spectra directly as integrated measures of plant phenotypes, and also how these two approaches are complementary. The motivations for choosing to use spectra directly can be multifaceted. In some cases, it may be motivated by a lack of suitable models for a wide range of traits, or concerns about existing models' accuracy or transferability (Dahlin 2016). It may also be motivated by the aspects of plants' function governing their impacts on or response to the environment in a given case being highly complex or uncertain. And just as there are aspects of trait variation that are not readily captured by reflectance spectra due to weak or absent optical features, there are aspects of spectral variation that have not yet been interpreted or modeled in terms of commonly measured traits. Because these aspects of spectral variation stand a good chance of being ecologically meaningful, spectra can complement standard traits in assessing the ways plants impact and respond to their environment.

With further methodological and conceptual improvements, both approaches could continue to develop and complement each other. Progress in the area of trait modeling and mapping may lead to ensembles of machine learning methods that allow the mapping of global trait distributions and associated uncertainties, combined perhaps with local scale models adjusted to specific site conditions. Global model repositories combined with cloud computing might allow dynamic trait maps to be produced by selecting appropriate models based on the scale of observation. Progress in the area of using spectra as their own entity may lead to better understanding regarding the separate contributions of suites of traits to spectral variation and their role in ecosystem functions and processes. A wider usage of spectral variance partitioning may also help us understand the ecological roles of rarely measured traits that affect spectra, including leaf anatomy, surface hairs and waxes, defense compounds, and—at the canopy level—leaf angle distribution.

As our examples illustrate, plant reflectance spectra serve as integrative measures of plant phenotypes that can be used to address long-standing ecological questions at the community or ecosystem scales. At the leaf scale, reflectance spectra can be measured quickly and at low marginal cost from many samples. At the canopy scale, imaging spectroscopy allows aspects of plant phenotypes to be mapped continuously over entire landscapes. These advantages could help to lower the barriers to incorporating intraspecific variation and a broader range of plant functions in our research, among other objectives.

In summary, spectroscopy of plants and vegetation is fundamentally connected to functional ecology. Using spectra to draw ecological inferences can benefit from a combination of spectral and trait analyses. We hope that the growing incorporation of spectroscopy into the standard toolkit of plant ecology will spur the advancement of ecological remote sensing—in other words, remote sensing grounded in ecological theory and praxis. Ultimately, this will allow targeted and explicit assessment and monitoring of plant biodiversity, ecosystem functions and processes, providing the basis for meaningful actions to counteract negative effects of environmental change from local to global scales.

Acknowledgements

CABO was funded through NSERC Discovery Frontiers grant 509190-2017. We thank the CABO PIs (especially Etienne Laliberté) and many researchers, including Rosalie Beauchamp-Rioux, Florence Blanchard, Anna Crofts, Sabrina Demers-Thibeault, Alizée Girard, Xavier Guilbeault-Mayers, Paul W. Hacker, and Maria Juliana Pardo Losada for coordinating the collection of CABO data with the help of countless workers. Data collection in BioDIV was supported by the NSF and NASA grants (DEB-1342872, DEB-1342823, DEB-1342778, DEB-1342827), by the Cedar Creek NSF Long-Term Ecological Research program (grant DEB-1234162), by the ASCEND Biology Integration Institute (NSF grant DBI 2021898), by iCORE/AITF (grants G224150012 and 200700172), NSERC (grant RGPIN-2015–05129) and CFI (grant 26793). We thank Brett Fredericksen, Ian Carriere and Cathleen Lapadat for help with spectral data collection and chemical analyses. Jeannine Cavender-Bares provided valuable feedback on this article. A.K.S. acknowledges the support of the Research Priority Program in Global Change and Biodiversity (URPP GCB) at the University of Zurich.

Conflict of interest

No conflict of interest to declare.

Authors’ contributions

S.K. and A.K.S. contributed equally to writing the review paper. A.K.S. conceptualized the review paper, analyzed BioDIV data and led the writing for section 4.1. S.K. analyzed CABO data and led the writing for sections 4.2 and 4.3.

Data availability statement

The BioDIV data used in section 4.1 are available from the EcoSIS Spectral Library (doi: 10.21232/hCtqmGdo) and from the Digital Repository for University of Minnesota (DRUM) (dois: 10.13020/d74j-wd55, 10.13020/j7fw-2g91, 10.13020/0476-9m60). The R code for PLSDA can be found on GitHub (https://github.com/annakat/speciesID). A portion of the CABO data used in sections 4.2 and 4.3 is available at the CABO Data Portal (https://data.caboscience.org/leaf/) and the rest will be made available upon publication. The R scripts for sections 4.2 and 4.3 can be found on GitHub (https://github.com/ShanKothari/CABO-trait-models).

References

American Society for Testing and Materials. (2006). Standard Tables for Reference Solar Spectral Irradiances: Direct Normal and Hemispherical on 37° Tilted Surface (Standard G173-06).

Anderegg, L. D. L., Berner, L. T., Badgley, G., Sethi, M. L., Law, B. E., & HilleRisLambers, J. (2018). Within-species patterns challenge our understanding of the leaf economics spectrum. Ecology Letters, 21(5), 734–744. https://doi.org/10.1111/ele.12945

Antonarakis, A. S., Saatchi, S. S., Chazdon, R. L., & Moorcroft, P. R. (2011). Using Lidar and Radar measurements to constrain predictions of forest ecosystem structure and function. Ecological Applications, 21(4), 1120–1137. https://doi.org/10.1890/10-0274.1

Asner, G. P., Brodrick, P. G., Anderson, C. B., Vaughn, N., Knapp, D. E., & Martin, R. E. (2016). Progressive forest canopy water loss during the 2012–2015 California drought. Proceedings of the National Academy of Sciences, 113(2), E249–E255. https://doi.org/10.1073/pnas.1523397113

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 Phytologist, 204(1), 127–139. https://doi.org/10.1111/nph.12895

Asner, G. P., Martin, R. E., Knapp, D. E., Tupayachi, R., Anderson, C. B., Sinca, F., Vaughn, N. R., & Llactayo, W. (2017). Airborne laser-guided imaging spectroscopy to map forest trait diversity and guide conservation. Science, 355(6323), 385–389. https://doi.org/10.1126/science.aaj1987

Asner, G. P., Martin, R. E., Knapp, D. E., Tupayachi, R., Anderson, C., Carranza, L., Martinez, P., Houcheime, M., Sinca, F., & Weiss, P. (2011). Spectroscopy of canopy chemicals in humid tropical forests. Remote Sensing of Environment, 115(12), 3587–3598. https://doi.org/10.1016/j.rse.2011.08.020

Asner, G. P., & Vitousek, P. M. (2005). Remote analysis of biological invasion and biogeochemical change. Proceedings of the National Academy of Sciences of the United States of America, 102(12), 4383–4386. https://doi.org/10.1073/pnas.0500823102

Bergen, K. M., Goetz, S. J., Dubayah, R. O., Henebry, G. M., Hunsaker, C. T., Imhoff, M. L., Nelson, R. F., Parker, G. G., & Radeloff, V. C. (2009). Remote sensing of vegetation 3-D structure for biodiversity and habitat: Review and implications for lidar and radar spaceborne missions. Journal of Geophysical Research: Biogeosciences, 114(G2). https://doi.org/10.1029/2008JG000883

Billings, W. D., & Morris, R. J. (1951). Reflection of Visible and Infrared Radiation from Leaves of Different Ecological Groups. American Journal of Botany, 38(5), 327–331. https://doi.org/10.1002/j.1537-2197.1951.tb14831.x

Blonder, B., Escobar, S., Kapás, R. E., & Michaletz, S. T. (2020). Low predictability of energy balance traits and leaf temperature metrics in desert, montane and alpine plant communities. Functional Ecology, 34(9), 1882–1897. https://doi.org/10.1111/1365-2435.13643

Blonder, B., Kapas, R. E., Dalton, R. M., Graae, B. J., Heiling, J. M., & Opedal, Ø. H. (2018). Microenvironment and functional-trait context dependence predict alpine plant community dynamics. Journal of Ecology, 106(4), 1323–1337. https://doi.org/10.1111/1365-2745.12973

Bongalov, B., Burslem, D. F. R. P., Jucker, T., Thompson, S. E. D., Rosindell, J., Swinfield, T., Nilus, R., Clewley, D., Phillips, O. L., & Coomes, D. A. (2019). Reconciling the contribution of environmental and stochastic structuring of tropical forest diversity through the lens of imaging spectroscopy. Ecology Letters, 22(10), 1608–1619. https://doi.org/10.1111/ele.13357

Burnett, A. C., Anderson, J., Davidson, K. J., Ely, K. S., Lamour, J., Li, Q., Morrison, B. D., Yang, D., Rogers, A., & Serbin, S. P. (2021). A best-practice guide to predicting plant traits from leaf-level hyperspectral data using partial least squares regression. Journal of Experimental Botany, erab295. https://doi.org/10.1093/jxb/erab295

Canham, C. D., Finzi, A. C., Pacala, S. W., & Burbank, D. H. (1994). Causes and consequences of resource heterogeneity in forests: Interspecific variation in light transmission by canopy trees. Canadian Journal of Forest Research, 24(2), 337–349. https://doi.org/10.1139/x94-046

Cattell, R. B. (1966). The Scree Test For The Number Of Factors. Multivariate Behavioral Research, 1(2), 245–276. https://doi.org/10.1207/s15327906mbr0102_10

Cavender-Bares, J., G. Fontes, C., & Pinto-Ledezma, J. (2020). Open questions in understanding the adaptive significance of plant functional trait variation within a single lineage. New Phytologist, 227(3), 659–663. https://doi.org/10.1111/nph.16652

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. American Journal of Botany, 104(7), 966–969. https://doi.org/10.3732/ajb.1700061

Cavender-Bares, J., Meireles, J. E., Couture, J. J., Kaproth, M. A., Kingdon, C. C., Singh, A., Serbin, S. P., Center, A., Zuniga, E., Pilz, G., & Townsend, P. A. (2016). Associations of Leaf Spectra with Genetic and Phylogenetic Variation in Oaks: Prospects for Remote Detection of Biodiversity. Remote Sensing, 8(3), 221. https://doi.org/10.3390/rs8030221

Cavender-Bares, J., Reich, P., Townsend, P., Banerjee, A., Butler, E., Desai, A., Gevens, A., Hobbie, S., Isbell, F., Laliberté, E., Meireles, J. E., Menninger, H., Pavlick, R., Pinto-Ledezma, J., Potter, C., Schuman, M., Springer, N., Stefanski, A., Trivedi, P., … Yang, Y. (2021). BII-Implementation: The causes and consequences of plant biodiversity across scales in a rapidly changing world. Research Ideas and Outcomes, 7, e63850. https://doi.org/10.3897/rio.7.e63850

Cavender-Bares, J., Schweiger, A. K., Gamon, J. A., Gholizadeh, H., Helzer, K., Lapadat, C., Madritch, M. D., Townsend, P. A., Wang, Z., & Hobbie, S. E. (2022). Remotely detected aboveground plant function predicts belowground processes in two prairie diversity experiments. Ecological Monographs, 92(1), e01488. https://doi.org/10.1002/ecm.1488

Chadwick, K. D., & Asner, G. P. (2016). Organismic-Scale Remote Sensing of Canopy Foliar Traits in Lowland Tropical Forests. Remote Sensing, 8(2), 87. https://doi.org/10.3390/rs8020087

Chadwick, K. D., & Asner, G. P. (2018). Landscape evolution and nutrient rejuvenation reflected in Amazon forest canopy chemistry. Ecology Letters, 21(7), 978–988. https://doi.org/10.1111/ele.12963

Chave, J., Coomes, D., Jansen, S., Lewis, S. L., Swenson, N. G., & Zanne, A. E. (2009). Towards a worldwide wood economics spectrum. Ecology Letters, 12(4), 351–366. https://doi.org/10.1111/j.1461-0248.2009.01285.x

Corner, E. J. H. (1949). The Durian Theory or the Origin of the Modern Tree. Annals of Botany, 13(52), 367–414.

Couture, J. J., Singh, A., Rubert-Nason, K. F., Serbin, S. P., Lindroth, R. L., & Townsend, P. A. (2016). Spectroscopic determination of ecologically relevant plant secondary metabolites. Methods in Ecology and Evolution, 7(11), 1402–1412. https://doi.org/10.1111/2041-210X.12596

Curran, P. J. (1989). Remote sensing of foliar chemistry. Remote Sensing of Environment, 30(3), 271–278. https://doi.org/10.1016/0034-4257(89)90069-2

Deacon, N. J., Grossman, J. J., Schweiger, A. K., Armour, I., & Cavender-Bares, J. (2017). Genetic, morphological, and spectral characterization of relictual Niobrara River hybrid aspens (Populus × smithii). American Journal of Botany, 104(12), 1878–1890. https://doi.org/10.3732/ajb.1700268

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., … Gorné, L. D. (2016). The global spectrum of plant form and function. Nature, 529(7585), 167–171. https://doi.org/10.1038/nature16489

Draper, F. C., Baraloto, C., Brodrick, P. G., Phillips, O. L., Martinez, R. V., Honorio Coronado, E. N., Baker, T. R., Zárate Gómez, R., Amasifuen Guerra, C. A., Flores, M., Garcia Villacorta, R., V. A. Fine, P., Freitas, L., Monteagudo-Mendoza, A., J. W Brienen, R., & Asner, G. P. (2019). Imaging spectroscopy predicts variable distance decay across contrasting Amazonian tree communities. Journal of Ecology, 107(2), 696–710. https://doi.org/10.1111/1365-2745.13067

Durán, S. M., Martin, R. E., Díaz, S., Maitner, B. S., Malhi, Y., Salinas, N., Shenkin, A., Silman, M. R., Wieczynski, D. J., Asner, G. P., Bentley, L. P., Savage, V. M., & Enquist, B. J. (2019). Informing trait-based ecology by assessing remotely sensed functional diversity across a broad tropical temperature gradient. Science Advances, 5(12), eaaw8114. https://doi.org/10.1126/sciadv.aaw8114

Ehleringer, J., Mooney, H. A., Gulmon, S. L., & Rundel, P. W. (1981). Parallel evolution of leaf pubescence in Encelia in coastal deserts of North and South America. Oecologia, 49(1), 38–41. https://doi.org/10.1007/BF00376895

Ellsworth, D. S., & Reich, P. B. (1993). Canopy structure and vertical patterns of photosynthesis and related leaf traits in a deciduous forest. Oecologia, 96(2), 169–178. https://doi.org/10.1007/BF00317729

Elvidge, C. D. (1990). Visible and near infrared reflectance characteristics of dry plant materials. International Journal of Remote Sensing, 11(10), 1775–1795. https://doi.org/10.1080/01431169008955129

Ely, K. S., Burnett, A. C., Lieberman-Cribbin, W., Serbin, S. P., & Rogers, A. (2019). Spectroscopy can predict key leaf traits associated with source–sink balance and carbon–nitrogen status. Journal of Experimental Botany, 70(6), 1789–1799. https://doi.org/10.1093/jxb/erz061

Farahmand, A. M., Szepesvári, C., & Audibert, J.-Y. (2007). Manifold-adaptive dimension estimation. Proceedings of the 24th International Conference on Machine Learning, 265–272. https://doi.org/10.1145/1273496.1273530

Feilhauer, H., Faude, U., & Schmidtlein, S. (2011). Combining Isomap ordination and imaging spectroscopy to map continuous floristic gradients in a heterogeneous landscape. Remote Sensing of Environment, 115(10), 2513–2524. https://doi.org/10.1016/j.rse.2011.05.011

Féret, J.-B., & Asner, G. P. (2013). Tree Species Discrimination in Tropical Forests Using Airborne Imaging Spectroscopy. IEEE Transactions on Geoscience and Remote Sensing, 51(1), 73–84. https://doi.org/10.1109/TGRS.2012.2199323

Féret, J.-B., & Asner, G. P. (2014). Mapping tropical forest canopy diversity using high-fidelity imaging spectroscopy. Ecological Applications, 24(6), 1289–1296. https://doi.org/10.1890/13-1824.1

Féret, J.-B., Berger, K., de Boissieu, F., & Malenovský, Z. (2021). PROSPECT-PRO for estimating content of nitrogen-containing leaf proteins and other carbon-based constituents. Remote Sensing of Environment, 252, 112173. https://doi.org/10.1016/j.rse.2020.112173

Féret, J.-B., Gitelson, A. A., Noble, S. D., & Jacquemoud, S. (2017). PROSPECT-D: Towards modeling leaf optical properties through a complete lifecycle. Remote Sensing of Environment, 193, 204–215. https://doi.org/10.1016/j.rse.2017.03.004

Fisher, J. B., Sweeney, S., Brzostek, E. R., Evans, T. P., Johnson, D. J., Myers, J. A., Bourg, N. A., Wolf, A. T., Howe, R. W., & Phillips, R. P. (2016). Tree–mycorrhizal associations detected remotely from canopy spectral properties. Global Change Biology, n/a-n/a. https://doi.org/10.1111/gcb.13264

Frye, H. A., Aiello-Lammens, M. E., Euston-Brown, D., Jones, C. S., Mollmann, H. K., Merow, C., Slingsby, J. A., Merwe, H. van der, Wilson, A. M., & Silander, J. A. (2021). Plant spectral diversity as a surrogate for species, functional and phylogenetic diversity across a hyper-diverse biogeographic region. Global Ecology and Biogeography, 30(7), 1403–1417. https://doi.org/10.1111/geb.13306

Gamon, J. A., & Berry, J. A. (2012). Facultative and constitutive pigment effects on the Photochemical Reflectance Index (PRI) in sun and shade conifer needles. Israel Journal of Plant Sciences, 60(1–2), 85–95. https://doi.org/10.1560/IJPS.60.1-2.85

Gamon, J. A., Somers, B., Malenovský, Z., Middleton, E. M., Rascher, U., & Schaepman, M. E. (2019). Assessing Vegetation Function with Imaging Spectroscopy. Surveys in Geophysics, 40(3), 489–513. https://doi.org/10.1007/s10712-019-09511-5

Gamon, J. A., Wang, R., Gholizadeh, H., Zutta, B., Townsend, P. A., & Cavender-Bares, J. (2020). Consideration of Scale in Remote Sensing of Biodiversity. In J. Cavender-Bares, J. A. Gamon, & P. A. Townsend (Eds.), Remote Sensing of Plant Biodiversity (pp. 425–447). Springer International Publishing. https://doi.org/10.1007/978-3-030-33157-3_16

Gates, D. M., Keegan, H. J., Schleter, J. C., & Weidner, V. R. (1965). Spectral Properties of Plants. Applied Optics, 4(1), 11–20. https://doi.org/10.1364/AO.4.000011

Gholizadeh, H., Gamon, J. A., Zygielbaum, A. I., Wang, R., Schweiger, A. K., & Cavender-Bares, J. (2018). Remote sensing of biodiversity: Soil correction and data dimension reduction methods improve assessment of α-diversity (species richness) in prairie ecosystems. Remote Sensing of Environment, 206, 240–253. https://doi.org/10.1016/j.rse.2017.12.014

Girard, A., Schweiger, A. K., Carteron, A., Kalacska, M., & Laliberté, E. (2020). Foliar Spectra and Traits of Bog Plants across Nitrogen Deposition Gradients. Remote Sensing, 12(15), 2448. https://doi.org/10.3390/rs12152448

Gower, S. T., Kucharik, C. J., & Norman, J. M. (1999). Direct and Indirect Estimation of Leaf Area Index, fAPAR, and Net Primary Production of Terrestrial Ecosystems. Remote Sensing of Environment, 70(1), 29–51. https://doi.org/10.1016/S0034-4257(99)00056-5

Grassberger, P., & Procaccia, I. (1983). Characterization of Strange Attractors. Physical Review Letters, 50(5), 346–349. https://doi.org/10.1103/PhysRevLett.50.346

Heberling, J. M., & Fridley, J. D. (2012). Biogeographic constraints on the world-wide leaf economics spectrum. Global Ecology and Biogeography, 21(12), 1137–1146. https://doi.org/10.1111/j.1466-8238.2012.00761.x

Hollinger, D. Y., Ollinger, S. V., Richardson, A. D., Meyers, T. P., Dail, D. B., Martin, M. E., Scott, N. A., Arkebauer, T. J., Baldocchi, D. D., Clark, K. L., Curtis, P. S., Davis, K. J., Desai, A. R., Dragoni, D., Goulden, M. L., Gu, L., Katul, G. G., Pallardy, S. G., Paw U, K. T., … Verma, S. B. (2010). Albedo estimates for land surface models and support for a new paradigm based on foliage nitrogen concentration. Global Change Biology, 16(2), 696–710. https://doi.org/10.1111/j.1365-2486.2009.02028.x

Jacquemoud, S., & Baret, F. (1990). PROSPECT: A model of leaf optical properties spectra. Remote Sensing of Environment, 34(2), 75–91. https://doi.org/10.1016/0034-4257(90)90100-Z

Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P. J., Asner, G. P., François, C., & Ustin, S. L. (2009). PROSPECT+SAIL models: A review of use for vegetation characterization. Remote Sensing of Environment, 113, S56–S66. https://doi.org/10.1016/j.rse.2008.01.026

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. Nature Plants, 2(3), 1–5. https://doi.org/10.1038/nplants.2016.24

Joswig, J. S., Wirth, C., Schuman, M. C., Kattge, J., Reu, B., Wright, I. J., Sippel, S. D., Rüger, N., Richter, R., Schaepman, M. E., van Bodegom, P. M., Cornelissen, J. H. C., Díaz, S., Hattingh, W. N., Kramer, K., Lens, F., Niinemets, Ü., Reich, P. B., Reichstein, M., … Mahecha, M. D. (2022). Climatic and soil factors explain the two-dimensional spectrum of global plant trait variation. Nature Ecology & Evolution, 6(1), 36–50. https://doi.org/10.1038/s41559-021-01616-8

Jucker, T., Bongalov, B., Burslem, D. F. R. P., Nilus, R., Dalponte, M., Lewis, S. L., Phillips, O. L., Qie, L., & Coomes, D. A. (2018). Topography shapes the structure, composition and function of tropical forest landscapes. Ecology Letters, 21(7), 989–1000. https://doi.org/10.1111/ele.12964

Jucker, T., Bouriaud, O., & Coomes, D. A. (2015). Crown plasticity enables trees to optimize canopy packing in mixed-species forests. Functional Ecology, 29(8), 1078–1086. https://doi.org/10.1111/1365-2435.12428

Karabourniotis, G., Liakopoulos, G., Bresta, P., & Nikolopoulos, D. (2021). The Optical Properties of Leaf Structural Elements and Their Contribution to Photosynthetic Performance and Photoprotection. Plants, 10(7), 1455. https://doi.org/10.3390/plants10071455

Kattge, J., Bönisch, G., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Tautenhahn, S., Werner, G. D. A., Aakala, T., Abedi, M., Acosta, A. T. R., Adamidis, G. C., Adamson, K., Aiba, M., Albert, C. H., Alcántara, J. M., C, C. A., Aleixo, I., Ali, H., … Wirth, C. (2020). TRY plant trait database – enhanced coverage and open access. Global Change Biology, 26(1), 119–188. https://doi.org/10.1111/gcb.14904

Kemp, J. E., & Ellis, A. G. (2019). Cryptic petal coloration decreases floral apparency and herbivory in nocturnally closing daisies. Functional Ecology, 33(11), 2130–2141. https://doi.org/10.1111/1365-2435.13423

Knipling, E. B. (1970). Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation. Remote Sensing of Environment, 1(3), 155–159. https://doi.org/10.1016/S0034-4257(70)80021-9

Knyazikhin, Y., Schull, M. A., Stenberg, P., Mõttus, M., Rautiainen, M., Yang, Y., Marshak, A., Carmona, P. L., Kaufmann, R. K., Lewis, P., Disney, M. I., Vanderbilt, V., Davis, A. B., Baret, F., Jacquemoud, S., Lyapustin, A., & Myneni, R. B. (2013). Hyperspectral remote sensing of foliar nitrogen content. Proceedings of the National Academy of Sciences, 110(3), E185–E192. https://doi.org/10.1073/pnas.1210196109

Kothari, S., Beauchamp-Rioux, R., Laliberté, E., & Cavender-Bares, J. (2021). Reflectance spectroscopy allows rapid, accurate, and non-destructive estimates of functional traits from pressed leaves. bioRxiv. https://doi.org/10.1101/2021.04.21.440856

Kothari, S., Cavender-Bares, J., Bitan, K., Verhoeven, A. S., Wang, R., Montgomery, R. A., & Gamon, J. A. (2018). Community-wide consequences of variation in photoprotective physiology among prairie plants. Photosynthetica, 1–13. https://doi.org/10.1007/s11099-018-0777-9

Kothari, S., Montgomery, R. A., & Cavender-Bares, J. (2021). Physiological responses to light explain competition and facilitation in a tree diversity experiment. Journal of Ecology, 109(5), 2000–2018. https://doi.org/10.1111/1365-2745.13637

Kraft, N. J. B., Godoy, O., & Levine, J. M. (2015). Plant functional traits and the multidimensional nature of species coexistence. Proceedings of the National Academy of Sciences, 201413650. https://doi.org/10.1073/pnas.1413650112

Kuhn, M. (2008). Building Predictive Models in R Using the caret Package. Journal of Statistical Software, 28, 1–26. https://doi.org/10.18637/jss.v028.i05

Kuusinen, N., Hovi, A., & Rautiainen, M. (2021). Contribution of woody elements to tree level reflectance in boreal forests. Silva Fennica, 55(5). https://doi.org/10.14214/sf.10600

Laliberté, E., & Legendre, P. (2010). A distance-based framework for measuring functional diversity from multiple traits. Ecology, 91(1), 299–305. https://doi.org/10.1890/08-2244.1

Laliberté, E., Schweiger, A. K., & Legendre, P. (2020). Partitioning plant spectral diversity into alpha and beta components. Ecology Letters, 23(2), 370–380. https://doi.org/10.1111/ele.13429

Laughlin, D. C. (2014). The intrinsic dimensionality of plant traits and its relevance to community assembly. Journal of Ecology, 102(1), 186–193. https://doi.org/10.1111/1365-2745.12187

Legendre, P., & De Cáceres, M. (2013). Beta diversity as the variance of community data: Dissimilarity coefficients and partitioning. Ecology Letters, 16(8), 951–963. https://doi.org/10.1111/ele.12141

Lenoir, J., Gril, E., Durrieu, S., Horen, H., Laslier, M., Lembrechts, J. J., Zellweger, F., Alleaume, S., Brasseur, B., Buridant, J., Dayal, K., De Frenne, P., Gallet-Moron, E., Marrec, R., Meeussen, C., Rocchini, D., Van Meerbeek, K., & Decocq, G. (2022). Unveil the unseen: Using LiDAR to capture time-lag dynamics in the herbaceous layer of European temperate forests. Journal of Ecology, 110(2), 282–300. https://doi.org/10.1111/1365-2745.13837

Levina, E., & Bickel, P. (2004). Maximum Likelihood Estimation of Intrinsic Dimension. Advances in Neural Information Processing Systems, 17. https://proceedings.neurips.cc/paper/2004/hash/74934548253bcab8490ebd74afed7031-Abstract.html

Marten, G. C., Shenk, J. S., & Barton, F. E. (1985). Near infrared reflectance spectroscopy (NIRS): Analysis of forage quality (No. 643). US Department of Agriculture, Agricultural Research Service.

Martens, H. (2001). Reliable and relevant modelling of real world data: A personal account of the development of PLS Regression. Chemometrics and Intelligent Laboratory Systems, 58(2), 85–95. https://doi.org/10.1016/S0169-7439(01)00153-8

Mathieson, I. (2021). The omnigenic model and polygenic prediction of complex traits. The American Journal of Human Genetics, 108(9), 1558–1563. https://doi.org/10.1016/j.ajhg.2021.07.003

McInnes, L., Healy, J., & Melville, J. (2020). UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. ArXiv:1802.03426 [Cs, Stat]. http://arxiv.org/abs/1802.03426

Meacham-Hensold, K., Montes, C. M., Wu, J., Guan, K., Fu, P., Ainsworth, E. A., Pederson, T., Moore, C. E., Brown, K. L., Raines, C., & Bernacchi, C. J. (2019). High-throughput field phenotyping using hyperspectral reflectance and partial least squares regression (PLSR) reveals genetic modifications to photosynthetic capacity. Remote Sensing of Environment, 231, 111176. https://doi.org/10.1016/j.rse.2019.04.029

Meireles, J. E., Cavender-Bares, J., Townsend, P. A., Ustin, S., Gamon, J. A., Schweiger, A. K., Schaepman, M. E., Asner, G. P., Martin, R. E., Singh, A., Schrodt, F., Chlus, A., & O'Meara, B. C. (2020). Leaf reflectance spectra capture the evolutionary history of seed plants. New Phytologist, 228(2), 485–493. https://doi.org/10.1111/nph.16771

Meireles, J., Schweiger, A., & Cavender-Bares, J. (2017). spectrolab: Class and Methods for Hyperspectral Data in R. R package version 0.0.10. https://CRAN.R-project.org/package=spectrolab

Mencuccini, M., Rosas, T., Rowland, L., Choat, B., Cornelissen, H., Jansen, S., Kramer, K., Lapenis, A., Manzoni, S., Niinemets, Ü., Reich, P. B., Schrodt, F., Soudzilovskaia, N., Wright, I. J., & Martínez-Vilalta, J. (2019). Leaf economics and plant hydraulics drive leaf: Wood area ratios. New Phytologist, 224(4), 1544–1556. https://doi.org/10.1111/nph.15998

Meyer, C., Weigelt, P., & Kreft, H. (2016). Multidimensional biases, gaps and uncertainties in global plant occurrence information. Ecology Letters, 19(8), 992–1006. https://doi.org/10.1111/ele.12624

Mooney, H. A., Ehleringer, J., & Björkman, O. (1977). The energy balance of leaves of the evergreen desert shrub Atriplex hymenelytra. Oecologia, 29(4), 301–310. https://doi.org/10.1007/BF00345804

Niinemets, Ü. (2007). Photosynthesis and resource distribution through plant canopies. Plant, Cell & Environment, 30(9), 1052–1071. https://doi.org/10.1111/j.1365-3040.2007.01683.x

Nunes, M. H., Davey, M. P., & Coomes, D. A. (2017). On the challenges of using field spectroscopy to measure the impact of soil type on leaf traits. Biogeosciences, 14(13), 3371–3385. https://doi.org/10.5194/bg-14-3371-2017

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., & Wagner, H. (2019). vegan: Community Ecology Package. R package version 2.5-5. https://CRAN.R-project.org/package=vegan

Ollinger, S. V., Richardson, A. D., Martin, M. E., Hollinger, D. Y., Frolking, S. E., Reich, P. B., Plourde, L. C., Katul, G. G., Munger, J. W., Oren, R., Smith, M.-L., U, K. T. P., Bolstad, P. V., Cook, B. D., Day, M. C., Martin, T. A., Monson, R. K., & Schmid, H. P. (2008). Canopy nitrogen, carbon assimilation, and albedo in temperate and boreal forests: Functional relations and potential climate feedbacks. Proceedings of the National Academy of Sciences, 105(49), 19336–19341. https://doi.org/10.1073/pnas.0810021105

Osnas, J. L. D., Katabuchi, M., Kitajima, K., Wright, S. J., Reich, P. B., Bael, S. A. V., Kraft, N. J. B., Samaniego, M. J., Pacala, S. W., & Lichstein, J. W. (2018). Divergent drivers of leaf trait variation within species, among species, and among functional groups. Proceedings of the National Academy of Sciences, 115(21), 5480–5485. https://doi.org/10.1073/pnas.1803989115

Pacala, S. W., & Tilman, D. (1994). Limiting Similarity in Mechanistic and Spatial Models of Plant Competition in Heterogeneous Environments. The American Naturalist, 143(2), 222–257. https://doi.org/10.1086/285602

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., … Cornelissen, J. H. C. (2013). New handbook for standardised measurement of plant functional traits worldwide. Australian Journal of Botany, 61(3), 167–234.

Pérez-Ramos, I. M., Matías, L., Gómez-Aparicio, L., & Godoy, Ó. (2019). Functional traits and phenotypic plasticity modulate species coexistence across contrasting climatic conditions. Nature Communications, 10(1), 2555. https://doi.org/10.1038/s41467-019-10453-0

Peterson, D. L., Aber, J. D., Matson, P. A., Card, D. H., Swanberg, N., Wessman, C., & Spanner, M. (1988). Remote sensing of forest canopy and leaf biochemical contents. Remote Sensing of Environment, 24(1), 85–108. https://doi.org/10.1016/0034-4257(88)90007-7

Pettis, K. W., Bailey, T. A., Jain, A. K., & Dubes, R. C. (1979). An Intrinsic Dimensionality Estimator from Near-Neighbor Information. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-1(1), 25–37. https://doi.org/10.1109/TPAMI.1979.4766873

Pontius, J., Hallett, R., & Martin, M. (2005). Using AVIRIS to assess hemlock abundance and early decline in the Catskills, New York. Remote Sensing of Environment, 97(2), 163–173. https://doi.org/10.1016/j.rse.2005.04.011

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Reich, P. B. (2014). The world-wide ‘fast–slow’ plant economics spectrum: A traits manifesto. Journal of Ecology, 102(2), 275–301. https://doi.org/10.1111/1365-2745.12211

Rocchini, D., Balkenhol, N., Carter, G. A., Foody, G. M., Gillespie, T. W., He, K. S., Kark, S., Levin, N., Lucas, K., Luoto, M., Nagendra, H., Oldeland, J., Ricotta, C., Southworth, J., & Neteler, M. (2010). Remotely sensed spectral heterogeneity as a proxy of species diversity: Recent advances and open challenges. Ecological Informatics, 5(5), 318–329. https://doi.org/10.1016/j.ecoinf.2010.06.001

Roddy, A. B. (2019). Energy Balance Implications of Floral Traits Involved in Pollinator Attraction and Water Balance. International Journal of Plant Sciences, 180(9), 944–953. https://doi.org/10.1086/705586

Roth, K. L., Roberts, D. A., Dennison, P. E., Alonzo, M., Peterson, S. H., & Beland, M. (2015). Differentiating plant species within and across diverse ecosystems with imaging spectroscopy. Remote Sensing of Environment, 167, 135–151. https://doi.org/10.1016/j.rse.2015.05.007

Rozza, A., Lombardi, G., Ceruti, C., Casiraghi, E., & Campadelli, P. (2012). Novel high intrinsic dimensionality estimators. Machine Learning, 89(1), 37–65. https://doi.org/10.1007/s10994-012-5294-7

Sapes, G., Lapadat, C., Schweiger, A. K., Juzwik, J., Montgomery, R., Gholizadeh, H., Townsend, P. A., Gamon, J. A., & Cavender-Bares, J. (2022). Canopy spectral reflectance detects oak wilt at the landscape scale using phylogenetic discrimination. Remote Sensing of Environment, 273, 112961. https://doi.org/10.1016/j.rse.2022.112961

Schimel, D. S., Asner, G. P., & Moorcroft, P. (2013). Observing changing ecological diversity in the Anthropocene. Frontiers in Ecology and the Environment, 11(3), 129–137. https://doi.org/10.1890/120111

Schmidtlein S., Feilhauer H. & Bruelheide H. (2012). Mapping plant strategy types using remote sensing. Journal of Vegetation Science 23(3), 395-405. https://doi.org/10.1111/j.1654-1103.2011.01370.x

Schneider, F. D., Morsdorf, F., Schmid, B., Petchey, O. L., Hueni, A., Schimel, D. S., & Schaepman, M. E. (2017). Mapping functional diversity from remotely sensed morphological and physiological forest traits. Nature Communications, 8(1), 1441. https://doi.org/10.1038/s41467-017-01530-3

Schweiger, A. K. (2020). Spectral Field Campaigns: Planning and Data Collection. In J. Cavender-Bares, J. A. Gamon, & P. A. Townsend (Eds.), Remote Sensing of Plant Biodiversity (pp. 385–423). Springer International Publishing. https://doi.org/10.1007/978-3-030-33157-3_15

Schweiger, A. K., Cavender-Bares, J., Kothari, S., Townsend, P. A., Madritch, M. D., Grossman, J. J., Gholizadeh, H., Wang, R., & Gamon, J. A. (2021). Coupling spectral and resource-use complementarity in experimental grassland and forest communities. Proceedings of the Royal Society B: Biological Sciences, 288(1958), 20211290. https://doi.org/10.1098/rspb.2021.1290

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. Nature Ecology & Evolution, 2(6), 976–982. https://doi.org/10.1038/s41559-018-0551-1

Schweiger, A. K. and Laliberté, E. (2022). Plant beta-diversity across biomes captured by imaging spectroscopy. Nature Communications (in press).

Schweiger, A. K., Schütz, M., Risch, A. C., Kneubühler, M., Haller, R., & Schaepman, M. E. (2017). How to predict plant functional types using imaging spectroscopy: Linking vegetation community traits, plant functional types and spectral response. Methods in Ecology and Evolution, 8(1), 86–95. https://doi.org/10.1111/2041-210X.12642

Segrestin, J., Sartori, K., Navas, M.-L., Kattge, J., Díaz, S., & Garnier, E. (2021). PhenoSpace: A Shiny application to visualize trait data in the phenotypic space of the global spectrum of plant form and function. Ecology and Evolution, 11(4), 1526–1534. https://doi.org/10.1002/ece3.6928

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. Ecological Applications, 24(7), 1651–1669. https://doi.org/10.1890/13-2110.1

Serbin, S. P., & Townsend, P. A. (2020). Scaling Functional Traits from Leaves to Canopies. In J. Cavender-Bares, J. A. Gamon, & P. A. Townsend (Eds.), Remote Sensing of Plant Biodiversity (pp. 43–82). Springer International Publishing. https://doi.org/10.1007/978-3-030-33157-3_3

Serbin, S. P., Wu, J., Ely, K. S., Kruger, E. L., Townsend, P. A., Meng, R., Wolfe, B. 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 Phytologist, 224(4), 1557–1568. https://doi.org/10.1111/nph.16123

Shiklomanov, A. N., Dietze, M. C., Viskari, T., Townsend, P. A., & Serbin, S. P. (2016). Quantifying the influences of spectral resolution on uncertainty in leaf trait estimates through a Bayesian approach to RTM inversion. Remote Sensing of Environment, 183, 226–238. https://doi.org/10.1016/j.rse.2016.05.023

Shull, C. A. (1929). A Spectrophotometric Study of Reflection of Light from Leaf Surfaces. Botanical Gazette, 87(5), 583–607.

Singh, A., Serbin, S. P., McNeil, B. E., Kingdon, C. C., & Townsend, P. A. (2015). Imaging spectroscopy algorithms for mapping canopy foliar chemical and morphological traits and their uncertainties. Ecological Applications, 25(8), 2180–2197. https://doi.org/10.1890/14-2098.1

Smith, A. B., Sandel, B., Kraft, N. J. B., & Carey, S. (2013). Characterizing scale-dependent community assembly using the functional-diversity–area relationship. Ecology, 94(11), 2392–2402. https://doi.org/10.1890/12-2109.1

Sousa, D., Fisher, J. B., Galvan, F. R., Pavlick, R. P., Cordell, S., Giambelluca, T. W., Giardina, C. P., Gilbert, G. S., Imran-Narahari, F., Litton, C. M., Lutz, J. A., North, M. P., Orwig, D. A., Ostertag, R., Sack, L., & Phillips, R. P. (2021). Tree Canopies Reflect Mycorrhizal Composition. Geophysical Research Letters, 48(10), e2021GL092764. https://doi.org/10.1029/2021GL092764

Strauss, S. Y., & Cacho, N. I. (2013). Nowhere to Run, Nowhere to Hide: The Importance of Enemies and Apparency in Adaptation to Harsh Soil Environments. The American Naturalist, 182(1), E1–E14. https://doi.org/10.1086/670754

Tenenbaum, J. B., Silva, V. de, & Langford, J. C. (2000). A Global Geometric Framework for Nonlinear Dimensionality Reduction. Science, 290(5500), 2319–2323. https://doi.org/10.1126/science.290.5500.2319

Townsend, P. A., Serbin, S. P., Kruger, E. L., & Gamon, J. A. (2013). Disentangling the contribution of biological and physical properties of leaves and canopies in imaging spectroscopy data. Proceedings of the National Academy of Sciences, 110(12), E1074–E1074. https://doi.org/10.1073/pnas.1300952110

Turner, W. (2014). Sensing biodiversity. Science, 346(6207), 301–302. https://doi.org/10.1126/science.1256014

Ustin, S. L., & Gamon, J. A. (2010). Remote sensing of plant functional types. New Phytologist, 186(4), 795–816. https://doi.org/10.1111/j.1469-8137.2010.03284.x

Ustin, S. L., & Jacquemoud, S. (2020). How the Optical Properties of Leaves Modify the Absorption and Scattering of Energy and Enhance Leaf Functionality. Remote Sensing of Plant Biodiversity, 349–384. https://doi.org/10.1007/978-3-030-33157-3_14

van der Kooi, C. J., Pen, I., Staal, M., Stavenga, D. G., & Elzenga, J. T. M. (2016). Competition for pollinators and intra-communal spectral dissimilarity of flowers. Plant Biology, 18(1), 56–62. https://doi.org/10.1111/plb.12328

Venables, W. N. and Ripley, B. D. (2002). Modern Applied Statistics with S. New York: Springer (4th ed)

Violle, C., Enquist, B. J., McGill, B. J., Jiang, L., Albert, C. H., Hulshof, C., Jung, V., & Messier, J. (2012). The return of the variance: Intraspecific variability in community ecology. Trends in Ecology & Evolution, 27(4), 244–252. https://doi.org/10.1016/j.tree.2011.11.014

Wang, R. & Gamon, J.A. (2019). Remote sensing of terrestrial plant biodiversity. Remote Sensing of Environment 231, 111218. https://doi.org/10.1016/j.rse.2019.111218

Wang, Z., Chlus, A., Geygan, R., Ye, Z., Zheng, T., Singh, A., Couture, J. J., Cavender-Bares, J., Kruger, E. L., & Townsend, P. A. (2020). Foliar functional traits from imaging spectroscopy across biomes in eastern North America. New Phytologist, 228(2), 494–511. https://doi.org/10.1111/nph.16711

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 Sensing of Environment, 221, 405–416. https://doi.org/10.1016/j.rse.2018.11.016

Wei, X., Savage, J. A., Riggs, C. E., & Cavender-Bares, J. (2017). An experimental test of fitness variation across a hydrologic gradient predicts willow and poplar species distributions. Ecology, 98(5), 1311–1323. https://doi.org/10.1002/ecy.1784

Wessman, C. A., Aber, J. D., Peterson, D. L., & Melillo, J. M. (1988). Remote sensing of canopy chemistry and nitrogen cycling in temperate forest ecosystems. Nature, 335(6186), 154–156. https://doi.org/10.1038/335154a0

Westoby, M. (1998). A leaf-height-seed (LHS) plant ecology strategy scheme. Plant and Soil, 199(2), 213–227. https://doi.org/10.1023/A:1004327224729

Wicklein, H. F., Ollinger, S. V., Martin, M. E., Hollinger, D. Y., Lepine, L. C., Day, M. C., Bartlett, M. K., Richardson, A. D., & Norby, R. J. (2012). Variation in foliar nitrogen and albedo in response to nitrogen fertilization and elevated CO2. Oecologia, 169(4), 915–925. https://doi.org/10.1007/s00442-012-2263-6

Williams, L. J., Cavender-Bares, J., Paquette, A., Messier, C., & Reich, P. B. (2020). Light mediates the relationship between community diversity and trait plasticity in functionally and phylogenetically diverse tree mixtures. Journal of Ecology, 108(4), 1617–1634. https://doi.org/10.1111/1365-2745.13346

Williams, L. J., Cavender-Bares, J., Townsend, P. A., Couture, J. J., Wang, Z., Stefanski, A., Messier, C., & Reich, P. B. (2021). Remote spectral detection of biodiversity effects on forest biomass. Nature Ecology & Evolution, 5(1), 46–54. https://doi.org/10.1038/s41559-020-01329-4

Williams, L. J., Paquette, A., Cavender-Bares, J., Messier, C., & Reich, P. B. (2017). Spatial complementarity in tree crowns explains overyielding in species mixtures. Nature Ecology & Evolution, 1, 0063. https://doi.org/10.1038/s41559-016-0063

Wold, S., Martens, H., & Wold, H. (1983). The multivariate calibration problem in chemistry solved by the PLS method. In B. Kågström & A. Ruhe (Eds.), Matrix Pencils (pp. 286–293). Springer. https://doi.org/10.1007/BFb0062108

Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: A basic tool of chemometrics. Chemometrics and Intelligent Laboratory Systems, 58(2), 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., … Villar, R. (2004). The worldwide leaf economics spectrum. Nature, 428(6985), 821–827. https://doi.org/10.1038/nature02403

Yan, Z., Guo, Z., Serbin, S. P., Song, G., Zhao, Y., Chen, Y., Wu, S., Wang, J., Wang, X., Li, J., Wang, B., Wu, Y., Su, Y., Wang, H., Rogers, A., Liu, L., & Wu, J. (2021). Spectroscopy outperforms leaf trait relationships for predicting photosynthetic capacity across different forest types. New Phytologist, 232(1), 134–147. https://doi.org/10.1111/nph.17579

Zhou, J., Cieraad, E., & van Bodegom, P. M. (2022). Global analysis of trait–trait relationships within and between species. New Phytologist, 233(4), 1643–1656. https://doi.org/10.1111/nph.17879

Supporting Information

Supplementary Methods

Species differentiation analysis

The Cedar Creek Biodiversity (BioDIV or e120) experiment is located at Cedar Creek Ecosystem Science Research in central Minnesota, USA. The experiment was established in 1994 with plots planted in various levels of richness from a pool of 18 prairie plant species (Tilman et al. 1997). For this analysis, we ignore the diversity treatments and focus on species identities. The species sampled were: *Achillea millefolium* L. (ACHMI, *n* = 49), *Amorpha canescens* Pursh (AMOCA, *n* = 28), *Andropogon gerardii* Vitman (ANDGE, *n* = 162), *Asclepias tuberosa* L. (ASCTU, *n* = 70), *Lespedeza capitata* Michx. (LESCA, *n* = 99), *Liatris aspera* Michx. (LIAAS, *n* = 49), *Lupinus perennis* L. (LUPPE, *n* = 121), *Panicum virgatum* L. (PANVI, *n* = 49), *Petalostemum candidum* (Willd.) Michx. (PETCA, *n* = 28), *Petalostemum purpureum* (Vent.) Rydb. (PETPU, *n* = 52), *Petalostemum villosum* Nutt. (PETVI, *n* = 42), *Schizachyrium scoparium* (Michx.) Nash (SCHSC, *n* = 76), *Solidago rigida* L. (SOLRI, *n* = 50) and *Sorghastrum nutans* (L.) Nash (SORNU, *n* = 27).

We measured leaf spectra using a leaf-clip assembly and two portable field spectrometers (SVC HR-1024i, Spectra Vista Corp., Poughkeepsie, NY, USA; and PSR+ 3500, Spectral Evolution Inc., Lawrence, MA, USA) covering the wavelength range from 350 nm to 2500 nm in 1024 spectral bands. We used the SVC instrument for measuring herbaceous species and the PSR+ for measuring tree species. To characterize one individual spectrally, we measured the reflectance of either three or five mature, healthy leaves per individual depending on plant height. Spectra were automatically corrected for dark current and stray light, and referenced to the white calibration disc of the leaf clip approximately every 10 minutes. Spectral data processing included correcting discontinuities at the sensor overlap regions between the Si and first InGaAs sensor (around 1000 nm) and between the first and second InGaAs sensor (around 1900 nm), removing noisy regions at the beginning and end of the spectrum, and interpolating spectra to 1 nm resolution. For all spectral processing, we used the package *spectrolab* v. 0.0.10 (Meireles et al. 2020) in R v. 3.6.2 (R Core Team 2020). For LDA we used the *R* package *MASS* v. 7.3.53 (Venables & Ripley 2002), for PLSDA we used the *R* package *caret* v. 6.0.86 (Kuhn 2018). In our case, linear discriminant axes (LDs) are linear combinations of all band-wise reflectance and trait values, respectively, which re-project observations into a new coordinate system while maximizing the differences between groups; our grouping variable was species identity.

We tested for phylogenetic signal of each trait using Blomberg et al. (2003)'s K statistic as implemented in the *R* package *picante* v. 1.7 (Kembel et al. 2010) and the phylogeny reconstructed by Kothari et al. (2018) with one missing species (*Petalostemum candidum*) added manually using *R* package *phytools* v. 0.6-44 (Revell 2012). We compared the observed K value (in Table S1 referred to as K_{trait}) to the distribution of the K statistics estimated from both a white noise (K_{null}) and a Brownian motion null model (K_{brown}). If observed K values are not significantly different ($P < 0.05$) from the Brownian motion null model, they can be considered phylogenetically conserved. If observed K values are not significantly different ($P < 0.05$) from the random expectation (white noise null model), they can be considered labile. We estimated the Brownian motion null model based on 1,000 simulations of Brownian motion evolution and the white noise model by randomly permuting traits values across the tips of the phylogeny 1,000 times.

We tested the degree to which plant species can be correctly identified based on spectra and traits using PLSDA as implemented in the *R* package *caret* (Kuhn 2008). We used random draws of 20 individuals per species for model training, the remaining data were used for validation and for evaluating model fit; all statistics and graphs are based on the validation results. We performed 100 PLSDA model iterations using new random draws of training samples, and selected the optimal number of components based on the minimum of the root mean squared error of prediction (RMSEP) for the test samples. We tested for significant differences in RMSEP values among the number of components using Tukey's HSD test as implemented in the R package *agricolae* v. 1.3-1 (de Mendiburu & Simon 2015), and used the smaller number of components when models performed similarly (p > 0.05).

Dimensionality analysis

We used data collected by the Canadian Airborne Biodiversity Observatory from a variety of sites and functional groups across temperate Canada, as well as one site in Australia. All spectral data were collected using an HR-1024i spectroradiometer equipped with a DC-R/T integrating sphere from Spectra Vista Corporation (Poughkeepsie, NY, USA). (See Laliberté & Soffer 2018a, Laliberté and Soffer 2018b, and Schweiger & Laliberté 2020 for the measurement and processing protocols. All trait data were collected using consistent methods (see Kothari et al. 2021 for trait protocols). We transformed all chemical traits (those besides LMA and EWT) to a normalization-independent basis following Osnas et al. (2013). The comparisons of fresh-leaf, pressed-leaf, and ground-leaf spectra are based on data from Kothari et al. (2021). For spectral processing we used the package *spectrolab v. 0.0.10* (Meireles et al. 2020) in *R v. 3.6.3* (R Core Team 2020).

We used PROSPECT-D to generate a synthetic dataset of 1000 spectra (Féret et al. 2017) using *R* package *hsdar v. 1.0.0* (Lehnert et al. 2018). For each spectrum, we sampled the leaf structure (N) from a uniform distribution between 1 and 2 and independently sampled chlorophyll, water, and dry matter content with replacement from measurements in the subsampled CABO dataset. For simplicity, we set brown pigments and anthocyanins to 0 and determined carotenoid content as chlorophyll divided by 6.24 (the mean chlorophyll : carotenoid in the full CABO dataset). This synthetic dataset thus had a known dimensionality of 4. We normalized and subsampled bands as in the real CABO data.

We used a number of methods to estimate the intrinsic dimensionality of our datasets. These included one simple, visual linear method (the PCA scree test) as well as several nonlinear methods. The methods include:

1. The correlation integral-based method of Grassberger & Procaccia (1983), a variant of the fractal dimension, as implemented in *R* package *ider v. 0.1.0* (Hino 2017). The parameter *p* (ambient dimension) was set to 20.
2. The maximum likelihood-based estimator of Levina & Bickel (2004) as corrected by MacKay and Ghahramani (2004) and implemented in *R* package *ider v. 0.1.0* (Hino 2017). The two nearest-neighbor parameters k_1 and k_2 were set to 5 and 10, respectively.
3. The manifold-adaptive local information dimension estimator of Farahmand et al. (2007), as implemented in *R* package *ider v. 0.1.0* (Hino 2017). The maximum candidate dimension was set at 10, and local dimension estimates were combined by averaging.
4. The nearest-neighbor information dimension estimator of Pettis et al. (1979) as implemented in *R* package *ider v. 0.1.0* (Hino 2017). The two nearest-neighbor parameters k_1 and k_2 were set to 5 and 30, respectively.
5. The minimum neighbor distance–maximum likelihood ($\mathrm{MIND}_{\mathrm{MLi}}$) method of Rozza et al. (2012) as implemented in *R* package *intrinsicDimension v. 1.2.0* (Johnsson 2019) with neighborhood parameter *k* set to 5.
6. Cattell (1966)'s scree test based on eigenvalues from principal components analysis (PCA).
7. An analog to the scree test based on a residual variance plot from Isomap (Tenenbaum et al. 2000), as implemented in *R* package *vegan v. 2.5.5* (Oksanen et al. 2019) with a nearest-neighbor parameter (*k*) of 5 and Manhattan distance. We used these same choices to visualize Isomap axes in section 4.2.

In the absence of clear guidance, we chose parameters based on avoidance of clear failures (e.g., fragmentation in Isomap), computational tractability, and closeness to values used in published implementations, among other factors. In most cases, neither the magnitudes of dimensionality estimates nor the rank-order of datasets for a given method was strongly sensitive to parameter choices. For the PCA-based scree test, we log-transformed N, chlorophyll, carotenoids, LMA, and EWT from the trait dataset prior to z-standardizing them to reduce their skewness.

Supplementary Figures

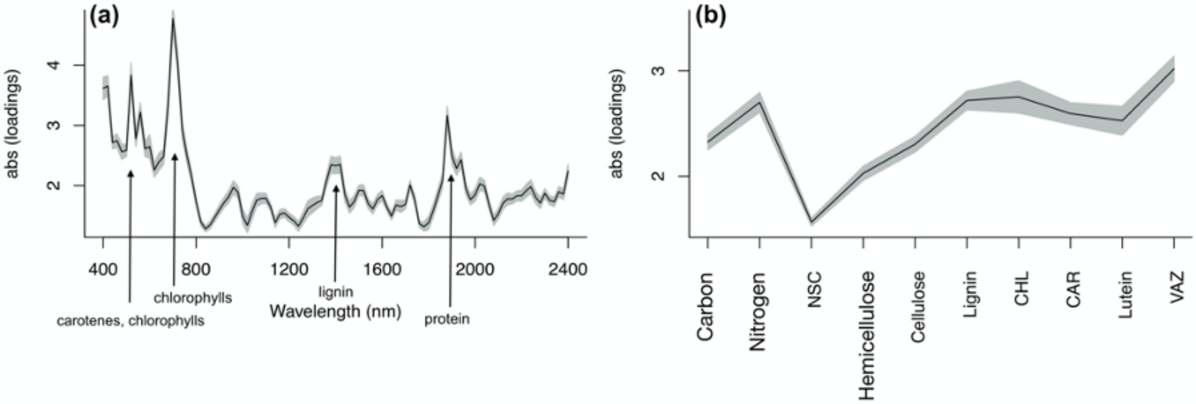

Figure S1. Influence of (a) spectral bands and (b) foliar traits for identifying 14 prairie-grassland perennials ($n = 902$) measured in the BioDIV experiment based on the absolute values of PLSDA loadings [abs(loadings)], one standard deviation is indicated in gray. The traits measured include non-structural carbohydrates (NSC), total chlorophylls (CHL), beta-carotene (CAR), and xanthophyll cycle pigments (violaxanthin + antheraxanthin + zeaxanthin = VAZ).

Supplementary Tables

Table S1. Phylogenetic signal of 10 foliar traits measured in the BioDIV experiment ($n = 902$) based on Blomberg's K. Summary statistics for Blomberg's K value (K_{trait}), the mean of the white noise model (K_{null} mean, 1000 simulations), the mean of the Brownian motion null model (K_{brown} mean, 1000 simulations), the number of simulations the white noise model K was greater than Blomberg's K ($K_{null} > K_{trait}$), the number of simulations the Brownian motion model K was greater than Blomberg's K ($K_{brown} > K_{trait}$), and the P values of observed vs. random variance of phylogenetic independent contrasts (PIC; $P < 0.05$ indicates non-random phylogenetic signal, shown in bold). See Supplementary Methods for details.

Foliar tait	**K_{trait}**	**K_{null} mean**	**K_{brown} mean**	**$K_{null} > K_{trait}$**	**$K_{brown} > K_{trait}$**	**PIC P**
Carbon %	0.4679	0.3212	1.0088	129	864	0.129
Nitrogen %	0.5342	0.3176	1.0205	53	802	**0.034**
Non-structural carbohydrates %	1.6535	0.3030	1.0022	0	143	**0.002**
Hemicellulose %	2.2057	0.3112	1.0167	1	48	**0.001**
Cellulose %	1.0187	0.3139	1.0075	7	394	**0.001**
Lignin %	0.5037	0.3144	0.9822	73	819	0.056
Chlorophyll μmol m^{-2}	0.3885	0.3114	0.9568	207	927	0.235
β-Carotene ratio	0.4310	0.3049	0.9859	103	896	0.113
Lutein ratio	0.3724	0.3070	0.9929	224	948	0.228
Xanthophyll pool ratio	0.5470	0.3104	1.0005	38	777	**0.044**

References for Supporting Information

Blomberg, S. P., Garland, T., & Ives, A. R. (2003). Testing for phylogenetic signal in comparative data: Behavioral traits are more labile. Evolution; International Journal of Organic Evolution, 57(4), 717–745.

Cattell, R. B. (1966). The Scree Test For The Number Of Factors. Multivariate Behavioral Research, 1(2), 245–276. https://doi.org/10.1207/s15327906mbr0102_10

Farahmand, A. M., Szepesvári, C., & Audibert, J.-Y. (2007). Manifold-adaptive dimension estimation. Proceedings of the 24th International Conference on Machine Learning, 265–272. https://doi.org/10.1145/1273496.1273530

Féret, J.-B., Gitelson, A. A., Noble, S. D., & Jacquemoud, S. (2017). PROSPECT-D: Towards modeling leaf optical properties through a complete lifecycle. Remote Sensing of Environment, 193, 204–215. https://doi.org/10.1016/j.rse.2017.03.004

Grassberger, P., & Procaccia, I. (1983). Characterization of Strange Attractors. Physical Review Letters, 50(5), 346–349. https://doi.org/10.1103/PhysRevLett.50.346

Hino, H. (2017). ider: Intrinsic Dimension Estimation with R. The R Journal, 9(2), 329. https://doi.org/10.32614/RJ-2017-054

Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., Blomberg, S. P., & Webb, C. O. (2010). Picante: R tools for integrating phylogenies and ecology. Bioinformatics, 26(11), 1463–1464. https://doi.org/10.1093/bioinformatics/btq166

Kothari, S., Beauchamp-Rioux, R., Laliberté, E., & Cavender-Bares, J. (2021). Reflectance spectroscopy allows rapid, accurate, and non-destructive estimates of functional traits from pressed leaves (p. 2021.04.21.440856). bioRxiv. https://doi.org/10.1101/2021.04.21.440856

Kothari, S., Cavender-Bares, J., Bitan, K., Verhoeven, A. S., Wang, R., Montgomery, R. A., & Gamon, J. A. (2018). Community-wide consequences of variation in photoprotective physiology among prairie plants. Photosynthetica, 1–13. https://doi.org/10.1007/s11099-018-0777-9

Kuhn, M. (2008). Building Predictive Models in R Using the caret Package. Journal of Statistical Software, 28, 1–26. https://doi.org/10.18637/jss.v028.i05

Laliberté, E., & Soffer, R. (2018a). Measuring spectral reflectance and transmittance (350-2500 nm) of large leaves using the Spectra Vista Corporation (SVC) DC-R/T Integrating Sphere. https://doi.org/10.17504/protocols.io.p8pdrvn

Laliberté, E., & Soffer, R. (2018b). Measuring spectral reflectance and transmittance (350-2500 nm) of small and/or narrow leaves using the Spectra Vista Corporation (SVC) DC-R/T Integrating Sphere. https://doi.org/10.17504/protocols.io.q56dy9e

Lehnert, L. W., Meyer, H., Obermeier, W. A., Silva, B., Regeling, B., & Bendix, J. (2019). Hyperspectral Data Analysis in R: The hsdar Package. Journal of Statistical Software, 89(12). https://doi.org/10.18637/jss.v089.i12

Levina, E., & Bickel, P. (2004). Maximum Likelihood Estimation of Intrinsic Dimension. Advances in Neural Information Processing Systems, 17. https://proceedings.neurips.cc/paper/2004/hash/74934548253bcab8490ebd74afed7031-Abstract.html

MacKay D. J.C., Ghahramani Z.(2005). Comments on “Maximum Likelihood Estimation of Intrinsic Dimension” by E. Levina and P. Bickel. http://www.inference.phy.cam.ac.uk/mackay/dimension/

Meireles, J., Schweiger, A., & Cavender-Bares, J. (2017). spectrolab: Class and Methods for Hyperspectral Data in R. R package version 0.0.10. https://CRAN.R-project.org/package=spectrolab

Mendiburu, F. D., & Simon, R. (2015). Agricolae—Ten years of an open source statistical tool for experiments in breeding, agriculture and biology (No. e1748). PeerJ Inc. https://doi.org/10.7287/peerj.preprints.1404v1

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O’Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., & Wagner, H. (2019). vegan: Community Ecology Package. R package version 2.5-5. https://CRAN.R-project.org/package=vegan

Osnas, J. L. D., Lichstein, J. W., Reich, P. B., & Pacala, S. W. (2013). Global Leaf Trait Relationships: Mass, Area, and the Leaf Economics Spectrum. Science, 340(6133), 741–744. https://doi.org/10.1126/science.1231574

Pettis, K. W., Bailey, T. A., Jain, A. K., & Dubes, R. C. (1979). An Intrinsic Dimensionality Estimator from Near-Neighbor Information. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-1(1), 25–37. https://doi.org/10.1109/TPAMI.1979.4766873

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Revell, L. J. (2012). phytools: An R package for phylogenetic comparative biology (and other things). Methods in Ecology and Evolution, 3(2), 217–223. https://doi.org/10.1111/j.2041-210X.2011.00169.x

Rozza, A., Lombardi, G., Ceruti, C., Casiraghi, E., & Campadelli, P. (2012). Novel high intrinsic dimensionality estimators. Machine Learning, 89(1), 37–65. https://doi.org/10.1007/s10994-012-5294-7

Schweiger, A., & Laliberté, E. (2020). Processing of leaf spectra. https://doi.org/10.17504/protocols.io.bhsdj6a6

Tenenbaum, J. B., Silva, V. de, & Langford, J. C. (2000). A Global Geometric Framework for Nonlinear Dimensionality Reduction. Science, 290(5500), 2319–2323. https://doi.org/10.1126/science.290.5500.2319

Tilman, D., Knops, J., Wedin, D., Reich, P., Ritchie, M., & Siemann, E. (1997). The Influence of Functional Diversity and Composition on Ecosystem Processes. Science, 277(5330), 1300–1302. https://doi.org/10.1126/science.277.5330.1300

Venables, W. N. and Ripley, B. D. (2002). Modern Applied Statistics with S. New York: Springer (4th ed)