New Phytologist

Research

From the Arctic to the tropics: multibiome prediction of leaf mass per area using leaf reflectance

Shawn P. Serbin[1], Jin Wu[1,2], Kim S. Ely[1], Eric L. Kruger[3], Philip A. Townsend[3], Ran Meng[1,4], Brett T. Wolfe[5,6], Adam Chlus[3], Zhihui Wang[3] and Alistair Rogers[1]

[1]Environmental and Climate Sciences Department, Brookhaven National Laboratory, Upton, NY 11973, USA; [2]School of Biological Sciences, The University of Hong Kong, Pokfulam, Hong Kong; [3]Department of Forest and Wildlife Ecology, University of Wisconsin-Madison, Madison, WI 53706, USA; [4]College of Natural Resources and Environment, Huazhong Agricultural University, Wuhan, Hubei, China; [5]Smithsonian Tropical Research Institute, Apartado 0843-03092, Balboa, Panama; [6]School of Renewable Natural Resources, Louisiana State University, Baton Rouge, LA, USA

Author for correspondence:
Shawn Paul Serbin
Tel: +1 631 344 3165
Email: sserbin@bnl.gov

Received: *20 March 2019*
Accepted: *28 July 2019*

New Phytologist (2019)
doi: 10.1111/nph.16123

Key words: leaf mass area, partial least-squares regression (PLSR), plant traits, remote sensing, specific leaf area, spectroscopy.

Summary

- Leaf mass per area (LMA) is a key plant trait, reflecting tradeoffs between leaf photosynthetic function, longevity, and structural investment. Capturing spatial and temporal variability in LMA has been a long-standing goal of ecological research and is an essential component for advancing Earth system models. Despite the substantial variation in LMA within and across Earth's biomes, an efficient, globally generalizable approach to predict LMA is still lacking.
- We explored the capacity to predict LMA from leaf spectra across much of the global LMA trait space, with values ranging from 17 to 393 g m^{-2}. Our dataset contained leaves from a wide range of biomes from the high Arctic to the tropics, included broad- and needleleaf species, and upper- and lower-canopy (i.e. sun and shade) growth environments.
- Here we demonstrate the capacity to rapidly estimate LMA using only spectral measurements across a wide range of species, leaf age and canopy position from diverse biomes. Our model captures LMA variability with high accuracy and low error ($R^2 = 0.89$; root mean square error (RMSE) = 15.45 g m^{-2}).
- Our finding highlights the fact that the leaf economics spectrum is mirrored by the leaf optical spectrum, paving the way for this technology to predict the diversity of LMA in ecosystems across global biomes.

Introduction

A key objective of plant ecology is to characterize the functional diversity of plants that have evolved different strategies for growth, reproduction and for coping with biotic and abiotic stressors (Wright *et al.*, 2004). Accurate characterization of this functional diversity in Earth system models (ESMs) will improve our ability to model the Earth system and understand the effect of global change on the cycling and storage of carbon (C), water and energy (Pavlick *et al.*, 2013; Fisher *et al.*, 2018). Therefore, an increasing number of ESMs are moving towards incorporating approaches that require a broader and more comprehensive representation of plant trait variation within and across biomes (Xu *et al.*, 2012; van Bodegom *et al.*, 2014; Wullschleger *et al.*, 2014; Fisher *et al.*, 2015, 2018). As a result, considerable effort has been invested in the measurement and monitoring of plant traits across a range of biomes, and the storage and synthesis of that information in global databases (Wright *et al.*, 2004; Kattge *et al.*, 2011; Lebauer *et al.*, 2013; Butler *et al.*, 2017). Yet, the high degree of plant functional diversity and plasticity makes this apparently simple goal extremely challenging (Reich *et al.*, 1997, 1999; Serbin *et al.*, 2014; Wu *et al.*, 2017; Osnas *et al.*, 2018). Consequently, the extent of global trait coverage is still woefully inadequate (Schimel *et al.*, 2015).

In recent decades, remote sensing has shown increasing promise as a means to capture plant traits across scales using spectroscopic approaches. For example, several recent studies have highlighted the capacity to connect remotely sensed spectra to characterize variation in a number of key functional traits across individual leaves, canopies and landscapes (Dahlin *et al.*, 2013; Asner *et al.*, 2015; Singh *et al.*, 2015; Shiklomanov *et al.*, 2016; Yang *et al.*, 2016; Wu *et al.*, 2017). Approaches have included empirical spectra–trait modeling, such as partial least-squares regression (PLSR; e.g. Serbin *et al.*, 2014; Singh *et al.*, 2015; Yendrek *et al.*, 2017) and spectral vegetation indices (SVIs; e.g. Feret *et al.*, 2011), as well as semimechanistic approaches, including the use of the PROSPECT leaf-level radiative transfer model (RTM; e.g. Shiklomanov *et al.*, 2016; Féret *et al.*, 2017). In general, all of these approaches rely on the fundamental biophysical connection between leaf chemistry and structure and the resultant optical properties of plants (Curran, 1989; Ustin *et al.*, 2004; Kokaly *et al.*, 2009; Ollinger, 2011), and, as a result, could be

No claim to US Government works
New Phytologist © 2019 New Phytologist Trust

used to fill critical gaps in our understanding of the variation in plant traits over landscapes and biomes (Schimel *et al.*, 2015; Jetz *et al.*, 2016).

However, most studies that have illustrated the strong promise of remote sensing to estimate foliar traits across scales have focused on a relatively narrow subsample of geographic regions (e.g. Townsend *et al.*, 2003; Feret *et al.*, 2011; Dahlin *et al.*, 2013; Asner *et al.*, 2015; Yang *et al.*, 2016), capturing only a small fraction of total trait and spectral space. In addition, semimechanistic methods such as PROSPECT may have other assumptions or limitations that could inhibit their broad application, for example, a limited number of supported traits (Féret *et al.*, 2017) or the generally poorer performance of needleleaf species (Malenovský *et al.*, 2006; Shiklomanov *et al.*, 2016). As a result, the general applicability of existing spectra–trait models and approaches across biomes and in the wider trait space is not known. In many cases, those models trained across a limited trait space have been shown to break down when applied more broadly (e.g. Sims & Gamon, 2002; Gitelson *et al.*, 2003; le Maire *et al.*, 2008).

Among plant traits, foliar morphology is commonly characterized using leaf mass per area (LMA) – the ratio of a leaf's dry mass to its surface area (g dry mass m^{-2} leaf area) – or its reciprocal, specific leaf area. LMA captures the tradeoff of a plant's investment in leaf structure and robustness vs leaf surface area and light harvesting for photosynthesis (Wright *et al.*, 2004; Shipley *et al.*, 2006; Poorter *et al.*, 2009). Given its strong linkage with overall plant functioning (Reich *et al.*, 1997; Wright *et al.*, 2004; Serbin *et al.*, 2012; Osnas *et al.*, 2018), LMA is a critical parameter in plant ecology. Illustrations of the importance of LMA include its use as a basis for monitoring biodiversity (Skidmore *et al.*, 2015), its role in modeling canopy radiation transfer (Jacquemoud *et al.*, 2009; Ollinger, 2011), and its widespread use as an input in ecosystem process models (Fisher *et al.*, 2014; Xu *et al.*, 2016; Ricciuto *et al.*, 2018).

Importantly, the substantial global variation in LMA, which ranges from 14 to 1500 g m^{-2} globally (Wright *et al.*, 2004), exists within and across species (Castro-Díez *et al.*, 2000; Wright *et al.*, 2004; Paula & Pausas, 2006; Poorter *et al.*, 2009) and is affected by local gradients in light, water and nutrient availability (Niinemets, 2007; de la Riva *et al.*, 2016; Liu *et al.*, 2017), leaf age (Wu *et al.*, 2016), as well as acclimation and adaptation to short- and long-term climate dynamics (Volin *et al.*, 2002; Paula & Pausas, 2006; Poorter *et al.*, 2009). LMA variation is attributed primarily to differences in leaf density and volume-to-area ratio (Poorter *et al.*, 2009; John *et al.*, 2017). The fundamental information on these attributes is found in the reflectance spectrum of a leaf, which captures its physical properties (e.g. thickness, density, the depth of palisade layers, albedo and elemental composition) (Ollinger, 2011). Given the coordination between leaf traits and optical properties, we expect that, using a spectroscopy approach, we can collapse the vast structural and functional diversity of leaves from different plant species, leaf types (e.g. grasses, forbs, broadleaf and needleleaf evergreen trees), across a wide range of environments, that is expressed as variation of LMA into a single generalizable model.

Here we present a broad, multibiome analysis linking fundamental covariation in LMA and spectroscopic (also known as hyperspectral) reflectance. We use a large dataset to develop a robust statistical model to infer LMA from corresponding leaf optical properties, and then validate this cross-biome model using independent datasets of additional spectra and LMA observations from a similar range of plant material as well as from external validation sources. The core training and validation datasets include leaves from the high Arctic to the tropics, and contain measurements from grasses, forbs, deciduous and evergreen shrubs, deciduous and evergreen broadleaf trees, needleleaf trees and crop species. They span a highly diverse range of leaf morphologies, including glabrous, highly reflective and waxy leaf types; microenvironments, including measurements from upper-canopy, sunlit leaves, and lower-canopy, shaded leaves; developmental stages, including recently emerged, mature and old leaves; and elevations, including measurements from sea level to > 2000 m above sea level. Thus, our analysis is based on data representing a large fraction of the global trait space for LMA.

Materials and Methods

Plant material

We collected and assembled a large dataset ($n = 2478$ leaves from > 176 species) of combined LMA and leaf reflectance spectra for model development from the high Arctic in northern Alaska to the tropics in central America and Brazil (Fig. 1). Our sites are distributed across a large proportion of the Earth's habitable climate space, encompassing a *c.* 40°C range of mean annual temperatures and a *c.* 200–2400 mm yr^{-1} range in mean annual precipitation. Our sites in Alaska include coastal tundra vegetation within the Barrow Environmental Observatory, near Barrow (now Utqiaġvik), Alaska (Brown *et al.*, 1980; Rogers *et al.*, 2017b) and dwarf and tall shrub vegetation on the Seward Peninsula (Rogers *et al.*, 2016; Serbin & Rogers, 2019). Our sites located in the Upper Midwest and northeastern US are dominated by northern temperate forest species, including deciduous broadleaf hardwoods and evergreen, needleleaf conifers (Serbin *et al.*, 2014). Measurements on Mediterranean and agricultural plants were conducted in the Coachella and Central valleys of California (Serbin *et al.*, 2015), and across an elevation gradient of sites ranging from low elevation woodlands to alpine forests in the Sierra Mountains (Goulden *et al.*, 2012; DuBois *et al.*, 2018). The data for tropical species were collected in several separate locations; a seasonal, wet evergreen Amazonian forest near Santarem, Brazil (Wu *et al.*, 2017); a seasonal, wet evergreen forest in the San Lorenzo Protected Area; a seasonal, dry forest in the Parque Natural Metropolitano near Panama City in the Republic of Panama (Wright *et al.*, 2003); trees within the Dona Inés Park urban forest site on the grounds of the Luis Muñoz Marín Foundation as well as other locations on the island of Puerto Rico (specific sites can be found in associated dataset metdata, Table S1), and a collection of plants grown in an artificial tropical forest within the Biosphere 2 facility (Walter & Carmen Lambrecht, 2004). All the data used in this study are

No claim to US Government works
New Phytologist © 2019 New Phytologist Trust

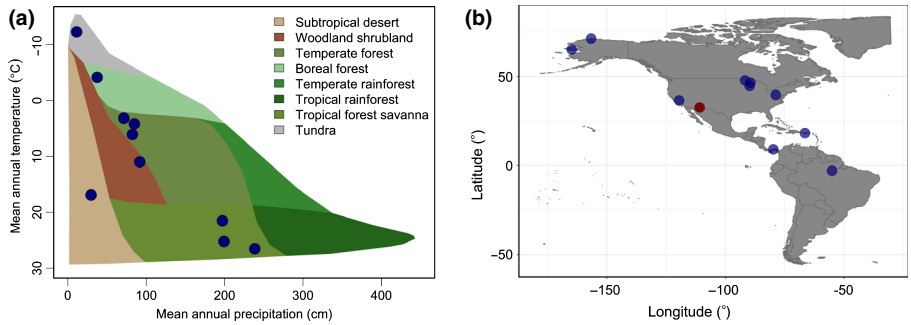

Fig. 1 General location of the 11 sites where leaf reflectance and leaf mass per area were measured in our model development dataset. Blue circles show field sites, and the red circle shows the location of Biosphere 2 from which we sampled tropical species within the glasshouse environment. (a) Sites plotted in climate space, binned by major climatic biomes (Whittaker, 1975); (b) general location in geographic space.

publicly available, and the published datasets include detailed descriptions of the sampling locations (Table S1).

Spectroscopic measurements

The approach, instrumentation and personnel used to measure leaf spectral reflectance varied across the different sites and projects included in this study. The leaf reflectance of Arctic plants was measured using a FieldSpec 3 (Analytical Spectral Devices (ASD) Inc., Longmont, CO, USA) with the ASD leaf clip assembly, or an HR-1024i full-spectrum spectroradiometer (Spectra Vista Corporation (SVC), Poughkeepsie, NY, USA) together with the standard leaf clip assembly as well as the more recent Leaf-Clip Reflectance-Probe Pro (LC-RP-Pro; SVC). Leaf spectra collected in the Upper Midwest and northeastern USA were measured with a FieldSpec pro or FieldSpec 3 (both ASD Inc.) with the ASD leaf clip assembly. The data collected in California were measured using either a FieldSpec 3 (ASD Inc.) or a PSR-3500 full-range spectrometer (Spectral Evolution, Lawrence, MA, USA) with an attached Spectral Evolution leaf clip assembly. Data from Biosphere2 and Brazil were measured with a FieldSpec Pro (ASD Inc.) and ASD leaf clip, the data from Puerto Rico with a PSR+ (Spectral Evolution) with an attached custom fiber optic connected to a LC-RP-Pro foreoptic (SVC), and data from the Republic of Panama were measured with an HR-1024i with the LC-RP-Pro attachment (SVC). All instruments had leaf clip assemblies that contained an internal calibrated light source, and all reflectance measurements were referenced against a 99% Spectralon reflectance standard. For needleleaf or very small leaf species, we used needle or leaf mats (Serbin, 2012). Needles of a similar age (i.e. current year, previous year, and older) were laid out edge to edge, creating a single layer and taped at the ends to hold the needles tightly together before inserting the mat into the leaf clip. No tape was visible to the fiberoptic inside the leaf clip. All spectral measurements were processed using the R-FieldSpectra package (https://github.com/serbinsh/R-FieldSpectra). For SVC data, however, we first corrected the discontinuities in the spectra in the detector overlap areas using the vendor-provided software. Measurements at the edges of the spectral range of these spectrometers (350–500 nm, > 2400 nm) suffer from low signal-to-noise and were thus excluded from the analysis.

Measurements of LMA

Measurements of LMA also varied among the different studies and biomes. Various methods were used to determine leaf area, including measurement of length and width with a ruler and a hand lens (graminoid species), leaf disk punches (broadleaf species) and optical approaches. For small or compound leaves where it was not possible to take disks, area was measured with a leaf area meter (LI-3100C Area Meter; Li-Cor, Lincoln, NE, USA) calibrated for use with high edge-to-area ratio leaves and operated at high (0.1 mm^2) resolution, or a flatbed scanner (evergreen needleleaf trees) followed by area estimation using ImageJ (Schneider *et al.*, 2012; as reported in Serbin *et al.*, 2014). Leaves and leaf sections of known area were then dried to constant mass in a ventilated oven (60–70°C) and leaf dry mass was measured on a top-pan balance. All leaf area data were provided on a projected leaf area basis.

LMA partial least-squares regression modeling

To relate the variability in LMA across sites, species, and environments we utilized a PLSR modeling approach (Geladi & Kowalski, 1986; Wold *et al.*, 2001) using the pls package (Mevik & Wehrens, 2007) in the R open source statistical environment (R Core Team, 2017). PLSR is widely utilized in spectroscopy and chemometric analyses given its ability to handle high predictor collinearity as well as a large number of predictor variables that may exceed the number of observations. PLSR minimizes the implications of these circumstances by reducing the number of predictor variables down to a relatively few, orthogonal latent components (Geladi & Kowalski, 1986; Wold *et al.*, 2001). Moreover, PLSR does not assume the measurement of predictor variables (reflectance values at given wavelengths in this case) was made without error.

Our PLSR model development has been described previously (e.g. Serbin *et al.*, 2014; Ely *et al.*, 2019) and is briefly summarized here. We first applied a square root transformation to the LMA data to reduce the right skewness distribution of the original data (Fig. S1) and satisfy the required normal distribution for the PLSR analysis. We then randomly split the full dataset into calibration (80%, $n = 1978$) and independent validation (20%,

No claim to US Government works
New Phytologist © 2019 New Phytologist Trust

$n = 500$) subsets (Fig. S2), ensuring that both subsets spanned the full range of observations and included measurements from each study and sample location. Using just the calibration dataset, we developed the multibiome leaf spectra–LMA model and then tested this final model using the validation data (500 observations) derived from our original dataset (2478 leaves). To avoid the potential to overfit the spectra–LMA calibration model, we optimized the number of PLSR components in the model by minimizing the prediction residual sum of squares (PRESS) statistic (Chen *et al.*, 2004). We calculated the PRESS statistic of successive model components through a cross-validation analysis where we minimized the PRESS statistic until successive PLSR components did not reduce model predictive error as assessed using a *t*-test (Serbin *et al.*, 2014). Lastly, we calculated the variable influence on projection (VIP) metric (Wold *et al.*, 2001) of the final model to identify the regions of the spectrum that contributed significantly to the prediction of LMA (Serbin *et al.*, 2014).

In addition to the general model development, we also conducted a PLSR model uncertainty analysis to characterize the predictive uncertainty, given the variability and error in measured LMA, spectra, and the relationship between the two. This was done by splitting the original calibration dataset into a balanced 70–30% via 1000 permutations and generating the same number of new model coefficient vectors, following Serbin *et al.* (2014). The result of this uncertainty analysis is an ensemble of PLSR models that can be used to predict new values of LMA based on spectral measurements plus the predictive uncertainty for each new value.

Finally, we quantified the performance of the multibiome LMA PLSR model using the independent validation dataset (500 observations). In this step, we validated the model and examined the residuals for model bias and predictive performance. We used four main evaluation metrics: the coefficient of determination (R^2), RMSE, the residual, and regression bias. All model and error results presented in the following are shown in original LMA units, not the square-root-transformed units that are the initial output of the PLSR model.

Examples of applying the PLSR model to new observations

We provide a simple R script (Methods S1) to illustrate the utility of our model and how it can be used to estimate LMA values from leaf spectral reflectance observations not used for model development. This script can be run to automatically download the foundational Leaf Optical Properties Experiment (LOPEX) and ANGERS spectral datasets (Hosgood *et al.*, 1994; Jacquemoud *et al.*, 2003), apply the PLSR model, and provide the results. The LMA model coefficients are provided through GitHub (https://github.com/serbinsh/SSerbin_etal_2019_NewPhytologist) and the leaf spectral data are provided from the Ecological Spectral Information System (EcoSIS) database (https://ecosis.org).

Furthermore, we also provide an extensive, external validation of our multibiome PLSR model with data collected in the Upper Midwest, US, and nine National Ecological Observatory Network (NEON) locations spanning seven NEON domains. The first dataset consisted of fully expanded, peak-growing-season samples collected in and around the Madison, Wisconsin, area and the University of Notre Dame Environmental Research Center (UNDERC) between June and September 2017 (A. Chlus *et al.*, unpublished). A mix of sunlit and shaded foliar samples were collected from broadleaf trees ($n = 7446$), graminoids ($n = 74$), forbs ($n = 2017$) and vine ($n = 218$) species across the growing season. All samples were immediately scanned using a FieldSpec 3 (ASD Inc.), HR-1024i (SVC) and/or PSR+ (Spectral Evolution) with their respective leaf contact probes and external light sources. A Spectralon white reference was scanned before each sample to calculate the relative reflectance. After spectral measurement, the leaf area of each sample was immediately recorded using the LI-3100 benchtop leaf scanner (Li-Cor). Samples were subsequently flash-frozen in liquid nitrogen, freeze-dried in a VirTis lyophilizer (SP Scientific, Gardiner, NY, USA), and weighed with a precision balance.

The second dataset consisted of peak-growing-season foliar samples collected during the summer of 2017 at NEON sites in Wisconsin, Alabama, Georgia, Florida, Virginia, Maryland, Tennessee, Kansas and North Dakota, and also included trees ($n = 2584$), graminoids ($n = 381$) and forb ($n = 195$) species (Z. Wang *et al.*, unpublished). Spectra were collected immediately using the FieldSpec and/or PSR+ with their leaf contact probes. Leaves were scanned on a 600 dpi flatbed scanner (Epson, Nagano, Japan), oven-dried at 65°C for 48 h to a constant mass and weighed on a precision balance. The only processing applied to the spectra was the removal of spectral discontinuities (following Serbin *et al.*, 2012) at *c.* 1000 and 1900 nm where 'jumps' sometimes occur at overlapping wavelengths between detectors within the instruments. The final multibiome PLSR model was applied to all spectra and compared with laboratory measurements of LMA.

Data availability

All data used in this manuscript are publicly available through online data portals, including the U.S. Department of Energy NGEE-Arctic and NGEE-Tropics data portals as well as the EcoSIS spectral database (Table S1).

Results and Discussion

We employed an extensive dataset of leaf spectral reflectance and LMA across multiple biomes and spanning large environmental gradients (Fig. 1) to develop a generalized approach to model the variation in LMA (Fig. 2) using only leaf optical properties (Fig. 3). We found a very strong capacity for the empirical PLSR spectra–trait modeling approach to accurately model multibiome variation in LMA using spectral reflectance (Fig. 4). Our results show that spectra alone can explain 89% of the variation in LMA with a low model bias (0.96 $g\,m^{-2}$) and RMSE (15.45 $g\,m^{-2}$) when compared with our core validation dataset. We provide more details on our input datasets, PLSR modeling, and validation in the following, as well as a discussion of our work in the larger functional trait and remote-sensing research communities.

No claim to US Government works
New Phytologist © 2019 New Phytologist Trust

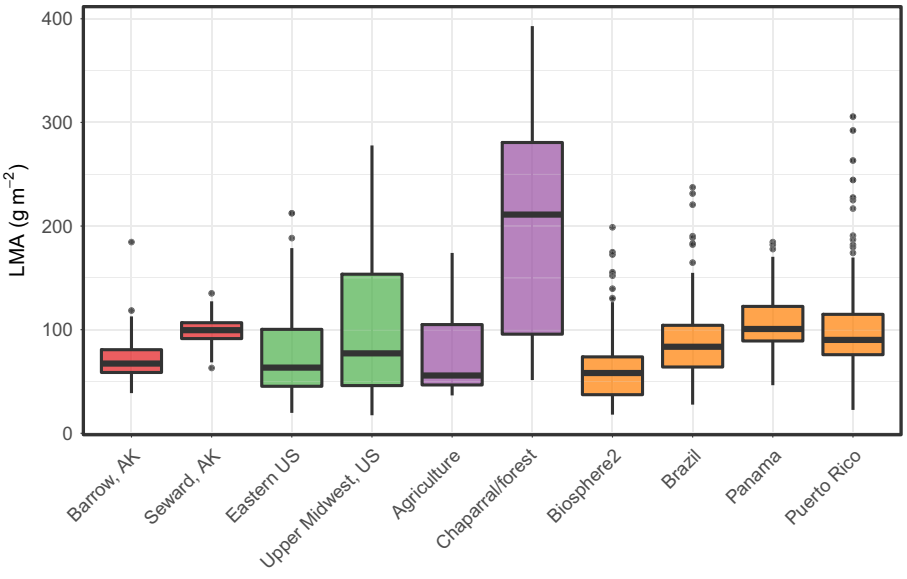

Fig. 2 Tukey box plots showing the range of leaf mass per area (LMA, g m^{-2}) in our model development dataset calculated from measurements of leaf area and dry mass for our 11 regions, including the Biosphere2 glasshouse location. Sites are color-coded by broad biome class (red, Arctic; green, boreal/temperate; purple, Mediterranean; orange, tropical). These are binned into four main biomes for clarity but correspond to the six Whittaker classes shown in Fig. 1. Box plots show the interquartile range (box), and median (solid horizontal line). The whiskers show lowest and highest datum still within 1.5× interquartile range of the lower and upper quartiles. Outliers are shown as black dots. Sample sizes by region: 609 Arctic leaves, 935 boreal/temperate leaves, 102 Mediterranean leaves (including 33 agricultural samples), and 832 tropical leaves (including 72 Biosphere2 samples)

Biotic and abiotic variation in LMA

Across our sample sites (Fig. 1), we observed a broad range of LMA in our model development dataset, with values ranging from 17 to 393 g m^{-2} (Figs 2, S1). As expected, mean (± SD) top-of-canopy LMA of the needleleaf evergreen conifer species (165 ± 62 g m^{-2}) was significantly larger than those of broadleaf deciduous species, as well as grasses and forbs (67 ± 24 g m^{-2}). The LMAs of evergreen broadleaf species from arid environments (169 ± 72 g m^{-2}) were similar to those of needleleaf evergreen species, consistent with past observations (Paula & Pausas, 2006; de la Riva *et al.*, 2016), and owing to the well-documented differences in leaf life spans that typically require higher resource investment in leaf construction (Wright *et al.*, 2004; Poorter *et al.*, 2009). For the tropical species, LMA averaged 106 ± 37 g m^{-2} and ranged from 22 to 306 g m^{-2}.

Within a biome, LMA variation (Fig. S3) was large and related to biotic differences across species as well as factors such as leaf and plant age, and abiotic factors such as the leaf growth environment, including position in the canopy (Serbin *et al.*, 2014; Wu *et al.*, 2017), consistent with other studies (Niinemets *et al.*, 2015; John *et al.*, 2017; Osnas *et al.*, 2018). Notably, variation within species was often as large as that across species (data not shown), as has been noted in previous work (Butler *et al.*, 2017). On average, LMA in upper-canopy, fully sunlit leaves were 36 ± 20% higher than in fully shaded leaves of the same species, but this difference varied with growth form and leaf habit, probably owing to known differences in growth strategies, seasonality and the difference in the light penetration through a broadleaf or needleleaf canopy (Wright *et al.*, 2004; Ollinger, 2011; Butler *et al.*, 2017). Overall, the range in LMA within and across biomes showed similarities and overlap from the Arctic to the tropics, despite the vastly different leaf morphologies resulting from environmental and biological drivers, and these similarities were probably related to the underlying variation in leaf density and volume : area ratios, which can lead to similar values of LMA with significantly different leaf geometries (Poorter *et al.*, 2009; John *et al.*, 2017).

Variation in leaf reflectance across biomes

As with LMA, we observed significant variation in the measured leaf-level spectral reflectance in our model development dataset within and across biomes (Figs 3a,b, S4). The biome-level spectral reflectance displayed a similar shape and comparable magnitude across the spectral range examined here (500–2400 nm; Fig. S4); however, the within-biome variation (presented as the coefficient of variation in reflectance by wavelength) showed some key differences (Fig. 3b) in the visible (i.e. 500–700 nm) and shortwave infrared (SWIR) regions (e.g. 1900–2500 nm). The biome-level mean reflectances appeared similar owing to the large variation in reflectance across wavelengths within each biome (Fig. S4), which mirrored the considerable within-biome variation in LMA (Figs 2, S3). This pattern has been observed in past research for similar vegetation types (Asner & Martin, 2008; Feret *et al.*, 2011; Yang *et al.*, 2016). Across the entire shortwave spectral region (i.e. 500–2400 nm), we also found that reflectance displayed the highest variation (Fig. 3b) in the visible region between about 500 and 700 nm (23–32% across biomes), in the SWIR region between 1300 and 1700 nm (15–17% across biomes), and in the far-SWIR region between 1900 and 2500 nm (24–37% across biomes). The near-infrared (NIR) region displayed only minor variation of *c.* 10% for all biome-level leaf spectra (Fig. 3b).

Exploring the connection between LMA and leaf spectral reflectance in the visible, NIR, and SWIR regions highlighted the strength and direction of the relationships between reflectance and LMA in different parts of the spectrum (Fig. S5). In the visible portion of the spectrum (Fig. S5a–c), the relationship between LMA and spectra is generally positive, as with the NIR (Fig. S5d) despite the much lower coefficient of variation in that region (Fig. 3b). There was a strong negative relationship

No claim to US Government works
New Phytologist © 2019 New Phytologist Trust

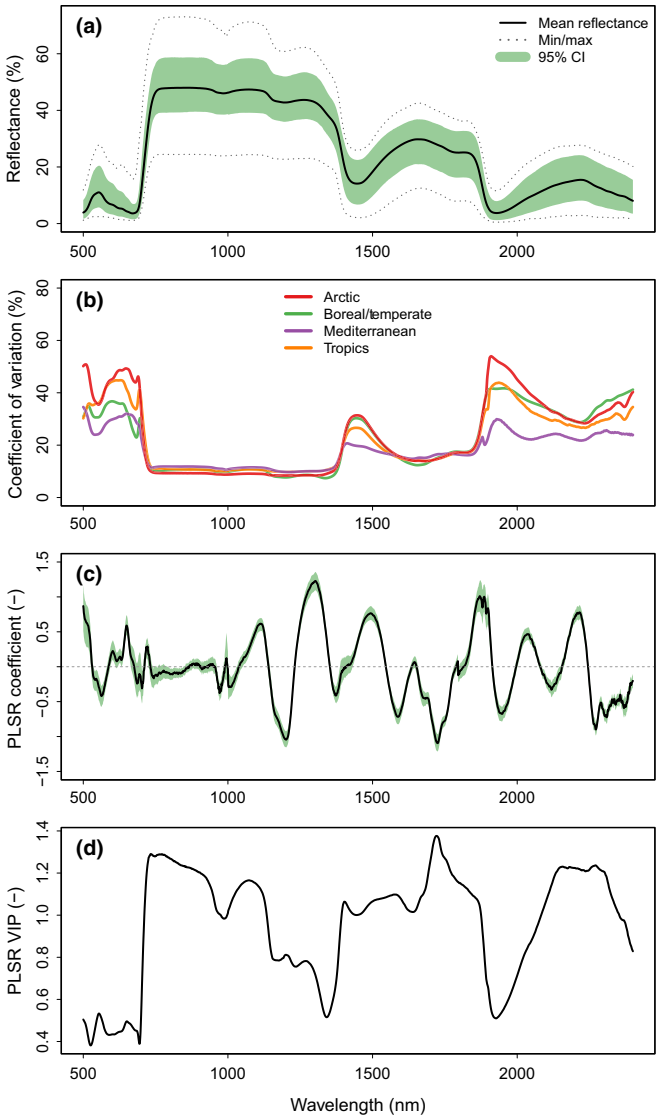

Fig. 3 Leaf reflectance and associated statistics. (a) Mean leaf reflectance, 95% confidence interval (green shading) and minimum and maximum reflectances (dotted lines) in our model development dataset. (b) Coefficient of variation for spectra from the four biomes represented in this study (red, Arctic; green, boreal/temperate; purple, Mediterranean; orange, tropical). (c) Plot of the partial least-squares regression (PLSR) model coefficients. (d) PLSR variable influence on projection (VIP).

between LMA and reflectance in the two SWIR (Fig. S5e,f) wavelengths (1800 and 2200 nm). Overall, these patterns are consistent with previous studies (le Maire *et al.*, 2008; Asner *et al.*, 2011; Feret *et al.*, 2011). For example, Feret *et al.* (2008) showed that spectral absorption of leaf dry matter content (the Cm or LMA parameter in the PROSPECT model) is generally near zero in the visible and NIR, but increases substantially when moving further out into the SWIR wavelengths. On the other hand, Slaton *et al.* (2001) illustrated the relationships of leaf morphology, thickness and other structural characteristics to changes in NIR reflectance across a range of leaf types and also found that there was a generally positive relationship between NIR reflectance (at 800 nm) and leaf structure parameters, including cuticle thickness and mesophyll properties.

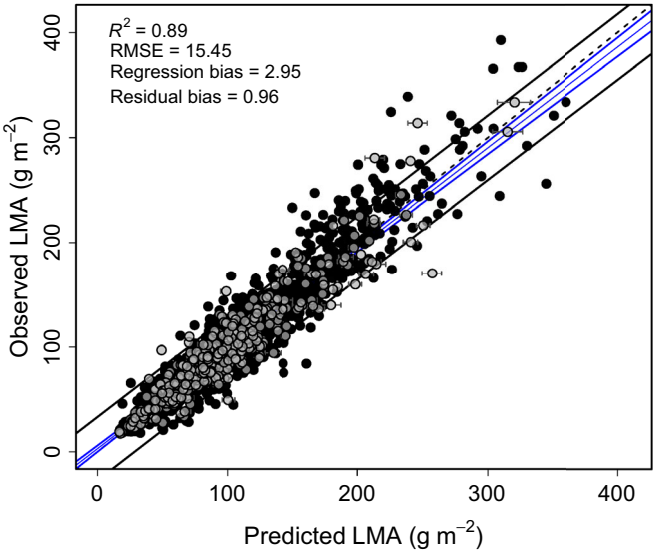

Fig. 4 Observed leaf mass per area (LMA) calculated from measured leaf area and dry mass vs LMA predicted using our spectral model. Our model development dataset (n = 2478) was split into two groups that were used to calibrate (black circles, n = 1978) and validate (gray circles, n = 500) the model. The validation points are shown with ± 95% confidence interval (CI) error bars. For clarity, validation points are shown layered on top of calibration points. The 1 : 1 line is shown as a broken black line and the predictive interval of the model is shown as solid black lines. The regression between observed and predicted LMA is shown in blue (regression, thin blue line; 95% CI, thick blue line). The R^2, root mean square error (RMSE) and regression bias (y-axis intercept) for the validation dataset are shown inside the panel. Data and statistics are presented in back-transformed standard LMA units.

Modeling the variation in LMA across biomes using reflectance spectroscopy

Using the observed variation in measured LMA (Figs 2, S1), leaf reflectance data (Figs 3a, S4), and the covariation between the two (Fig. S5) within our model development dataset, we then evaluated the capacity to build a widely applicable, multibiome model to estimate LMA from leaf spectral reflectance across the broad range of plant species, growth environments, leaf ages, measurement approaches, and spectrometer instrumentation reflected in our dataset. Our PLSR results (Figs 3c,d, 4) showed that a model based on leaf reflectance data was able to explain 89% of the variation in LMA in the independent validation data (and 91% in the calibration data using a cross-validation analysis; Fig. S6a). Our multibiome model also displayed a low overall independent validation error (RMSE = 16 g m^{-2}) and bias (residual bias = 0.96 g m^{-2}, regression bias = 2.95 g m^{-2}), with the model calibration and validation residuals both centered around zero (Fig. S6e,f). Within each biome in the core calibration and validation datasets, the model fit also showed low error, ranging from a minimum RMSE of 12 g m^{-2} in the Arctic plants to 30 g m^{-2} in the Mediterranean plants (Fig. S7). We further explored whether there were significant model biases attributable to variation with canopy position (Fig. S8) and across leaf ages (Fig. S9). Our results demonstrated that the model performance was very consistent across these different axes of variation,

No claim to US Government works
New Phytologist © 2019 New Phytologist Trust

although the leaf age evaluation displayed some bias in the young (expanding leaves or current year foliage) and mature (1 yr old) and mature (> 1 yr old) leaf age classes. Combined, the predictive error (dark black lines, 95% CI error bars on the points in Fig. 4) for the estimated LMA is small, particularly for the most commonly observed values of LMA (e.g. LMA ≤ 150 g m^{-2}; Fig. S1).

Our multibiome LMA PLSR model has a performance similar to or better than previous studies using leaves from a much smaller range of species, locations and growth environments (e.g. Asner *et al.*, 2011; Serbin *et al.*, 2014; Yang *et al.*, 2016). In addition, the model is able to cover a larger range of the global LMA trait-space (Figs 2, S1), habitable climates (Fig. 1) and variation in leaf optical properties (Fig. S4) than earlier work. This includes previous studies utilizing semimechanistic leaf RTMs, including the PROSPECT model (e.g. Feret *et al.*, 2008, 2011; Shiklomanov *et al.*, 2016). For example, Féret *et al.* (2018) found that performance of PROSPET-D LMA inversion varied based on the input dataset (and species within), inversion approach and spectral domain, but that the overall PROSPECT-D inversion results across a smaller set of tropical to boreal samples were comparable to those shown here. Similarly, Shiklomanov *et al.* (2016) also found comparable results using a Bayesian inversion of PROSPECT-5b. However, in both cases, and with previous research (Malenovský *et al.*, 2006), thicker leaves and needles have hindered the inversion accuracy of semimechanistic models like PROSPECT (e.g. Shiklomanov *et al.*, 2016) and, as a result, needleleaf species are typically removed before analysis (e.g. Féret *et al.*, 2018). Although approaches like PROSPECT are attractive to the remote-sensing and plant trait ecology communities, issues surrounding the handling of more complex leaf morphologies and absorption properties need to be addressed in order to facilitate the confident use of these models' global applications. On the other hand, our empirical approach was able to account for a broad range of morphologies and other drivers of leaf optical variation to produce a widely applicable multibiome model (Fig. 4).

The evaluation of the global model PLSR coefficient and VIP plots (Fig. 3c,d) highlights the regions of the spectrum that provide the strongest predictive power, many of which correspond to the areas showing the highest CV values, but note that the NIR region was also important. The SWIR spectral regions are known to contain absorption features related to structure, dry matter content, C compounds and internal leaf water content (Curran, 1989) which also covary with LMA (Elvidge, 1990; Poorter *et al.*, 2009; de la Riva *et al.*, 2016). Our global model showed relatively high VIPs in the NIR region (Fig. 3c,d), consistent with previous studies that also highlighted the importance of this region (Asner *et al.*, 2011; Yendrek *et al.*, 2017), particularly in the transition between visible and NIR reflectance (*c.* 750–800 nm) and the region between 950 and 1200 nm. Corresponding PLSR coefficient values in the NIR were highest in the *c.* 1000–1200 nm region. In general, the NIR has been shown to contain information connected to leaf internal scatter related to mesophyll layer thickness and water content, and varies with water, structural C, leaf thickness and variation in the epidermis layer (Ollinger, 2011), which also varies strongly with LMA (Castro-Díez *et al.*, 2000; Jacquemoud *et al.*, 2009; Poorter *et al.*, 2009; de la Riva *et al.*, 2016).

Our results show the powerful capability of the spectral approach to estimate a key leaf trait (LMA) across a high diversity of plant species, growth environments, leaf ages, leaf morphology and biomes. Our synthesis represents the first time multiple datasets collected from different locations, by different groups, with different instrumentation across such large climatic and geographical ranges and from such a wide diversity of leaf types, including needleleaf species (Fig. 1), have been combined to test the ability to generalize the spectral PLSR modeling approach. Importantly, there appears to be a general pattern in the PLSR models shown for this and previous studies where similar portions of the visible, NIR and SWIR regions display high importance in the estimation of LMA with spectra (Asner *et al.*, 2011; Serbin *et al.*, 2014; Yang *et al.*, 2016). This strongly suggests that the coordination between leaf optical properties and traits can be used to develop generalized global models for leaf traits using the spectroscopic approach.

External validation of the multibiome LMA model

We further tested the predictive capability of our global leaf spectra–LMA model using three additional, completely independent datasets of measured leaf reflectance and LMA from trees, forbs, shrubs and grasses. These datasets include the foundational LOPEX and ANGERS spectral datasets (Hosgood *et al.*, 1994; Jacquemoud *et al.*, 2003), as well as data collected across NEON locations in the continental US. The first two datasets have been heavily used in remote-sensing literature to develop and test the PROSPECT model (Feret *et al.*, 2008). We applied our model coefficients (as shown in Fig. 3c) to the first two leaf reflectance datasets (i.e. LOPEX and ANGERS) and our results showed strong model performance ($R^2 = 0.66$) and low predictive error (RMSE = 21 g m^{-2}; Fig. S10). Similarly, using the datasets collected in the Upper Midwest and at select NEON sites, we applied our model to the measured reflectance data to estimate LMA and found similar results (RMSE = 16.1 and 12.5 g m^{-2} for the Midwest and NEON locations, respectively; Fig. S11), suggesting the high model performance and generality of the spectra–trait approach for estimating LMA across a range of species and environments. As such, the development of globally applicable models like that presented here will reduce time and money spent developing site- or project-specific models to focus more on the rapid collection of much larger datasets by utilizing widely applicable models, like that presented here, to estimate traits such as LMA using only spectral measurements. The spectral approach has the added benefit of providing the capacity to repeat measurements of the same leaves during development or over a season, and during a manipulation or stress event, because it does not require destructive harvesting.

Main applications of our results

Our study of multibiome convergence in the leaf-level spectra–LMA relationships has four major implications. First, our

No claim to US Government works
New Phytologist © 2019 New Phytologist Trust

framework and approach can largely be extended to other leaf traits. In addition to LMA, some other plant functional traits (e.g. pigments (Chl*a*,*b* and carotenoid), chemical concentration (water%, N%, C%, isotopic N and C), and carboxylation capacity (V_{cmax})) are also very important inputs for ESMs, and critical measures of plant form and function (Xu *et al.*, 2012; van Bodegom *et al.*, 2014; Rogers, 2014; Butler *et al.*, 2017; Rogers *et al.*, 2017a; Ricciuto *et al.*, 2018). As these plant functional traits have been shown to be connected to leaf spectral reflectance (Asner & Martin, 2008; Feret *et al.*, 2008; Kokaly *et al.*, 2009; Serbin *et al.*, 2012, 2014, 2015; Wu *et al.*, 2019), we thus expect that similar globally convergent relationships with leaf spectra could be developed for these and other traits. To explore these additional generalities, future work can leverage emerging leaf spectra and trait databases, such as EcoSIS, which will facilitate much faster and easier development, testing and refinement of spectra–trait models.

Second, our work highlights the feasibility of using vegetation spectroscopy to advance large-scale monitoring of plant functional traits. This study evaluated the power of the spectroscopy approach at the leaf level, but an important next step is to explore the ability to use a range of imaging spectroscopy platforms to map traits, such as LMA, at the canopy and landscape scales using airborne and spaceborne instruments across larger regions than currently explored, and ultimately globally (Schimel *et al.*, 2015; Stavros *et al.*, 2017). Here we show the capacity to generalize the spectral approach at the leaf scale and there is strong evidence that similar approaches can be used at the canopy scale with imaging systems (e.g. Asner *et al.*, 2015; Singh *et al.*, 2015). This could augment or supplant the need for direct sampling and measurement of leaf traits at the global scale, especially given important upcoming satellite missions (e.g. Surface Biology and Geology Mission (National Academies of Sciences, Engineering, and Medicine, 2018), previously named *HyspIRI* (Lee *et al.*, 2015), and EnMAP (Guanter *et al.*, 2015)). Furthermore, functional trait mapping from imaging spectroscopy could supplement methods using remote sensing combined with climatology (Butler *et al.*, 2017; Moreno-Martínez *et al.*, 2018). These capabilities would significantly enhance the use of trait observations to inform ESMs and would address critical needs in biodiversity monitoring (Skidmore *et al.*, 2015; Jetz *et al.*, 2016). However, lack of global coverage, inconsistent processing workflows and other challenges with the use of imaging spectroscopy have limited the ability to derive consistent global trait maps. Furthermore, approaches are needed to separate trait retrieval from climatology to allow for the characterization of biotic and abiotic drivers, and for future mapping under novel climates (Fisher *et al.*, 2015). This can only be realized with a spaceborne mapping imaging spectroscopy mission (National Academies of Sciences, 2018).

Third, we hypothesized that a single, multibiome leaf reflectance model of LMA could be developed using datasets across diverse growth environments. We supported this hypothesis by showing a global convergence in the spectra–LMA relationship across samples representing a large portion of the global trait-space for LMA (Figs 4, S1), which, when applied to external datasets from new, independent locations (Figs S10, S11) showed similar model performance for estimating LMA. Importantly, as leaf traits, particularly LMA, display adaptations and acclimation to their growth environment (Poorter *et al.*, 2009; Osnas *et al.*, 2018), the success of a generalized spectra–LMA model shown in this study further suggests that the spectroscopic approach could be an important, nondestructive means to help quantify and understand how plant traits acclimate to climatic variability and global change through rapid collection of traits with spectra (Shiklomanov *et al.*, 2019). Moreover, while the one-time costs of the spectrometer instrumentation utilized in this work can be high, this cost is low compared with the overall costs associated with traditional measurements of LMA (e.g. labor and supply costs over the lifetime of a project) and provides a much more rapid means for the retrieval of LMA.

Finally, the resulting multi-biome LMA PLSR model (see Methods S1 and associated links within) developed here can be downloaded and used by the scientific community. As measurement of leaf spectral reflectance is the only input required, our model enables the rapid, accurate and nondestructive estimation of LMA and can therefore be used for a broad range of additional applications, such as monitoring plant response to an emerging stress or evaluating physiological traits of interest to breeders in high-throughput phenotyping experiments (Silva-Perez *et al.*, 2017; Yendrek *et al.*, 2017; Ely *et al.*, 2019). As a result, databases such as EcoSIS can now be mined for increased coverage of important plant traits by applying this and other spectra–trait models, potentially increasing the amount of available data for modeling and ecological research.

Conclusions

In this work, we present a detailed synthesis of an extensive dataset containing coincident measurements of full-spectrum (i.e. 0.3–2.5 μm) leaf reflectance and LMA across multiple biomes and spanning large environmental gradients (Fig. 1). The variation in our dataset covered a large portion of the global variation observed for LMA (Fig. S1) and similarly contained a high degree of variation in leaf reflectance spectra (e.g. Figs 3, S4). Despite this high degree of variation in LMA and leaf reflectance observed across biomes, species, environment, leaf age and canopy position, our spectra–trait approach was able to collapse this variation into a single, widely applicable model (Fig. 4) that can be used to estimate LMA rapidly and accurately using only reflectance data (Figs S10, S11). These results are important because they highlight that the spectral approach presented here can be used to develop algorithms using datasets from a range of instruments, groups and locations to develop generalized models. Furthermore, this work illustrates the utility of the spectral approach in providing rapid, relatively low-cost and nondestructive (when the leaves are accessible) measurements of key plant traits across diverse plant species. The use of models such as that presented here will help to rapidly expand trait databases in order to address the known bias in observational datasets (Schimel *et al.*, 2015).

No claim to US Government works
New Phytologist © 2019 New Phytologist Trust

Acknowledgements

We thank Wil Lieberman-Cribbin and Jennifer Liebig for assistance with data collection in Barrow. We thank Neill Prohaska, Moira Hough and Anthony John Junqueira Garnello for assistance with data collection in Biosphere 2, and Christian Torres for assistance in the collection of tree samples within the Dona Inés Park, on the grounds of the Luis Muñoz Marín Foundation in Puerto Rico. We also gratiously thank Alexey Shiklomanov who provided valuable feedback on an earlier draft of this manuscript. This work and associated field data collection campaigns were supported by the Next-Generation Ecosystem Experiments (NGEE Arctic and NGEE Tropics), projects that are supported by the Office of Biological and Environmental Research in the Department of Energy, Office of Science, and through the United States Department of Energy contract no. DE-SC0012704 to Brookhaven National Laboratory, by NASA Earth and Space Sciences Fellowship (NNX08AV07H) provided to SPS, Forest Functional Types (NNX12AQ28G) and HyspIRI grants (NNX12AQ28G), NSF Macrosystems Biology grant (1638720) to PAT and ELK, as well as a USDA McIntire-Stennis grant (WIS01809) to PAT and ELK.

Author contributions

SPS, JW and AR conceived the study, and defined the scope and focus of the manuscript. SPS developed the spectroscopic model of LMA. KSE, JW, RM, BTW, AR, and SPS participated in the collection of data in the Arctic and tropical field sites (including the Biosphere 2 experiment). SPS, PAT, ELK, AC and ZW collected the datasets in California, the Upper Midwest, and at the NEON locations. SPS, JW and AR wrote the first draft of the manuscript. All authors contributed to the final version.

ORCID

Adam Chlus https://orcid.org/0000-0001-6719-9956
Kim S. Ely https://orcid.org/0000-0002-3915-001X
Eric L. Kruger https://orcid.org/0000-0003-2463-9231
Ran Meng https://orcid.org/0000-0003-4756-9934
Alistair Rogers https://orcid.org/0000-0001-9262-7430
Shawn P. Serbin https://orcid.org/0000-0003-4136-8971
Philip A. Townsend https://orcid.org/0000-0001-7003-8774
Zhihui Wang https://orcid.org/0000-0003-1064-7820
Brett T. Wolfe https://orcid.org/0000-0001-7535-045X
Jin Wu https://orcid.org/0000-0001-8991-3970

References

Asner GP, Martin RE. 2008. Spectral and chemical analysis of tropical forests: scaling from leaf to canopy levels. *Remote Sensing of Environment* **112**: 3958–3970.

Asner GP, Martin RE, Anderson CB, Knapp DE. 2015. Quantifying forest canopy traits: imaging spectroscopy versus field survey. *Remote Sensing of Environment* **158**: 15–27.

Asner GP, Martin RE, Tupayachi R, Emerson R, Martinez P, Sinca F, Powell GVN, Wright SJ, Lugo AE. 2011. Taxonomy and remote sensing of leaf mass per area (LMA) in humid tropical forests. *Ecological Applications* **21**: 85–98.

van Bodegom PM, Douma JC, Verheijen LM. 2014. A fully traits-based approach to modeling global vegetation distribution. *Proceedings of the National Academy of Sciences, USA* **111**: 13733–13738.

Brown J, Everett KR, Webber PJ, Maclean SF, Murray DF. 1980. The coastal tundra at Barrow. In: Brown J, Miller PC, Tiezen LL, Bunnell FL, eds. *An Arctic ecosystem: the coastal tundra at Barrow, Alaska*. Stroudsburg, PA: Dowden, Hutchinson & Ross, Inc., 571.

Butler EE, Datta A, Flores-Moreno H, Chen M, Wythers KR, Fazayeli F, Banerjee A, Atkin OK, Kattge J, Amiaud B *et al.* 2017. Mapping local and global variability in plant trait distributions. *Proceedings of the National Academy of Sciences, USA* **114**: E10937–E10946.

Castro-Díez P, Puyravaud JP, Cornelissen JHC. 2000. Leaf structure and anatomy as related to leaf mass per area variation in seedlings of a wide range of woody plant species and types. *Oecologia* **124**: 476–486.

Chen S, Hong X, Harris CJ, Sharkey PM. 2004. Spare modeling using orthogonal forest regression with PRESS statistic and regularization. *IEEE Transaction on Systems, Man and Cybernetics* **34**: 898–911.

R Core Team. 2017. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

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

Dahlin KM, Asner GP, Field CB. 2013. Environmental and community controls on plant canopy chemistry in a Mediterranean-type ecosystem. *Proceedings of the National Academy of Sciences, USA* **110**: 6895–6900.

DuBois S, Desai AR, Singh A, Serbin SP, Goulden ML, Baldocchi DD, Ma S, Oechel WC, Wharton S, Kruger EL *et al.* 2018. Using imaging spectroscopy to detect variation in terrestrial ecosystem productivity across a water-stressed landscape. *Ecological Applications* **28**: 1313–1324.

Elvidge CD. 1990. Visible and near-infrared reflectance characteristics of dry plant materials. *International Journal of Remote Sensing* **11**: 1775–1795.

Ely KS, Burnett AC, Lieberman-Cribbin W, Serbin S, Rogers A. 2019. Spectroscopy can predict key leaf traits associated with source-sink balance and carbon-nitrogen status. *Journal of Experimental Botany* **70**: 1789–1799.

Feret JB, Francois C, Asner GP, Gitelson AA, Martin RE, Bidel LPR, Ustin SL, le Maire G, Jacquemoud S. 2008. PROSPECT-4 and 5: advances in the leaf optical properties model separating photosynthetic pigments. *Remote Sensing of Environment* **112**: 3030–3043.

Feret J-B, Francois C, Gitelson A, Asner GP, Barry KM, Panigada C, Richardson AD, Jacquemoud S. 2011. Optimizing spectral indices and chemometric analysis of leaf chemical properties using radiative transfer modeling. *Remote Sensing of Environment* **115**: 2742–2750.

Féret JB, Gitelson AA, Noble SD, Jacquemoud S. 2017. PROSPECT-D: towards modeling leaf optical properties through a complete lifecycle. *Remote Sensing of Environment* **193**: 204–215.

Féret JB, le Maire G, Jay S, Berveiller D, Bendoula R, Hmimina G, Cheraiet A, Oliveira JC, Ponzoni FJ, Solanki T *et al.* 2018. Estimating leaf mass per area and equivalent water thickness based on leaf optical properties: potential and limitations of physical modeling and machine learning. *Remote Sensing of Environment* **231**: 110959.

Fisher JB, Huntzinger DN, Schwalm CR, Sitch S. 2014. Modeling the terrestrial biosphere. *Annual Review of Environment and Resources* **39**: 91–123.

Fisher RA, Koven CD, Anderegg WRL, Christoffersen BO, Dietze MC, Farrior CE, Holm JA, Hurtt GC, Knox RG, Lawrence PJ *et al.* 2018. Vegetation demographics in Earth System Models: a review of progress and priorities. *Global Change Biology* **24**: 35–54.

Fisher RA, Muszala S, Verteinstein M, Lawrence P, Xu C, McDowell NG, Knox RG, Koven C, Holm J, Rogers BM *et al.* 2015. Taking off the training wheels: the properties of a dynamic vegetation model without climate envelopes, CLM4.5(ED). *Geoscientific Model Development* **8**: 3593–3619.

Geladi P, Kowalski BR. 1986. Partial least-squares regression – a tutorial. *Analytica Chimica Acta* **185**: 1–17.

Gitelson AA, Gritz Y, Merzlyak MN. 2003. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive

No claim to US Government works
New Phytologist © 2019 New Phytologist Trust

chlorophyll assessment in higher plant leaves. *Journal of Plant Physiology* **160**: 271–282.

Goulden ML, Anderson RG, Bales RC, Kelly AE, Meadows M, Winston GC. 2012. Evapotranspiration along an elevation gradient in California's Sierra Nevada. *Journal of Geophysical Research-Biogeosciences* 117.

Guanter L, Kaufmann H, Segl K, Foerster S, Rogass C, Chabrillat S, Kuester T, Hollstein A, Rossner G, Chlebek C *et al.* 2015. The EnMAP spaceborne imaging spectroscopy mission for Earth observation. *Remote Sensing* 7: 8830.

Hosgood B, Jacquemoud S, Andreoli G, Verdebout J, Pedrini G, Schmuck G. 1994. *Leaf Optical Properties EXperiment 93 (LOPEX93)*. Ispra, Italy: European Commission—Joint Research Centre.

Jacquemoud S, Bidel L, Francois C, Pavan G. 2003. ANGERS Leaf Optical Properties Database. [WWW document] URL http://opticleaf.ipgp.fr/index.php?page=database: OPTICLEAF [accessed 7 February 2018].

Jacquemoud S, Verhoef W, Baret F, Bacour C, Zarco-Tejada PJ, Asner GP, Francois C, Ustin SL. 2009. PROSPECT + SAIL models: a review of use for vegetation characterization. *Remote Sensing of Environment* **113**(Suppl. 1, Sp. Iss. SI): S56–S66.

Jetz W, Cavender-Bares J, Pavlick R, Schimel D, Davis FW, Asner GP, Guralnick R, Kattge J, Latimer AM, Moorcroft P *et al.* 2016. Monitoring plant functional diversity from space. *Nature Plants* **2**: 16024.

John GP, Scoffoni C, Buckley TN, Villar R, Poorter H, Sack L, Maherali H. 2017. The anatomical and compositional basis of leaf mass per area. *Ecology Letters* **20**: 412–425.

Kattge J, DÍaz S, Lavorel S, Prentice IC, Leadley P, Bönisch G, Garnier E, Westoby M, Reich PB, Wright IJ *et al.* 2011. TRY – a global database of plant traits. *Global Change Biology* **17**: 2905–2935.

Kokaly RF, Asner GP, Ollinger SV, Martin ME, Wessman CA. 2009. Characterizing canopy biochemistry from imaging spectroscopy and its application to ecosystem studies. *Remote Sensing of Environment* **113**: S78–S91.

Lebauer DS, Wang D, Richter KT, Davidson CC, Dietze MC. 2013. Facilitating feedbacks between field measurements and ecosystem models. *Ecological Monographs* **83**: 133–154.

Lee CM, Cable ML, Hook SJ, Green RO, Ustin SL, Mandl DJ, Middleton EM. 2015. An introduction to the NASA Hyperspectral InfraRed Imager (HyspIRI) mission and preparatory activities. *Remote Sensing of Environment* **167**: 6–19.

Liu M, Wang Z, Li S, Lü X, Wang X, Han X. 2017. Changes in specific leaf area of dominant plants in temperate grasslands along a 2500-km transect in northern China. *Scientific Reports* 7: 10780.

le Maire G, François C, Soudani K, Berveiller D, Pontailler J-Y, Bréda N, Genet H, Davi H, Dufrêne E. 2008. Calibration and validation of hyperspectral indices for the estimation of broadleaved forest leaf chlorophyll content, leaf mass per area, leaf area index and leaf canopy biomass. *Remote Sensing of Environment* **112**: 3846–3864.

Malenovský Z, Albrechtová J, Lhotáková Z, Zurita-Milla R, Clevers JGPW, Schaepman ME, Cudlín P. 2006. Applicability of the PROSPECT model for Norway spruce needles. *International Journal of Remote Sensing* **27**: 5315–5340.

Mevik B-H, Wehrens R. 2007. The pls package: principal component and partial least squares regression in R. *Journal of Statistical Software* **18**: 1–24.

Moreno-Martínez Á, Camps-Valls G, Kattge J, Robinson N, Reichstein M, van Bodegom P, Kramer K, Cornelissen JHC, Reich P, Bahn M *et al.* 2018. A methodology to derive global maps of leaf traits using remote sensing and climate data. *Remote Sensing of Environment* **218**: 69–88.

National Academies of Sciences, Engineering, and Medicine. 2018. *Thriving on our changing planet: a decadal strategy for Earth observation from space*. Washington, DC: The National Academies Press.

Niinemets U. 2007. Photosynthesis and resource distribution through plant canopies. *Plant, Cell & Environment* **30**: 1052–1071.

Niinemets Ü, Keenan TF, Hallik L. 2015. A worldwide analysis of within-canopy variations in leaf structural, chemical and physiological traits across plant functional types. *New Phytologist* **205**: 973–993.

Ollinger SV. 2011. Sources of variability in canopy reflectance and the convergent properties of plants. *New Phytologist* **189**: 375–394.

Osnas JLD, Katabuchi M, Kitajima K, Wright SJ, Reich PB, Van Bael SA, Kraft NJB, Samaniego MJ, Pacala SW, Lichstein JW. 2018. Divergent drivers of leaf trait variation within species, among species, and among functional groups. *Proceedings of the National Academy of Sciences, USA* **115**: 5480–5485.

Paula S, Pausas JG. 2006. Leaf traits and resprouting ability in the Mediterranean basin. *Functional Ecology* **20**: 941–947.

Pavlick R, Drewry DT, Bohn K, Reu B, Kleidon A. 2013. The Jena Diversity-Dynamic Global Vegetation Model (JeDi-DGVM): a diverse approach to representing terrestrial biogeography and biogeochemistry based on plant functional trade-offs. *Biogeosciences* **10**: 4137–4177.

Poorter H, Niinemets U, Poorter L, Wright IJ, Villar R. 2009. Causes and consequences of variation in leaf mass per area (LMA): a meta-analysis. *New Phytologist* **182**: 565–588.

Reich PB, Ellsworth DS, Walters MB, Vose JM, Gresham C, Volin JC, Bowman WD. 1999. Generality of leaf trait relationships: a test across six biomes. *Ecology* **80**: 1955–1969.

Reich PB, Walters MB, Ellsworth DS. 1997. From tropics to tundra: global convergence in plant functioning. *Proceedings of the National Academy of Sciences, USA* **94**: 13730–13734.

Ricciuto D, Sargsyan K, Thornton P. 2018. The impact of parametric uncertainties on biogeochemistry in the E3SM land model. *Journal of Advances in Modeling Earth Systems* **10**: 297–319.

de la Riva EG, Olmo M, Poorter H, Ubera JL, Villar R. 2016. Leaf mass per area (LMA) and its relationship with leaf structure and anatomy in 34 Mediterranean woody species along a water availability gradient. *PLoS ONE* **11**: e0148788.

Rogers A. 2014. The use and misuse of Vcmax in Earth System Models. *Photosynthesis Research* **119**: 15–29.

Rogers A, Medlyn BE, Dukes JS, Bonan G, von Caemmerer S, Dietze MC, Kattge J, Leakey ADB, Mercado LM, Niinemets Ü *et al.* 2017a. A roadmap for improving the representation of photosynthesis in Earth system models. *New Phytologist* **213**: 22–42.

Rogers A, Serbin SP, Ely K. 2016. Leaf Mass Area, Leaf Carbon and Nitrogen Content, Kougarok Road and Teller Road, Seward Peninsula, Alaska, 2016. *Next Generation Ecosystem Experiments Arctic Data Collection*. Oak Ridge National Laboratory: U.S. Department of Energy, Oak Ridge, Tennessee, USA.

Rogers A, Serbin SP, Ely KS, Sloan VL, Wullschleger SD. 2017b. Terrestrial biosphere models underestimate photosynthetic capacity and CO2 assimilation in the Arctic. *New Phytologist* **216**: 1090–1103.

Schimel D, Pavlick R, Fisher JB, Asner GP, Saatchi S, Townsend P, Miller C, Frankenberg C, Hibbard K, Cox P. 2015. Observing terrestrial ecosystems and the carbon cycle from space. *Global Change Biology* **21**: 1762–1776.

Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**: 671–675.

Serbin SP. 2012. *Spectroscopic determination of leaf nutritional, morphological, and metabolic traits*. PhD Dissertation, University of Wisconsin – Madison, Madison, WI, USA.

Serbin SP, Dillaway DN, Kruger EL, Townsend PA. 2012. Leaf optical properties reflect variation in photosynthetic metabolism and its sensitivity to temperature. *Journal of Experimental Botany* **63**: 489–502.

Serbin SP, Rogers A. 2019. NGEE Arctic Leaf Spectral Reflectance, Kougarok Road, Seward Peninsula, Alaska, 2016. *Next Generation Ecosystem Experiments Arctic Data Collection*: US Department of Energy, Oak Ridge, Tennessee, USA.

Serbin SP, Singh A, Desai AR, Dubois SG, Jablonski AD, Kingdon CC, Kruger EL, Townsend PA. 2015. Remotely estimating photosynthetic capacity, and its response to temperature, in vegetation canopies using imaging spectroscopy. *Remote Sensing of Environment* **167**: 78–87.

Serbin SP, Singh A, McNeil BE, Kingdon CC, Townsend PA. 2014. Spectroscopic determination of leaf morphological and biochemical traits for northern temperate and boreal tree species. *Ecological Applications* **24**: 1651–1669.

Shiklomanov AN, Bradley BA, Dahlin KM, Fox A, Gough CM, Hoffman FM, Middleton E, Serbin SP, Smallman L, Smith WK. 2019. Enhancing global change experiments through integration of remote-sensing techniques. *Frontiers in Ecology and the Environment* **17**: 215–224.

Shiklomanov AN, Dietze MC, Viskari T, Townsend PA, Serbin SP. 2016. Quantifying the influences of spectral resolution on uncertainty in leaf trait

No claim to US Government works
New Phytologist © 2019 New Phytologist Trust

estimates through a Bayesian approach to RTM inversion. *Remote Sensing of Environment* **183**: 226–238.

Shipley B, Lechowicz MJ, Wright I, Reich PB. 2006. Fundamental trade-offs generating the worldwide leaf economics spectrum. *Ecology* **87**: 535–541.

Silva-Perez V, Molero G, Serbin SP, Condon AG, Reynolds MP, Furbank RT, Evans JR. 2017. Hyperspectral reflectance as a tool to measure biochemical and physiological traits in wheat. *Journal of Experimental Botany* **69**: 483–496.

Sims DA, Gamon JA. 2002. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sensing of Environment* **81**: 337–354.

Singh A, Serbin SP, McNeil BE, Kingdon CC, Townsend PA. 2015. Imaging spectroscopy algorithms for mapping canopy foliar chemical and morphological traits and their uncertainties. *Ecological Applications* **25**: 2180–2197.

Skidmore AK, Pettorelli N, Coops NC, Geller GN, Hansen M, Lucas R, Mucher CA, O'Connor B, Paganini M, Pereira HM *et al.* 2015. Agree on biodiversity metrics to track from space. *Nature* **523**: 403–405.

Slaton MR, Raymond Hunt Jr E, Smith WK. 2001. Estimating near-infrared leaf reflectance from leaf structural characteristics. *American Journal of Botany* **88**: 278–284.

Stavros EN, Schimel D, Pavlick R, Serbin S, Swann A, Duncanson L, Fisher JB, Fassnacht F, Ustin S, Dubayah R *et al.* 2017. ISS observations offer insights into plant function. *Nature Ecology & Evolution* **1**: 0194.

Townsend PA, Foster JR, Chastain RA, Currie WS. 2003. Application of imaging spectroscopy to mapping canopy nitrogen in the forests of the central Appalachian Mountains using Hyperion and AVIRIS. *IEEE Transactions on Geoscience and Remote Sensing* **41**: 1347–1354.

Ustin SL, Roberts DA, Gamon JA, Asner GP, Green RO. 2004. Using imaging spectroscopy to study ecosystem processes and properties. *BioScience* **54**: 523–534.

Volin JC, Kruger EL, Lindroth RL. 2002. Responses of deciduous broadleaf trees to defoliation in a CO2 enriched atmosphere. *Tree Physiology* **22**: 435–448.

Walter A, Carmen Lambrecht S. 2004. Biosphere 2 Center as a unique tool for environmental studies. *Journal of Environmental Monitoring* **6**: 267–277.

Whittaker RH. 1975. *Communities and ecosystems.* New York, NY: MacMillan Publishing Co.

Wold S, Sjostrom M, Eriksson L. 2001. PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems* **58**: 109–130.

Wright SJ, Horlyck V, Basset Y, Barrios H, Bethancourt A, Bohlman SA, Gilbert CS, Goldstein G, Graham EA, Kitajima K. 2003. *Tropical canopy biology program, Republic of Panama. Studying forest canopies from above: the International Canopy Crane Network.* Balbao, Ancon, Republic of Panama: Smithsonian Tropical Research Institute and UNEP.

Wright IJ, Reich PB, Westoby M, Ackerly DD, Baruch Z, Bongers F, Cavender-Bares J, Chapin T, Cornelissen JHC, Diemer M *et al.* 2004. The worldwide leaf economics spectrum. *Nature* **428**: 821–827.

Wu J, Albert LP, Lopes AP, Restrepo-Coupe N, Hayek M, Wiedemann KT, Guan K, Stark SC, Christoffersen B, Prohaska N *et al.* 2016. Leaf development and demography explain photosynthetic seasonality in Amazon evergreen forests. *Science* **351**: 972–976.

Wu J, Chavana-Bryant C, Prohaska N, Serbin SP, Guan K, Albert LP, Yang X, Leeuwen WJD, Garnello AJ, Martins G *et al.* 2017. Convergence in relationships between leaf traits, spectra and age across diverse canopy environments and two contrasting tropical forests. *New Phytologist* **214**: 1033–1048.

Wu J, Rogers A, Albert LP, Ely K, Prohaska N, Wolfe BT, Oliveira RC Jr, Saleska SR, Serbin SP. 2019. Leaf reflectance spectroscopy captures variation in carboxylation capacity across species, canopy environment, and leaf age in lowland moist tropical forests. *New Phytologist* **224**: 663–674.

Wullschleger SD, Epstein HE, Box EO, Euskirchen ES, Goswami S, Iversen CM, Kattge J, Norby RJ, van Bodegom PM, Xu X. 2014. Plant functional types in Earth system models: past experiences and future directions for application of dynamic vegetation models in high-latitude ecosystems. *Annals of Botany* **114**: 1–16.

Xu C, Fisher R, Wullschleger SD, Wilson CJ, Cai M, McDowell NG. 2012. Toward a mechanistic modeling of nitrogen limitation on vegetation dynamics. *PLoS ONE* **7**: e37914.

Xu X, Medvigy D, Powers Jennifer S, Becknell Justin M, Guan K. 2016. Diversity in plant hydraulic traits explains seasonal and inter-annual variations of vegetation dynamics in seasonally dry tropical forests. *New Phytologist* **212**: 80–95.

Yang X, Tang J, Mustard JF, Wu J, Zhao K, Serbin S, Lee J-E. 2016. Seasonal variability of multiple leaf traits captured by leaf spectroscopy at two temperate deciduous forests. *Remote Sensing of Environment* **179**: 1–12.

Yendrek CR, Tomaz T, Montes CM, Cao Y, Morse AM, Brown PJ, McIntyre LM, Leakey ADB, Ainsworth EA. 2017. High-throughput phenotyping of maize leaf physiological and biochemical traits using hyperspectral reflectance. *Plant Physiology* **173**: 614–626.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Comparison of the leaf mass per area (LMA) distribution in this article vs that found within the TRY global trait database.

Fig. S2 Distributions of the untransformed leaf mass per area (LMA) calibration and validation datasets used for model development and testing.

Fig. S3 The coefficient of variation of the measured leaf mass per area (LMA) across the four main biomes included in this research.

Fig. S4 The patterns of leaf reflectance across the four main biomes.

Fig. S5 The relationship between leaf mass per area (LMA) and spectral reflectance within the model development dataset.

Fig. S6 Model calibration and validation results for the leaf mass per area (LMA) partial least-squares regression (PLSR) model.

Fig. S7 Partial least-squares regression (PLSR) model fit by biome.

Fig. S8 Combined calibration and validation partial least-squares regression (PLSR) model residuals by canopy position.

Fig. S9 Combined calibration and validation partial least-squares regression (PLSR) model residuals by leaf age.

Fig. S10 Evaluation of the leaf mass per area (LMA) partial least-squares regression (PLSR) model using the LOPEX and ANGERS foundational datasets.

Fig. S11 External validation of the leaf mass per area (LMA) partial least-squares regression (PLSR) model using data collected from NEON sites.

Methods S1 An example R script illustrating the use of the multibiome PLSR model to estimate leaf mass per area for

No claim to US Government works
New Phytologist © 2019 New Phytologist Trust

LOPEX and Angers datasets, which are provided through the EcoSIS spectral library.

Table S1 Leaf spectra and LMA data sources.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.

About *New Phytologist*

- *New Phytologist* is an electronic (online-only) journal owned by the New Phytologist Trust, a **not-for-profit organization** dedicated to the promotion of plant science, facilitating projects from symposia to free access for our Tansley reviews and Tansley insights.
- Regular papers, Letters, Research reviews, Rapid reports and both Modelling/Theory and Methods papers are encouraged. We are committed to rapid processing, from online submission through to publication 'as ready' via *Early View* – our average time to decision is <26 days. There are **no page or colour charges** and a PDF version will be provided for each article.
- The journal is available online at Wiley Online Library. Visit **www.newphytologist.com** to search the articles and register for table of contents email alerts.
- If you have any questions, do get in touch with Central Office (np-centraloffice@lancaster.ac.uk) or, if it is more convenient, our USA Office (np-usaoffice@lancaster.ac.uk)
- For submission instructions, subscription and all the latest information visit **www.newphytologist.com**

No claim to US Government works
New Phytologist © 2019 New Phytologist Trust