Rapid estimates of leaf litter chemistry using reflectance spectroscopy

Shan Kothari[1,2,4], Sarah E. Hobbie[3], Jeannine Cavender-Bares[1,3]

[1]Department of Plant and Microbial Biology, University of Minnesota, 1479 Gortner Ave, St. Paul, MN, United States, 55108
[2]Centre d'étude de la forêt, Université du Québec à Montréal, Montréal, QC, Canada H2X 3Y7
[3]Department of Ecology, Evolution, and Behavior, University of Minnesota, 1479 Gortner Ave, St. Paul, MN, United States, 55108
[4]Corresponding author: kotha020@umn.edu

Abstract

Measuring the chemical traits of leaf litter is important for understanding plants' roles in nutrient cycles, including through nutrient resorption and litter decomposition, but conventional leaf trait measurements are often destructive and labor-intensive. Here, we develop and evaluate the performance of partial least-squares regression (PLSR) models that use reflectance spectra of intact or ground leaves to estimate leaf litter traits, including carbon and nitrogen concentration, carbon fractions, and leaf mass per area (LMA). Our analyses included more than 300 samples of senesced foliage from 11 species of temperate trees, including needleleaf and broadleaf species. Across all samples, we could predict each trait with moderate-to-high accuracy from both intact-leaf litter spectra (validation R^2 = 0.543-0.941; %RMSE = 7.49-18.5) and ground-leaf litter spectra (validation R^2 = 0.491-0.946; %RMSE = 7.00-19.5). Ground-leaf spectra yielded better estimates of mass-based chemical traits, while intact-leaf spectra yielded better predictions of LMA. Our results support the feasibility of building models to estimate multiple chemical traits from leaf litter of a range of species. In particular, the success of intact-leaf spectral models allows non-destructive trait estimation in a matter of seconds, which could enable researchers to measure the same leaves over time in studies of nutrient resorption.

Keywords: decomposition; functional traits; nutrient cycling; nutrient resorption; partial least-squares regression (PLSR); reflectance spectroscopy

Introduction

Long-lived plants resorb and store a large fraction of the nutrients from their leaves before they senesce. These reserves support early growth in the following growing season and reduce the need for nutrient uptake from the soil (El Zein et al. 2011). But nutrient resorption (much like soil nutrient uptake) has metabolic and physiological costs, so the optimal level of nutrient resorption varies among environments (Wright & Westoby 2003). Accordingly, species vary broadly in their nutrient resorption efficiency—the fraction of a given nutrient that is resorbed (Vergutz et al. 2012). For example, plants often (Kobe et al. 2005; Hayes et al. 2014; Yuan & Chen 2015), though not always (Diehl et al. 2003), resorb a greater fraction of leaf nutrients when soil nutrients are scarce. This adjustment of resorption efficiency in response to the environment highlights the key role of resorption in plant nutrient economics.

Resorption often leaves the resulting leaf litter relatively nutrient-poor, which can limit the rate of decomposition of newly shed litter by soil microorganisms (Berg 2014). Hence, nutrient-poor litter decomposes more slowly than nutrient-rich litter at first (Cornwell et al. 2008), although this trend can reverse in later stages (Prescott 2010; Berg 2014; Gill et al. 2021). Because litter nutrient levels influence the nutrient flux into soil inorganic or microbial pools via decomposition, resorption may be a key process for explaining how plants influence their environment (Hobbie 2015). Despite the metabolic changes that occur during senescence, relationships between fresh-leaf and leaf litter composition may help explain why soil processes can often be inferred from remote sensing of aboveground vegetation (Madritch et al. 2020, Cavender-Bares et al. 2022).

Addressing many ecological questions about nutrient resorption and litter decomposition requires measuring litter chemistry. Because measuring resorption efficiency requires the nutrient content of leaves both before and after senescence, it can be expensive and labor-intensive. Furthermore, because measuring nutrient concentration through elemental analysis is destructive, the same exact leaves cannot be measured at each step. Decomposition also depends on litter traits beyond macronutrient

concentrations, including lignin, cellulose, condensed tannins, and various micronutrients (Schweitzer et al. 2004; Cornwell et al. 2008; Talbot & Treseder 2012; Keiluweit et al. 2015; Bourget et al. 2023). The number of traits and samples can add up to make even simple projects expensive and time-consuming.

Such complications may make it harder to understand and predict the fluxes of carbon and mineral nutrients between plants, litter, and soil—the determinants of which are still far from well-understood. For example, large-scale comparative tests of the role of nutrient resorption in plant life history or nutrient economics are still rare (e.g., Freschet et al. 2010; Rea et al. 2018). Meanwhile, research continues to reveal new paths through which a plant community's litter traits or diversity interact with microbial communities and the abiotic environment to influence decomposition rates (Liu et al. 2020; Grossman et al. 2020; Gill et al. 2021; Bourget et al. 2023). Fast, simple estimates of litter traits could make it easier both to study the functional significance of nutrient resorption and to estimate litter decomposability rapidly across taxa and ecosystems.

The need for faster trait estimates has led some plant ecologists to turn to reflectance spectroscopy, a technique that involves measuring the reflectance of light from a sample across many wavelengths. Many kinds of samples, including leaf tissue, are analyzed using the 400-2400 nm range. Reflectance in this range can be used to estimate many leaf structural and chemical traits because those traits influence how leaves scatter and absorb light across wavelengths.

Many widely studied functional traits nevertheless cannot be identified with specific structures or chemical constituents that have unique absorption features. Leaf nitrogen concentration (N_{mass}), for example, includes nitrogen in molecules as varied as RuBisCO and chlorophyll. This challenge is among the reasons that many plant scientists have adopted a multivariate empirical approach to linking traits and reflectance spectra using statistical techniques like partial least-squares regression (PLSR; Burnett et al. 2021). Researchers have built PLSR models predicting traits like N_{mass} and leaf mass per area (LMA) with

a high degree of accuracy from spectra, measured either while the leaf was fresh or after drying and grinding it into powder (Serbin et al. 2014; Serbin et al. 2019; Kothari et al. 2023a; Kothari et al. 2023b). Further research has shown that spectral models can predict traits as varied as defense compounds (Couture et al. 2016), water status (Cotrozzi et al. 2017; Kothari et al. 2023a), leaf age (Chavana-Bryant et al. 2017), and photosynthetic capacity (Dechant et al. 2017).

The vast majority of this research is based on measurements taken during the growing season, before the onset of leaf senescence. These models are not expected to accurately predict the traits of senesced leaves, which are well outside the scope of their training data. Some researchers have built models to predict decomposition-related traits (Joffre et al. 1992; McTiernan et al. 2003; Coûteaux et al. 2005; Hobbie 2005; Petit Bon et al. 2020) or the decomposition rate itself (Gillon et al. 1999; Fortunel et al. 2009) from ground leaf litter using near-infrared (NIR) reflectance spectroscopy. Many traits that influence decomposition also influence forage quality, and there is an extensive body of research predicting forage traits using NIR spectroscopy, including in senescent foliage (e.g., Norris et al. 1976). Such studies have often shown high accuracy ($R^2 > 0.9$) in predicting traits like N_{mass}, carbon fractions, and phenolics, albeit usually using a limited number (<120) of calibration samples from one or a few species.

Here, we expand on prior work by building models to predict traits related to litter quality and recalcitrance from reflectance spectra. Compared to previous studies, we calibrate and validate our models using more samples from more species. This breadth allows us to make predictions over a wide range of trait values and test how well models predict both intra- and interspecific trait variation. We also measured spectra from the same samples both while intact and ground. Existing models to predict leaf litter traits require the destructive and time-consuming step of grinding the tissue, which motivated us to test whether spectra of whole, intact litter could also yield accurate estimates. We predicted that ground-leaf litter spectra would provide more accurate estimates of chemical traits than intact-leaf litter spectra, as most comparisons on non-senesced leaves have shown (Serbin et al. 2014; Couture et al. 2016; Kothari

et al. 2023b). On the other hand, we predicted that intact-leaf litter spectra would outperform ground-leaf litter spectra for estimating structural traits like LMA because grinding destroys the leaf structure (Kothari et al. 2023b).

Finally, we verified whether our litter chemistry estimates could capture known differences among species. We predicted that conifer litter would have greater LMA, recalcitrant carbon, and C_{mass} and lower N_{mass} than broadleaf litter. These patterns are thought to explain much of why conifer litter tends to show slower initial decomposition rates than broadleaf litter (Cornelissen et al. 1999; Cornwell et al. 2008; Prescott 2010).

Methods

Samples

During fall 2018, we collected senescent leaf litter from the Forests and Biodiversity experiment (FAB1) at Cedar Creek Ecosystem Science Reserve (East Bethel, MN, USA; Grossman et al. 2017). In May 2013, one- to two-year-old trees were planted in a 0.5 × 0.5 m grid within plots that varied in their species richness (1, 2, 5, or 12 species) and composition. The species pool included eight broadleaf angiosperms and four needleleaf conifers (Table 1). We left out one species (*Juniperus virginiana*) from sampling because its scaly needle morphology meant that individual trees seldom produced enough litter at once for spectroscopic or chemical analyses. To get enough samples of another species (*Acer negundo*), we also collected material from mature trees in a residential yard in Minneapolis, MN, USA 50.0 km from FAB1.

We collected litter samples as close as we feasibly could to the time of senescence using both litterfall traps on the ground and hand-collection of senesced leaves from trees. We only took leaves from trees when the petiole had a complete abscission layer and the leaf could be detached with a gentle pull (Chapin & Moilanen 1991). Because tree species dropped their leaves at different times, we collected litter repeatedly from September to November 2018 to ensure that the litter we collected had recently

senesced. To avoid leaching we discarded litter samples from traps after rain, although nitrogen and carbon leaching tends to be low even under extreme wetting (Schreeg et al. 2013).

We collected 322 litter samples, seeking to represent the variation in leaf traits within each species (Table 1). While our sampling did not explicitly account for the species richness and composition treatments in FAB1, we collected leaves from each species across neighborhoods and crown positions to enhance intraspecific variation. Each sample included 1.5-3.0 g of tissue and could comprise one leaf or multiple similar leaves. When a single trap contained more than 2.5-3.0 g of leaves from a single species, we divided them into two or more samples by keeping together leaves that we judged by eye to be more visually similar. We dried each sample in a dark shed at 40 °C for at least three days, and then stored them in a cool, dark room until spectral and trait measurements.

Table 1: Details of species in the Forests and Biodiversity (FAB1) experiment. All broadleaf species listed are deciduous angiosperms, and all needleleaf species are evergreen conifers. We did not sample *J. virginiana*. Twelve of the *A. negundo* samples come from a residential yard rather than FAB1. One *B. papyrifera* sample included in the tally below was omitted from ground litter analyses due to lack of tissue.

Species	Code	Functional group	# samples
Acer negundo	ACNE	Broadleaf	27
Acer rubrum	ACRU	Broadleaf	28
Betula papyrifera	BEPA	Broadleaf	36
Juniperus virginiana	JUVI	Needleleaf	n/a
Pinus banksiana	PIBA	Needleleaf	20
Pinus resinosa	PIRE	Needleleaf	31
Pinus strobus	PIST	Needleleaf	26
Quercus alba	QUAL	Broadleaf	37
Quercus ellipsoidalis	QUEL	Broadleaf	33
Quercus macrocarpa	QUMA	Broadleaf	34
Quercus rubra	QURU	Broadleaf	33
Tilia americana	TIAM	Broadleaf	17
Total			322

Spectral data collection

We measured reflectance spectra of each dried sample, both intact and ground, using a PSR+ 3500 full-range (350-2500 nm) spectrometer (Spectral Evolution, Lawrence, MA, USA). For intact-leaf spectra, we used the Spectral Evolution leaf clip assembly. We used a built-in 99% Spectralon white reflectance standard to calibrate readings before each sample. Among broadleaf species, many of the leaves curled in on themselves during senescence and drying; when possible, we sought to measure spectra in relatively flat portions of the leaf to minimize specular reflection (Petibon et al. 2021). Among needleleaf species, we made mats by laying out needles in a single layer; because of their shape, there often remained small gaps or areas of overlap between needles. We discarded spectra with large discrepancies in sensor overlap regions, or (particularly for needleleaf species) with very high (>70%) or very low (<30%) peak reflectance, which often indicated that the needle mat was highly non-uniform. There remained at least three reflectance spectra for each sample, and more for most of the broadleaf samples that had multiple leaves.

After removing their petioles, we ground leaf samples by placing them in individual plastic vials with ball bearings and shaking them with a paint shaker for several hours. To measure ground-leaf spectra, we used the Spectral Evolution benchtop reflectance probe and glass-windowed sample trays. Before measuring each sample, we measured a white Spectralon panel placed in a sample tray for calibration. For sample measurements, we put at least 0.6 g of leaf powder into a sample tray. Preliminary tests showed that adding more material did not change the spectra, suggesting that transmittance was close to zero. The benchtop probe pressed the loose powder into an even pellet. We measured three spectra per sample—turning the sample tray 120° between the first two measurements, and loosening and mixing the powder between the second and third measurements.

The spectrometer automatically interpolated the spectra to 1 nm resolution. We performed all subsequent processing using *spectrolab v. 0.0.18* (Meireles et al. 2023) in *R v. 3.6.1* (R Core Team, 2019). Because

the 350-400 nm and 2400-2500 nm regions could be noisy, we trimmed each spectrum to 400-2400 nm. For intact leaves only, we corrected discontinuities at the overlap region between the Si and first InGaAs detectors (970-1000 nm) using the *match_sensors()* function, and spline-smoothed only this region using the *smooth_spline()* function. Finally, we averaged all intact spectra and all ground spectra for a given sample. As a metric of measurement consistency, we calculated the mean spectral angle between measurements of the same sample (Kruse et al. 1993). Intact measurements (median across samples: 6.8°; 2.5^{th}-97.5^{th} percentile: 2.9°-20.4°) showed much less consistency than ground measurements (median: 0.6°; 2.5^{th}-97.5^{th} percentile: 0.2°-3.2°).

Leaf mass per area measurements

Researchers risk underestimating resorption efficiency when they fail to account for leaf mass loss during resorption. To avoid this risk, many studies either express nutrient content of non-senesced and senesced leaves on a per-area basis, or equivalently include a mass loss correction factor (van Heerwaarden et al. 2003). We measured LMA of each litter sample to enable conversion between a per-mass and per-area basis. We did not attempt to quantify or correct for the smaller bias sometimes caused by leaf shrinkage (van Heerwaarden et al. 2003).

For broadleaf species, we could not measure the area of whole leaves because many were curled and brittle. Instead, we used a hole punch to take four to six 0.3 cm^2 disks from each dried leaf sample. We sought to capture the variation within the sample, including veins. We calculated LMA as the total mass divided by the total area of all disks. This method has generally been shown to provide accurate estimates of the LMA of whole leaves (Perez et al. 2020).

For needleleaf species, we measured the mass and area of five needles per sample. We estimated area as the length times the maximum width of the needle measured using digital calipers. We then measured the mass of each needle and calculated LMA as the total mass divided by the total area.

Chemical trait measurements

We measured leaf carbon and nitrogen concentration (C_{mass} and N_{mass}) on ground, oven-dried samples using dry combustion gas chromatography performed by an elemental analyzer (Costech ECS 4010 Analyzer, Valencia, California, USA). We converted C_{mass} and N_{mass} to an area basis (C_{area} and N_{area}) by multiplying them with LMA.

We measured carbon fractions on ground, oven-dried samples using an ANKOM 200 Fiber Analyzer (Ankom Technology, Macedon, New York, USA). The first digestion in a hot neutral detergent solution washes out plant solubles. The second digestion in a hot acidic detergent solution further washes out bound proteins and hemicellulose (henceforth 'hemicellulose' for simplicity). The remaining fraction comprises cellulose and acid-unhydrolyzable residue (AUR) such as lignin (henceforth 'recalcitrant carbon'). We estimated each component using the changes in sample mass between steps. Because the paint shaker ground our samples into a very fine powder, we had to use small-pored ANKOM F58 fiber bags, which cannot be used in the final acid detergent lignin digestion to measure AUR concentration alone. We also did not determine the ash concentration of the samples.

We removed some trait values prior to statistical analyses because of clear issues during trait data collection—for example, the rupture of sample bags while measuring carbon fractions. Unlike some other studies (e.g., Gillon et al. 1999), we only removed samples due to known issues with trait measurements, not simply because they had large prediction residuals during statistical analyses.

Statistical analyses

We modeled the relationship between traits and spectra using partial least-squares regression (PLSR), which is well-suited to handle datasets that include many collinear predictors (Burnett et al. 2021). PLSR reduces the full matrix of spectra to a smaller number of orthogonal latent components that best explain

the variation in the variables to be predicted. We implemented our PLSR models in *R* package *pls v. 2.7.1* (Mevik et al. 2019).

Previous studies using a PLSR modeling framework have often restricted the range of wavelengths used to predict certain traits in order to select bands that, based on prior research or known absorption features, may be most causally linked to the traits in question (e.g., Serbin et al. 2014). In our main set of analyses, we used the full spectrum (400-2400 nm) to predict each trait on the grounds that biochemical changes during senescence may alter the relationship between integrative functional traits like N_{mass} and specific optical features. However, we also explored two ways of restricting the spectral range for intact-leaf spectra only. First, we built models using only 400-1000 nm ('VIS/NIR models') given that many less-expensive spectrometers only measure this range. Second, we built models using only 1300-2400 nm ('SWIR models'). The pigment composition of leaves changes rapidly during senescence, including breakdown products ('brown pigments') which may absorb up to 1300 nm (Fourty et al. 1996). The SWIR models could perform better in cases where pigments are poor indicators of the overall composition of the leaf.

Researchers may transform reflectance of ground-leaf spectra to pseudoabsorbance (*-log R*; Blackburn 1998; Serbin et al. 2014) or its derivatives, which may somewhat linearize the relationship between traits and spectral features and make those features more prominent. We found in preliminary tests that these transformations did not substantially improve model performance. Likewise, brightness normalization of intact-leaf spectra (Feilhauer et al. 2010) did not have a strong influence on model performance.

We divided our samples into calibration (75%) and validation (25%) datasets, stratified such that the proportions of each species were identical in each dataset. We fit models for each trait on the complete calibration dataset, using 10-fold cross-validation to select the optimal number of model components to use in further analyses while avoiding overfitting. We selected the smallest number of components whose

cross-validation root mean squared error of prediction (RMSEP) was within one standard deviation of the very lowest RMSEP at any number of components. We used these models to calculate the variable importance in projection metric (VIP) for each trait (Wold et al. 1994) but we do not report their accuracy.

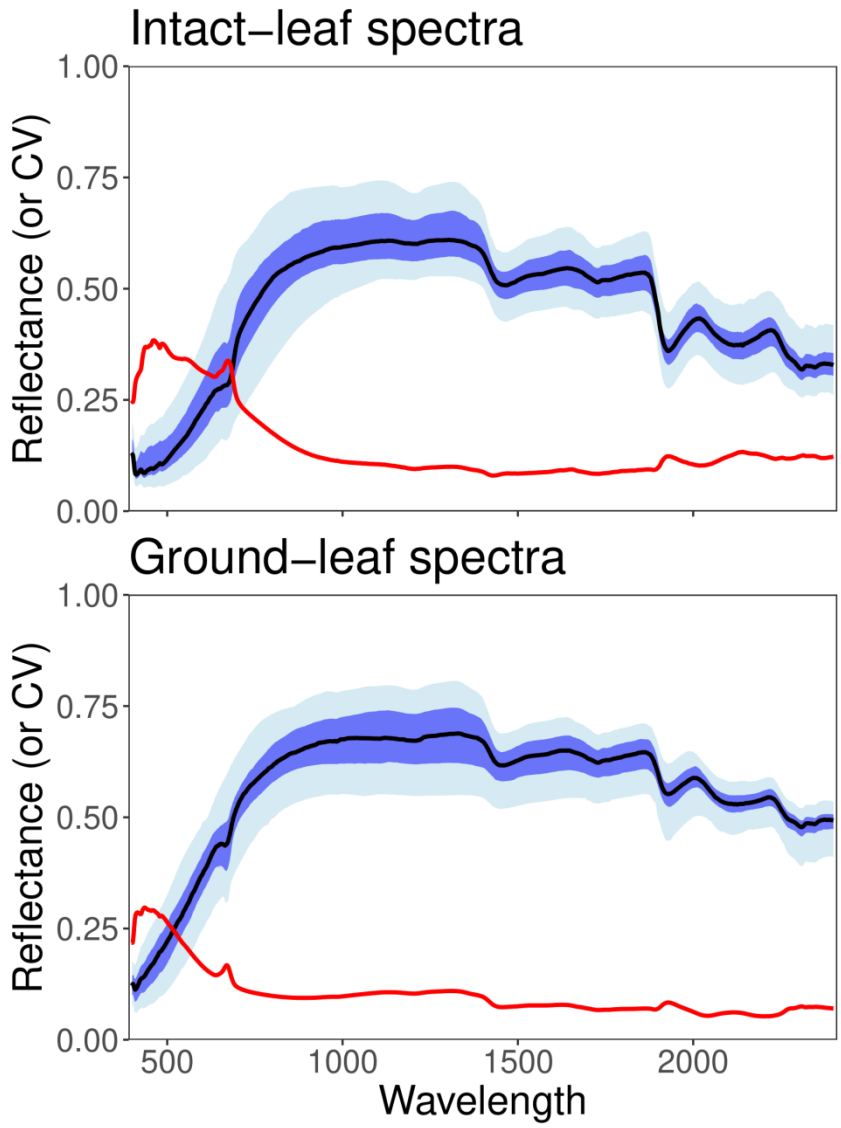
We developed our main set of models using a resampling procedure in which we divided the calibration data at random 200 times into 70% training and 30% testing subsets (Burnett et al. 2021). During each resampling iteration, we trained a model for each trait on the 70% and assessed its performance (RMSE and R^2) on the remaining 30%, using the previously determined number of components for that trait. This procedure left us with an ensemble of 200 models for each trait, which could be used both to make trait predictions and to examine the variability of model performance under random variation in training data.

We used these ensembles of models to predict traits from the validation (25%) dataset. We quantified model performance for each trait by calculating R^2 and root mean squared error (RMSE) between the predictions (averaged for each validation sample across the 200 estimates) and the measured values. We also report the %RMSE, calculated as the RMSE divided by the 2.5% trimmed range of the measured values in the validation data (Kothari et al. 2023a).

Table 2: Summary statistics of PLSR model validation based on full (400-2400 nm) intact- and ground-leaf spectra. The range column shows the complete trait range across the full dataset. %RMSE is calculated as RMSE divided by the 2.5% trimmed range of data within the validation data set.

		Intact				Ground			
Litter trait	Range	No. comps	R^2	RMSE	%RMSE	No. comps	R^2	RMSE	%RMSE
LMA (g m^{-1})	39.5-282	11	0.941	15.3	7.49	13	0.842	25.5	12.5
C_{mass} (%)	40.9-61.4	12	0.764	1.82	14.0	12	0.689	2.11	16.3
N_{mass} (%)	0.24-2.57	15	0.919	0.136	8.58	19	0.946	0.111	7.00
Solubles (%)	27.6-71.4	10	0.794	4.43	14.0	11	0.808	4.33	13.7
Hemicellulose (%)	5.72-26.0	8	0.543	2.79	18.5	8	0.491	2.94	19.5
Recalcitrant carbon (%)	20.7-58.7	11	0.881	2.96	10.0	13	0.887	3.03	10.2
C_{area} (g m^{-1})	16.2-172	10	0.940	9.20	8.02				
N_{area} (g m^{-1})	0.272-2.43	12	0.766	0.188	13.2				

Fig. 1: Distributions of spectral reflectance and its coefficient of variation (CV) among intact (top) and ground (bottom) leaf litter samples. The black line represents the median reflectance at each wavelength, flanked by the 25-75 percentile region (dark blue), and the 2.5-97.5 percentile region (light blue). The solid red line shows the coefficient of variation among spectra at each wavelength.



Results

There was a high level of variation in each trait both across the entire dataset and within species. For example, N_{mass} varied nearly 11-fold; in contrast, C_{mass} only varied about 1.5-fold (Table 2). In spite of the high level of intraspecific variation, especially among broadleaf species, species identity alone could account for between 46.0% (N_{area}) and 93.7% (LMA) of the variation in each trait. In general, needleleaf species had much higher LMA, C_{mass}, recalcitrant carbon, and lower N_{mass} than broadleaf species (Table S1), consistent with trends in non-senescent leaves (Serbin et al. 2014; Kothari et al. 2023a). Consequently, we see correlations among pairs of leaf economic traits across the whole dataset (for log(LMA) and log(N_{mass}), $R^2 = 0.592$; $p < 10^{-15}$), but a weaker correlation among just needleleaf samples ($R^2 = 0.255$; $p < 10^{-5}$) and none among just broadleaf samples ($R^2 = 0.011$; $p = 0.111$).

Among both ground-leaf and intact-leaf spectra, the coefficient of variation (CV) was highest in the visible range and decreased into the near-infrared range (NIR; Fig. 1). For intact-leaf spectra only, the CV once again increased into the short-wave infrared range (SWIR). There was a global maximum CV around 440-490 nm for intact spectra, shifted to lower wavelengths for ground spectra; there were also local maxima centered at around 670-675 nm for both kinds of spectra. Both maxima lie near absorption peaks of chlorophylls, and reflectance at 670 nm is particularly sensitive to variation in chlorophyll at low levels, such as during late senescence (Gitelson & Merzlyak 1996).

Each of our traits could be predicted with moderate-to-high success from full intact-leaf and ground-leaf spectra (Figs. 2-4; Table 2). The models for N_{mass} and (especially for intact-leaf spectra) LMA achieved the highest calibration and validation accuracy, followed by recalcitrant carbon, solubles, C_{mass}, and finally hemicellulose (Table 2). There was no general tendency for estimation of chemical traits to be better using intact- or ground-leaf spectra, but estimation of LMA was better using intact-leaf spectra. Because we used LMA to convert from a mass to an area basis, we built models predicting N_{area} and C_{area} (nitrogen and carbon on a per-area basis) only from intact-leaf spectra (Fig. 4). The C_{area} model (%RMSE

= 8.02) had better performance than the C_{mass} model (%RMSE = 14.0), but the N_{area} model (%RMSE = 13.2) had worse performance than the N_{mass} model (%RMSE = 8.58; Table 2). Jackknife analyses showed that traits for which the ensemble model performance was poorest—such as hemicellulose, C_{mass}, and N_{area}—also had the greatest variability in model performance (Figs. S1 and S2).

For every trait prediction model trained on the full spectrum, the VIP metric also showed a global maximum around 677 nm (intact) or 672 nm (ground; Fig. 5), next to one of the aforementioned chlorophyll absorption peaks. This pattern suggests that the amount of chlorophyll remaining at abscission may be closely related to many traits, especially N_{mass}. More generally, the visible range was most important for predicting all traits, and the NIR range was also important for LMA. There were further peaks in the SWIR around 1430-1450 nm (for intact leaves), 1910-1950 nm and 2130-2220 nm. In fresh tissue, variation in reflectance around 1910-1950 nm is often interpreted as a measure of absorption by water, but the trace amounts of water remaining in our dried samples likely have no direct biological relevance. In general, reflectance in the visible range and parts of the SWIR range seemed most informative in predicting nutrients and carbon fractions in litter.

Our VIS/NIR models for intact leaves had considerably worse performance than our full-range models for every trait, particularly N_{mass}, N_{area}, and solubles. By contrast, our SWIR models had very similar performance to the full-range models (Table S2). Both sets of restricted-range models had a similar rank-ordering of model performance across traits to each other and the full models. Moreover, the VIP metric for these models revealed that the regions that were most important for these predictions were mainly the same bands in their respective regions that were important for the full-range models; however, for the VIS/NIR models, very low wavelengths close to 400 nm also had high importance.

Fig. 2: Validation results for predictions of solubles, hemicellulose, and recalcitrant carbon from intact-leaf litter spectra (left) and ground-leaf litter spectra (right). In each panel, a separate OLS regression line is shown for each of the 11 species, overlaid on top of the thick dashed 1:1 line. The error bars for each data point are 95% intervals calculated from the distribution of predictions based on the model coefficients from the 200 jackknife iterations. See Table 1 for species codes.

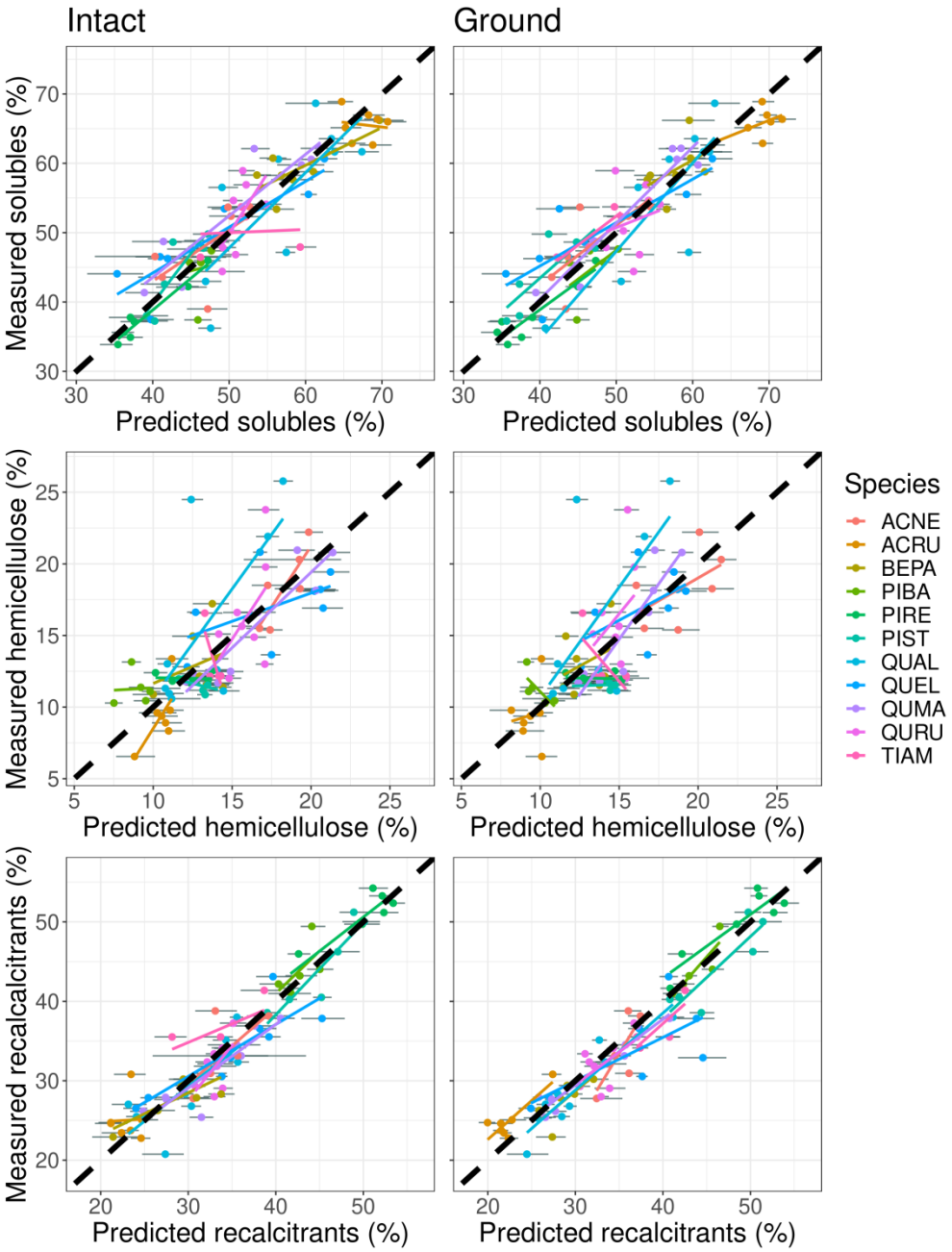


Fig. 3: Validation results for predictions of C_{mass}, N_{mass}, and LMA from intact-leaf litter spectra (left) and ground-leaf litter spectra (right). Regression lines and error bars are as in Fig. 2. See Table 1 for species codes.

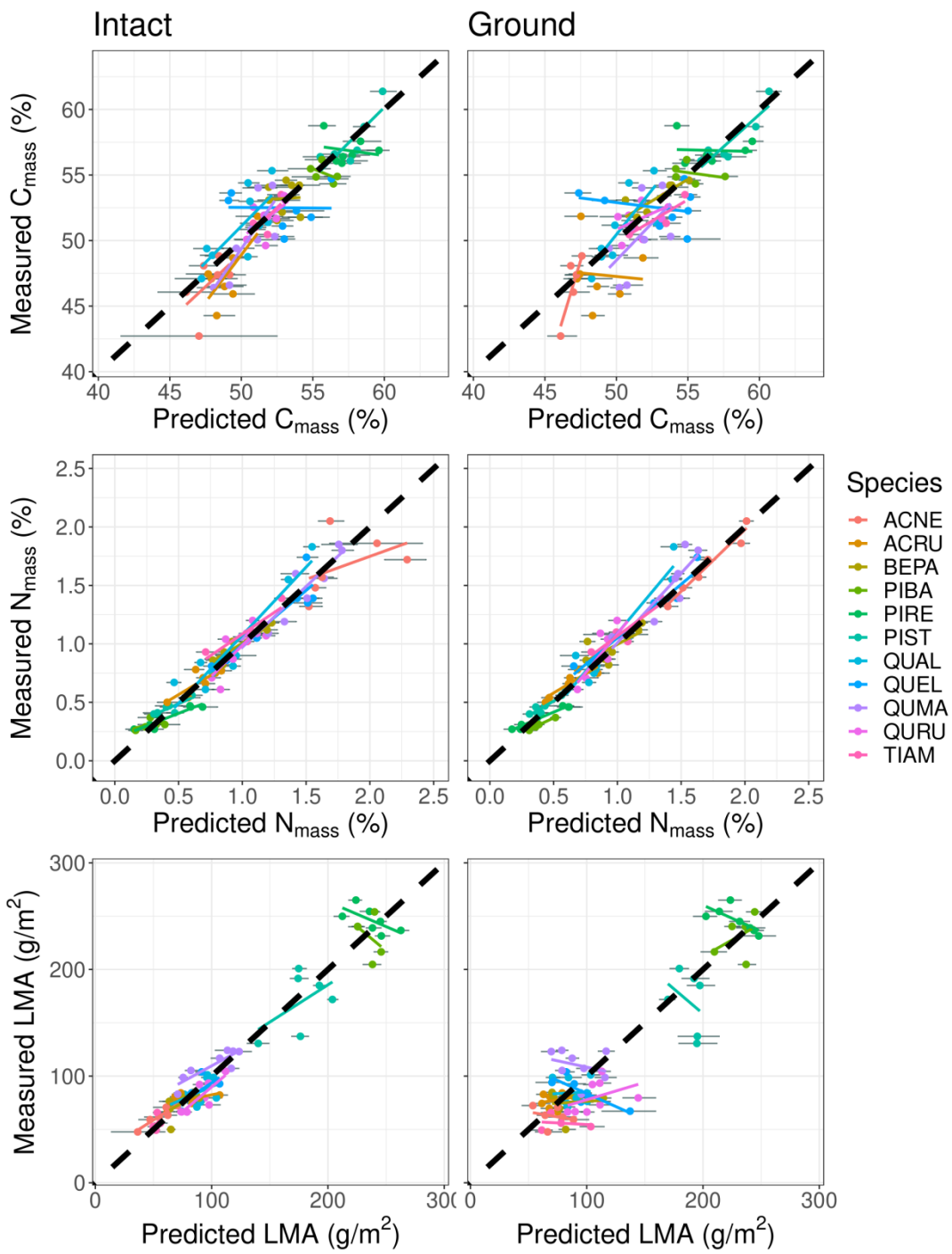


Fig. 4: Validation results for predictions of C_{area} and N_{area} from intact-leaf litter spectra. Regression lines and error bars are as in Fig. 2. See Table 1 for species codes.

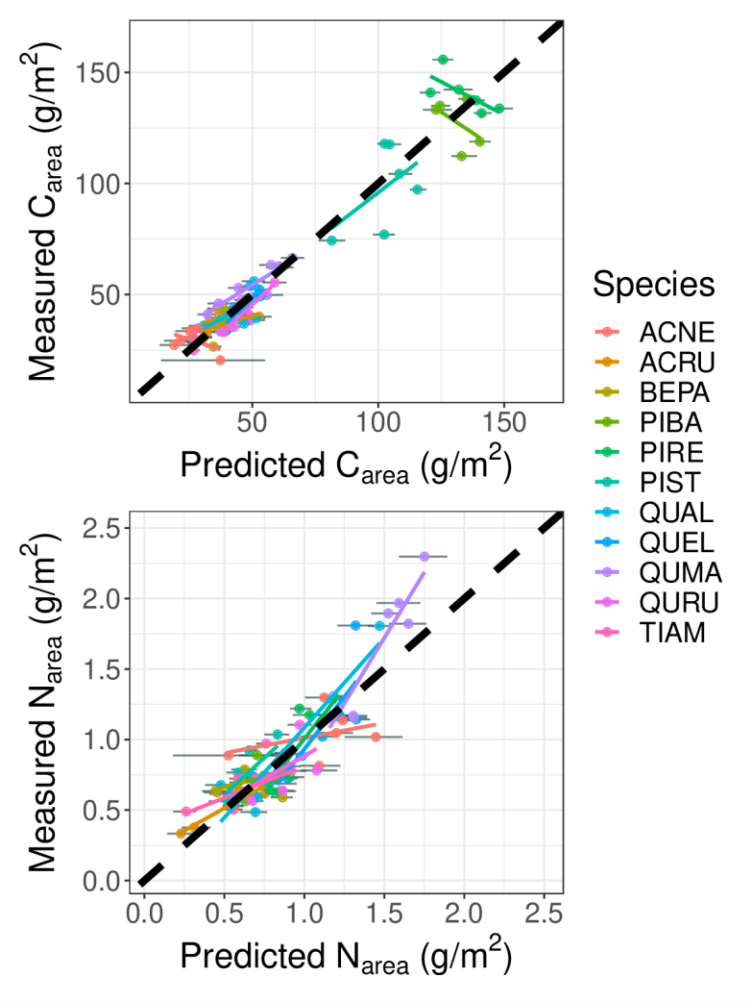
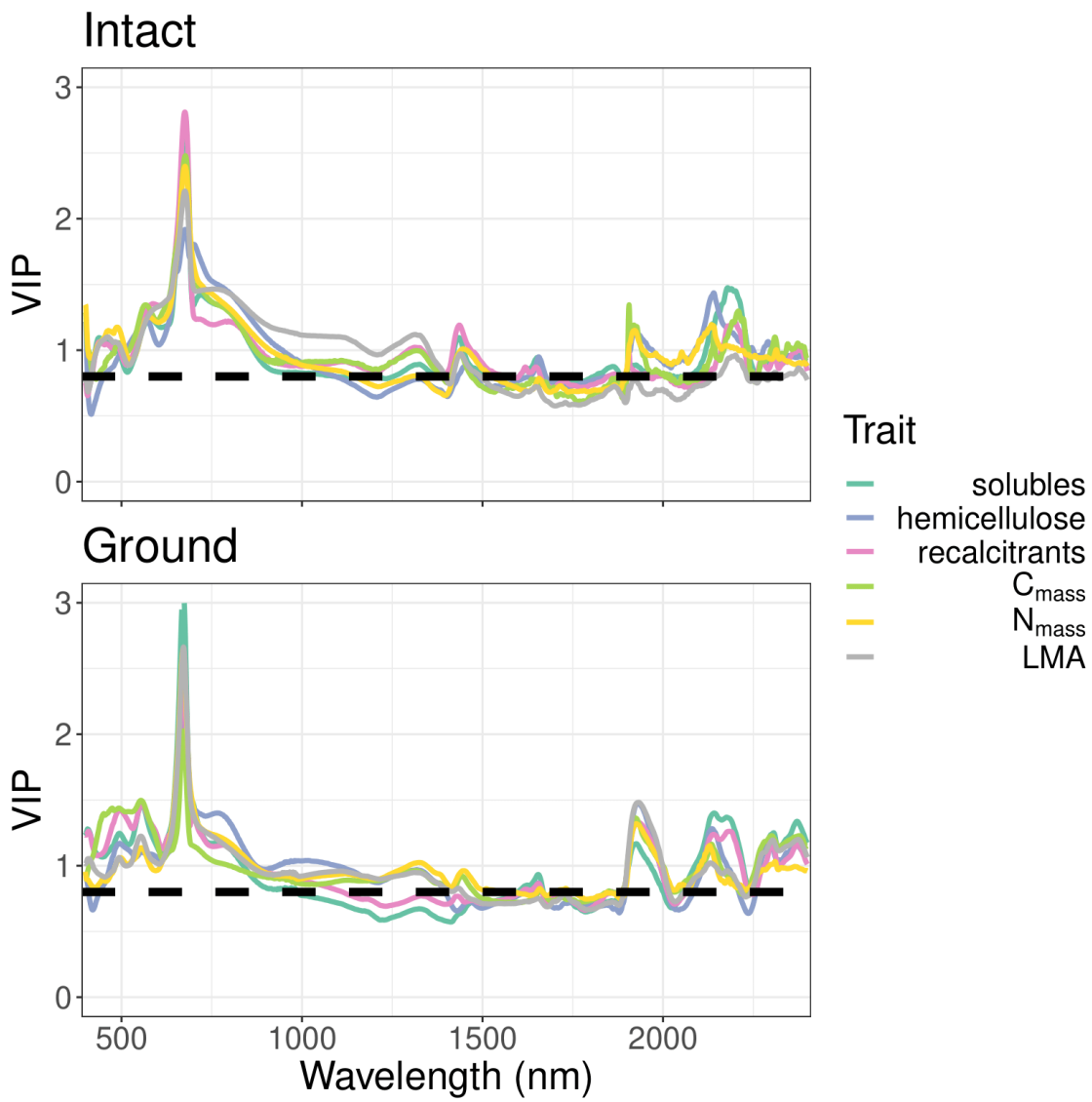


Fig. 5: The variable importance of prediction (VIP) metric in intact (top) and ground (bottom) models. The dashed black line represents a threshold of 0.8 suggested as a rough heuristic of importance by Burnett et al. (2021).



Discussion

We show that PLSR models trained on reflectance spectra can accurately predict several leaf litter traits related to nutrient cycling across temperate broadleaf and needleleaf tree species. For each trait, a single model captures the variation among and within our 11 species. The models' predictions capture well-known ecological differences among species—for example, that needleleaf species have more recalcitrant carbon and lower N_{mass} than broadleaf species. Importantly, intact-leaf models are about as accurate as ground-leaf models for most traits, which suggests that multiple traits can be estimated from a single measurement that takes only seconds.

Model performance and interpretation

The hierarchy of which traits can be predicted best mirrors results from non-senesced leaves, including fresh ones (Asner et al. 2011; Nunes et al. 2017; Kothari et al. 2023a). Leaf mass per area can be predicted very well from intact-leaf spectra in particular, which has been shown in both fresh leaves (Serbin et al. 2019) and in pressed, dried leaves (Costa et al. 2018; Kothari et al. 2023b). Predictions of LMA from ground leaves are worse, but still quite good (Serbin et al. 2014). N_{mass} can also be predicted nearly perfectly from intact- and ground-leaf spectra (Richardson & Reeves 2005; Serbin et al. 2014). Recalcitrant carbon, solubles, and C_{mass} all can be predicted with $R^2 > 0.65$, while the smallest carbon fraction (hemicellulose) shows the worst performance (Kothari et al. 2023b). For chemical traits, intact- and ground-leaf models had only small and idiosyncratic differences in performance.

Our comparisons of variation among replicate spectral measurements within samples imply that grinding homogenizes the variation in the leaf and allows more consistent measurements. This consistency may owe to several factors. Veins and other features cause intact leaves to vary in their chemical makeup across the laminar area, which may be hard to capture representatively. Likewise, if chemical traits change with depth through the laminar cross-section, the optical features of deeper layers may be obscured in part by shallower layers. The geometries of uneven needle mats and curled-up leaves may

also reduce consistency by increasing variation in anisotropic surface reflectance (Petibon et al. 2021). Nevertheless, there was no strong, consistent indication that ground-leaf spectra are necessarily better than intact-leaf spectra for predicting chemical traits. This result deviates from the few previous comparisons of intact- and ground-leaf spectra (Serbin et al. 2014; Couture et al. 2016; Kothari et al. 2023b). It is unclear why the theoretical considerations in favor of ground-leaf spectra do not amount to much in practice here; we speculate that information about structure might, depending on the context, either confound or indirectly aid the prediction of chemical traits.

In contrast to chemical traits, LMA was clearly better predicted from intact-leaf spectra. Given that grinding destroys the leaf structure, it may be perplexing that ground-leaf spectra predict structural traits like LMA at all (Serbin et al. 2014; Kothari et al. 2023b). The relationship between ground-leaf spectra and LMA is likely indirect, mediated through correlations between LMA and chemical traits that retain clear optical signatures in ground tissue (Nunes et al. 2017; Kothari & Schweiger 2022). Since trait correlations within species often differ from those among species (Anderegg et al. 2018), such effects can explain why a model trained across many species may fail to capture intraspecific variation (Kothari & Schweiger 2022). Indeed, in our dataset, ground-leaf spectra do a poor job of predicting intraspecific variation in LMA; species-specific slopes of measured vs. predicted values are often far from 1.

Changing the basis of nitrogen and carbon from mass to area also changes the accuracy of estimates. While N_{mass} is predicted more accurately than N_{area}, C_{mass} is predicted less accurately than C_{area}. Across samples, N_{mass} correlates negatively and C_{mass} correlates positively with LMA, driven in large part by the split between broadleaf (low LMA) and needleleaf (high LMA) species. As a result, multiplying by LMA to convert N and C to an area basis flattens interspecific variation in N, but enhances it in C. For example, the percentage of variation explained by species identity alone increases for carbon (R^2 from 0.721 for C_{mass} to 0.945 for C_{area}), but decreases for nitrogen (R^2 from 0.786 for N_{mass} to 0.460 N_{area}). Absorption features that differentiate species or functional groups can aid estimates of C_{area} much more than N_{area},

whether or not they are causally related to those traits. Area- and mass-based traits often have different uses: there is a strong justification for using area-based nutrients for studying resorption (van Heerwaarden et al. 2003), while mass-based traits are more often used for explaining decomposition (Cornwell et al. 2008), and a 'normalization-independent' basis has been proposed for making sense of trait covariance (Osnas et al. 2018). In general, the basis of estimation should be determined by these factors rather than the apparent accuracy of the models.

Visible-range wavelengths where chlorophyll and other pigments absorb had high importance in our main set of full-range models. However, removing these wavelengths ('SWIR models') did not decrease performance while removing the SWIR range ('VIS/NIR models') did. It seems that pigments are neither essential for predicting litter chemistry, nor do they confound it. Although the bands where they absorb are assigned high importance, much of the information may be redundant with that of the SWIR range, where many leaf macromolecules absorb. However, those SWIR features are not fully redundant with the information about pigments and leaf structure captured by the VIS/NIR range. The spectrum of leaf or litter samples could remain more stable above 1300 nm over the long term (Kothari et al. 2023b), so the SWIR models may also transfer better when sample preparation and storage practices are inconsistent.

Advancing spectroscopic estimation of litter traits

Compared to conventional measurements, spectroscopic estimates of plant tissue functional traits are easy and fast to generate and can be non-destructive. Spectrometers are costly, but measuring each spectrum is (on the margin) nearly free. Once samples have been collected and organized, measuring each sample's intact-leaf spectrum takes under a minute. In contrast, both ground-leaf spectra and conventional measurements require grinding leaves, which is time-consuming and destructive. Using our protocol, we could measure the spectrum of a single ground sample (including sample tray preparation and cleanup) in 5-7 minutes, not including grinding time. In contrast, we found that doing both elemental and carbon

fraction analysis takes about 25 minutes of operator time per sample. The time saved through spectroscopic trait estimation grows as the number of traits of interest increases.

Many studies that leverage spectroscopy to address ecological questions use models tailored to the particular species and tissues under study (e.g., McTiernan et al. 2003; Hobbie 2005; Fortunel et al. 2009). While this practice allows researchers more confidence in their trait estimates, it requires them to spend time and money doing the manual trait measurements needed for calibration, which they must repeat for any new study system they adopt. Spectroscopic trait estimation is unlikely to serve as a complete substitute for standard measurements unless we can create models that are demonstrably accurate across a wide range of species and conditions.

The good validation performance of our models does not guarantee accurate trait estimates from any set of litter samples. To take one example, consider the aforementioned pattern that the ground-leaf model for LMA often performs poorly within species, despite its strong performance across species. If it were used to study intraspecific variation in litter traits, the model's estimates could be useless or even misleading. Moreover, when models simply 'learn' to exploit any features that happen to correspond with LMA among species within a given dataset—without respect to their causal relationship with LMA—they may also transfer poorly to new species (Kothari & Schweiger 2022). This case is closely linked to the more general problem in machine learning of distinguishing spurious and invariant correlations (Arjovsky et al. 2019). As the volume of available spectral data increases, taking care to assess out-of-sample generalization and borrowing tools from machine learning could help us train better spectral models and make more realistic judgments about their performance (Kattenborn et al. 2022; Cherif et al. 2023).

Recent studies have attempted to tackle the challenge of building global trait-spectra models. For example, Serbin et al. (2019) show that a single model can accurately predict the LMA of fresh, non-senesced leaves across 11 sites from the tropics to the Arctic. It remains to be seen whether any single

model could predict traits like N_{mass} or LMA from leaf tissue across many species at all stages of senescence and decomposition. The predictive success of global models may rely on the constrained patterns of trait covariation among non-senesced leaves—for example, along continua from sun to shade leaves or conservative to acquisitive leaves (Osnas et al. 2018)—which may help in estimating traits that are not themselves associated with strong absorption features (Nunes et al. 2017; Kothari & Schweiger 2022). However, the dramatic shifts in leaf chemistry and cellular structure during senescence (Keskitalo et al. 2005)—including pigment degradation and nutrient resorption—could weaken some of these relationships among traits, or between traits and their optical features. If so, it could be hard to build a global model for senesced leaves that achieves the predictive success of global models for non-senesced leaves.

Some evidence suggests that litter traits do in fact predictably covary with each other, and with non-senesced leaf traits (Freschet et al. 2010; Freschet et al. 2012; Jackson et al. 2013). But while most fresh green leaves look alike, each species' leaves may senesce in its own way. For example, some of our species (*A. rubrum*, *Q. alba*, *Q. ellipsoidalis*, *Q. rubra*) tend to produce reddish anthocyanins during senescence, but the rest do not. The breakdown of certain metabolites during senescence also creates a complex mixture of brown pigments (Fourty et al. 1996), which are poorly described and may vary from species to species. Such variation could make the relationship between traits and optical features more contingent for senescent foliage.

Even setting aside these concerns, model generality may also be complicated by differences in sample preparation (e.g., grind size) and instruments, which may affect the spectrum in subtle but important ways (Foley et al. 1998; Petibon et al. 2021). Overcoming these challenges may require investing in protocol development, or in trying algorithms that are more flexible than PLSR in 'learning' the complex forms of biological variation in heterogeneous datasets. In the meantime, it would be prudent for researchers who

plan to use existing spectral models in new systems to take some conventional trait measurements in order to validate model performance.

Finally, we note that plants produce and shed many kinds of tissue besides leaves, and these other litter sources also contribute to nutrient cycling through resorption and decomposition (Lü et al. 2012). Reflectance spectroscopy may help measure nutrient-related traits in other tissues. For example, fine root decomposition is an important but little-understood part of nutrient cycling; fine roots may contribute nearly half of annual litter inputs in forests (Freschet et al. 2013) and vary considerably in resorption efficiency (Freschet et al. 2010). Elle et al. (2019) showed that PLSR models built using near-infrared reflectance spectroscopy can predict fine root lignin, which increases recalcitrance (See et al. 2019; but see, in contrast, Sun et al. 2018). Continuing to develop models for other plant tissues could make it easier to study how plants alter nutrient cycling in a holistic way. Likewise, nutrients other than nitrogen often limit or co-limit growth and decomposition, and tools to estimate them quickly and cheaply could advance the study of plant nutrient economies. However, estimates of some micronutrients from reflectance spectra may be unreliable (Nunes et al. 2017; Kothari et al. 2023a), and other methods may prove more useful (e.g., X-ray fluorescence spectroscopy; van der Ent et al. 2018).

Ultimately, whether accurate global models for predicting litter traits from spectra can be built is an empirical question that can only be settled by amassing a wide variety of data. We take an initial step towards this long-term goal by showing that reflectance spectroscopy can provide accurate litter trait estimates across many temperate tree species that vary widely in their litter traits. This line of research shows great promise in relieving some of the major limitations toward understanding plant functional ecology in an ecosystem context.

Acknowledgements

The University of Minnesota, including Cedar Creek, is located on the traditional and contemporary Dakota and Ojibwe land. Many members of the Cavender-Bares Lab and the UMN Physiological Ecology Discussion Group contributed to discussions that improved this paper. We give particular thanks to Laura Williams, Artur Stefanski, Sam Reed, and Habacuc Flores-Moreno for useful discussions and help in data collection, as well as Rebecca Montgomery for contributions to the FAB1 experiment. The research was supported by an Alexander & Lydia Anderson Grant from University of Minnesota and by the National Science Foundation's funding of Cedar Creek LTER under DEB #1234162. Spectral measurements were conducted as part of an NSF/NASA Dimensions of Biodiversity project (DEB #1342778) and the NSF ASCEND Biology Integration Institute (DBI #2021898). SK was supported by an NSF Graduate Research Fellowship (Grant No. 00039202) and a UMN Doctoral Dissertation Fellowship.

Data availability

All spectral and trait data will be shared on the EcoSIS repository upon acceptance for publication. Analysis code will be archived on Zenodo. (A current version of the analysis code is available at: https://github.com/ShanKothari/senesced-trait-models.) Models will be uploaded to the Data Repository of U of M (DRUM).

Author contributions

SK and JCB conceived the project with input from SEH. SK collected the spectral and trait data and conducted all data analysis. SK wrote the first draft of the manuscript, and all authors contributed to further revisions.

References

Arjovsky, M., Bottou, L., Gulrajani, I., & Lopez-Paz, D. (2020). *Invariant Risk Minimization* (arXiv:1907.02893). arXiv. https://doi.org/10.48550/arXiv.1907.02893

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.

Berg, B. (2014). Decomposition patterns for foliar litter – A theory for influencing factors. *Soil Biology and Biochemistry*, *78*, 222–232.

Blackburn, G. A. (1998). Quantifying Chlorophylls and Caroteniods at Leaf and Canopy Scales: An Evaluation of Some Hyperspectral Approaches. *Remote Sensing of Environment*, *66*(3), 273–285.

Bourget, M. Y., Fanin, N., Fromin, N., Hättenschwiler, S., Roumet, C., Shihan, A., Huys, R., Sauvadet, M., & Freschet, G. T. (2023). Plant litter chemistry drives long-lasting changes in the catabolic capacities of soil microbial communities. *Functional Ecology*, *37*(7), 2014–2028.

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*, *72*(18), 6175–6189.

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.

Chapin III, F. S., & Moilanen, L. (1991). Nutritional Controls Over Nitrogen and Phosphorus Resorption From Alaskan Birch Leaves. *Ecology*, *72*(2), 709–715.

Chavana-Bryant, C., Malhi, Y., Wu, J., Asner, G. P., Anastasiou, A., Enquist, B. J., Caravasi, E. G. C., Doughty, C. E., Saleska, S. R., Martin, R. E., & Gerard, F. F. (2017). Leaf aging of Amazonian canopy trees as revealed by spectral and physiochemical measurements. *New Phytologist*, *214*(3), 1049–1063.

Cherif, E., Feilhauer, H., Berger, K., Dao, P. D., Ewald, M., Hank, T. B., He, Y., Kovach, K. R., Lu, B., Townsend, P. A., & Kattenborn, T. (2023). From spectra to plant functional traits: Transferable multi-trait models from heterogeneous and sparse data. *Remote Sensing of Environment*, *292*, 113580.

Cornelissen, J. H. C., Pérez-Harguindeguy, N., Díaz, S., Grime, J. P., Marzano, B., Cabido, M., Vendramini, F., & Cerabolini, B. (1999). Leaf structure and defence control litter decomposition rate across species and life forms in regional floras on two continents. *New Phytologist*, *143*(1), 191–200.

Cornwell, W. K., Cornelissen, J. H. C., Amatangelo, K., Dorrepaal, E., Eviner, V. T., Godoy, O., Hobbie, S. E., Hoorens, B., Kurokawa, H., Pérez-Harguindeguy, N., Quested, H. M., Santiago, L. S., Wardle, D. A., Wright, I. J., Aerts, R., Allison, S. D., Bodegom, P. V., Brovkin, V., Chatain, A., … Westoby, M. (2008). Plant species traits are the predominant control on litter decomposition rates within biomes worldwide. *Ecology Letters*, *11*(10), 1065–1071.

Costa, F. R. C., Lang, C., Almeida, D. R. A., Castilho, C. V., & Poorter, L. (2018). Near-infrared spectrometry allows fast and extensive predictions of functional traits from dry leaves and branches. *Ecological Applications*, *28*(5), 1157–1167.

Cotrozzi, L., Couture, J. J., Cavender-Bares, J., Kingdon, C. C., Fallon, B., Pilz, G., Pellegrini, E., Nali, C., & Townsend, P. A. (2017). Using foliar spectral properties to assess the effects of drought on plant water potential. *Tree Physiology*, *37*(11), 1582–1591.

Coûteaux, M.-M., Sarmiento, L., Hervé, D., & Acevedo, D. (2005). Determination of water-soluble and total extractable polyphenolics in biomass, necromass and decomposing plant material using near-infrared reflectance spectroscopy (NIRS). *Soil Biology and Biochemistry*, *37*(4), 795–799.

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.

Dechant, B., Cuntz, M., Vohland, M., Schulz, E., & Doktor, D. (2017). Estimation of photosynthesis traits from leaf reflectance spectra: Correlation to nitrogen content as the dominant mechanism. *Remote Sensing of Environment*, *196*, 279–292.

Diehl, P., Mazzarino, M. J., Funes, F., Fontenla, S., Gobbi, M., & Ferrari, J. (2003). Nutrient conservation strategies in native Andean-Patagonian forests. *Journal of Vegetation Science*, *14(1)*, 63–70.

El Zein, R., Bréda, N., Gérant, D., Zeller, B., & Maillard, P. (2011). Nitrogen sources for current-year shoot growth in 50-year-old sessile oak trees: An in situ 15N labeling approach. *Tree Physiology*, *31*(12), 1390–1400.

Elle, O., Richter, R., Vohland, M., & Weigelt, A. (2019). Fine root lignin content is well predictable with near-infrared spectroscopy. *Scientific Reports*, *9*(1), 6396.

Feilhauer, H., Asner, G. P., Martin, R. E., & Schmidtlein, S. (2010). Brightness-normalized Partial Least Squares Regression for hyperspectral data. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *111*(12), 1947–1957.

Foley, W. J., McIlwee, A., Lawler, I., Aragones, L., Woolnough, A. P., & Berding, N. (1998). Ecological applications of near infrared reflectance spectroscopy – a tool for rapid, cost-effective prediction of the composition of plant and animal tissues and aspects of animal performance. *Oecologia*, *116*(3), 293–305.

Fortunel, C., Garnier, E., Joffre, R., Kazakou, E., Quested, H., Grigulis, K., Lavorel, S., Ansquer, P., Castro, H., Cruz, P., DoleŽal, J., Eriksson, O., Freitas, H., Golodets, C., Jouany, C., Kigel, J., Kleyer, M., Lehsten, V., Lepš, J., … Zarovali, M. (2009). Leaf traits capture the effects of land use changes and climate on litter decomposability of grasslands across Europe. *Ecology*, *90*(3), 598–611.

Fourty, Th., Baret, F., Jacquemoud, S., Schmuck, G., & Verdebout, J. (1996). Leaf optical properties with explicit description of its biochemical composition: Direct and inverse problems. *Remote Sensing of Environment*, *56*(2), 104–117.

Freschet, G. T., Aerts, R., & Cornelissen, J. H. C. (2012). A plant economics spectrum of litter decomposability. *Functional Ecology*, *26*(1), 56–65.

Freschet, G. T., Cornelissen, J. H. C., Logtestijn, R. S. P. van, & Aerts, R. (2010). Substantial nutrient resorption from leaves, stems and roots in a subarctic flora: What is the link with other resource economics traits? *New Phytologist*, *186*(4), 879–889.

Freschet, G. T., Cornwell, W. K., Wardle, D. A., Elumeeva, T. G., Liu, W., Jackson, B. G., Onipchenko, V. G., Soudzilovskaia, N. A., Tao, J., & Cornelissen, J. H. C. (2013). Linking litter decomposition of above- and below-ground organs to plant–soil feedbacks worldwide. *Journal of Ecology*, *101*(4), 943–952.

Gill, A. L., Schilling, J., & Hobbie, S. E. (2021). Experimental nitrogen fertilisation globally accelerates, then slows decomposition of leaf litter. *Ecology Letters*, *24*(4), 802–811.

Gillon, D., Joffre, R., & Ibrahima, A. (1999). Can Litter Decomposability Be Predicted by Near Infrared Reflectance Spectroscopy? *Ecology*, *80*(1), 175–186.

Gitelson, A. A., & Merzlyak, M. N. (1996). Signature Analysis of Leaf Reflectance Spectra: Algorithm Development for Remote Sensing of Chlorophyll. *Journal of Plant Physiology*, *148*(3), 494–500.

Grossman, J. J., Cavender-Bares, J., & Hobbie, S. E. (2020). Functional diversity of leaf litter mixtures slows decomposition of labile but not recalcitrant carbon over two years. *Ecological Monographs*, *90*(3), e01407.

Grossman, J. J., Cavender-Bares, J., Hobbie, S. E., Reich, P. B., & Montgomery, R. A. (2017). Species richness and traits predict overyielding in stem growth in an early-successional tree diversity experiment. *Ecology*, *98*(10), 2601–2614.

Hayes, P., Turner, B. L., Lambers, H., & Laliberté, E. (2014). Foliar nutrient concentrations and resorption efficiency in plants of contrasting nutrient-acquisition strategies along a 2-million-year dune chronosequence. *Journal of Ecology*, *102*(2), 396–410.

Heerwaarden, L. M. V., Toet, S., & Aerts, R. (2003). Current measures of nutrient resorption efficiency lead to a substantial underestimation of real resorption efficiency: Facts and solutions. *Oikos*, *101*(3), 664–669.

Hobbie, S. E. (2005). Contrasting Effects of Substrate and Fertilizer Nitrogen on the Early Stages of Litter Decomposition. *Ecosystems*, *8*(6), 644–656.

Hobbie, S. E. (2015). Plant species effects on nutrient cycling: Revisiting litter feedbacks. *Trends in Ecology & Evolution*, *30*(6), 357–363.

Jackson, B. G., Peltzer, D. A., & Wardle, D. A. (2013). Are functional traits and litter decomposability coordinated across leaves, twigs and wood? A test using temperate rainforest tree species. *Oikos*, *122*(8), 1131–1142.

Joffre, R., Gillon, D., Dardenne, P., Agneessens, R., & Biston, R. (1992). The use of near-infrared reflectance spectroscopy in litter decomposition studies. *Annales des Sciences Forestières*, *49*(5), 481–488.

Kattenborn, T., Schiefer, F., Frey, J., Feilhauer, H., Mahecha, M. D., & Dormann, C. F. (2022). Spatially autocorrelated training and validation samples inflate performance assessment of convolutional neural networks. *ISPRS Open Journal of Photogrammetry and Remote Sensing*, *5*, 100018.

Keiluweit, M., Nico, P., Harmon, M. E., Mao, J., Pett-Ridge, J., & Kleber, M. (2015). Long-term litter decomposition controlled by manganese redox cycling. *Proceedings of the National Academy of Sciences*, *112*(38), E5253–E5260.

Keskitalo, J., Bergquist, G., Gardeström, P., & Jansson, S. (2005). A Cellular Timetable of Autumn Senescence. *Plant Physiology*, *139*(4), 1635–1648.

Kobe, R. K., Lepczyk, C. A., & Iyer, M. (2005). Resorption Efficiency Decreases with Increasing Green Leaf Nutrients in a Global Data Set. *Ecology, 86(10)*, 2780–2792.

Kothari, S., Beauchamp-Rioux, R., Blanchard, F., Crofts, A. L., Girard, A., Guilbeault-Mayers, X., Hacker, P. W., Pardo, J., Schweiger, A. K., Demers-Thibeault, S., Bruneau, A., Coops, N. C., Kalacska, M., Vellend, M., & Laliberté, E. (2023a). Predicting leaf traits across functional groups using reflectance spectroscopy. *New Phytologist*, *238*(2), 549–566.

Kothari, S., Beauchamp-Rioux, R., Laliberté, E., & Cavender-Bares, J. (2023b). Reflectance spectroscopy allows rapid, accurate and non-destructive estimates of functional traits from pressed leaves. *Methods in Ecology and Evolution*, *14*, 385–401.

Kothari, S., & Schweiger, A. K. (2022). Plant spectra as integrative measures of plant phenotypes. *Journal of Ecology*, *110*(11), 2536–2554.

Kruse, F. A., Lefkoff, A. B., Boardman, J. W., Heidebrecht, K. B., Shapiro, A. T., Barloon, P. J., & Goetz, A. F. H. (1993). The spectral image processing system (SIPS)—Interactive visualization and analysis of imaging spectrometer data. *Remote Sensing of Environment*, *44*(2), 145–163.

Liu, J., Liu, X., Song, Q., Compson, Z. G., LeRoy, C. J., Luan, F., Wang, H., Hu, Y., & Yang, Q. (2020). Synergistic effects: A common theme in mixed-species litter decomposition. *New Phytologist*, *227*(3), 757–765.

Lü, X.-T., Freschet, G. T., Flynn, D. F. B., & Han, X.-G. (2012). Plasticity in leaf and stem nutrient resorption proficiency potentially reinforces plant–soil feedbacks and microscale heterogeneity in a semi-arid grassland. *Journal of Ecology, 100(1)*, 144–150.

Madritch, M., Cavender-Bares, J., Hobbie, S. E., & Townsend, P. (2020). Linking foliar traits to belowground processes. In J. Cavender-Bares, J. Gamon, & P. Townsend (Eds.), *Remote Sensing of Plant Biodiversity* (pp. 173–198). Springer, New York.

McTiernan, K. B., Coûteaux, M.-M., Berg, B., Berg, M. P., Calvo de Anta, R., Gallardo, A., Kratz, W., Piussi, P., Remacle, J., & Virzo De Santo, A. (2003). Changes in chemical composition of Pinus sylvestris needle litter during decomposition along a European coniferous forest climatic transect. *Soil Biology and Biochemistry*, *35*(6), 801–812.

Meireles, J. E., Schweiger, A. K., & Cavender-Bares, J. (2023). spectrolab: Class and Methods for Spectral Data in R. R package v.0.0.18. [WWW document] URL https:// CRAN.R-project.org/package=spectrolab

Mevik, B.-H., Wehrens, R., & Liland, K. H. (2019). PLS: partial least squares and principal component regression. R package v.2.7-1. [WWW document] URL https:// CRAN.R-project.org/package=pls

Norris, K. H., Barnes, R. F., Moore, J. E., & Shenk, J. S. (1976). Predicting Forage Quality by Infrared Reflectance Spectroscopy. *Journal of Animal Science*, *43*(4), 889–897.

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.

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.

Perez, T. M., Rodriguez, J., & Heberling, J. M. (2020). Herbarium-based measurements reliably estimate three functional traits. *American Journal of Botany*, *107*(10), 1457–1464.

Petibon, F., Czyż, E. A., Ghielmetti, G., Hueni, A., Kneubühler, M., Schaepman, M. E., & Schuman, M. C. (2021). Uncertainties in measurements of leaf optical properties are small compared to the biological variation within and between individuals of European beech. *Remote Sensing of Environment*, *264*, 112601.

Petit Bon, M., Böhner, H., Kaino, S., Moe, T., & Bråthen, K. A. (2020). One leaf for all: Chemical traits of single leaves measured at the leaf surface using near-infrared reflectance spectroscopy. *Methods in Ecology and Evolution*, *11*(9), 1061–1071.

Prescott, C. E. (2010). Litter decomposition: What controls it and how can we alter it to sequester more carbon in forest soils? *Biogeochemistry*, *101*(1), 133–149.

R Core Team. 2020. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. [WWW document] URL https://www.R-project.org/ [accessed 9 March 2020].

Rea, A. M., Mason, C. M., & Donovan, L. A. (2018). Evolution of nutrient resorption across the herbaceous genus Helianthus. *Plant Ecology*, *219*(8), 887–899.

Richardson, A. D., & Reeves III, J. B. (2005). Quantitative reflectance spectroscopy as an alternative to traditional wet lab analysis of foliar chemistry: Near-infrared and mid-infrared calibrations compared. *Canadian Journal of Forest Research*, *35*(5), 1122–1130.

Schreeg, L. A., Mack, M. C., & Turner, B. L. (2013). Nutrient-specific solubility patterns of leaf litter across 41 lowland tropical woody species. *Ecology*, *94*(1), 94–105.

Schweitzer, J. A., Bailey, J. K., Rehill, B. J., Martinsen, G. D., Hart, S. C., Lindroth, R. L., Keim, P., & Whitham, T. G. (2004). Genetically based trait in a dominant tree affects ecosystem processes. *Ecology Letters*, *7*(2), 127–134.

See, C. R., McCormack, M. L., Hobbie, S. E., Flores-Moreno, H., Silver, W. L., & Kennedy, P. G. (2019). Global patterns in fine root decomposition: Climate, chemistry, mycorrhizal association and woodiness. *Ecology Letters*, *22*(6), 946–953.

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.

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.

Sun, T., Hobbie, S. E., Berg, B., Zhang, H., Wang, Q., Wang, Z., & Hättenschwiler, S. (2018). Contrasting dynamics and trait controls in first-order root compared with leaf litter decomposition. *Proceedings of the National Academy of Sciences*, *115*(41), 10392–10397.

Talbot, J. M., & Treseder, K. K. (2012). Interactions among lignin, cellulose, and nitrogen drive litter chemistry–decay relationships. *Ecology*, *93*(2), 345–354.

van der Ent, A., Przybyłowicz, W. J., de Jonge, M. D., Harris, H. H., Ryan, C. G., Tylko, G., Paterson, D. J., Barnabas, A. D., Kopittke, P. M., & Mesjasz-Przybyłowicz, J. (2018). X-ray elemental mapping techniques for elucidating the ecophysiology of hyperaccumulator plants. *New Phytologist*, *218*(2), 432–452.

Vergutz, L., Manzoni, S., Porporato, A., Novais, R. F., & Jackson, R. B. (2012). Global resorption efficiencies and concentrations of carbon and nutrients in leaves of terrestrial plants. *Ecological Monographs*, *82*(2), 205–220.

Wold, S. (1994). PLS for multivariate linear modeling. In H. van de Waterbeemd, editor. Chemometric methods in molecular design, methods and principles in medicinal chemistry, Verlag-Chemie (Weinheim, Germany), 195–218.

Wright, I. J., & Westoby, M. (2003). Nutrient concentration, resorption and lifespan: Leaf traits of Australian sclerophyll species. *Functional Ecology*, *17*(1), 10–19.

Yuan, Z. Y., & Chen, H. Y. H. (2015). Negative effects of fertilization on plant nutrient resorption. *Ecology, 96(2)*, 373–380.

Supplementary Material

Table S1: Species means of predicted values from intact-leaf models (I) and measured values (M) within the validation dataset for each of the eight leaf litter traits under consideration. See Table 1 in main text for species codes.

	LMA (g m^{-1})		C_{mass} (%)		N_{mass} (%)		Solubles (%)		Hemicellulose (%)		Recalcitrant carbon (%)		C_{area} (g m^{-1})		N_{area} (g m^{-1})	
Species	I	M	I	M	I	M	I	M	I	M	I	M	I	M	I	M
ACNE	55.3	62.5	47.8	46.7	1.79	1.67	46.9	48.2	18.3	18.4	34.1	33.5	26.3	29.4	1.11	1.03
ACRU	75.6	77.2	49.0	47.4	0.664	0.673	67.6	65.6	10.5	9.42	23.0	25.0	36.8	36.7	0.52	0.526
BEPA	75.1	75.2	53.0	53.2	0.939	0.964	58.8	59.0	13.0	13.1	28.9	27.9	41.0	40.0	0.686	0.718
PIBA	235	231	55.8	55.1	0.280	0.308	45.9	44.8	8.94	11.3	42.8	44.0	131	127	0.724	0.713
PIRE	238	246	57.5	56.9	0.423	0.373	39.3	38.2	12.3	12.0	49.0	49.8	135	140	0.944	0.925
PIST	177	170	57.7	57.7	0.424	0.452	42.9	44.0	12.9	11.5	45.3	44.5	102	98.1	0.652	0.768
QUAL	87.8	84.3	50.0	51.0	0.912	0.971	57.0	55.4	13.4	15.8	29.5	28.8	44.7	43.1	0.801	0.831
QUEL	88.6	85.5	52.8	52.5	1.23	1.21	47.4	49.1	17.7	17.1	35.1	33.9	46.8	45.0	1.07	1.04
QUMA	101	110	50.6	50.2	1.45	1.44	50.9	53.2	16.4	15.6	32.6	31.2	51.8	55.6	1.43	1.60
QURU	89.7	80.2	51.9	51.8	0.975	0.96	51.1	51.7	16.0	16.3	33.1	31.9	46.3	41.6	0.813	0.761
TIAM	52.8	55.9	51.6	51.6	1.04	1.11	52.3	50.1	13.8	13.3	33.9	36.6	27.1	28.8	0.581	0.623

Table S2: Summary statistics of PLSR model validation based on intact- and ground-leaf spectra restricted to the VIS/NIR range (400-1000 nm) or to part of the SWIR range (1300-2400 nm). %RMSE is calculated as RMSE divided by the 2.5% trimmed range of data within the validation data set.

	VIS/NIR				SWIR			
Trait	No. comps	R^2	RMSE	%RMSE	No. comps	R^2	RMSE	%RMSE
LMA (g m^{-1})	10	0.871	23.1	11.3	7	0.939	15.6	7.63
C_{mass} (%)	10	0.663	2.15	16.6	6	0.647	2.21	17.0
N_{mass} (%)	12	0.712	0.25	15.8	11	0.939	0.12	7.57
Solubles (%)	8	0.613	5.92	18.7	6	0.782	4.58	14.4
Hemicellulose (%)	8	0.371	3.30	21.9	6	0.555	2.76	18.3
Recalcitrant carbon (%)	12	0.789	3.91	13.2	9	0.869	3.07	10.4
C_{area} (g m^{-1})	10	0.873	13.6	11.9	7	0.942	9.12	7.95
N_{area} (g m^{-1})	9	0.588	0.249	17.4	10	0.782	0.186	13.0

Fig. S1: Intact-leaf litter spectral model performance statistics for each trait based on 200 jackknife iterations from the calibration data set. %RMSE is calculated as RMSE divided by the 2.5% trimmed range of data within the testing data set. In this case, the testing data set for each iteration is the 30% of data randomly chosen from within the calibration set not used for training the model in that iteration. Abbreviations: sol = solubles, hemi = hemicellulose, recalc = recalcitrant carbon.

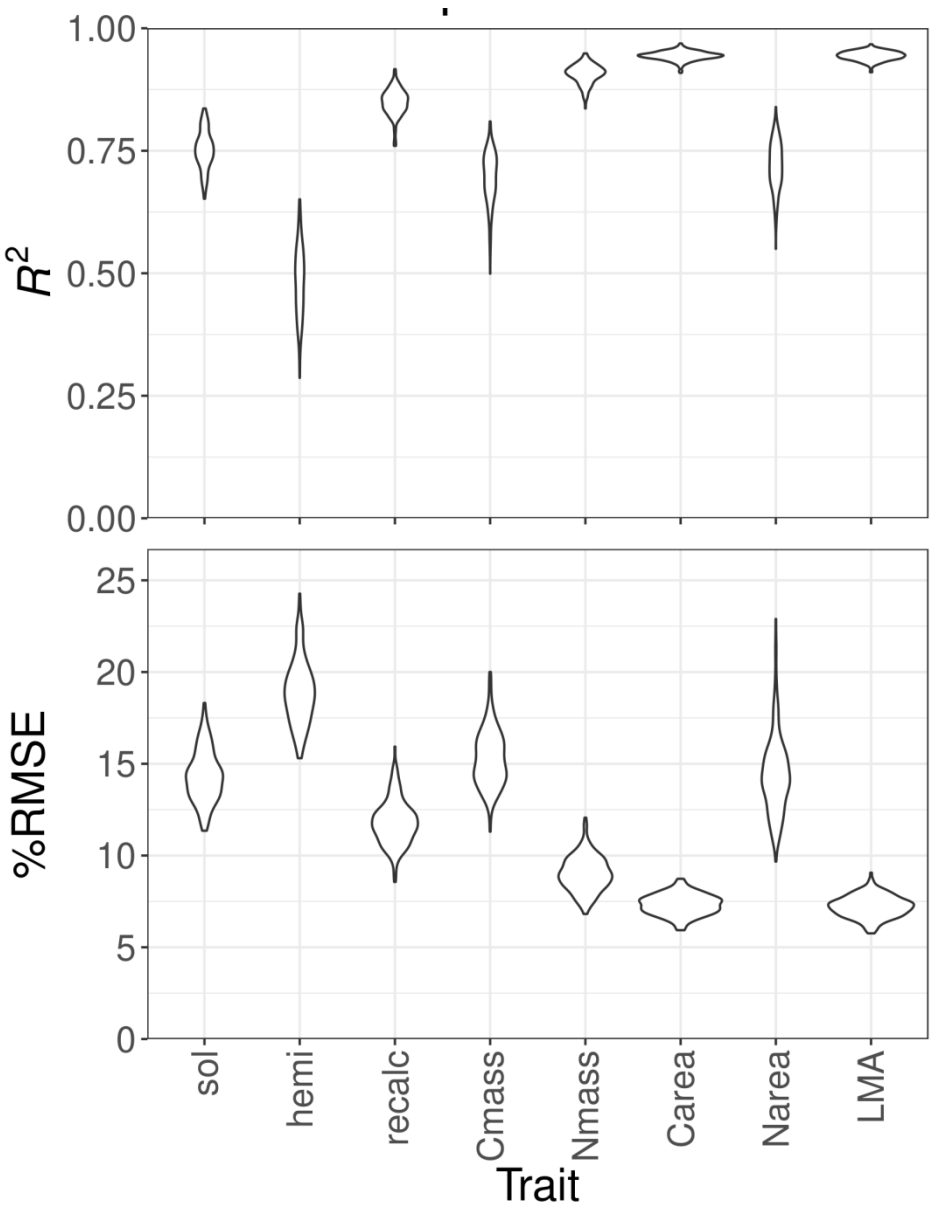


Fig. S2: Ground-leaf litter spectral model performance statistics for each trait based on 200 jackknife iterations from the calibration data set. %RMSE is calculated as RMSE divided by the 2.5% trimmed range of data within the testing data set. In this case, the testing data set for each iteration is the 30% of data randomly chosen from within the calibration set not used for training the model in that iteration. Abbreviations: sol = solubles, hemi = hemicellulose, recalc = recalcitrant carbon.

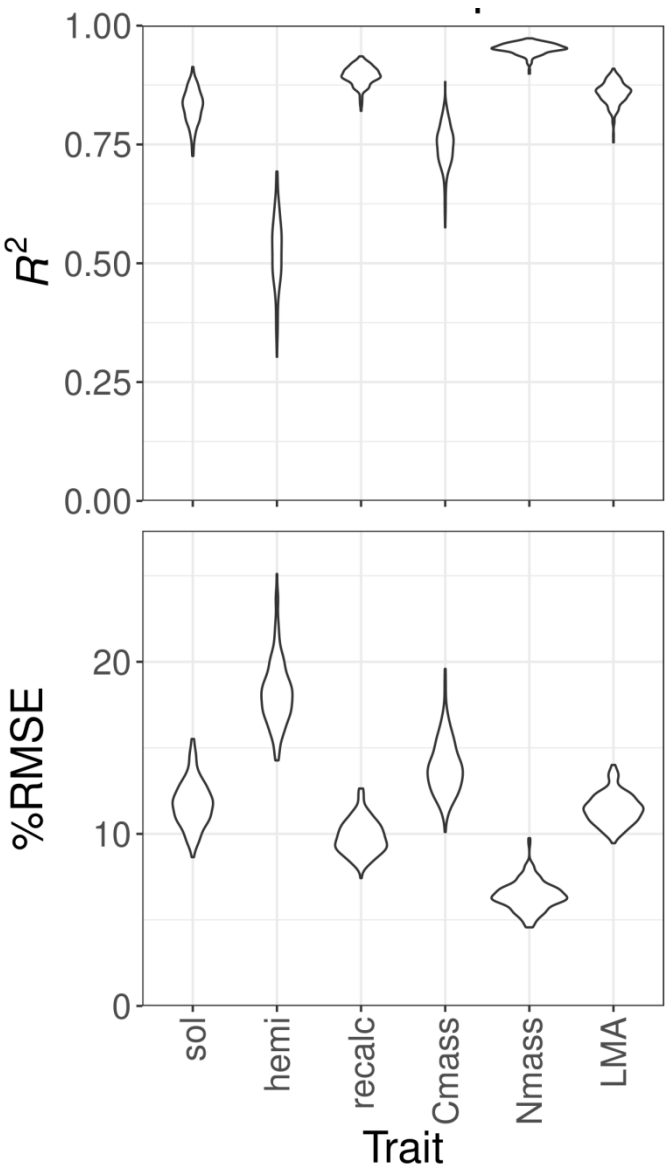


Fig. S3: The variable importance of prediction (VIP) metric in intact-leaf VIS/NIR models. The dashed black line represents a threshold of 0.8 suggested as a rough heuristic of importance by Burnett et al. (2021). Abbreviations: sol = solubles, hemi = hemicellulose, recalc = recalcitrant carbon.

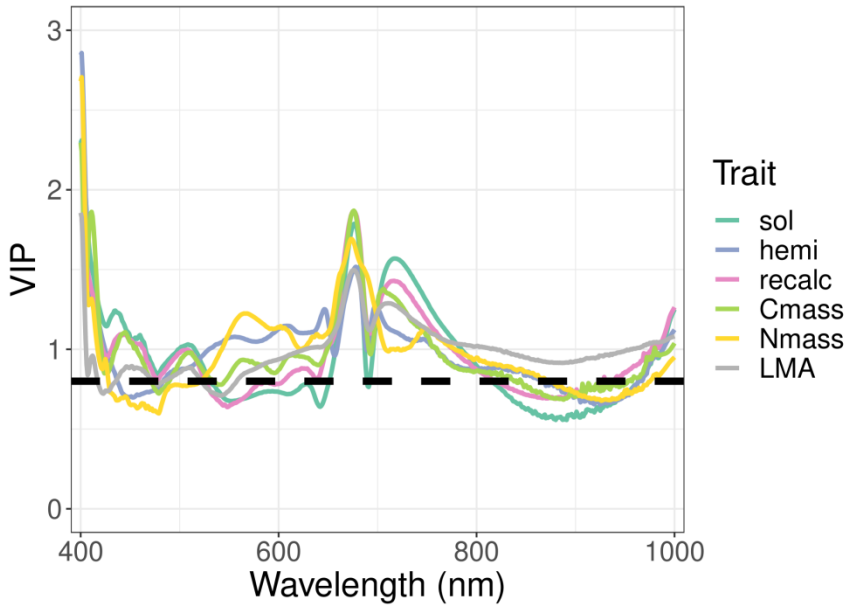
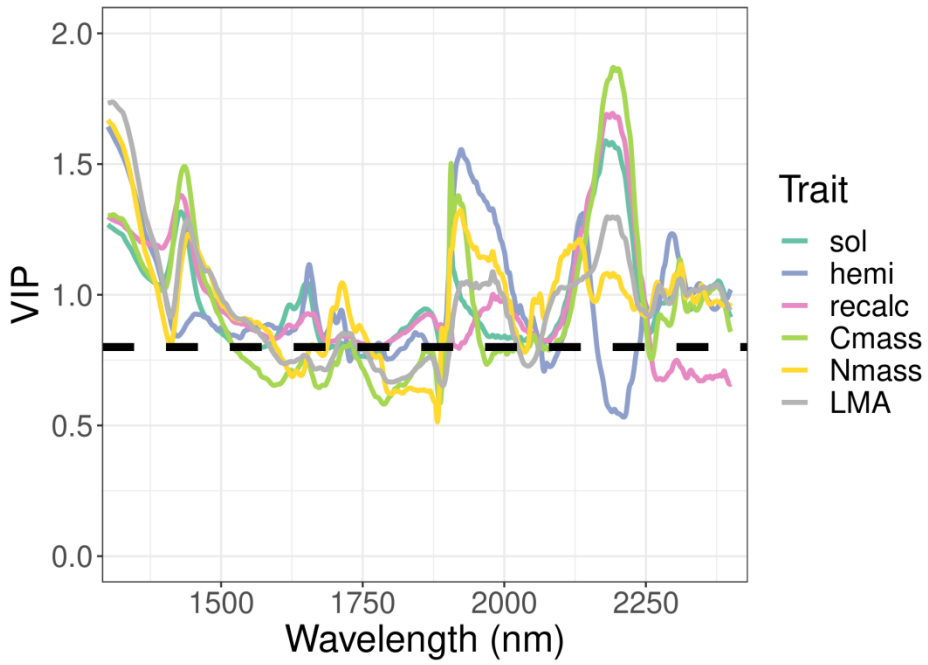



Fig. S4: The variable importance of prediction (VIP) metric in intact-leaf SWIR models. The dashed black line represents a threshold of 0.8 suggested as a rough heuristic of importance by Burnett et al. (2021). Abbreviations: sol = solubles, hemi = hemicellulose, recalc = recalcitrant carbon.