

# Tree-mycorrhizal associations detected remotely from canopy spectral properties

JOSHUA B. FISHER<sup>1,2</sup>, SEAN SWEENEY<sup>3</sup>, EDWARD R. BRZOSTEK<sup>4</sup>, TOM P. EVANS<sup>3,5</sup>, DANIEL J. JOHNSON<sup>6</sup>, JONATHAN A. MYERS<sup>7</sup>, NORMAN A. BOURG<sup>8,9</sup>, AMY T. WOLF<sup>10</sup>, ROBERT W. HOWE<sup>10</sup> and RICHARD P. PHILLIPS<sup>11</sup>

<sup>1</sup>Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA, <sup>2</sup>Joint Institute for Regional Earth System Science and Engineering, University of California at Los Angeles, 607 Charles E Young Drive East, Young Hall #4242, Los Angeles, CA 90095-7228, USA, <sup>3</sup>Center for the Study of Institutions, Populations, and Environmental Change (CIPEC), Indiana University, Bloomington, Indiana 47408, USA, <sup>4</sup>Department of Biology, West Virginia University, 53 Campus Drive, Morgantown, WV 26506, USA, <sup>5</sup>Department of Geography, Indiana University, Student Building 120, Bloomington, IN 47405, USA, <sup>6</sup>Yale School of Forestry and Environmental Studies, Kroon Hall, 195 Prospect St., New Haven, CT 06511, USA, <sup>7</sup>Department of Biology, Washington University in St. Louis, Saint Louis, MO 63130, USA, <sup>8</sup>Conservation Ecology Center, Smithsonian Conservation Biology Institute, National Zoological Park, 1500 Remount Road, Front Royal, VA 22630, USA, <sup>9</sup>U.S. Geological Survey, National Research Program - Eastern Branch, 12201 Sunrise Valley Drive, Reston, VA MS430, USA, <sup>10</sup>Department of Natural and Applied Sciences and Cofrin Center for Biodiversity, University of Wisconsin-Green Bay, 2420 Nicolet Drive, Green Bay, WI 54311, USA, <sup>11</sup>Department of Biology, Indiana University, 247 Jordan Hall, 1001 E. Third St., Bloomington, IN 47405, USA

## Abstract

A central challenge in global ecology is the identification of key functional processes in ecosystems that scale, but do not require, data for individual species across landscapes. Given that nearly all tree species form symbiotic relationships with one of two types of mycorrhizal fungi – arbuscular mycorrhizal (AM) and ectomycorrhizal (ECM) fungi – and that AM- and ECM-dominated forests often have distinct nutrient economies, the detection and mapping of mycorrhizae over large areas could provide valuable insights about fundamental ecosystem processes such as nutrient cycling, species interactions, and overall forest productivity. We explored remotely sensed tree canopy spectral properties to detect underlying mycorrhizal association across a gradient of AM- and ECM-dominated forest plots. Statistical mining of reflectance and reflectance derivatives across moderate/high-resolution Landsat data revealed distinctly unique phenological signals that differentiated AM and ECM associations. This approach was trained and validated against measurements of tree species and mycorrhizal association across ~130 000 trees throughout the temperate United States. We were able to predict 77% of the variation in mycorrhizal association distribution within the forest plots ( $P < 0.001$ ). The implications for this work move us toward mapping mycorrhizal association globally and advancing our understanding of biogeochemical cycling and other ecosystem processes.

**Keywords:** canopy, landscape, mycorrhizae, nutrients, remote sensing, species, spectral

Received 6 April 2015; revised version received 27 January 2016 and accepted 19 February 2016

## Introduction

Nearly all tree species form symbiotic relationships with one of two types of mycorrhizal fungi – arbuscular mycorrhizal (AM) or ectomycorrhizal (ECM) fungi – exchanging carbon from trees for nutrients from fungi. Recently, several investigators have hypothesized that the type of mycorrhizal fungi with which a tree species forms a symbiosis (i.e., the mycorrhizal association) represents an integration of multiple plant and soil microbial functional traits (Read & Perez-Moreno, 2003), such that forests dominated by AM or ECM asso-

ciations exhibit distinct nutrient economies (Chapman *et al.*, 2006; Phillips *et al.*, 2013). AM-associated trees generally have relatively high foliar phosphorus and leaf mass area (Koele *et al.*, 2012), early leaf out (Key *et al.*, 2001; McCormack *et al.*, 2014), and leaf litters that decompose rapidly (Cornelissen *et al.*, 2001; Hobbie, 2006), resulting in the predominance of inorganic forms of nutrients characterized by accelerated carbon and nutrient cycling (Phillips *et al.*, 2013; Midgley & Phillips, 2014; Brzostek *et al.*, 2015; Waring *et al.*, 2015) (but, see also, Koele *et al.*, 2012). In contrast, ECM-associated trees have relatively low foliar phosphorus and leaf mass area, late leaf out, more slowly decomposing leaf litter, a greater proportion of nutrients in organic forms, and tend to slow carbon and nutrient cycling (Vesterdal

Correspondence: Joshua B. Fisher, tel. 1 323 540 4569, fax 1 818 354 9476, e-mail: joshbfisher@gmail.com

*et al.*, 2013; Averill *et al.*, 2014; Soudzilovskaia *et al.*, 2015).

The relative abundance of AM- and ECM-associated trees in an ecosystem or region may provide an index of differences in nutrient economics within and among forests. As such, a broad classification approach that considers mycorrhizal associations in forests from the plot scale to the regional scale can facilitate improved projections of the biogeochemical cycling and impacts of tree species changes. However, detecting and mapping AM- and ECM-associated trees across the landscape represents a grand challenge in global ecology. While AM- and ECM-dominated forests occur relatively homogeneously across some biome types (e.g., AM-associated trees in tropical forests, ECM-associated trees in boreal forests), most forests in temperate regions contain highly mixed assemblages of AM- and ECM-associated trees (Allen *et al.*, 1995; Read & Perez-Moreno, 2003; Phillips *et al.*, 2013). Complicating this classification, temperate forests have experienced numerous disturbances over the last century, leading to highly heterogeneous biogeochemical regimes across disturbance and successional histories (Abrams, 1992; Nowacki & Abrams, 2008). Still, given that mycorrhizal associations of most temperate tree species are known (Brundrett *et al.*, 1990; Wang & Qiu, 2006), unique spectral signatures between AM- and ECM-associated trees would enable large-scale remote sensing mapping of the mycorrhizal framework with which to track ecosystem functions across landscapes.

Mapping tree species from remote sensing has a long and varied history (Martin *et al.*, 1998; Turner *et al.*, 2003). High-resolution hyperspectral and Light Detection And Ranging (LiDAR) approaches have demonstrated a high degree of success in accurate classification of multiple species across heterogeneous assemblages (Brandtberg *et al.*, 2003; Holmgren & Persson, 2004; Moffiet *et al.*, 2005; Heinzl & Koch, 2011). However, these studies typically require an airborne platform to increase the spatial resolution necessary to detect individual trees, thereby limiting the ability to map mycorrhizal associations across large geographic areas.

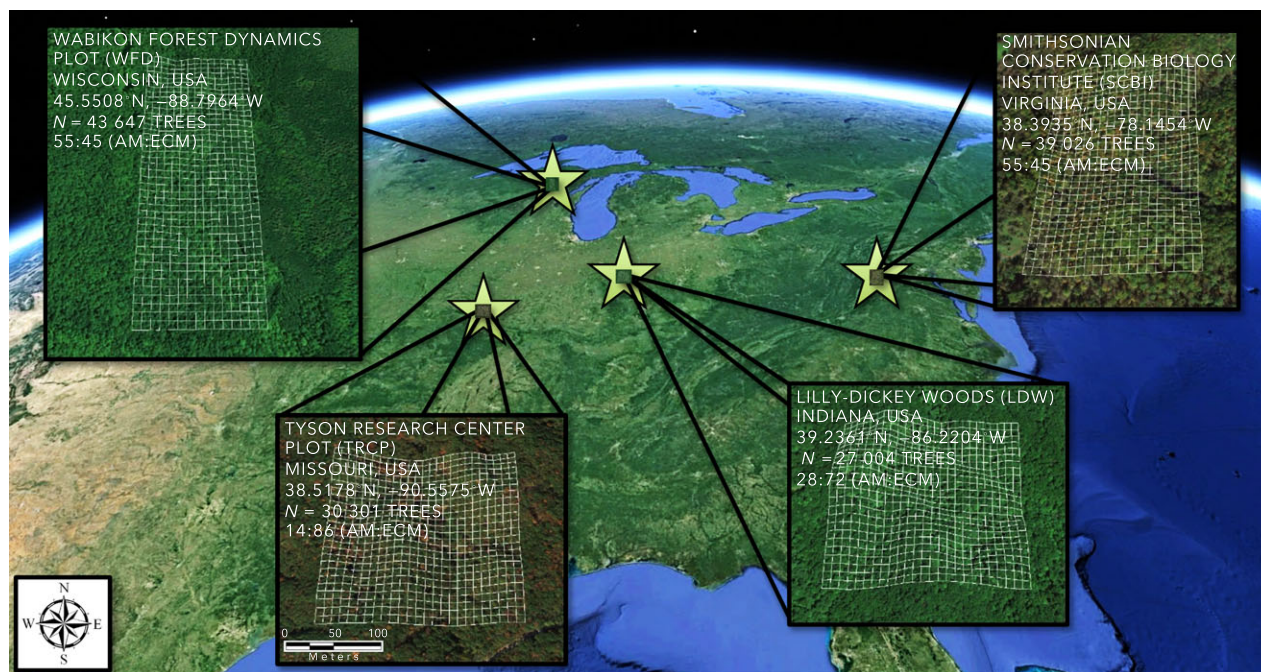
Larger scale, medium-resolution satellite data (e.g., Landsat) have been used successfully to create species-level forest classifications (Walsh, 1980; Schriever & Congalton, 1995; Wolter *et al.*, 1995; Mickelson *et al.*, 1998; Key *et al.*, 2001; Wittmann *et al.*, 2002; Saatchi *et al.*, 2008; Ohmann *et al.*, 2011; Wilson *et al.*, 2012). These approaches are particularly effective for homogeneous species assemblages; however, increases in species diversity often lead to proportional decreases in accuracy. Refinements to reduce accuracy loss can be obtained through integration of aerial photography or

spectral mixture analysis (Martin *et al.*, 1998; Key *et al.*, 2001; Plourde *et al.*, 2007). Still, with these additions/refinements, the complexity of processing increases by more than an order of magnitude, including rapidly rising poorly constrained degrees of freedom, through necessary processing steps such as co-registration, vignetting, atmospheric correction, solar orientation, illumination, radiometric aerial triangulation, endmember identification and differentiation, and model reliance (Key *et al.*, 2001; Underwood *et al.*, 2003; Tuominen & Pekkarinen, 2004; Plourde *et al.*, 2007; Chandelier & Martinoty, 2009). Therefore, large-scale mapping of individual tree species, especially in mixed-species environments, is highly labor intensive with inconsistent accuracies. Thus, the grand challenge of large-scale mapping of mycorrhizal association would seem to remain a grand challenge, that is, if we remained within the paradigm of 'tree species equals mycorrhizal association'.

Here, we present a paradigm shift in understanding how to map and detect the distribution of mycorrhizal associations. We extend the large body of work in remote sensing of individual tree species to mapping mycorrhizal associations among groups of species. We ask, are tree species that associate with one type of mycorrhizal fungi more similar, in terms of their spectral characteristics, than tree species that associate with the other type of mycorrhizal fungi? For instance, maples, tulip poplars, and white ashes may all be unique species with unique spectral signatures, but they all associate with AM fungi only; oaks, American beeches, and hickories likewise all associate with ECM fungi only, again each with unique spectral signatures. But, are there spectral signatures that are both in common with the first group that differ from spectral signatures that may be in common with the second group? Like individual fingers with unique fingerprints – i.e., tree species – acting in concert as part of the left or right hand (AM or ECM), actions of the respective hand may be detectable and distinguishable. We test a new paradigm in mapping mycorrhizal association on large scales that does not require mapping of individual species. This framework can both circumvent problems of individual tree identification requirements in moderate-scale satellite data, and provide direct and consistent measures of mixed pixel – and mixed below-ground associations – information.

Our objectives were as follows:

1. To identify spectral and phenological characteristics in 30 m resolution Landsat multi-band data across a gradient of AM to ECM dominant association ecosystems (Fig. 1);
2. To develop and test a statistical model to predict AM vs. ECM dominance, as well as mixed-dominance areas;



**Fig. 1** Study sites spanned mycorrhizal dominance throughout the temperate United States, encompassing ~130 000 trees. Site locations included Wabikon Forest Dynamics Plot (WFD) in Wisconsin, Tyson Research Center Plot (TRCP) in Missouri, Lilly-Dickey Woods (LDW) in Indiana, and Smithsonian Conservation Biology Institute (SCBI) in Virginia. Arbuscular mycorrhizal (AM)-associated trees comprise the majority of basal area in WFD and SCBI; whereas, ectomycorrhizal (ECM)-associated trees are predominantly prevalent in TRCP and LDW.

3. To classify and map mycorrhizal associations over large areas based on spaceborne observations.

## Materials and methods

### Study area

Forest plot data were collected from long-term research sites throughout the temperate United States that are part of the Smithsonian Institution's Center for Tropical Forest Science-Forest Global Earth Observatory (CTFS-ForestGEO) network of forest dynamics plots (Anderson-Teixeira *et al.*, 2015): south-central Indiana (39.2361°N, -86.2204°W; 0.25 km<sup>2</sup>), northwest Virginia (38.8935°N, -78.1454°W; 0.25 km<sup>2</sup>), north-east Wisconsin (45.5508°N, -88.7964°W; 0.25 km<sup>2</sup>), and east-central Missouri (38.5178°N, -90.5575°W; 0.20 km<sup>2</sup>) (Fig. 1). The Indiana site is the Lilly-Dickey Woods (LDW) forest dynamics plot, which is a part of Indiana University's Research and Teaching Preserves. The Virginia site is part of the Smithsonian Conservation Biology Institute (SCBI) forest dynamics plot at the Smithsonian National Zoological Park's Conservation Biology Institute. The Wisconsin site includes the Wabikon Forest Dynamics (WFD) plot and is situated within the Chequamegon-Nicolet Forest. The Missouri site includes the Tyson Research Center Plot (TRCP), located at Washington University in St. Louis' Tyson Research Center.

Elevations range from 172 to 233 m at TRCP, 192 to 311 m at LDW, 233 to 301 m at SCBI, and 488 to 512 m at WFD.

Topography varies between sites with slopes averaging 4.3° at WFD, 9.3° at LDW, 11.3° at SCBI, and 16.5° at TRCP. The climates are similar – temperate continental with minimum temperatures occurring in January and maximum temperatures in July. Total precipitation averages 78 cm at WFD, 94 cm at SCBI, 98 cm at TRCP, and 104 cm at LDW with May generally the wettest month and January/February the driest.

Each site contains a rich assemblage of hardwood tree species colonized by both AM and ECM fungi. All tree stems equal to or larger than 1 cm diameter at breast height (DBH) were measured, mapped, and identified for species (27 004 in LDW, 30 301 in TRCP, 39 026 in SCBI, and 43 647 in WFD) (Bourg *et al.*, 2013; Johnson *et al.*, 2014; Spasojevic *et al.*, 2014). Tree species were then classified as either AM- or ECM-associated based on known fungal associations (also referred to as 'mycorrhizal composition') (Brundrett *et al.*, 1990; Wang & Qiu, 2006). These data provided the ground truth data with which to compare our model. Our study included 77 tree species, 43 of which were AM-associated and 34 of which were ECM-associated (Table S1). The LDW and TRCP sites are predominantly ECM-associated species, 72.3% and 86.3% of total basal area, respectively. Conversely, AM-associated species comprise the majority at the more mixed SCBI and WFD sites, 54.9% and 55.2% of total basal area, respectively.

High-precision GPS coordinates were collected at all 20 m subplot corners and subsequently converted to polygon shapefiles with attributed plot data. To minimize geolocation error between field data and satellite data, a 3 × 3 aggregation was applied to each site's matrix of polygon cells, combining

adjacent features into new  $60 \times 60$  m polygons without overlap. For example, in the  $20 \times 20$  m LDW plot matrix, the aggregated cells  $i,j$  were centered at  $x(3i+2, 3j+2)$  for  $i = 0 \dots 7$  and  $j = 0 \dots 7$ . Total tree basal area, total AM-associated tree basal area, and total ECM-associated tree basal area were updated for the enlarged areas of the aggregated polygons by  $\text{sum}(x(k, m))$ , for  $k = 3i+1$  to  $3i+3$  and  $m = 3j+1$  to  $3j+3$  for  $i = 0 \dots 7$  and  $j = 0 \dots 7$ , expressed as  $\sum_{k=3i+1}^{3i+3} \sum_{m=3j+1}^{3j+3} x_{k,m}$  for  $i = 0 \dots 7$  and  $j = 0 \dots 7$ . Our study encompassed 64 of these 60 m subplots at the LDW site, 60 at SCBI, 70 at WFD, and 56 at TRCP. At SCBI, 12 plots encompassed a deer enclosure, and were thus excluded from this analysis (and grayed out in figures). High-resolution topographic data were available for each study location from which we generated topographic characteristics, including elevation, slope, and aspect.

We analyzed the carbon and nitrogen content of green leaves, leaf litter, and soil from ECM- and AM-dominated plots ( $12 \times 12$  m) at the LDW site (15 plots total). During the peak of the growing season in 2012, leaves were collected from the top of the canopy from four dominant trees in each plot. In 2012, soils were collected from the top 15 cm of mineral soil at four times across the growing season. Litterfall was collected during the fall of 2013 using litter traps in each plot. The litter was sorted to mycorrhizal association. All samples were dried at  $60^\circ\text{C}$ , ground, and run on an elemental analyzer for C and N content (ECS 4010; Costech Analytics, Valencia California). For the green leaves only, we estimated specific leaf area from the area (LI-3100C; LiCor Biosciences, Lincoln, NE, USA) and the mass. For the soils only, we measured pH after preparing soil/DI water slurries for each sample. To test for differences between mycorrhizal type for all response variables, we performed a one-way ANOVA using the 'aov' function in the R statistical package.

### Image processing

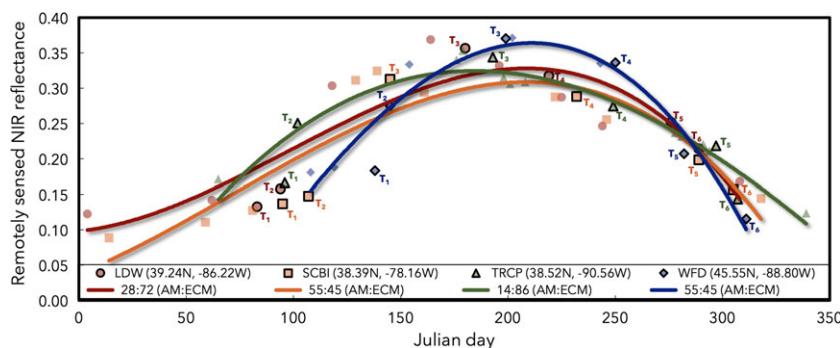
Landsat 5 Thematic Mapper (TM) data were downloaded from the USGS Global Visualization Viewer (GLOVIS;

glovis.usgs.gov) over each site from 2008 to 2011. The total of cloud-free images available for each site included the following: 14 for LDW, 15 for SCBI, 19 for TRCP, and 11 for WFD. Nominal scene center locations are as follows: LDW (Path: 021/Row: 033), SCBI (Path: 016/Row: 033), TRCP (Path: 024/Row: 033), and WFD (Path: 024/Row: 028). Quantized calibrated pixel values were converted to at-sensor radiance based on the minimum and maximum spectral radiance for each band (Markham & Barker, 1987). At-sensor radiance values were then converted to top of atmosphere (TOA) reflectance using the apparent reflectance model (Chavez, 1989):  $R_\lambda = \text{PI} * D^2 * L_{\lambda\text{sensor}} / (\text{ESun}_\lambda * \cos(\theta))$ , where  $R_\lambda$  = surface reflectance,  $D$  = distance between the Earth and the sun (AU),  $L_{\lambda\text{sensor}}$  = apparent at-sensor radiance,  $\text{ESun}_\lambda$  = exo-atmospheric solar irradiance, and  $\theta$  = solar zenith angle.

### Statistical model

We constructed an ordinary least squares (OLS) model to predict the proportion of AM-associated (Perc\_AM) or ECM-associated (Perc\_ECM) tree species (dependent variable) within the  $60 \times 60$  m plots. Either percent AM or ECM can be assigned the dependent variable because they sum to unity; the difference in model output would simply be a change in sign of the regression coefficients. Seasonal near-infrared (TM<sub>4</sub>) reflectance profiles were constructed for each site using the library of acquired images. Six periods were identified that matched phenological events across sites: leaf flush (T<sub>1</sub>), green-up (T<sub>2</sub>), peak green (T<sub>3</sub>), early leaf senescence (T<sub>4</sub>), late leaf senescence (T<sub>5</sub>), and leaf abscission (T<sub>6</sub>) (Fig. 2).

The pool of candidate predictor variables included temporal reflectance of Landsat bands 1–5 and 7 (herein referred to as TM<sub>n</sub>), seasonal normalized difference vegetation index (NDVI), intraband differencing of temporal reflectance, and topographic characteristics of elevation and slope (Table 1). Mean predictor values were extracted for each  $60 \times 60$  m plot, where the reflectance contribution of each pixel was weighted by the area of intersection between the pixel and the 60 m plot.



**Fig. 2** Near-infrared (NIR) reflectance at the top of atmosphere from Landsat TM<sub>4</sub> time series shows phenological evolution as a function of both climate and ecosystem properties for each of the four sites: Lilly-Dickey Woods (LDW), Smithsonian Conservation Biology Institute (SCBI), Tyson Research Center Plot (TRCP), and Wabikon Forest Dynamics (WFD); the legend also includes the latitude/longitude and the overall site proportion between arbuscular mycorrhizal (AM)- and ectomycorrhizal (ECM)-associated trees. Third-order polynomials are fit for visualization. Six periods are highlighted with black outline for individual points that correspond to phenological events: leaf flush (T<sub>1</sub>), green-up (T<sub>2</sub>), peak green (T<sub>3</sub>), early leaf senescence (T<sub>4</sub>), late leaf senescence (T<sub>5</sub>), and leaf abscission (T<sub>6</sub>).

**Table 1** Candidate predictors for the unified model to predict percentage of arbuscular mycorrhizal (AM)- vs. ectomycorrhizal (ECM)-associated tree occurrence typically used different reflectance signals weighted by phenological events. T<sub>1</sub> is temporal acquisition point (see Fig. 2). TM<sub>n</sub> is the Landsat Thematic Mapper band number

Predictor	Spectral window (μm)	Description
Elevation	n/a (meters)	Height above sea level
Slope	n/a (degrees)	Slope
T <sub>1</sub> TM <sub>1</sub>	0.45–0.52	Visible blue
T <sub>1</sub> TM <sub>2</sub>	0.52–0.60	Visible green
T <sub>1</sub> TM <sub>3</sub>	0.63–0.69	Visible red (related to chlorophyll absorption)
T <sub>1</sub> TM <sub>4</sub>	0.76–0.90	Near-infrared (related to plant health)
T <sub>1,3,5,6</sub> TM <sub>4</sub>	0.76–0.90	Average reflectance of band 4 over T <sub>1</sub> , T <sub>3</sub> , T <sub>5</sub> , T <sub>6</sub>
T <sub>1</sub> TM <sub>5</sub>	1.55–1.75	Mid-infrared (related to turgidity)
T <sub>1</sub> TM <sub>7</sub>	2.08–2.35	Mid-infrared
T <sub>1</sub> TM <sub>4</sub> /T <sub>1</sub> TM <sub>5</sub>	TM <sub>4</sub> , TM <sub>5</sub>	Ratio of band 4 and band 5 (related to moisture content)
T <sub>1</sub> NDVI	TM <sub>3</sub> , TM <sub>4</sub>	Normalized Difference Vegetation Index (related to canopy density)

**Table 2** Statistical influence of the 10 most significant predictor variables on estimation of percent arbuscular mycorrhizal (AM) association for the unified model. See Table 1 for descriptions of the variables. All predictor variables contributed significantly to the model at  $\alpha = 0.05$ . Normalized difference vegetation index (NDVI) measured during peak green (T<sub>3</sub>NDVI) resulted in the greatest change in percent AM association by unit increase, followed by T<sub>3</sub>TM<sub>3</sub> and T<sub>6</sub>TM<sub>2</sub>. Six of ten variables were positively correlated with percent AM association, and four were negatively correlated

Model variable	Unstandardized coefficients		Std. coefficients		Sig.	Correlations			Collinearity statistics	
	B	Std. error	Beta	t		Zero-Order	Partial	Part	Tolerance	VIF
(Constant)	−992.6	126.1		−7.9	0.000					
Slope	0.3	0.2	0.20	2.2	0.026	−0.38	0.15	0.07	0.14	7.16
T <sub>5</sub> TM <sub>2</sub>	−622.7	189.7	−0.97	−3.3	0.001	0.24	−0.21	−0.11	0.01	75.00
T <sub>6</sub> TM <sub>2</sub>	1589.4	177.8	1.34	8.9	0.000	0.06	0.51	0.29	0.05	21.96
T <sub>1</sub> TM <sub>3</sub>	120.5	30.1	0.25	4.0	0.000	0.21	0.26	0.13	0.25	3.94
T <sub>3</sub> TM <sub>3</sub>	3804.7	769.4	1.67	5.0	0.000	−0.11	0.31	0.16	0.01	92.14
T <sub>1,3,5,6</sub> TM <sub>4</sub>	−687.9	64.3	−0.66	−10.7	0.000	−0.34	−0.58	−0.34	0.27	3.71
T <sub>3</sub> TM <sub>4</sub> /T <sub>3</sub> TM <sub>5</sub>	−109.8	11.3	−0.66	−9.8	0.000	−0.01	−0.54	−0.31	0.22	4.46
T <sub>4</sub> TM <sub>4</sub> /T <sub>4</sub> TM <sub>5</sub>	65.1	9.6	0.41	6.8	0.000	0.19	0.41	0.22	0.27	3.65
T <sub>1</sub> NDVI	−96.3	25.3	−0.77	−3.8	0.000	−0.36	−0.25	−0.12	0.03	40.19
T <sub>3</sub> NDVI	1366.9	127.7	2.63	10.7	0.000	0.20	0.58	0.34	0.02	59.25

Candidate predictors were categorized into three groups: (i) TM<sub>n</sub> reflectance during the growth phase (t<sub>1</sub> through t<sub>3</sub>); (ii) TM<sub>n</sub> reflectance during the senescence phase (t<sub>4</sub> through t<sub>6</sub>); and (iii) NDVI and the TM<sub>4</sub>:TM<sub>5</sub> ratio for the entire year, mean TM<sub>4</sub> reflectance, elevation, and slope. A total of 18 variables comprised the first group, 18 in the second group, and 15 in the third group. Each group of candidate variables was separately entered into a linear regression model with bootstrap aggregation, where information was aggregated across iterations and the predictors scored, retaining the five highest-ranking predictors from each model ensemble. The 15 retained predictors were used in a stepwise regression model producing a final model of 10 predictors (Table 2).

The OLS regression equation was initially computed using all variables. The least significant variables were iteratively

eliminated using backwards exclusion, evaluating the partial *F*-statistic of each remaining variable against  $F_{\alpha} = 0.10$ , until all remaining variables significantly contributed to the prediction of Perc\_AM or Perc\_ECM. The partial *F*-statistic (*F* to remove) was calculated as follows: Partial *F*-statistic =  $(RSS_1 - RSS_2)/MSE_1$ , where  $RSS_1$  is the residual sum of squares for all variables currently in the equation,  $RSS_2$  is the residual sum of squares with one variable removed from the equation, and  $MSE_1$  is the mean-squared error for all variables remaining in the equation. Ten explanatory variables were selected from the subset pool of candidate predictors.

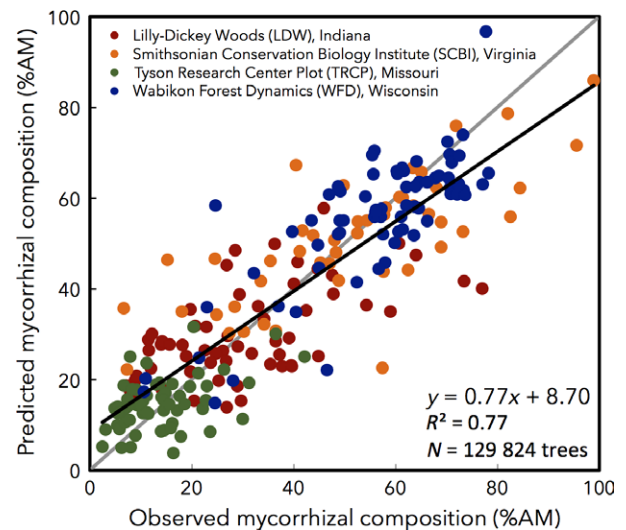
For statistical independence and robustness, we also performed multiple iterations of the model construction by splitting the training dataset into calibration and independent validation sets, holding all predictors constant. In total, five

additional sets were constructed: the model calibrated to each site individually then tested against the other three sites (a total of four sets), and 50% of the data randomly across all sites used for calibration with the other half of the data used for validation.

## Results

Seasonal reflectance patterns varied by site both in terms of reflectance magnitude and phenological timing (Fig. 2). The overall duration and magnitude of the growth and decline phases ( $t_1$ – $t_3$ ) were similar among the lower latitude sites, and much longer in duration than in the higher latitude Wisconsin site (WFD). Timing and duration of the growth and decline phases were similar in the Indiana (LDW) and Missouri (TRCP) sites, but near-infrared reflectance magnitudes were markedly different, particularly during the growth phase. The temporal window from leaf flush ( $t_1$ ) to peak green ( $t_3$ ) was most narrow (i.e., rapid green-up) in the Virginia (SCBI) and Wisconsin (WFD) sites, although the duration of the decline phase ( $t_3$ – $t_6$ ) was longer in SCBI. The magnitude of peak green reflectance was also least in SCBI.

Using a unified predictive model across all sites, 77% of the variation in mycorrhizal association was explained ( $r^2_{\text{adj}} = 0.76$ ;  $\sigma_{\text{est}} = 11.35$ ;  $F(10, 226) = 75.5$ ;  $P < 0.05$ ) (Fig. 3). Predicted per-plot percent AM or ECM association ranged from 4% to 97%, while measured percent AM or ECM association ranged from 3% to 98%, indicating that both the data and the methodological approach can capture nearly a full range. The amount of common variance in the model shared by all predictors with the dependent variable was found by examining the semipartial correlations by subtracting the sum of squared semipartial correlations from the overall  $r^2$ , equal to 0.22. Variance inflation factors (VIFs) were greater than 2 for all predictors, indicating collinearity among the variables. Collinearity diagnostics revealed four variable eigenvalues approaching 0 with condition index values exceeding 30. However, collinearity does not adversely affect the overall fit of the regression model, and its presence does not impact the predictions made by the model so long as the sample is representative of the area to which the model is being applied. Residuals vs. predicted values indicated no systematic deviation from the reference line that would suggest nonlinearity in one or more variables. No serial correlation was detected in residuals; Durbin–Watson ( $d$ ) = 1.8. The data exhibited homoscedasticity with error variance constant, while varying values in the predicted variable. Finally, the normal probability plot of the residuals revealed minimal departure from normality, but not significant to warrant modification

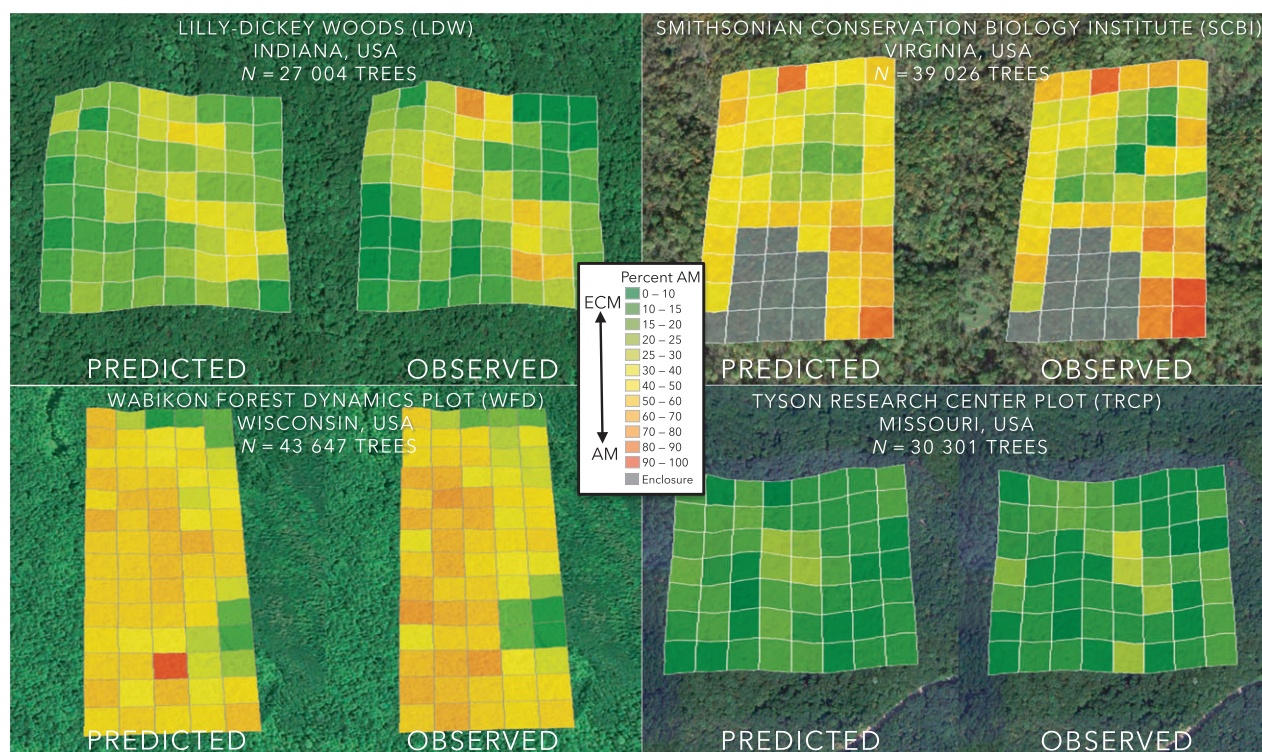


**Fig. 3** Predicted vs. observed mycorrhizal association forest composition across ~130 000 trees (aggregated to 60 m plots). Four megaplots are color-differentiated: Lilly-Dickey Woods (LDW; red), Smithsonian Conservation Biology Institute (SCBI; orange), Tyson Research Center Plot (TRCP; green), and Wabikon Forest Dynamics (WFD; blue). Arbuscular mycorrhizal (AM)-associated trees comprise the majority of basal area in WFD and SCBI, whereas ectomycorrhizal (ECM)-associated trees are predominantly prevalent in TRCP and LDW. The plot is given in percentage AM association composition (or, inversely, ECM association composition). The black best-fit line is across all sites with statistics shown in lower right; the gray line is the 1 : 1 line.

to the model. The statistics for the unified model are shown in Table 2.

The model was also tested by splitting the dataset into multiple sets for calibration and validation. In the test whereby we withheld half of the data across all sites for calibration and the other half for validation, holding the unified model (all sites, all data) predictors constant, 75% of the variation in mycorrhizal association was explained across the validation set (76% across the calibration set; 2% degradation in predictive power from the unified model) (Fig. S1). Nine of the ten predictor variables remained statistically significant; however,  $T_1B_3$  was no longer significant. We also tested the robustness of each individual site to predict the patterns of all of the other sites. No individual site's variability was able to robustly capture the variability in the other sites (10–31% predictive power), although the within-site predictive power (68–79%) was similar to the half-data and unified models.

The predicted (unified model) and observed surfaces of percent AM are shown spatially for all four sites in Fig. 4. At the ECM-dominated LDW and TRCP sites, the model tended to overestimate the percent AM by



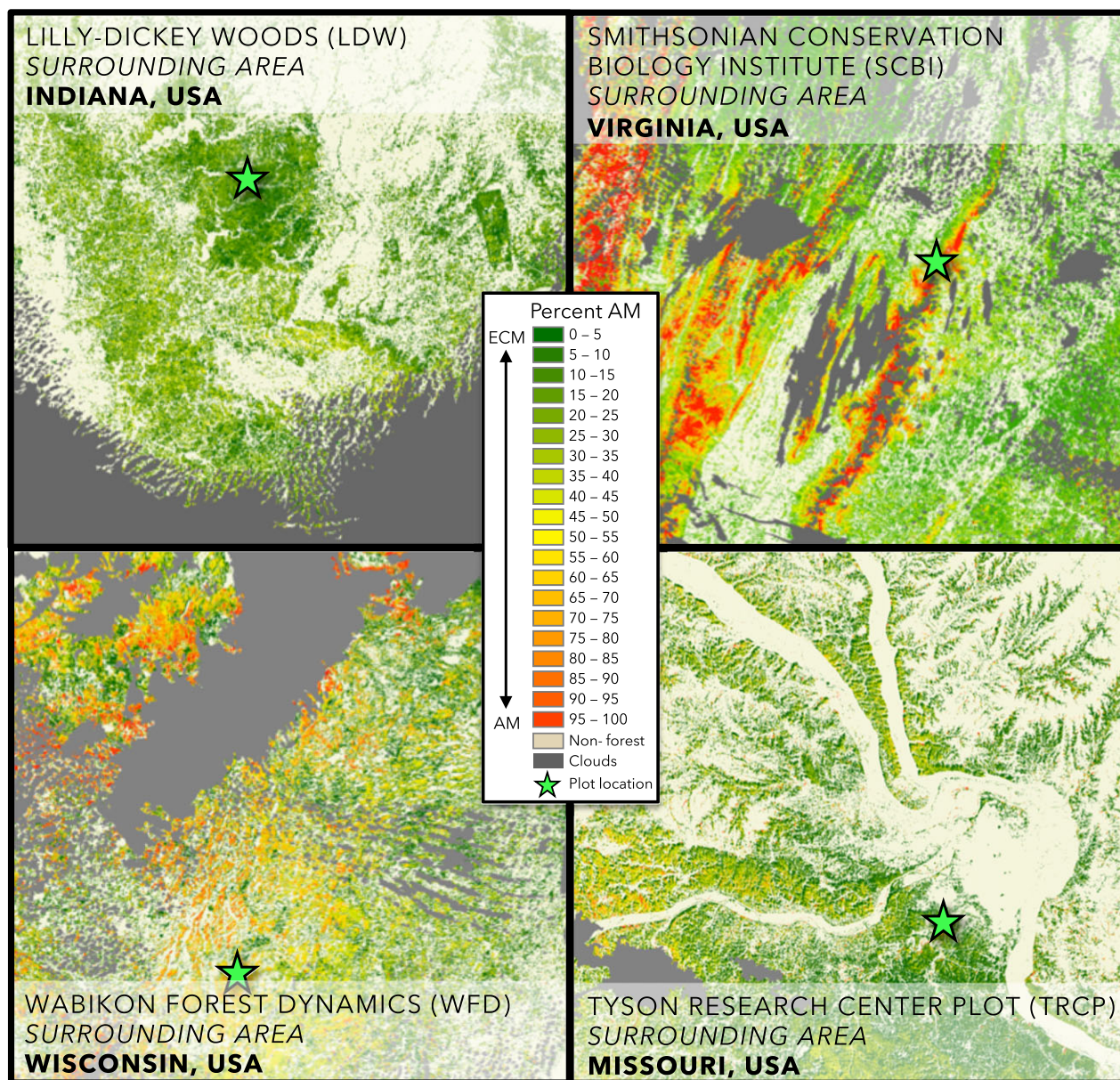
**Fig. 4** Predicted and observed mycorrhizal association forest composition across ~130 000 trees, aggregated and mapped to 60 m plots, at four megaplot sites: Lilly-Dickey Woods (LDW; Indiana), Smithsonian Conservation Biology Institute (SCBI; Virginia), Tyson Research Center Plot (TRCP; Missouri), and Wabikon Forest Dynamics (WFD; Wisconsin). Arbuscular mycorrhizal (AM)-associated trees (warm colors) comprise the majority of basal area in WFD and SCBI, whereas ectomycorrhizal (ECM)-associated trees (cool colors) are predominantly prevalent in TRCP and LDW.

1.5% on average for subplots mostly comprised of ECM-associated trees (<50% AM), and underestimate the percent AM in the few subplots that were mostly AM-dominated (>50% AM) by 19% on average. At the AM-dominated SCBI and WFD sites, the model tended to underestimate the percent AM by 5% on average for subplots mostly comprised by AM-associated trees (>50% AM), and overestimate the percent AM by 7% on average for subplots mostly comprised by ECM-associated trees (<50% AM).

We show the predicted mycorrhizal association regional coverage for the entire Landsat scenes over the Indiana, Virginia, Missouri, and Wisconsin sites overlaid on thematic land cover images in Fig. 5. The LDW and TRCP sites are surrounded primarily by ECM-weighted mixed forests, increasing in ECM dominance closer to the sites. Forests surrounding the SCBI and WFD are very heterogeneous, with multiple areas of strong AM or ECM dominance near SCBI; the areas closest to the sites are relatively mixed and AM-dominant. The entire Landsat scenes are not fully populated by mycorrhizal designation due to exclusion of non-forest pixels and clouds.

## Discussion

Temperate forests have experienced dramatic changes in species composition and spatial distribution over the past century due to land use change (Hurt *et al.*, 2006), timber harvesting (Schuler, 2004), invasive insects (Twery & Patterson, 1984), and altered disturbance regimes (Abrams, 1992). Predicting the biogeochemical consequences of these shifts is challenging, as tree species differ widely in their aboveground and belowground functional traits (Binkley & Menyailo, 2005). Hence, an important priority for understanding how shifts in forest composition influence ecosystem functioning is to develop a classification scheme that integrates the functional traits of tree species at the plot, stand, and landscape scales. However, given that species-specific approaches are not generalizable, an alternative approach is to classify forests based on the functional traits of the dominant trees, such as mycorrhizal association, an approach that strikes the balance between abstraction and detail to facilitate 'scalability' in considering the biogeochemical attributes of the system.



**Fig. 5** Full Landsat scene coverage of mycorrhizal composition across the larger regions surrounding each of our four study sites illustrates the diversity and heterogeneity of belowground dynamics. Site locations included Wabikon Forest Dynamics Plot (WFD) in Wisconsin, Tyson Research Center Plot (TRCP) in Missouri, Lilly-Dickey Woods (LDW) in Indiana, and Smithsonian Conservation Biology Institute (SCBI) in Virginia. Arbuscular mycorrhizal (AM)-associated trees (warm colors) comprise the majority of basal area in WFD and SCBI, whereas ectomycorrhizal (ECM)-associated trees (cool colors) are predominantly prevalent in TRCP and LDW.

Mapping forest mycorrhizal composition at large scales has substantial practical and science applications. New global terrestrial biosphere model developments of plant–nutrient dynamics are actively being developed (Thornton *et al.*, 2007; Wang *et al.*, 2007; Sokolov *et al.*, 2008; Xu-Ri & Prentice, 2008; Fisher *et al.*, 2010, 2014; Gerber *et al.*, 2010; Zaehle *et al.*, 2010; Shi *et al.*, 2016), as their impact on global climate trajectories is

large due to reductions in biospheric carbon uptake (Hungate *et al.*, 2003; Fisher *et al.*, 2012). The impact of mycorrhizae on nutrient cycling is significant (Cornelissen *et al.*, 2001; Read & Perez-Moreno, 2003; Chapman *et al.*, 2006; Phillips *et al.*, 2013; Averill *et al.*, 2014), although only recently have mycorrhizae been incorporated into such models (Brzostek *et al.*, 2014; Shi *et al.*, 2016). For these models to succeed, boundary

conditions of mycorrhizal distribution are required, yet none exist beyond broad biome groupings (Read, 1991; Allen *et al.*, 1995).

The Fixation & Uptake of Nitrogen (FUN) model (Fisher *et al.*, 2010), for instance, is a cutting-edge plant nitrogen uptake model based on carbon economics and resource optimization; recently, Brzostek *et al.* (2014) incorporated mycorrhizal-based nutrient uptake into FUN with different carbon costs assigned to AM vs. ECM associations. FUN is now coupled into larger terrestrial biosphere/land surface models, including the Community Land Model (CLM) (Shi *et al.*, 2016), the Noah Multi-Parameterization model (Noah-MP) (Cai *et al.*, 2016), and the Joint UK Land Environment Simulator (JULES) (Fisher *et al.*, 2013); CLM and JULES are coupled to larger Earth System Models used in assessments of future climate changes (Ipcc, 2007). Shi *et al.* (2016) showed that including carbon costs to nitrogen uptake in CLM reduced annual net primary production by 13% globally; 1238.5 Tg C yr<sup>-1</sup> was allocated to mycorrhizae, representing 52% of total global C used for nitrogen uptake. But Shi *et al.* (2016) acknowledge that their spatial distribution of AM vs. ECM associations by plant functional types was very coarse (Read, 1991; Allen *et al.*, 1995) and did not capture heterogeneity across landscapes. Thus, incorporation of a remote sensing/observational-based large-scale boundary condition map of mycorrhizal associations into these large-scale models will improve the accuracy and representation of mycorrhizal and nutrient feedbacks to carbon cycling and climate impacts. The results described here present the first step toward closing that gap.

We were able to predict 77% of the variation in mycorrhizal composition across ~130 000 trees with wide variation in mycorrhizal associations and species assemblages using equivalent predictor variables from moderate resolution canopy spectral signatures. This is a significant advancement toward gleaning biophysical understanding in how belowground processes influence canopy characteristics. Nonetheless, more work is necessary to pinpoint the underlying mechanisms controlling these properties. While there is evidence in the literature that AM and ECM forests vary in nutrient status (Phillips *et al.*, 2013; Midgley & Phillips, 2014; Brzostek *et al.*, 2015; Waring *et al.*, 2015; Rosling *et al.*, 2016), evidence on phenological differences is limited to few tree species (Key *et al.*, 2001; McCormack *et al.*, 2014), making it difficult to identify the mechanisms that underlie the phenological patterns across landscapes. Spring leaf out generally relates to a tree species' stem anatomy: diffuse porous trees often leaf out earlier than ring porous trees owing to their reduced susceptibility of early season cavitation (Polgar & Primack, 2011). Interestingly, most of the dominant AM

tree species at our sites (e.g., sugar maple, tulip poplar, red maple, black tupelo) are diffuse porous. In contrast, most of the dominant ECM trees at our sites (e.g., white oak, red oak, pignut hickory) are ring porous. However, there are some exceptions, particularly at the WFD site. Ash, which is one of the most dominant AM trees at WFD, is ring porous, and leafs out significantly later than the other AM trees. Similarly, basswood and birch, which are common ECM trees at WFD, are diffuse porous; these species leaf out earlier than the other ECM trees. Thus, while there is some evidence that AM and ECM trees differ in stem anatomy, more research is needed to determine the generality of this pattern.

Soil characteristics and leaf traits could also help elucidate canopy patterns, although such information may not be readily available from moderate resolution remotely sensed observations. Leaf size, morphology, longevity, color, nutrients, and structure are traits likely to be influenced by mycorrhizal association, with some being more easily detectable from moderate resolution remote sensing than others. We gathered available soil and leaf data at the LDW site to probe this question further (Table S2). Within LDW, there was little difference in the carbon and nitrogen content or structure of green leaves between AM- and ECM-associated trees. By contrast, the leaf litter of AM-associated trees had greater nitrogen content and lower C:N ratios than ECM-associated tree litter. This suggests that ECM-associated trees retranslocate more nitrogen from their litter than AM-associated trees (Cornelissen *et al.*, 2001; Brzostek *et al.*, 2014); this is a plausible mechanism to explain differences in the spectral signatures that we observed during leaf abscission. The differences in litter chemistry coupled with the primary pathways of soil decomposition (i.e., scavenging in AM soils vs. priming in ECM soils) has been hypothesized to promote the greater carbon storage per unit nitrogen in ECM soils than AM soils observed here (Phillips *et al.*, 2013; Averill *et al.*, 2014). However, given the marginal differences in peak foliar nitrogen, it appears that soil properties may only indirectly influence the spectral signatures we observed through their impact on retranslocation. We acknowledge that further investigation is needed to understand how differences in soil properties vs. retranslocation of nutrients contribute to the linkages between belowground processes and canopy traits. Regardless, this work sets up a hypothesis-testing framework to assess the predictive or explanatory power of mycorrhizal association vs. other plant and soil traits on canopy spectral characteristics.

The prediction of mycorrhizal composition may be improved by incorporating more information in our model. Major areas of statistical improvement should focus on the characteristics of over- and

underestimation described in the Results. The model consistently overestimated percent AM in plots with low percent AM, and underestimated percent AM in plots with high percent AM; factors underlying these patterns are likely to provide an important source of model improvements. Additional exploration of instruments with additional finer spectral resolution may be fruitful. New NASA missions being developed for the International Space Station will capture ecosystem functioning and structure at relatively high spatial resolutions and over the diurnal cycle (land surface temperature and evapotranspiration from ECOSTRESS; chlorophyll fluorescence from OCO-3; and 3-D structure and biomass from GEDI), potentially revealing even greater distinctions between AM- and ECM-associated trees (Dubayah, 2015; Eldering *et al.*, 2015; Fisher *et al.*, 2015).

A representative time series of cloud-free imagery is necessary for our approach to be successful, particularly in detection of nuances in phenology. An abundance of medium-resolution data is available, particularly Landsat, and is easily accessible, particularly historical data beyond the time frame analyzed here that can help refine our approach, including characteristics of interannual variability. There may be potential for such analyses by normalizing reflectance using pseudo-invariant features and applying the model coefficients to the normalized reflectance values of the historical data.

Understanding the factors that control species responses to global changes is a central challenge in the field of ecosystem science, and the use of functional categorizations – particularly ones that can be remotely sensed – can simplify this challenge. Characterizing forests based on the type of tree-fungal associations provides an integrative index of both plant and microbial traits, and should enable the mapping of biogeochemical syndromes across diverse landscapes. Differences in AM- and ECM-dominated forests have recently been used to make predictions about ecosystem services such as soil carbon storage (Averill *et al.*, 2014; Soudzilovskaia *et al.*, 2015; Waring *et al.*, 2015) and nitrogen retention (Phillips *et al.*, 2013; Midgley & Phillips, 2014). Moreover, there is increasing evidence that AM- and ECM-associated trees respond differently to global change drivers such as water availability (Querejeta *et al.*, 2009), N deposition (Thomas *et al.*, 2010; Midgley & Phillips, 2014), and elevated atmospheric CO<sub>2</sub> (Norby *et al.*, 2010; Drake *et al.*, 2011). The ability to detect tree mycorrhizal associations using remotely sensed data will enable researchers to generate maps of mycorrhizal distributions, ecosystem services, and putative biogeochemical syndromes that can be validated at regional scales (e.g., US Forest Service's Forest Inventory and Analysis Program) or globally (e.g., the Smithsonian

Institution's CTFs-ForestGEO; Anderson-Teixeira *et al.*, 2015). This detection ability with mapping capability may provide a valuable resource to track the impacts of the ongoing increase in the dominance of AM-associated trees, particularly in temperate forests due to management and global change (Nowacki & Abrams, 2008, 2015). At the global scale, this would facilitate the development of new plant functional types to improve the predictive power of terrestrial biosphere models, and enable forecasting the biogeochemical consequences of species gains and losses, as well as shifts in mycorrhizal association and forest composition as a result of global change.

### Acknowledgements

Tree data were collected by many investigators at the Smithsonian Institution's CTFs-ForestGEO, a global network of forest dynamics plots coordinated by Stuart Davies, Richard Condit, Sean McMahon, and many others. We thank Jeremy Degler, Mark Sheehan, Anthony Sipes, and Andrew Quebbeman for mapping the plot data for LDW; funding for LDW was provided by the Indiana Academy of Sciences and CTFs-ForestGEO. We thank William McShea for providing the plot data for SCBI. Major contributors to the WFD data include Kathryn Corio, Juniper Sundance, and Gary Fewless, with funding from the 1923 Fund and Smithsonian Institution. TRCP is supported by Washington University in St. Louis' Tyson Research Center, with funding provided by the International Center for Advanced Renewable Energy and Sustainability (I-CARES); more than 60 high school students, undergraduate students, and researchers contributed to the TRCP. Anonymous reviewers provided useful suggestions. Funding for the remote sensing analysis was provided by the US Department of Energy Office of Biological and Environmental Research Terrestrial Ecosystem Science Program; the US National Science Foundation Ecosystem Science Program; Indiana University's Research and Teaching Preserve; and Indiana University's Office of the Vice Provost for Research. JBF carried out the research at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration, and at the University of California at Los Angeles. Government sponsorship acknowledged.

### Author contributions

TPE and RPP formulated idea; TPE, SS, and JBF designed research; SS and JBF performed research; ATW, DJJ, JAM, NAB, and RWH collected field data; all authors contributed to the writing of the manuscript.

### Conflict of interest

The authors declare no conflict of interest.

### References

- Abrams MD (1992) Fire and the development of oak forests. *BioScience*, **42**, 346–353.

- Allen EB, Allen MF, Helm DJ, Trappe JM, Molina R, Rincon E (1995) Patterns and regulation of mycorrhizal plant and fungal diversity. *Plant and Soil*, **170**, 47–62.
- Anderson-Teixeira KJ, Davies SJ, Bennett AC *et al.* (2015) CTFs-ForestGEO: a worldwide network monitoring forests in an era of global change. *Global Change Biology*, **21**, 528–549.
- Averill C, Turner BL, Finzi AC (2014) Mycorrhiza-mediated competition between plants and decomposers drives soil carbon storage. *Nature*, **505**, 543–545.
- Binkley D, Menyailo O (2005) Gaining insights on the effects of tree species on soils. In: *Tree Species Effects on Soils: Implications for Global Change* (eds. Binkley D, Menyailo O), pp. 1–16. Springer, the Netherlands.
- Bourg NA, McShea WJ, Thompson JR, Mcgarvey JC, Shen X (2013) Initial census, woody seedling, seed rain, and stand structure data for the SCBI SIGEO Large Forest Dynamics Plot: ecological archives E094-195. *Ecology*, **94**, 2111–2112.
- Brandtberg T, Warner TA, Landenberger RE, McGraw JB (2003) Detection and analysis of individual leaf-off tree crowns in small footprint, high sampling density lidar data from the eastern deciduous forest in North America. *Remote Sensing of Environment*, **85**, 290–303.
- Brundrett M, Murase G, Kendrick B (1990) Comparative anatomy of roots and mycorrhizae of common Ontario trees. *Canadian Journal of Botany*, **68**, 551–578.
- Brzostek ER, Fisher JB, Phillips RP (2014) Modeling the carbon cost of plant nitrogen acquisition: mycorrhizal trade-offs and multi-path resistance uptake improve predictions of retranslocation. *Journal of Geophysical Research: Biogeosciences*, **119**, 1684–1697.
- Brzostek ER, Dragoni D, Brown ZA, Phillips RP (2015) Mycorrhizal type determines the magnitude and direction of root-induced changes in decomposition in a temperate forest. *New Phytologist*, **206**, 1274–1282.
- Cai X, Yang ZL, Fisher JB, Zhang X, Barlage M, Chen F (2016) Integration of nitrogen dynamics into the Noah-MP land surface model v1.1 for climate and environmental predictions. *Geoscientific Model Development*, **9**, 1–15.
- Chandelier L, Martinoty G (2009) Radiometric aerial triangulation for the equalization of digital aerial images and orthoimages. *Photogrammetric Engineering and Remote Sensing*, **75**, 193–200.
- Chapman SK, Langley JA, Hart SC, Koch GW (2006) Plants actively control nitrogen cycling: uncorking the microbial bottleneck. *New Phytologist*, **169**, 27–34.
- Chavez PS (1989) Radiometric calibration of Landsat Thematic Mapper multispectral images. *Photogrammetric Engineering and Remote Sensing*, **55**, 1285–1294.
- Cornelissen J, Aerts R, Cerabolini B, Werger M, Van Der Heijden M (2001) Carbon cycling traits of plant species are linked with mycorrhizal strategy. *Oecologia*, **129**, 611–619.
- Drake JE, Gallet-Budynek A, Hofmockel KS *et al.* (2011) Increases in the flux of carbon belowground stimulate nitrogen uptake and sustain the long-term enhancement of forest productivity under elevated CO<sub>2</sub>. *Ecology Letters*, **14**, 349–357.
- Dubayah R (2015) Crowd-sourced calibration: the GEDI strategy for empirical biomass estimation using spaceborne Lidar. *American Geophysical Union Fall Meeting*. American Geophysical Union, San Francisco.
- Eldering A, Basilio R, Bennett M (2015) The OCO-3 Mission: overview of science objectives and status. *American Geophysical Union Fall Meeting*. American Geophysical Union, San Francisco.
- Fisher JB, Sitch S, Malhi Y, Fisher RA, Huntingford C, Tan S-Y (2010) Carbon cost of plant nitrogen acquisition: a mechanistic, globally-applicable model of plant nitrogen uptake and fixation. *Global Biogeochemical Cycles*, **24**, 1–17.
- Fisher JB, Badgley G, Blyth E (2012) Global nutrient limitation in terrestrial vegetation. *Global Biogeochemical Cycles*, **26**, GB3007.
- Fisher JB, Clark D, Smith P (2013) Nitrogen cycling in the Hadley Centre land surface model (JULES). *CESM Land Model and Biogeochemistry Working Group Meeting*, National Center for Atmospheric Research (NCAR), Boulder, Colorado.
- Fisher JB, Huntzinger DN, Schwalm CR, Sitch S (2014) Modeling the terrestrial biosphere. *Annual Review of Environment and Resources*, **39**, 91–123.
- Fisher JB, Hook S, Allen R *et al.* (2015) ECOSTRESS: NASA's next-generation mission to measure evapotranspiration from the International Space Station. *American Geophysical Union Fall Meeting*. American Geophysical Union, San Francisco.
- Gerber S, Hedin LO, Oppenheimer M, Pacala SW, Shevliakova E (2010) Nitrogen cycling and feedbacks in a global dynamic land model. *Global Biogeochemical Cycles*, **24**, 1–15.
- Heinzel J, Koch B (2011) Exploring full-waveform LiDAR parameters for tree species classification. *International Journal of Applied Earth Observation and Geoinformation*, **13**, 152–160.
- Hobbie EA (2006) Carbon allocation to ectomycorrhizal fungi correlates with below-ground allocation in culture studies. *Ecology*, **87**, 563–569.
- Holmgren J, Persson Å (2004) Identifying species of individual trees using airborne laser scanner. *Remote Sensing of Environment*, **90**, 415–423.
- Hungate BA, Dukes JS, Shaw MR, Luo Y, Field CB (2003) Nitrogen and climate change. *Science*, **302**, 1512–1513.
- Hurt G, Frolking S, Fearon MG *et al.* (2006) The underpinnings of land-use history: three centuries of global gridded land-use transitions, wood-harvest activity, and resulting secondary lands. *Global Change Biology*, **12**, 1208–1229.
- IPCC (2007) Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change. In: *Climate Change 2007: The Physical Science Basis* (eds Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL), pp. 1–996. Cambridge University Press, Cambridge, United Kingdom and New York.
- Johnson DJ, Bourg NA, Howe R, McShea WJ, Wolf A, Clay K (2014) Conspecific negative density-dependent mortality and the structure of temperate forests. *Ecology*, **95**, 2493–2503.
- Key T, Warner TA, McGraw JB, Fajvan MA (2001) A comparison of multispectral and multitemporal information in high spatial resolution imagery for classification of individual tree species in a temperate hardwood forest. *Remote Sensing of Environment*, **75**, 100–112.
- Koele N, Dickie IA, Oleksyn J, Richardson SJ, Reich PB (2012) No globally consistent effect of ectomycorrhizal status on foliar traits. *New Phytologist*, **196**, 845–852.
- Markham B, Barker J (1987) Thematic Mapper bandpass solar exoatmospheric irradiances. *International Journal of Remote Sensing*, **8**, 517–523.
- Martin M, Newman S, Aber J, Congalton R (1998) Determining forest species composition using high spectral resolution remote sensing data. *Remote Sensing of Environment*, **65**, 249–254.
- McCormack ML, Gaines KP, Pastore M, Eissenstat DM (2014) Early season root production in relation to leaf production among six diverse temperate tree species. *Plant and Soil*, **389**, 121–129.
- Mickelson JG, Civco DL, Silander J (1998) Delineating forest canopy species in the northeastern United States using multi-temporal TM imagery. *Photogrammetric Engineering and Remote Sensing*, **64**, 891–904.
- Midgley MG, Phillips RP (2014) Mycorrhizal associations mediate nitrate leaching responses to N deposition: a meta-analysis. *Biogeochemistry*, **117**, 241–253.
- Moffett T, Mengersen K, Witte C, King R, Denham R (2005) Airborne laser scanning: exploratory data analysis indicates potential variables for classification of individual trees or forest stands according to species. *ISPRS Journal of Photogrammetry and Remote Sensing*, **59**, 289–309.
- Norby RJ, Warren JM, Iversen CM, Medlyn BE, Mcmurtrie RE (2010) CO<sub>2</sub> enhancement of forest productivity constrained by limited nitrogen availability. *Proceedings of the National Academy of Sciences*, **107**, 19368–19373.
- Nowacki GJ, Abrams MD (2008) The demise of fire and “mesophication” of forests in the eastern United States. *BioScience*, **58**, 123–138.
- Nowacki GJ, Abrams MD (2015) Is climate an important driver of post-European vegetation change in the Eastern United States? *Global Change Biology*, **21**, 314–334.
- Ohmann JL, Gregory MJ, Henderson EB, Roberts HM (2011) Mapping gradients of community composition with nearest-neighbour imputation: extending plot data for landscape analysis. *Journal of Vegetation Science*, **22**, 660–676.
- Phillips RP, Brzostek E, Midgley MG (2013) The mycorrhizal-associated nutrient economy: a new framework for predicting carbon–nutrient couplings in temperate forests. *New Phytologist*, **199**, 41–51.
- Plourde LC, Ollinger SV, Smith M-L, Martin ME (2007) Estimating species abundance in a northern temperate forest using spectral mixture analysis. *Photogrammetric Engineering & Remote Sensing*, **73**, 829–840.
- Polgar CA, Primack RB (2011) Leaf-out phenology of temperate woody plants: from trees to ecosystems. *New Phytologist*, **191**, 926–941.
- Querejeta JL, Egerton-Warburton LM, Allen MF (2009) Topographic position modulates the mycorrhizal response of oak trees to interannual rainfall variability. *Ecology*, **90**, 649–662.
- Read DJ (1991) Mycorrhizas in ecosystems. *Experientia*, **47**, 376–391.
- Read DJ, Perez-Moreno J (2003) Mycorrhizas and nutrient cycling in ecosystems – a journey towards relevance? *New Phytologist*, **157**, 475–492.
- Rosling A, Midgley MG, Cheeke T, Urbina H, Fransson P, Phillips RP (2016) Phosphorus cycling in deciduous forest soil differs between stands dominated by ecto- and arbuscular mycorrhizal trees. *New Phytologist*, **209**, 1184–1195.
- Saatchi S, Buermann W, Ter Steege H, Mori S, Smith TB (2008) Modeling distribution of Amazonian tree species and diversity using remote sensing measurements. *Remote Sensing of Environment*, **112**, 2000–2017.
- Schriever JR, Congalton RG (1995) Evaluating seasonal variability as an aid to cover-type mapping from Landsat Thematic Mapper data in the Northeast. *Photogrammetric Engineering and Remote Sensing*, **61**, 321–327.

- Schuler TM (2004) Fifty years of partial harvesting in a mixed mesophytic forest: composition and productivity. *Canadian Journal of Forest Research*, **34**, 985–997.
- Shi M, Fisher JB, Brzostek ER, Phillips RP (2016) Carbon cost of plant nitrogen acquisition: global carbon cycle impact from an improved plant nitrogen cycle in the Community Land Model. *Global Change Biology*, **22**, 1299–1314.
- Sokolov AP, Kicklighter DW, Melillo JM, Felzer BS, Schlosser CA, Cronin TW (2008) Consequences of considering carbon–nitrogen interactions on the feedbacks between climate and the terrestrial carbon cycle. *Journal of Climate*, **21**, 3776–3796.
- Soudzilovskaia NA, Heijden MG, Cornelissen JH *et al.* (2015) Quantitative assessment of the differential impacts of arbuscular and ectomycorrhiza on soil carbon cycling. *New Phytologist*, **208**, 280–293.
- Spasojevic MJ, Yablon EA, Oberle B, Myers JA (2014) Ontogenetic trait variation influences tree community assembly across environmental gradients. *Ecosphere*, **5**, 1–20 art129.
- Thomas RD, Canham CD, Weathers KC, Goodale CL (2010) Increased tree carbon storage in response to nitrogen deposition in the US. *Nature Geoscience*, **3**, 13–17.
- Thornton PE, Lamarque J-F, Rosenbloom NA, Mahowald NM (2007) Influence of carbon–nitrogen cycle coupling on land model response to CO<sub>2</sub> fertilization and climate variability. *Global Biogeochemical Cycles*, **21**, 1–15.
- Tuominen S, Pekkarinen A (2004) Local radiometric correction of digital aerial photographs for multi source forest inventory. *Remote Sensing of Environment*, **89**, 72–82.
- Turner W, Spector S, Gardiner N, Fladeland M, Sterling E, Steininger M (2003) Remote sensing for biodiversity science and conservation. *Trends in Ecology and Evolution*, **18**, 306–314.
- Twery MJ, Patterson WA III (1984) Variations in beech bark disease and its effects on species composition and structure of northern hardwood stands in central New England. *Canadian Journal of Forest Research*, **14**, 565–574.
- Underwood E, Ustin S, Dipietro D (2003) Mapping nonnative plants using hyperspectral imagery. *Remote Sensing of Environment*, **86**, 150–161.
- Vesterdal L, Clarke N, Sigurdsson BD, Gundersen P (2013) Do tree species influence soil carbon stocks in temperate and boreal forests? *Forest Ecology and Management*, **309**, 4–18.
- Walsh SJ (1980) Coniferous tree species mapping using Landsat data. *Remote Sensing of Environment*, **9**, 11–26.
- Wang B, Qiu YL (2006) Phylogenetic distribution and evolution of mycorrhizas in land plants. *Mycorrhiza*, **16**, 299–363.
- Wang YP, Houlton B, Field CB (2007) A model of biogeochemical cycles of carbon, nitrogen and phosphorus including symbiotic nitrogen fixation and phosphatase production. *Global Biogeochemical Cycles*, **21**, 1–15.
- Waring BG, Adams R, Branco S, Powers JS (2015) Scale-dependent variation in nitrogen cycling and soil fungal communities along gradients of forest composition and age in regenerating tropical dry forests. *New Phytologist*, **209**, 845–854.
- Wilson BT, Lister AJ, Riemann RI (2012) A nearest-neighbor imputation approach to mapping tree species over large areas using forest inventory plots and moderate resolution raster data. *Forest Ecology and Management*, **271**, 182–198.
- Wittmann F, Anhu D, Funk WJ (2002) Tree species distribution and community structure of central Amazonian várzea forests by remote-sensing techniques. *Journal of Tropical Ecology*, **18**, 805–820.
- Wolter PT, Mladenoff DJ, Host GE, Crow TR (1995) Improved forest classification in the Northern Lake States using multi-temporal Landsat imagery. *Photogrammetric Engineering and Remote Sensing*, **61**, 1129–1144.
- Xu-Ri X, Prentice IC (2008) Terrestrial nitrogen cycle simulation with a dynamic global vegetation model. *Global Change Biology*, **14**, 1745–1764.
- Zaehle S, Friend AD, Friedlingstein P, Dentener F, Peylin P, Schulz M (2010) Carbon and nitrogen cycle dynamics in the O-CN land surface model, II: the role of the nitrogen cycle in the historical terrestrial C balance. *Global Biogeochemical Cycles*, **24**, 1–14.

## Supporting Information

Additional Supporting Information may be found in the online version of this article:

**Table S1.** Tree species composition, abundance (%), and mycorrhizal associations (arbuscular mycorrhizal, AM; ectomycorrhizal, ECM) at each site (Lilly-Dickey Woods, LDW; Smithsonian Conservation Biology Institute, SCBI; Tyson Research Center Plot, TRCP; Wabikon Forest Dynamics, WFD). Only species with greater than 0.01% basal area are included in this table. A dash indicates the species was not present at that site.

**Table S2.** Properties of leaves ( $n = 110$ ) and soils in AM and ECM dominated plots ( $n = 15$ ) at the LDW site.

**Figure S1.** A sensitivity analysis whereby the predictive model was trained on half the data then validated against the other half of the data showed robust predictive power relative to the original model (2% degradation).