Imaging spectroscopy algorithms for mapping canopy foliar chemical and morphological traits and their uncertainties

Aditya Singh,1,3 Shawn P. Serbin,1,4 Brenden E. McNeil,2 Clayton C. Kingdon,1 and Philip A. Townsend1

1Department of Forest and Wildlife Ecology, University of Wisconsin, 1630 Linden Drive, Madison, Wisconsin 53706 USA
2Department of Geology and Geography, West Virginia University, Morgantown, West Virginia 26506 USA

Abstract. A major goal of remote sensing is the development of generalizable algorithms to repeatedly and accurately map ecosystem properties across space and time. Imaging spectroscopy has great potential to map vegetation traits that cannot be retrieved from broadband spectral data, but rarely have such methods been tested across broad regions. Here we illustrate a general approach for estimating key foliar chemical and morphological traits through space and time using NASA’s Airborne Visible/Infrared Imaging Spectrometer (AVIRIS-Classic). We apply partial least squares regression (PLSR) to data from 237 field plots within 51 images acquired between 2008 and 2011. Using a series of 500 randomized 50/50 subsets of the original data, we generated spatially explicit maps of seven traits (leaf mass per area (Marea), percentage nitrogen, carbon, fiber, lignin, and cellulose, and isotopic nitrogen concentration, δ15N) as well as pixel-wise uncertainties in their estimates based on error propagation in the analytical methods. Both Marea and %N PLSR models had a R² > 0.85. Root mean square errors (RMSEs) for both variables were less than 9% of the range of data. Fiber and lignin were predicted with R² > 0.65 and carbon and cellulose with R² > 0.45. Although R² of %C and cellulose were lower than Marea and %N, the measured variability of these constituents (especially %C) was also lower, and their RMSE values were beneath 12% of the range in overall variability. Model performance for δ15N was the lowest (R² = 0.48, RMSE = 0.95), but within 15% of the observed range. The resulting maps of chemical and morphological traits, together with their overall uncertainties, represent a first-of-its-kind approach for examining the spatiotemporal patterns of forest functioning and nutrient cycling across a broad range of temperate and sub-boreal ecosystems. These results offer an alternative to categorical maps of functional or physiognomic types by providing non-discrete maps (i.e., on a continuum) of traits that define those functional types. A key contribution of this work is the ability to assign retrieval uncertainties by pixel, a requirement to enable assimilation of these data products into ecosystem modeling frameworks to constrain carbon and nutrient cycling projections.

Key words: AVIRIS, Airborne Visible/Infrared Imaging Spectrometer; carbon cycling; deciduous and coniferous forests; foliar traits; imaging spectroscopy; leaf chemistry; leaf morphology; nutrient cycling; PLS regression; spatial mapping; uncertainty.

INTRODUCTION

Terrestrial ecosystems play an important role in the global carbon cycle by sequestering 3–5 Pg of carbon (C) per year from the atmosphere via photosynthesis (see Schimel 1995, Cramer et al. 2001, 2007, Le Quere et al. 2009). Characterization of the factors that influence terrestrial C uptake is required to develop an improved process-based understanding of ecosystem dynamics and accurately model vegetation response to global change. However, there remains a large uncertainty in the magnitude, seasonality, and location of this flux (Friedlingstein et al. 2014).

Chemical, structural, and morphological properties of the foliage in vegetation canopies correlate strongly with plant function, including ecosystem-wide nutrient cycling rates (Scott and Binkley 1997, Craine et al. 2002, Santiago et al. 2004, Meier and Bowman 2008) and photosynthetic capacity (Reich et al. 1997, 1999, Shipley et al. 2005, Kergoat et al. 2008). Thus, to map ecosystem productivity, modeling efforts to estimate the dynamics of global carbon stocks should make use of “functional biogeographic” (Violle et al. 2014) information on the covariance between plant traits and patterns of biogeochemical cycling (Enquist et al. 2007). The use of remote imaging spectroscopy to map key plant trait patterns—those that govern carbon and nutrient dynamics—will greatly enhance our understanding of ecosystem functioning (Ustin and Gamon 2010).
Global patterns of nutrient cycling and primary productivity in forested ecosystems are driven, in large part, by a tractable suite of foliar structural and biochemical traits representing a trade-off between leaf construction costs and photosynthetic carbon uptake potential (Reich et al. 1992, Shipley and Lechowicz 2000, Wright et al. 2004, Shipley et al. 2005, 2006). This defines a spectrum of plant function from slow-growing, nutrient-poor species to fast-growing, nutrient-rich, albeit “leaky,” plants. Specifically, coordination between foliar nitrogen content and specific leaf area (SLA, or its reciprocal, leaf mass per area, M\text{area}) maximizes C fixation (Shipley et al. 2005), with lower SLA resulting in lower foliage photosynthetic potential per unit dry mass (Niinemets 2001), but greater leaf structural strength and consequently longer leaf life spans (Wright and Westoby 2002, Wright et al. 2005a, Violle et al. 2009). These patterns have been found to be consistent among broadleaf, needleleaf, and herbaceous species (Reich et al. 1991, 1992, Grime 2006) and have been found to operate independently of growth form or phylogeney (Shipley et al. 2006). Foliar nitrogen content (i.e., nitrogen on scaled to an area basis, g/m²) scales with the content of RuBisCo, the key protein responsible for the carboxylation of RuBP in the initial steps of CO₂ fixation (Long 1991, Collatz et al. 1992, Ainsworth and Rogers 2007). Thus, foliar nitrogen reflects a plant’s investment in photosynthetic enzymes, structure, and light-harvesting complexes that modulate photosynthetic parameters such as the maximum rates of carboxylation (V\text{cmax}) and electron transport (J\text{max}) (Long 1991, Ripullone et al. 2003).

Foliar nitrogen concentration and M\text{area}/SLA also covary strongly with decomposition and nitrogen mineralization rates across ecosystems (Chapin 2003, Kazakou et al. 2006, Quested et al. 2007, Santiago 2007). Decomposition releases plant nutrients and drives the largest flux of terrestrial C to the atmosphere (Meier and Bowman 2008), and rates of decomposition are positively correlated with nitrogen content of foliage and litter, SLA, and leaf water content (Bonan 1993, Hattenschwiler et al. 2005, Fortunel et al. 2009) and negatively with leaf dry matter content (Quested et al. 2007, Fortunel et al. 2009, Kazakou et al. 2009) and foliar lignin concentrations (Hobbit et al. 2007, Johnson et al. 2007, Carrera and Bertiller 2010). Foliar lignin-to-nitrogen ratios are thus strong indicators of litter quality and consistent predictors of litter decomposability (Melillo et al. 1982, Knorr et al. 2005, Fortunel et al. 2009) and, together with cellulose concentration, regulate rates of late-stage forest litter decomposition (McClaugherty and Berg 1987, Berg 2000, Johnson et al. 2007).

Foliar isotopic N concentration (δ¹⁵N) is an integrator of the nitrogen cycle (Robinson 2001, Amundson et al. 2003) and has been used as an indicator of relative N availability to plants and a basis to infer N cycling rates (Austin and Vitousek 1998, Martinelli et al. 1999, Craine et al. 2009). Specifically, variations in N fractionation in foliar matter can indicate effects of climate on organic/inorganic nitrogen availability (Amundson et al. 2003, Craine et al. 2009) or local-scale variations in plant dependence on mycorrhizal associations (Hogberg 1997, Hobbit et al. 2000, Lilleskov et al. 2002, Craine et al. 2009). Spatial patterns of foliar isotopic N concentrations are also associated with plant response to disturbances (Amundson et al. 2003, McLauchlan et al. 2007) or elevated external inputs such as N deposition (Hogberg 1997, Emmett et al. 1998, McLauchlan et al. 2007).

Synoptic, repeatable, and consistent retrievals of key foliar traits such as foliar nitrogen, SLA (or M\text{area}), foliar lignin and cellulose, and isotopic nitrogen are invaluable for characterizing determinants of ecosystem function across large regions (Chapin et al. 1996, Diaz et al. 2004, Ustin and Gamon 2010), and may provide valuable inputs for modeling nutrient fluxes and vegetation range shifts under changing land use and climate scenarios (Wright et al. 2004, Townsend et al. 2008). Remote sensing, specifically imaging spectroscopy, offers the capacity to map the primary drivers of plant trait variability from the landscape to the global scale (Ustin and Gamon 2010). Imaging spectrometers measure reflected light in narrow, contiguous wavebands (typically ≤10–15 nm) over a large portion of the incident solar spectrum (e.g., 350–2500 nm for NASA’s Airborne Visible/Infrared Imaging Spectrometer, AVI-RIS; see Vane et al. 1993, Green et al. 1998). It has long been demonstrated that reflectance spectroscopy is sensitive to foliar chemistry as well as the structural and morphological characteristics of plant canopies, and indeed, this is the basis for standard spectroscopic approaches to estimate chemistry on dried samples (Kokal 2001, Richardson and Reeves 2005, Petisco et al. 2006) as well as to nondestructively estimate chemistry or photosynthetic properties of fresh leaves (Curran et al. 1992, Sims and Gamon 2002, Blackburn 2007, Menesatti et al. 2010). The ability to do this is a consequence of well-known biochemical absorption features at different wavelengths (Curran 1989, Curran et al. 1992, Fourry et al. 1996, Curran et al. 2001) related to electron transitions and bending and stretching of chemical bonds (Curran 1989, Fourry et al. 1996).

The ability to retrieve chemistry from remotely sensed imagery is complicated by a number of factors, including crown architecture and seasonality, although it is clear that chemical features related to leaf chlorophyll (Curran et al. 1992, Yoder and Pettigrewroosby 1995, Curran et al. 2001) and water content are clearly expressed in both narrowband and broadband imagery (Gates et al. 1965, Tucker 1980, Asner and Vitousek 2005, Sanchez-Azofeifa et al. 2009, Ustin et al. 2012). The effects of canopy structure on reflectance may confound the identification of absorption features in foliage (Knyazikhin et al. 2013), for example, in wavelengths with high leaf scattering, but there is also
strong evidence that features identifiable at the leaf level can be detected in measurements at the canopy level (Baret et al. 1994). Ultimately, the coordination of traits at both the leaf and canopy level facilitates using spectral information to map traits of interest, even if the biophysical drivers of strong correlations at some wavelengths have not yet been fully characterized (Townsend et al. 2003, Ollinger 2011, Ollinger et al. 2013).

An extensive body of literature has demonstrated the capacity of imaging spectroscopy to map foliar traits relevant to photosynthesis (Wessman et al. 1989, Matson et al. 1994, Curran et al. 1997, Martin and Aber 1997, Coops et al. 2003, Smith et al. 2003, Townsend et al. 2003, Asner and Martin 2008, Martin et al. 2008, McNeil et al. 2008, Kokaly et al. 2009). In particular, spectra in the shortwave infrared (especially 1400–1850 nm and 2000–2400 nm) are critical to retrieval of parameters related to photosynthesis on account of strong absorption related to nitrogen bonds in proteins that are expressed in those regions (Fourry et al. 1996). Studies also show that radiative transfer models (Asner 1998, Asner et al. 2008, 2011a, Asner and Martin 2008, 2009) can be used in conjunction with imaging spectroscopy to effectively map foliar traits. In general, existing efforts have been site specific (Martin and Aber 1997, Ollinger et al. 2002, Coops et al. 2003, Townsend et al. 2003) and, apart from Martin et al. (2008), studies that incorporate information from multiple scenes across multiple locations are rare (Appendix: Table A1). Here, we report a set of generalizable spectroscopic calibrations for the determination of leaf chemical composition (nitrogen, carbon, and fiber constituents), morphology (leaf mass per area, $M_{area}$), and isotopic composition ($^{15}N$) of temperate and boreal tree species using imaging spectroscopy. Moreover, by demonstrating techniques to explicitly propagate uncertainties from the leaf to stand to image level, we show how our algorithms can be directly integrated into global change models that require uncertainty estimates of input parameters.

**Materials and Methods**

**Study sites**

We established 237 plots across a broad swath of the Upper Midwest and Northeast USA (Fig. 1). These plots sampled a range of broadleaf and needleleaf forests, including 36 dominant species (see Appendix: Table A2) reported by Serbin et al. (2014). Each plot consisted of two crossing 60 × 60 m transects, defining five points at which a metric factor-2 prism was used to tally trees by species (Townsend et al. 2003). Foliar biomass by species was estimated from diameter–biomass relationships in Jenkins et al. (2003), and was subsequently used to scale foliar trait estimates by species to the canopy of the entire plot.

We collected foliage from the top, middle, and bottom of the canopy of the predominant species on each plot (>5% of relative biomass) using a shotgun, line launcher, or pole pruner. Plot-level values for each trait were estimated from the relative proportion of foliar biomass by species on a plot. We measured leaf-level percentage nitrogen (%N), carbon (%C), fiber (ADF), lignin (ADL), and cellulose, isotopic $^{15}N$: $^{14}N$ ratio ($^{15}N$), and leaf mass per area ($M_{area}$), as well as their uncertainties, by species and by plot from reflectance spectroscopy on both dry and fresh samples (Serbin et al. 2014). Leaf-level trait measurements employed chemometric partial least-squares regression (PLSR) models that link contact reflectance spectra and measurements made using standard laboratory techniques. This approach facilitated estimation of per sample (i.e., per leaf) uncertainties of trait estimates. We scaled measurement uncertainty from leaf-level traits to estimates of canopy traits based on proportional foliar biomass. Leaf-to-canopy trait scaling used 1000 replicates drawn from the distributions of leaf-level estimates and their uncertainties from Serbin et al. (2014) to generate plot-level trait estimates and uncertainties. Note that we had trait estimates for multiple levels of the canopy at each plot, which facilitated testing the ability to map “whole-canopy” trait characteristics (see Ollinger et al. 2002) vs. top-of-canopy traits (presumably all that is visible to the sensor). Using data from Green et al. (2003) on vertical variation in foliar traits within a canopy profile, we weighed plot-level estimates according to four different schemes: (1) top weighted, 90% weight for top-of-canopy foliage, 9% weight for mid-canopy foliage, and 1% weight for bottom-canopy foliage; (2) top-to-middle weighted, 64.5%, 32.3%, and 3.2% weights for top, middle, and bottom canopy foliage; (3) whole canopy, 40%, 40%, and 20% weights for top, middle, and bottom canopy foliage; and (4) top only, where only the top-of-canopy foliage was considered (100% for top-of-canopy). We report results from top weighted scaling (option 1), but results and coefficients for the rest of the scaling options are provided in the Appendix and Supplement). For each of these options, uncertainty in relative foliar biomass was incorporated by randomly varying species-level biomass estimates by ±25% in each of the 1000 permutations. The Supplement provides plot locations, plot-level canopy trait estimates, and associated uncertainties.

**Image processing**

145 AVIRIS-Classic images were acquired from NASA’s ER-2 platform during the summers of 2008–2011 at flight altitudes ranging from 14 000–20 000 m (pixel sizes of 12–18 m; see Fig. 1). Data were provided by the Jet Propulsion Laboratory as orthorectified, calibrated radiance images (Vane et al. 1993, Green et al. 1998), and all images were processed using a common processing stream to ensure comparability. We atmospherically corrected images using ATREM (TAFKAA algorithm; Gao et al. 2000, Montes and Gao 2004). Images were topographically corrected using the mod-
ified sun-canopy-sensor topographic method (Soenen et al. 2005) and were BRDF (bidirectional reflectance distribution function; i.e., cross-track)-corrected using a quadratic function of the volumetric scattering term of the Ross-Thick BRDF model (Roujean et al. 1992, Lucht et al. 2000). Post hoc georectification was performed on a small number of misaligned images, with final pixel registration errors < 0.5 pixel across all scenes. We used 51 images for model development and the remainder for testing and analysis.
The general approach to the estimation of canopy traits (including chemistry) from imaging spectroscopy (IS) is illustrated in Fig. 2. We build upon the approach taken by Martin et al. (2008), in which they used data from eight study sites and 137 field plots to develop a general, cross-scene predictive equation for foliar nitrogen. We extend their effort by additionally deriving maps of carbon, leaf mass per area ($M_{\text{area}}$), fiber, lignin, and cellulose, as well as mapping uncertainties in all traits of interest. For each trait, we developed models using 500 permutations of our data, drawing randomly from the plot-level trait estimates and perturbing those estimates based on uncertainties propagated at every level of measurement (spectra, trait, foliar biomass, etc.). For each of the 500 permutations, we randomly split our data 50/50 for calibration and validation. Because spectra for a given plot could have been sourced from multiple images available for each location, we further constrained the selection such that an observation could appear only once in the training set (i.e., training spectra for each plot came from a single image). This procedure was repeated 500 times and model coefficients, model fits, and predictions were saved for all iterations. We subsequently applied the 500 models to each pixel in all images. This allowed us to report the range of expected model performance and to calculate pixel-by-pixel uncertainty for the mapped models as the standard deviation of model predictions for all 500 permutations.

Statistical analysis

The general approach to the estimation of canopy traits (including chemistry) from imaging spectroscopy (IS) is illustrated in Fig. 2. We build upon the approach taken by Martin et al. (2008), in which they used data from eight study sites and 137 field plots to develop a general, cross-scene predictive equation for foliar nitrogen. We extend their effort by additionally deriving maps of carbon, leaf mass per area ($M_{\text{area}}$), fiber, lignin, and cellulose, as well as mapping uncertainties in all traits of interest. For each trait, we developed models using 500 permutations of our data, drawing randomly from the plot-level trait estimates and perturbing those estimates based on uncertainties propagated at every level of measurement (spectra, trait, foliar biomass, etc.). For each of the 500 permutations, we randomly split our data 50/50 for calibration and validation. Because spectra for a given plot could have been sourced from multiple images available for each location, we further constrained the selection such that an observation could appear only once in the training set (i.e., training spectra for each plot came from a single image). This procedure was repeated 500 times and model coefficients, model fits, and predictions were saved for all iterations. We subsequently applied the 500 models to each pixel in all images. This allowed us to report the range of expected model performance and to calculate pixel-by-pixel uncertainty for the mapped models as the standard deviation of model predictions for all 500 permutations.

To test if our models were stable across space and time,
we conducted additional cross-validations by reformulating PLSR models by leaving out data from individual geographic locations (areas with groups of plots; Fig. 1), or from years of image acquisition and evaluating model fits on the withheld data.

We employed partial least-squares regression, PLSR (Wold et al. 1984, Geladi and Kowalski 1986, Wolter et al. 2008), to predict canopy traits from imaging spectroscopy, as done in many previous studies (Coops et al. 2003, Smith et al. 2003, Townsend et al. 2003, Martin et al. 2008, McNeil et al. 2008, Deel et al. 2012). In practice, PLSR iteratively transforms predictor and response variables to find latent vectors and subsequently produce calibration factors and a linear model. PLS techniques achieve computational efficiency by maximizing the covariance between independent and dependent variables while simultaneously maintaining the constraint of being orthogonal to the previously determined factors (Wold et al. 1984, Geladi and Kowalski 1986, Frank and Friedman 1993, Wold et al. 2001). PLSR is better suited than traditional multiple linear regression methods for handling multi-collinear and over-sampled data, i.e., when the number of independent variables approaches or exceeds the number of observations (Wold et al. 1984, 2001, Geladi and Kowalski 1986). In particular, variations of unconstrained linear models (such as stepwise variable selection approaches; see Appendix: Table A1), violate cardinal assumptions of independent predictors and uncorrelated errors. At best, this results in biased estimates of beta coefficients (Tibshirani 1996) and underestimated errors and P values (Altman and Andersen 1989). Specific to chemometric data, stepwise selection procedures have been shown to perform poorly in calibration experiments (see Grossman et al. 1996) and may not generate results based on known absorption characteristics of materials being examined (Townsend et al. 2003). PLSR avoids difficulties in interpretation of synthetic PCA variables or the potential for modeling spurious relationships using stepwise regression (Grossman et al. 1996). The Predicted REmidual Sum of Squares (PRESS) statistic was used to select the appropriate number of PLS model components (Chen et al. 2004). We used the Variable Importance of Projection (VIP) statistic (Wold 1994) and standardized coefficient, by waveband, to evaluate the contribution of each wavelength and associated absorption features to trait predictions. The VIP score reflects each predictor’s (i.e., waveband) influence in fitting the PLSR model. It is based on the predictor’s absolute coefficient size and partial $R^2$ within the overall PLSR model. Standardized coefficients are employed for graphing to account for scale differences in reflectance between different wavelength regions (i.e., higher reflectance and greater variability in the NIR equate to smaller coefficients, but not necessarily less important coefficients), as well as in the dependent variables (traits).

We averaged the $3 \times 3$ image pixel array surrounding each field plot’s center, leaving us with one composite AVIRIS spectrum per 60 m plot. We postulate that averaging surrounding pixels is desirable because field plots in closed-canopy forests can rarely be identified to a specific pixel.

Finally, maps of the traits and their uncertainties were made for all AVIRIS images as the mean and standard deviation of the 500 permutations. Geographic patterns of trait distribution were analyzed with respect to gradients of latitude, precipitation, and temperature (derived from PRISM 2004), and nitrogen deposition (from NADP 2014).

**Results**

**Geographic variation in foliar traits**

Canopy-averaged ranges of plot-level trait estimates stratified by leaf physiognomy are presented in Table 1 (data on plot-level foliar traits by species are in the Appendix: Table A2; the entire data set is in the Supplement). There was a nearly twofold difference in $M_{\text{area}}$ between needleleaf and broadleaf species ($T = 28.88, P < 0.0001$), and mass-based nitrogen content of broadleaf species was around 1.5 times that of needleleaf species ($T = 24.15, P < 0.0001$). Carbon content, fiber, lignin, and cellulose exhibited less variation between needleleaf and broadleaf species, but all differences were significant ($P < 0.0001$). Deciduous-dominated plots showed slightly higher isotopic N content than coniferous-dominated plots ($T = 2.32, P = 0.02$). Correlation analyses revealed a strong negative relationship between leaf nitrogen and $M_{\text{area}}$ ($r = -0.766, P < 0.0001$) as well as...
as significant negative relationships with other leaf structural compounds (Table 2). Leaf nitrogen content decreased with latitude, while $M_{area}$ increased, a reflection of an increasing dominance of needleleaf species in the north. We also found a significant positive correlation between foliar nitrogen and patterns of nitrogen deposition across our sites ($r = 0.392$, $P < 0.0001$). To further explore this relationship, we fit a mixed-effects model to predict foliar nitrogen concentrations while controlling for latitude with Gaussian random effects specified for species (not shown). The model explained 88.7% of the variation in foliar nitrogen, of which 59% was attributed to differences between species alone. The effects of elevated foliar nitrogen concentrations corresponding to the N deposition gradient were apparent ($P < 0.0001$), even after taking into consideration latitudinal gradients and inherent differences between species. Foliar $\delta^{15}N$ concentrations were highly correlated with mean annual temperature and declined with mean annual precipitation (MAT, $r = 0.413$; MAP, $r = -0.455$; both $P < 0.0001$). We also found evidence of foliar isotopic N enrichment along the N deposition gradient ($r = 0.398$, $P < 0.0001$).

Canopy spectra

Canopy spectra were highly variable (Fig. 3), with nearly complete separation ($\pm 1$ SD) between spectra of all broadleaf and all needleleaf plots in the NIR (700–1400 nm) and the 1400–1900 nm SWIR1 regions. Spectra of the two physiognomic types partially overlapped in the visible (400–700 nm) and SWIR2 (1900–2400 nm) wavelengths. The highest spectral variability (in terms of the coefficient of variation; data

![Image](image_url)

**FIG. 3.** Means (with standard deviation bands) of AVIRIS spectra obtained from 51 images and 237 plots stratified by dominant leaf habit (deciduous, needleleaf).
not observed in the visible regions of the spectrum (>23–50%) and was generally constant across the rest of the spectrum (~22%). Across all plots, needleleaf-dominated plots had ~10% less spectral variability than plots dominated by broadleaf species.

**PLSR model results**

Results for models built using the four aggregation schemes (Appendix: Table A3) were highly consistent. We provide model comparisons (Appendix: Figs. A1–A5) and PLSR coefficients from all canopy aggregation schemes and randomized models (Supplement), but focus here on results obtained using the top-weighted scheme. In our results, we show the model predictions from AVIRIS vs. field-based measurements (Fig. 4). Note that we show error bars in both directions: error on the \(y\)-axis is due to uncertainty of trait values from the scaling from leaf-level spectra to plot-level traits, whereas uncertainty on the \(x\)-axis is due to uncertainty from the image-based PLSR model. To our knowledge, this is the first time both uncertainties in the training data and mapped estimates have been provided. We also show standardized coefficients (Fig. 5), to compare wavebands important for mapping different traits. The shaded uncertainties in the standardized coefficients are based on the 500 permutations of the data used to generate the final maps. Additional diagnostics, such as VIP by wavelength and PRESS statistic by number of PLS coefficients (used to select the number coefficients in the PLS model), are graphed in Fig. 6. The Supplement lists the 500 raw PLS coefficients for the traits reported in this paper. These are the values that should be applied to make maps from new images that have been processed consistent with this study. We also show wavelength regions that correspond to known absorption features (from Curran 1989, Fourty et al. 1996, Curran et al. 2001). The standardized coefficients for mapping all of the traits from AVIRIS-Classic (except isotopic N, which has not previously been studied in depth) align in at least some wavelengths with known absorption features (Fig. 5; dotted bars).

Both \(M_{area}\) and %N had model fits with \(R^2 > 0.85\), which are comparable to the results reported by Martin et al. (2008); see Appendix: Table A2). Root mean square errors (RMSEs) for both variables were less than 9% of the range of data (Table 3). There was considerable spatial variation in retrieval uncertainties for both traits. High uncertainties (e.g., >2% for nitrogen or >10% for \(M_{area}\)) typically occurred at forest edges, in regions that were probably disturbed and where the dominant species were not represented in our database (e.g., crops or grasses and forbs in recently disturbed areas.). Fiber and lignin were predicted with \(R^2 > 0.60\), and carbon and cellulose with \(R^2 \sim 0.5\). Although \(R^2\) values of fiber and lignin were lower than those of \(M_{area}\) and %N, their RMSE values were within 10% of the range of data. The comparatively lower \(R^2\) values for %C and cellulose are related to the low amount of natural variability in these constituents. Isotopic \(\delta^{15}N\) had the worst fit of all models (\(R^2 = 0.48\), but the RMSE was within 15% of the range in the data. Cross-validations conducted to test model stability across years and sites were in agreement with overall model results. Averaged across all strata (by year or by geographic area) and traits, more than 90% of held-out samples were predicted within the prediction error bounds of calibration models. In addition, validation errors of prediction (RMSEs) were well within the observed range of modeled traits, i.e., with RMSE values lower than 15% of the range of variation in the trait (and usually with %RMSE < 10%; complete results are in the Supplement).

**Mapping and application**

We used the PLSR-derived coefficients (Supplement) to generate spatially explicit maps of traits and respective uncertainties. We show a subset from one of the many scenes used in this study (Green Ridge State Forest, Maryland, USA; Fig. 7). With these maps, we synthesized spatial patterns in foliar traits into images of functional variation using false color composites of key foliar traits such as nitrogen, lignin, and leaf mass per area (Appendix: Fig. A6). Such maps illustrate gradients of canopy biochemical and morphological traits that are likely to be related to species associations and provide greater detail than traditional land cover maps. The trait maps expose patterns of plant investment in foliar nutrients vs. structural compounds across environmental gradients on the landscape. Comparison of trait maps among regions reveals variability in relative abundance of traits inferred from, but not captured by, land cover classifications (Appendix: Fig. A6). Maps of functional traits from a single hyperspectral image can reveal differences in species associations across gradients that are otherwise only apparent in multi-date imagery (Appendix: Fig. A7), and, moreover, can help to identify ecosystem-level consequences of disturbance events through changes in functional traits and mapped uncertainty (Appendix: Fig. A8).

**Discussion**

In this study we confirm the capacity of imaging spectroscopy to accurately and repeatedly map functional traits important for describing landscape patterns in nutrient cycling (e.g., Chapin 2003, Kazakou et al. 2006, Quested et al. 2007, Santiago 2007) and photosynthetic capacity (e.g., Kattge et al. 2009). The approach was robust across multiple traits, using data from numerous sites, vegetation types, and years. Our method for estimating canopy chemistry integrated three scales of observation—leaf spectra, plot structure, and image spectra—thus allowing the propagation of uncertainty from the leaf-level dependent variables to the plot and canopy levels, and finally to maps created from image spectra. Although imaging spectroscopy has been used in several studies to map foliar traits (Wessman et al.
Fig. 4. PLSR model fits using AVIRIS-derived spectra fitted to foliar traits scaled from leaves to the canopy. Model performance was evaluated by averaging 500 randomized models built with a 50/50 calibration/validation split. Horizontal bars indicate uncertainty (1 SD) in predictions; vertical bars indicate uncertainty (1 SD) in plot-level estimates propagated from leaf-level measurements. The shading scheme of dots indicates increasing relative biomass of needleleaf foliage (white to black). Abbreviations are: $M_{\text{area}}$, leaf mass per area; %ADF, acid detergent fiber; %ADL, acid detergent lignin; RMSE, root mean square error.

1989, Curran et al. 1997, Serrano et al. 2002, Townsend et al. 2003, Huber et al. 2008, Martin et al. 2008), ours demonstrates the capacity to map multiple functional traits across a diversity of sites, years, images, and geographic complexity, as well as to explicitly account for the propagation of measurement, scaling, and modeling uncertainty through to the end product.

The per trait waveband calibration coefficients provide a basis for physiological interpretation of the results as well as the assessment of the results' generality.
for mapping to other data sets. The important wavebands align with fresh-leaf results for $M_{\text{area}}$ and dry-sample results for chemical constituents reported by Serbin et al. (2012, 2014), as well as with wavebands indicated by Martin et al. (2008) as important for foliar nitrogen from AVIRIS (Fig. 5). In comparison with leaf-level results, standardized coefficients matched across most of the spectrum, but there was a reversal of sign in coefficients across the red edge (500–700 nm).

Leaf-level NIR reflectance results from scattering at the interface between cell walls and intercellular air spaces, whereas canopy (i.e., image) reflectance is strongly dominated by canopy structural properties such as leaf clumping, angle, and density (Jacquemoud et al. 2009, Sullivan et al. 2013). For instance, leaf-level NIR reflectance generally increases with thickness of needle leaves (Ehleringer and Mooney 1978, Lin and Ehleringer 1983, DeLucia et al. 1996, Slaton et al. 2001), but

---

**Fig. 5.** Standardized PLSR coefficients indicating the magnitude and direction of influence of each wavelength (with ± 1 SD bands in gray). Note that raw coefficients are applied to image spectra to derive spatial predictions. Short vertical gray bars are absorption features at wavelengths known from previous studies. Dashed gray curves indicate standardized coefficients from Serbin et al. (2014); the white-filled line in the %N panel represents coefficients from Martin et al. (2008).
canopy-level NIR reflectance decreases with greater abundance of needleleaf species because absorption-inducing canopy properties (e.g., greater needle clumping) overwhelm leaf traits in affecting NIR canopy reflectance (Ollinger 2011, Sullivan et al. 2013). Considering recent critiques of imaging spectroscopy-based trait mapping (Knyazikhin et al. 2013), this switch in the direction of NIR coefficients between leaf and canopy-level models of $M_{area}$ provides added confidence that imaging spectroscopy captures sufficient information about the canopy and constituent foliage for purposes of mapping (Ollinger et al. 2013, Townsend et al. 2013). Moreover, the striking similarity of results obtained from the four schemes to scale canopy-height-related differences in measured leaf traits (Appendix A) further supports recent work emphasizing the mass basis for measuring variability in the global leaf economics spectrum (Niinemets et al. 2014), and again underscores

Fig. 6. (A) Profile of the PRESS (predicted residual sum of squares) statistic (with ±1 SD bars) obtained by fitting 500 randomized models, each using an increasing number of PLS components ($h$, on the x-axis). The final model was fit using the components that minimized the PRESS statistic. (B) Profile of the VIP (variable importance of prediction) statistic with ±1 SD gray bands. The VIP describes the relative importance of each wavelength in predicting the quantity of interest, and is considered significant where VIP > 0.8 (dotted line).
the robust ability of imaging spectroscopy to map spatial patterns from canopy-scale measurements.

We compared coefficients obtained from our analysis of %N with those published by Martin et al. (2008), the only other study using comparable multi-scene and multi-location data. Loadings of standardized coefficients for our models generally agreed with the Martin et al. models (2008); see Fig. 5. Discrepancies in locations of dips and spikes in coefficients result primarily from our use of a standardized data set (i.e., centered and scaled) compared to raw coefficients from Martin et al. (2008). Martin et al. (2008) employed 11 images (5 AVIRIS, 6 Hyperion) covering a wide variety of temperate and tropical forest ecosystems in North America and Australia; theirs was the first study to show the potential for building general models to map foliar traits across large regions.

While not as global in scope as Martin et al. (2008), our analysis incorporates multiple forest types dominated by a range of broadleaf and conifer species (Appendix: Table A2; Supplement; see Serbin et al. 2014), and covers the breadth of forest functional variability within North American eastern temperate forests. Any missing species and forest types probably are bracketed by the range of our data, and we expect that spatial patterns of uncertainties (e.g., Fig. 7) are indicative of gaps in our database for different forest optical types (sensu Ustin and Gamon 2010) rather than weaknesses in model performance. The largest uncertainties were observed in locations where (1) the vegetation was not represented in our database of foliar traits due to our focus on forests (i.e., grasses, forbs, crops, wetlands); (2) there were strong terrain effects in the imagery not fully accounted for in the topographic normalization; or (3) along forest edges and disturbed areas. In particular, the uncertainty mapping helped to identify areas that had been recently disturbed (Appendix: Fig A8). Logged areas showed a consistent pattern of higher lignin and $M_{area}$ and lower foliar nitrogen content (and also higher $\delta^{15}$N, indicating increased N availability) compared to adjacent undisturbed areas. Although such patterns are not unexpected, those areas also had consistently high uncertainties due to optical properties of disturbances being on the edge of the range of variability of optical properties for the intact forests we sampled. Future analysis of vegetation traits in disturbed areas using imaging spectroscopy may provide the opportunity to better understand changes in ecosystem dynamics brought about by human or natural perturbations.

Composite maps of functional traits (e.g., Appendix: Fig A6) revealed a greater amount of variability in forest canopies than apparent in individual maps, and in particular, land cover classifications. For example, in the Central Appalachian sites, codominant conifers in ridges and valleys were revealed by the higher lignin content and $M_{area}$ of conifers (Fig. 7). Comparisons between maps of individual traits (Fig. 7) also revealed gradients in functional traits that follow the inverse relationship between investments in leaf structural mass in contrast to leaf nutrients (Wright and Westoby 2002, Reich et al. 2003, Wright et al. 2004). Regions showing higher foliar nitrogen concentrations were associated with lower values of $M_{area}$ and vice versa. Although this is widely known from multiple species inhabiting diverse ecosystems (Reich et al. 1998, 1999, Wright et al. 2001, Sanchez-Azofeifa et al. 2009, Asner et al. 2011b), imaging spectroscopy allowed these associations to be mapped explicitly and across multiple ecoregions.

The geographic patterns in trait variability matched expectations associated with climate and other environmental drivers (Table 2). For example, the negative correlation between $M_{area}$ and leaf nitrogen is in agreement with the broader concept of the leaf economic
The positive correlation of foliar N with regional-scale N deposition corroborates evidence of foliar N enrichment reported in numerous studies in northeastern forests (McNeil et al. 2007, 2012, Davis et al. 2009, Crowley et al. 2012). Correlations of foliar $\delta^{15}$N concentrations with mean annual temperature (positive) and precipitation (negative) agree with results in Amundson et al. (2003), Craine et al. (2009), and Serbin et al. (2014). Mapped evidence of increased isotopic N along the N deposition gradient corroborates Emmett et al. (1998) and McLauchlan et al. (2007), showing that greater N availability yields less preferential retention of lighter $^{14}$N in the soil system.

It has been recently demonstrated that estimates of canopy structure and composition obtained from imaging spectroscopy are invaluable in constraining regional carbon flux models (Antonarakis et al. 2014). Maps of foliar traits obtained from this study can be leveraged to drive ecosystem process models requiring parameterization of foliar nitrogen content and leaf mass per area, which are coordinated in leaves to maximize carbon fixation (Shipley et al. 2005) in patterns independent of species-level physiognomy (Reich et al. 1991, 1992, Grime 2006). This offers the potential to develop data sets needed to drive spatially explicit ecosystem process models at large scales from proposed forthcoming satellites (e.g., HyspIRI, EnMAP; see Stufler et al. 2007, Middleton et al. 2013). Although spatial patterns of canopy foliar N, lignin, and $M_{\text{area}}$ are largely interpretable in terms of species associations, patterns of $\delta^{15}$N will require considerable follow-up field

---

**Fig. 7.** Spatial predictions of $M_{\text{area}}$, %N, %C, $\delta^{15}$N, %Cellulose, %ADL, and %ADF means (left panels) and uncertainties, i.e., standard deviations (right panels), obtained by applying 500 sets of coefficients obtained from randomized PLSR models to AVIRIS imagery acquired over the Green Ridge State Forest, Maryland (Fig. 1, box F). Each model was built with a 50/50 calibration/validation split. See Table 3 for model statistics and the Supplement for raw PLSR coefficients.
research. In locations where δ^{15}N concentrations appear elevated without associated increased uncertainties or changes in foliar N and or M_{area} (indicating an obvious change in plant functional associations), localized patterns of δ^{15}N enrichment (for e.g., see Fig. 7, SE corner in δ^{15}N map) could be an indicator of elevated N availability via ephemeral disturbances such as insect defoliation or logging. The availability of spatial predictions of foliar traits may therefore open up new investigations into biotic or abiotic nutrient cycling controls across both small regions and large landscapes.

Imaging spectroscopy offers the opportunity for applications ranging from assessment of biochemical effects of invasive species (Glenn et al. 2005, Asner et al. 2008, He et al. 2011), to characterizing photosynthetic down-regulation (Gamon et al. 1990, 1992, 1997) and measurement of the inductance of plant defense to perturbations (Couture et al. 2013). By directly measuring the chemical and physiological consequences of disturbances and other environmental drivers on vegetation, space-borne imaging spectroscopy should provide the ability to directly quantify the consequences of environmental change rather than inferring it from land cover maps or vegetation indices. This will provide agencies the tools to measure impacts of invasions or help to better guide management activities.

**Conclusions**

Using data from a wide range of forest ecosystems, we rigorously tested the capacity of imaging spectroscopy to map multiple forest functional traits. The resulting maps enable us to describe fundamental axes of variability in plant physiology along a spectrum ranging from “fast” (high nutrients and low lignin, thin, low M_{area} leaves) to “slow” patterns of forest nutrient cycling (Reich et al. 1991, 1992, 1999, Craine et al. 2002, Wright et al. 2005a, b, Sanchez-Azofeia et al. 2009). It is exciting to begin answering questions on the spatial patterns in these traits, but ultimately, trait mapping from imaging spectroscopy is likely to have its greatest impact by enabling researchers to pose more applied questions. For example: how does functional biogeography (Violle et al. 2014) mediate ecosystem response to environmental change? By helping to answer such questions, multiple-trait mapping, coupled with robust assessments of its uncertainty, opens up the potential for making defensible, tractable, and scalable predictions of the delivery of key ecosystem services from complex forest environments that are increasingly subjected to multiple agents of global environmental change.

Our work accomplishes two objectives: we provide a set of equations (Supplement) that can be used with imaging spectroscopy data to maps traits in similar environments (temperate forests). As well, we outline a workflow that could be implemented in other ecosystem types. An important next step for trait mapping is the integration of data collected from across the globe (Asner and Martin 2009, Asner et al. 2012). This integration will further determine the extent of the generality of these methods, with respect to sensor types, measurement or analytical strategies, and environmental variability. Coupled with other studies, this research represents a step toward the development of global approaches to mapping plant functional properties from imaging spectroscopy, and thus helps to provide a more solid foundation for the development of “trait” products from the proposed HyspIRI mission and NEON’s Aerial Observatory Platform (Kampe et al. 2010, Middleton et al. 2013, Sims et al. 2013).

**Acknowledgments**

A. Singh and S. P. Serbin were supported by graduate assistantships provided by NASA Terrestrial Ecology grant NNX08AN31G to P. A. Townsend and B. E. McNeil, and a NASA Earth and Space Science Fellowship grant NNX08AV07H to S. P. Serbin. S. P. Serbin was also supported in part by DOE contract No. DE-SC00112704 to Brookhaven National Laboratory. Thanks also to E. Kruger for helpful comments on earlier versions of this paper. The authors thank two anonymous reviewers for very constructive comments on the manuscript.

**Literature Cited**


December 2015

FOLIAR TRAITS FROM IMAGING SPECTROSCOPY


Meier, C. L., and W. D. Bowman. 2008. Links between plant litter chemistry, species diversity, and below-ground ecosystem...


