Submitted To: New Phytologist

Type of Paper: Original Research

Title: Spectroscopy outperforms leaf trait relationships for predicting photosynthetic capacity across different forest types

Author List: Zhengbing Yan¹, Zhengfei Guo¹, Shawn P. Serbin², Guangqin Song¹, Yingyi Zhao¹, Yang Chen¹, Shengbiao Wu¹, Jing Wang¹, Xin Wang³, Jing Li³,⁴, Bin Wang³,⁴, Yuntao Wu³,⁴, Yanjun Su³,⁴, Han Wang⁵,⁶, Alistair Rogers², Lingli Liu³,⁴, Jin Wu¹*

Author Affiliations:
(1) Division for Ecology and Biodiversity, School of Biological Sciences, The University of Hong Kong, Hong Kong, China
(2) Environmental & Climate Sciences Department, Brookhaven National Laboratory, Upton, New York, NY 11973, USA
(3) State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Xiangshan, Beijing, China
(4) University of Chinese Academy of Sciences, Yuquanlu, Beijing, China
(5) Ministry of Education Key Laboratory for Earth System Modelling, Department of Earth System Science, Tsinghua University, Beijing, China
(6) Joint Centre for Global Change Studies, Tsinghua University, Beijing, China

* Corresponding Author: Jin Wu
Division for Ecology and Biodiversity, School of Biological Sciences, The University of Hong Kong, Pokfulam Road, Hong Kong (email: jinwu@hku.hk; phone: +852 2299-0655)

Words Record:
Summary: 199 words
Maintext Total: 6498 words
  Introduction: 1340 words
Materials and Methods: 2479 words
Results: 612 words
Discussion: 2067 words

Figures/Tables Record:
- Figures: 6
- Tables: 1
- Supplementary Figures/Tables: 10 figures + 3 tables

Others:
- Supplementary Method S1

ORCID ID
Zhengbing Yan: http://orcid.org/0000-0002-5028-8167
Shawn P. Serbin: https://orcid.org/0000-0003-4136-8971
Shengbiao Wu: https://orcid.org/0000-0003-4569-6366
Xin Wang: http://orcid.org/0000-0002-5274-7566
Bin Wang: https://orcid.org/0000-0003-0893-0539
Yuntao Wu: https://orcid.org/0000-0001-6505-1851
Yanjun Su: https://orcid.org/0000-0001-7931-339X
Han Wang: https://orcid.org/0000-0003-2482-1818
Alistair Rogers: https://orcid.org/0000-0001-9262-7430
Lingli Liu: https://orcid.org/0000-0002-5696-3151
Jin Wu: https://orcid.org/0000-0001-8991-3970
Summary

• Leaf trait relationships are widely used to predict ecosystem function in Terrestrial Biosphere Models (TBMs), in which leaf maximum carboxylation capacity ($V_{c,\text{max}}$), an important trait for modelling photosynthesis, can be inferred from other easier-to-measure traits. However, whether trait-$V_{c,\text{max}}$ relationships are robust across different forest types remains unclear.

• Here we used measurements of leaf traits, including one morphological trait (leaf mass per area), three biochemical traits (leaf water content, area-based leaf nitrogen content, and leaf chlorophyll content), one physiological trait ($V_{c,\text{max}}$), as well as leaf reflectance spectra, and explored their relationships within and across three contrasting forest types in China.

• We found weak and forest type-specific relationships between $V_{c,\text{max}}$ and the four morphological and biochemical traits ($R^2\leq0.15$), indicated by significantly changing slopes and intercepts across forest types. In contrast, reflectance spectroscopy effectively collapsed the differences in the trait-$V_{c,\text{max}}$ relationships across three forest biomes into a single robust model for $V_{c,\text{max}}$ ($R^2=0.77$), and also accurately estimated the four traits ($R^2=0.75-0.94$).

• These findings challenge the traditional use of empirical trait-$V_{c,\text{max}}$ relationships in TBMs for estimating terrestrial plant photosynthesis, but also highlight spectroscopy as an efficient alternative for characterizing $V_{c,\text{max}}$ and multi-trait variability, with critical insights into ecosystem modeling and functional trait ecology.

Key Words
gas exchange, plant functional traits, multi-trait covariance, vegetation spectroscopy, leaf hyperspectral reflectance, partial least-squares regression (PLSR), maximum carboxylation capacity
**Introduction**

An accurate understanding and representation of the relationships among plant traits has been an essential prerequisite to quantify their roles in determining key ecological processes, ranging from instantaneous physiological response and resource use to carbon allocation, long-term adaptation and community assembly (Wright et al., 2004; Diaz et al., 2016). Given their importance in shaping emergent ecosystem function, the process modelling community has started to include these processes and mechanisms by simultaneously representing multiple important traits and trait variability in ecosystem models (Rogers et al., 2017a; Bonan & Doney, 2018; Berzaghi et al., 2020). These representations of trait variability have been shown to improve simulations of ecosystem-scale resilience and the sensitivity to climate change (Bonan & Doney, 2018). Since it is difficult to measure the full suite of leaf traits at the same time, but also redundant to incorporate many of these traits as they are highly correlated, empirical and well-established leaf trait relationships, including the worldwide leaf economics spectrum (Wright et al., 2004; Diaz et al., 2016), are often adopted in the ecosystem models. With leaf trait relationships, several important but difficult-to-measure traits (e.g. leaf life span and light saturated photosynthetic rate) can be indirectly inferred from other relatively easier-to-measure traits (e.g. leaf mass per area (LMA) and leaf nitrogen (N) content) with more extensive spatial coverage (Wright et al., 2005; Reich et al., 2007; Berzaghi et al., 2020).

One of the most important but difficult-to-measure traits is the leaf maximum carboxylation rate of the enzyme Rubisco standardized to a reference temperature of 25°C ($V_{c,max25}$; Rogers et al., 2017a). $V_{c,max25}$ directly mediates biotic controls on whole-plant to canopy photosynthetic carbon uptake and interactions with climate (Kattge et al., 2009; Wu et al., 2016). In terrestrial biosphere models (TBMs), $V_{c,max25}$ is an important biochemical parameter that regulates modelled leaf photosynthetic rate, and its parameterization largely determines the accuracy of TBM simulations of terrestrial photosynthesis and carbon uptake (Bernacchi et al., 2013; Rogers et al., 2017a). However, $V_{c,max25}$ also displays large variability in nature and is affected by multiple abiotic and biotic factors, such as climate conditions, soil properties, plant species, leaf morphology and leaf
age (e.g. Walker et al., 2014; Ali et al., 2015; Albert et al., 2018; Smith & Dukes, 2018; Wu et al., 2019). Therefore, efficient and accurate estimation of $V_{c,\text{max}25}$ variability across scales from individual organisms to forest ecosystems remains an essential step to improve the characterization of the drivers of $V_{c,\text{max}25}$ variability and the subsequent parameterization and evaluation of TBMs.

At present, $V_{c,\text{max}25}$ is parameterized in TBMs in a few different ways (Rogers, 2014). For example, the default approach for some models is to use a single static value for each discrete plant functional type (PFT), often parameterized with mean values for the PFT derived from the literature (Kattge et al., 2009; Rogers, 2014; Berzaghi et al., 2020). As an alternative to the fixed parameterizations, some groups have explored the capacity to develop custom parameterizations based on trait databases and statistical methods to derive the best estimate for each PFT (e.g. Lebauer et al., 2013; Dietze et al., 2014). Other approaches include a more direct parameterization through model inversion, for example by using eddy covariance-derived photosynthesis (gross primary productivity, GPP) (Zheng et al., 2017; Schimel et al., 2019) or satellite-based solar-induced chlorophyll fluorescence data (Zhang et al., 2018; Camino et al., 2019; He LM et al., 2019) to tune $V_{c,\text{max}25}$. Others may instead build in a broad range of trait variance into simulations to explore which trait value combinations result in simulation more closely related to benchmarks, such as eddy covariance-derived GPP (Koven et al., 2020). More recently, researchers also explored the use of optimality approaches (e.g. Prentice et al., 2014) that rely on the photosynthetic least-cost theory (Wright et al., 2003) to infer $V_{c,\text{max}25}$ given ambient climate conditions (Wang et al., 2017; Smith et al., 2019; Jiang et al., 2020). These various approaches vary in their accuracy and remain limited to the estimation of $V_{c,\text{max}25}$ at the ecosystem-scale rather than at the individual organism level.

Alternatively, to capture the variability at the finer scale, $V_{c,\text{max}25}$ is often indirectly inferred through its empirical relationships with other leaf traits, such as LMA (Kattge et al., 2009; Walker et al., 2014), leaf N content per unit leaf area ($N_a$) (Kattge et al., 2009; Walker et al., 2014), chlorophyll (Chl) content (Croft et al., 2017), and leaf phosphorus (P) content (Walker et al., 2014; Norby et al., 2017). However, as shown by many field-based studies,
the empirical trait-$V_{c,\text{max}25}$ relationships can generate large uncertainties in characterizing $V_{c,\text{max}25}$ variability at the finer scale both within and across sites, because the trait-$V_{c,\text{max}25}$ relationships often vary with plant species, seasons, and plant growth environments (Domingues et al., 2010; Serbin et al., 2012; Prentice et al., 2014; Dechant et al., 2017; Norby et al., 2017; Rogers et al., 2017b). In addition, whether the trait-$V_{c,\text{max}25}$ relationships are robust across diverse PFTs remains unclear, and the combined measurements of both traits and $V_{c,\text{max}25}$ over large geographical areas for this evaluation remain lacking. An exploration of the potential breakdown of trait-$V_{c,\text{max}25}$ relationships across various scales from tree individuals to contrasting forest types as well as an accurate and scalable alternative for consistent estimation of $V_{c,\text{max}25}$ across scales remains critically needed.

Importantly, leaf reflectance spectroscopy, or the measurement of reflected radiance from leaves in many narrow, continuous spectral channels across a portion of the electromagnetic spectrum (Serbin & Townsend, 2020), may fill an important role in enhancing our understanding of trait variation within and across Earth’s terrestrial ecosystems. Leaf reflectance spectra are a collection of optical properties that are linked to a large number of leaf morphological and biochemical characteristics via electronic and vibrational absorption (Curran, 1989; Elvidge, 1990; Kokaly et al., 2009). Given the strong connection between traits and reflectance spectra, a range of studies have demonstrated that leaf reflectance spectra can accurately estimate a broad suite of plant traits (e.g. leaf water content (LWC), LMA, $N_a$, leaf Chl and P contents, and leaf age) with high precision and accuracy, even despite strong variation in traits due to biotic and abiotic factors (Serbin et al., 2014, 2019; Singh et al., 2015; Asner et al., 2016; Chavana-Bryant et al., 2017, 2019). Recent studies have also demonstrated that spectroscopy can accurately predict $V_{c,\text{max}25}$ variability at some specific ecosystems/PFTs, but present research has been limited to relatively narrow sampling of woody plants or crop species that are often found in a single ecosystem (Serbin et al., 2012; Dechant et al., 2017; Meacham-Hensold et al., 2019; Wu et al., 2019). This leaves a fundamental but unanswered question: Is the spectroscopic approach an accurate and scalable means for $V_{c,\text{max}25}$ estimation across various forest types spanning large environmental gradients?
The goal of this study was to evaluate the different approaches for predicting finer-scale $V_{c,\text{max}25}$ variability both within and across diverse forest types, as well as to explore the possibility of a cross-site approach for inferring $V_{c,\text{max}25}$ using leaf reflectance spectroscopy. Specifically, we addressed the following two questions: (1) Can empirical trait-$V_{c,\text{max}25}$ relationships hold up for predicting $V_{c,\text{max}25}$ across forest types? and (2) Can leaf spectroscopy offer an efficient alternative to predict $V_{c,\text{max}25}$ and represent trait-$V_{c,\text{max}25}$ relationships by accurately inferring both $V_{c,\text{max}25}$ and other traits from leaf reflectance spectra? To address these questions, we collected a set of leaf traits and reflectance spectra of canopy trees from the three forest types, namely a temperate broadleaved-coniferous forest, a subtropical evergreen broadleaved forest, and a tropical evergreen broadleaved forest. We focused on four key leaf traits (i.e. LMA, LWC, $N_a$, and leaf Chl content) that are of interest to the broad plant ecology community and have been shown to drive a large fraction of model simulation uncertainty in current ecosystem models (Wright et al., 2004; Diaz et al., 2016; Bonan & Doney, 2018; Ricciuto et al., 2018). Through answering these questions, we hope it can advise a practical approach to capture the $V_{c,\text{max}25}$ variability across various scales and improve the understanding of the relationships among leaf $V_{c,\text{max}25}$, traits, and spectra both within and across forest types.

Materials and Methods

Study sites and plant materials

This study was conducted at three forest sites that represent contrasting vegetation types in China, including Mountain Changbai (CB; 42°24' N, 128°06' E), Mountain Dinghu (DH; 23°10' N, 112°32' E), and Xishuangbanna (XSBN; 21°47' N, 101°03' E) (Fig. 1a). At each site, the Chinese Academy of Sciences (CAS) maintains a canopy crane facility enabling access to a 1-hectare area of each forest, with the crane tower height of 40 m in CB, 60 m in DH and 81 m in XSBN. We selected these sites for two reasons. First, they represent dominant forest types (i.e. temperate, subtropical and tropical) in China. These sites cover large gradients in mean annual temperature and precipitation (i.e. 2.8°C and 691 mm yr$^{-1}$ in CB, 20.9°C and 1927 mm yr$^{-1}$ in DH, and 21.8°C and 1493 mm yr$^{-1}$ in XSBN) (Ye et al., 2008; Shen et al., 2018; He NP et al., 2019) and soil types (i.e. from dark brown forest soil in CB to laterite soil in both DH and XSBN; see Table S1 for details on the variation...
in soil properties) (Cao et al., 2006; Wu et al., 2006; Gui et al., 2019). These ensure our collected leaf traits and reflectance spectra span sufficient ranges caused by plant species and growth environments. Second, these sites have canopy crane facilities that allow easy access to the abundant canopy tree species that are representative of each forest type. Specifically, the abundant canopy tree species include *Pinus koraiensis*, *Tilia amurensis*, *Juglans mandshurica*, *Ulmus davidiana*, *Fraxinus mandschurica*, and other interspersed deciduous species in CB (Wu et al., 2006; Liang & Liu, 2019); *Schima superba*, *Pinus massoniana*, *Castanopsis chinensis*, *Machilus breviflora*, and other interspersed evergreen broad-leaved species in DH (Ye et al., 2008; Gui et al., 2019); and *Parashorea chinensis*, *Canarium album*, *Pometia tomentosa*, *Sloanea tomentosa* and *Semecarpus reticulate* and other interspersed evergreen broad-leaves species in XSBN (Cao et al., 2006; Shen et al., 2018).

Within the forest canopy crane footprint, 19 trees from 7 dominant canopy tree species in CB, 31 trees from 12 dominant canopy tree species in DH, and 57 trees from 28 dominant canopy tree species in XSBN were selected (Table S2). Only upper canopy, sunlit leaves of these tree species were sampled and measured. Field measurements were conducted in the peak growing season (July-August) of 2019, including the measurements of leaf gas exchange, leaf reflectance spectra and four morphological and biochemical traits (i.e. LMA, LWC, $N_a$, and leaf Chl content). Details of measurement protocols are shown below and the results of these measurements are summarized in Table S2. It is noteworthy that even if field measurements were conducted in the peak growing season, the leaf samples were still mixed with different leaf ages, especially in the subtropical and tropical forests. We thus adopted the same approach as Wu et al. (2019) and classified the leaves into the two age categories (i.e. immature and mature leaves) according to the color, size and rigidity of the leaves.

**Field measurements**

*Leaf gas exchange and $V_{c,max25}$*

Three portable gas exchange systems (LI-6400XT; Li-COR Inc., Lincoln, NE, USA) were used for leaf gas exchange measurements. The response of net CO$_2$ assimilation rate ($A$) to
intracellular CO₂ concentration (C_i) (commonly known as the A-C_i curve; e.g. Fig. S1) was measured on leaves from cut branches, which were sampled before dawn with water cut to avoid inducing xylem embolism when collecting branches (Wu et al., 2019). The A-C_i curves were measured closely following the protocol of Rogers et al. (2017b). Also see Supplementary-Method S1 for details. For each tree, two branches were sampled and normally 1-3 leaves of all age classes (if present) per branch were measured. These field-derived A-C_i curves were then fit to a biochemical photosynthesis model (Farquhar et al., 1980), in which the modeled A best matched with the field-measured A, and thus the leaf maximum carboxylation capacity (V_c,max) was derived for each A-C_i curve. More details regarding the fitting procedure have been shown previously (Bernacchi et al., 2013; Rogers et al., 2017b; Wu et al., 2019). We used the same code developed in MATLAB (Mathworks, Natick, MA, USA) by Wu et al. (2019) for the A-C_i curve fitting, with example demonstrations shown in Fig. S1. Finally, the same kinetic constants and temperature response functions as Bernacchi et al. (2013) were used to standardize V_c,max to a reference temperature of 25°C (V_c,max25). Admittedly, there is also an alternative approach for deriving V_c,max25 that accounts for the environmental acclimation of photosynthetic machinery (Kumarathunge et al., 2019). We thus cross-compared the two approaches, and found that V_c,max25 derived using the two approaches was almost identical (Fig. S2). Since Bernacchi et al. (2013) has been widely used in previous trait-V_c,max25 and spectral-V_c,max25 analysis (e.g. Walker et al., 2014; Rogers et al., 2017b; Scafaro et al., 2017; Wu et al., 2019), we present our results based on that approach.

Leaf reflectance spectra

Upon finishing gas exchange measurements, leaves were immediately measured for leaf reflectance spectra. A portable spectroradiometer Spectra Vista Corporation (SVC) HR-1024i (SVC, Poughkeepsie, NY, USA; spectral full-range: 350-2500 nm; spectral resolution: ≤3.3 nm at 700 nm, ≤9.5 nm at 1500 nm, and ≤6.5 nm at 2100 nm; sampling internal: linearly interpolated to 1 nm; Ely et al., 2019) together with the SVC LC-RP-Pro foreoptic was used to measure leaf reflectance. During the spectral collection, the leaf reflectance probe used the internal calibrated light source to illuminate the samples with a black background, following the protocol of Wu et al. (2019). A 99% reflective Spectralon
white panel by Labsphere Inc. (North Dutton, NH, USA) was used as the reference standard. The SVC was set to 1s collection time (i.e. the integral scanning time for each spectrum collected was 1 second) with the spectrometer’s automatic integration optimization, in order to avoid the impacts of excessive heat loads on the data quality of leaf spectra collected (Serbin et al., 2019). For each leaf, we measured the spectra on 3-6 different parts of the leaf adaxial surface (with only one spectrum of each part) depending on the leaf size, and used the averaged spectra to indicate the optical properties across all wavelengths. Additionally, we corrected the discontinuities in the spectra in the detector overlap areas using the vendor-provided SVC instrument software (Ely et al., 2019). The examples of collected leaf reflectance spectra are shown in Fig. S3 for demonstration.

Leaf biochemical and morphological traits

After leaf spectral measurements, fresh leaf chlorophyll (Chl) content was estimated with a portable optical chlorophyll meter (SPAD-502 Plus; Konika-Minolta, Inc., Tokyo, Japan). The SPAD value is an index based on the absorbance of the leaf at 650 and 940 nm (Uddling et al., 2007), which has been shown as a good proxy for leaf Chl content in both crops and woody plants (Uddling et al., 2007; Coste et al., 2010; Silva-Perez et al., 2018). We also compared the SPAD-based leaf Chl content with that inversed from a leaf-level radiative transfer model PROSPECT-5 with field-based spectral measurement as the model input, following the same approach as shown in Jacquemoud et al. (2009). These two proxies of leaf Chl content showed tight relationships ($R^2=0.75$; $p<0.001$; Fig. S4), providing additional confidence on the use of the SPAD value to approximate fresh leaf Chl content in our leaf samples. We thus used Chl$_{SPAD}$ (unitless) that is short for SPAD-based leaf Chl content in this study.

Additionally, other three leaf traits were also derived, including LWC ($g$ $g^{-1}$), LMA ($g$ $m^{-2}$) and area-based leaf N content ($N_a$; $g$ $m^{-2}$). Leaf area was measured using the cork borers, and leaf fresh mass was weighted using a precision balance (precision at 0.001g; Meilen; Meifu Electronics Co. Ltd, Shenzhen, China) in a fixed location to reduce the impact of air movements. After that, the samples were oven dried to constant mass at 65°C for over 72 h, and then the dry mass was determined with the precision balance. LWC was calculated
using the difference between leaf fresh and dry mass divided by leaf fresh mass. LMA was calculated using the leaf dry mass divided by leaf area. Dried leaves were then ground using a ball mill (NM200, Retsch, Haan, Germany) before measuring leaf N content. Mass-based leaf N content was analyzed by the Dumas combustion method using an elemental analyzer (Eurovector EA3028) coupled to a stable isotope ratio mass spectrometer (Nu Instruments Perspective) in continuous flow mode (EAIRMS). Area-based leaf N content was then calculated with the mass-based leaf N content multiplied by LMA. In terrestrial biosphere modelling researches, leaf N and $V_{c,\text{max}}$ are generally expressed on an area basis given that the main function of leaves is to intercept light (Osnas et al., 2013; Walker et al., 2014). Thus, we restricted our following analyses to area-based leaf trait measurements.

**Data analysis**

In this study, we focused on exploring the relationships of $V_{c,\text{max}}$ with the four leaf traits and leaf spectra. Thus, we chose leaf samples including all these measurements across the three forest sites. As a result, the dataset used for the analyses had 72 measurements in CB, 91 measurements in DH, and 173 measurements in XSBN, respectively. The species-specific mean and range of $V_{c,\text{max}}$ for all leaves across the three forest sites are presented in Table S2. We noted that our leaf samples from the tropical (XSBN) and subtropical (DH) forest types were mixed with different leaf age classes, with mature leaves ($n=281$) dominating over young leaves ($n=55$). In order to examine whether leaf age would affect the trait-$V_{c,\text{max}}$ relationships, we performed a sensitivity analysis on the full dataset that respectively included (i.e. mature leaves only) or excluded (i.e. all leaves) the leaf age control. Our sensitivity analysis demonstrated that the results remained consistent regardless of with or without leaf age control (Figs S5, S6; Table S3). For clarity, we primarily focused on presenting the data analyses for the entire dataset without leaf age control thereafter.

**Exploring trait-$V_{c,\text{max}}$ relationships within and across forest types**

To explore the relationships between the four leaf traits and $V_{c,\text{max}}$ within and across forest types, we performed the three analyses. First, to explore the separate contribution of each trait on $V_{c,\text{max}}$ prediction, we analyzed the relationships between $V_{c,\text{max}}$ and each of the
four traits (i.e. LMA, LWC, ChlSPAD and Na) using ordinary least-squares regression within
and across forest types. Second, to explore the joint contribution of all the four traits on
$V_{c,max25}$ prediction, we performed a multiple linear regression analysis both within and
across forest types. Third, to rank the independent, relative importance of each trait on
$V_{c,max25}$ prediction, we performed a principal component analysis (PCA) on all traits and
$V_{c,max25}$ with data from all forest types using the R package ‘prcomp’. All trait data were
standardized using the zero-mean approach before the PCA analysis. The first two PCA
axes were then plotted to visualize the trait-space.

Developing spectral models of $V_{c,max25}$ and the four leaf traits

The partial least square regression (PLSR) approach (Serbin et al., 2014) was used here to
model the $V_{c,max25}$ and the four traits (LWC, LMA, Na and ChlSPAD) from leaf spectra within
and across forest types. This PLSR method is commonly used in spectroscopic and
chemometric analyses and has been shown superior to those classical regression
approaches, given that it has the ability to handle high predictor collinearity and allow the
number of predictor variables to be much higher than the number of observations (Ollinger
& Smith, 2005; Asner et al., 2016; Serbin et al., 2014, 2019).

We adopted the same PLSR modelling approach as has been developed previously
(Dechant et al., 2017), in which the standard PLSR regression analysis was integrated with
a repeated double cross-validation (rdCV) method (Filzmoser et al., 2009). rdCV first splits
the full dataset into calibration and independent validation subsets repeatedly using a cross-
validation (outer CV loop), and then performed additional splits of each calibration subset
into training and testing components using a cross-validation procedure (inner CV loop) to
ensure that the optimal number of latent variables could be determined independently from
the actual performance evaluation. Prediction performance is evaluated on the independent
validation subsets (outer loop) over many possible random splits ($n=200$ in our case). This
method thus has the advantage to avoid the occurrence of good or bad results purely by
chance.
Given that the model detail and settings have been shown in Dechant et al. (2017), we briefly summarized the five major steps of this approach below. First, we performed a square-root-transformation on $V_{c,max25}$ and the four traits to reduce the right-skewed frequency distribution of the original data for the PLSR analysis (Serbin et al., 2019; Wu et al., 2019). Second, we performed 200 repetitions for the rdCV and ten-fold cross-validation for both the outer and inner CV loops (Filzmoser et al., 2009; Dechant et al., 2017). Third, to avoid over-fitting, we determined the optimal latent variable number by maximizing the averaged coefficient of determination ($R^2$) and minimizing the averaged root mean squares of error (RMSE) of the inner CV loop (Fig. S7). With the optimal PLSR latent variable number, we further derived the regression coefficients and variable importance in projection (VIP) metric (Wold et al., 2001) for each permutation in the inner loop, and then averaged them to obtain the mean regression coefficients and VIP spectrum for each repetition. Fourth, we presented the mean and 95% confidence interval of PLSR regression coefficients and VIP metric generated by the 200 repetitions in the outer loop. This ensemble of PLSR model coefficients were ultimately used as the final model to predict $V_{c,max25}$ or each trait plus the 95% confidence interval over the 200 repetitions. Fifth, the performance of the final PLSR model was evaluated using the independent validation subset for each of the 200 repetitions in the outer loop, measured by both $R^2$ and RMSE (Fig. S8). For both $V_{c,max25}$ and the four traits, all results were presented in the original units rather than the square-root-transformed units as the initial output of the PLSR model.

Exploring spectral-$V_{c,max25}$ relationships under three spectral modeling scenarios

We performed the PLSR modeling of the three scenarios, with an aim to explore whether the spectral-$V_{c,max25}$ relationship can be extended to the cross-site scale. These three scenarios are summarized in Table 1 and illustrated as follows. First, a ‘site-specific’ spectral model of $V_{c,max25}$ was developed and evaluated using all the data from each forest site. Second, a ‘XSBN’ spectral model of $V_{c,max25}$ was developed and evaluated using all the data from the tropical forest site of XSBN only, and then the developed model was applied to the independent data from CB and DH. Third, a “cross-site” spectral model of $V_{c,max25}$ was developed and evaluated using the data from all three forest sites. By performing these modeling scenarios, it would not only improve our understanding of
cross-site generalizability of spectral-$V_{c,\text{max}25}$ relationships, but also help identify the potential scenarios in which spectra-$V_{c,\text{max}25}$ relationships might break down.

**Results**

**Weak and decoupled relationships between $V_{c,\text{max}25}$ and leaf traits across forest types**

To investigate our first question—trait-based approaches for $V_{c,\text{max}25}$ predictions, we analyzed the distribution of each leaf trait and $V_{c,\text{max}25}$ (Fig. 1b-f) and then explored the trait-$V_{c,\text{max}25}$ relationships within and across forest types (Fig. 2). We found that the leaf morphological and biochemical traits and $V_{c,\text{max}25}$ displayed a high degree of variation within and across forest types, with the tropical forest (XSBN) showing the largest spread in values, followed by the subtropical (DH) and temperate (CB) forests (Fig. 1b-f). We also found that the trait ranges of XSBN covered the full ranges of DH and CB in all traits except $V_{c,\text{max}25}$ (Fig. 1b-f), and thus XSBN had a significantly lower $V_{c,\text{max}25}$ than CB ($t$=19.85, $p<0.001$; two-tailed Student’s t-test) and DH ($t$=4.84, $p<0.001$; two-tailed Student’s t-test).

Exploring the trait-$V_{c,\text{max}25}$ relationships further, we found that these relationships were weak overall with a high degree of variation in forest type-specific slope and intercept of the relationships (Fig. 2a-d; Table S3). Among these relationships, $N_a$ showed the highest $V_{c,\text{max}25}$ prediction across forest types ($R^2=0.13$, $p<0.001$), followed by ChlSPAD ($R^2=0.05$, $p<0.001$), LWC ($R^2=0.01$, $p=0.030$), and LMA ($R^2=0.01$, $p=0.179$). PCA confirmed these patterns and illustrated that leaf N content had the strongest relationship with $V_{c,\text{max}25}$, followed by ChlSPAD (Fig. 3). LMA and LWC were near-orthogonal to $V_{c,\text{max}25}$ (Fig. 3), suggesting no clear relationships between these two traits and $V_{c,\text{max}25}$. Besides the univariate exploration, we also used a multiple linear regression model for explaining variation in $V_{c,\text{max}25}$ (Fig. 2e). The results showed that these traits together explained slightly higher degrees of variation in $V_{c,\text{max}25}$ across all the three contrasting forest types ($R^2=0.15$; Fig. 2e) than models that relied on the univariate exploration (Fig. 2a-d). The results also showed higher predictive power of the multivariate model within each forest type, with $R^2$ of 0.34, 0.30 and 0.19 for CB, DH and XSBN, respectively (Fig. 2e). In addition, our results showed that the trait-based model when including the site effect (i.e. site-specific trait-
based model) had the predictive power of $R^2=0.61$, which was much higher than its correspondence using the cross-site trait-based model ($R^2=0.15$) (Fig. 2e).

**Spectroscopy outperforms leaf trait relationships for predicting $V_{c,max25}$ across forest types**

To investigate our second question—spectral-based approach for reconciling the lower performance of the trait-based approach in explaining cross-site $V_{c,max25}$ variability, we analyzed the relationships between leaf traits/$V_{c,max25}$ and spectra. We found that the cross-site spectral models accurately capture the variations in all traits, including $V_{c,max25}$ ($R^2=0.77$, RMSE=9.7 μmol CO$_2$ m$^{-2}$ s$^{-1}$), $N_a$ ($R^2=0.79$, RMSE=0.19 g m$^{-2}$), Chl$_{SPAD}$ ($R^2=0.92$, RMSE=3.2), LMA ($R^2=0.93$, RMSE=6.6 g m$^{-2}$), and LWC ($R^2=0.93$, RMSE=0.021) (Figs 4, S9). These results demonstrated that leaf spectroscopy accurately predicted not only the four leaf biochemical and morphological traits but also the physiological trait of $V_{c,max25}$ across the three contrasting forest types.

To explore the spectral-$V_{c,max25}$ relationships further, we conducted two additional tests, in which the spectral model of $V_{c,max25}$ was calibrated for 1) each forest site, and 2) tropical forest site only (XSBN). The site-specific spectral model of $V_{c,max25}$ had the comparable predictive power with the cross-site model ($R^2=0.76$ in Fig. 5a vs. $R^2=0.77$ in Fig. 4). The XSBN model had the worst performance ($R^2=0.66$ in Fig. 5b), with a significant model bias when applied to the other two forest sites that were not involved in the model built (Fig. 5b; Table 1). Further, we found the cross-site model yielded the lowest prediction uncertainty as indicated by the horizontal error bars (Fig. 4 vs. Fig. 5). These results collectively demonstrated that an accurate and cross-site scalable spectra-$V_{c,max25}$ relationship could be derived when sufficient leaf samples were involved in the model development.

**Discussion**

This study has two main findings. First, we demonstrated that there were weak, forest type-specific trait-$V_{c,max25}$ relationships (Figs 1, 2). Second, we revealed that leaf spectroscopy collapsed this variability into a single spectral model that accurately predicted $V_{c,max25}$ both
within and across forest types (Fig. 4). Collectively, our study suggests that leaf spectroscopy outperforms trait-$V_{c,\text{max}25}$ relationships for predicting leaf photosynthetic capacity across three contrasting forest types in China.

Our observations of weak and forest type-specific trait-$V_{c,\text{max}25}$ relationships are consistent with many previous field-based studies, which also report moderate-to-weak trait-$V_{c,\text{max}25}$ relationships at the inter-species, -site and -PFT level (Domingues et al., 2010; Serbin et al., 2012; Prentice et al., 2014; Walker et al., 2014; Rogers et al., 2017b). Meanwhile, our observed forest type-specific $N_a-V_{c,\text{max}25}$ relationship also agrees with a previous global-scale synthesis study, which shows that the slopes and intercepts of $N_a-V_{c,\text{max}25}$ relationships vary with PFTs (i.e. tropical trees, temperate broadleaved trees and coniferous trees) at the global scale (Kattge et al., 2009). There are several possible explanations for this including variation in the fraction of $N_a$ allocated to Rubisco and variation in mesophyll conductance which both affect the ratio between $N_a$ and $V_{c,\text{max}25}$ (Ghimire et al., 2017; Evans, 2000). As a result, the $N_a-V_{c,\text{max}25}$ relationship has been observed to vary with species, growth environments, and PFTs, with $N_a$ alone explaining a relatively low degree of $V_{c,\text{max}25}$ variance at the global scale (Walker et al., 2014; Ali et al., 2015; Smith et al., 2019).

Despite the weak $N_a-V_{c,\text{max}25}$ relationship observed here and previously, this relationship remains widely used in TBMs to first infer $V_{c,\text{max}25}$ and then to simulate plant photosynthesis (Kattge et al., 2009; Rogers et al., 2017a). Our results, together with these previous studies, thus raise concerns over implementing $N_a-V_{c,\text{max}}$ relationships in TBMs for modelling large-scale terrestrial photosynthesis and associated vegetative responses to the current and changing climate.

In addition to the weak $N_a-V_{c,\text{max}25}$ relationships, we observed the similarly weak and forest type-specific Chl-$V_{c,\text{max}25}$ relationships (using either Chl$_{\text{SPAD}}$ or PROSPECT-5 inverted Chl; Figs 2b, S4b). Our result is consistent with Luo et al. (2019) who also found the PFT-dependent Chl-$V_{c,\text{max}25}$ relationships based on previous studies with direct field measurements of Chl and $V_{c,\text{max}25}$ across PFTs and use these PFT-specific relationships, together with satellite-derived leaf chlorophyll content, to estimate $V_{c,\text{max}25}$ at the global-scale. Taken together, these results suggests that it is important to account for PFT-specific
Chl-$V_{c,max25}$ relationships when leveraging field- and/or satellite-derived Chl measurements for large-scale $V_{c,max25}$ prediction. Admittedly, both Chl$_{SPAD}$ and PROSPECT-5 inverted Chl used in this study are indirect Chl measurements, and thus our Chl$_{SPAD}$-$V_{c,max25}$ relationship might not be directly applicable to those Chl-$V_{c,max25}$ studies relying on direct Chl measurements (e.g. Croft et al., 2017; Chou et al., 2020). However, our results suggest that global-scale studies in which satellite-derived leaf Chl content (comparable with our PROSPECT-5 inverted Chl) is used in combination with ground-derived Chl-$V_{c,max25}$ relationships may mask marked variation in $V_{c,max25}$ within and among forest types (Alton, 2017; Luo et al., 2019; Croft et al., 2020). We recommend that direct Chl measurements, satellite-derived Chl measurements, combined with $V_{c,max25}$ and spectra data, across different PFTs are still needed to confirm our observations which relied on the two indirect Chl metrics.

If empirical trait-$V_{c,max25}$ relationships do not hold up for accurate and scalable predictions of $V_{c,max25}$ across forest types, what is an alternative? Our answer is leaf spectroscopy, as a single spectral model accurately predicted $V_{c,max25}$ across all three forest types ($R^2=0.77$; Fig. 4). To the best of our knowledge, this represents the first study demonstrating the cross-site scalability of $V_{c,max25}$ prediction using leaf spectroscopy. Also because Wu et al. (2019) demonstrated leaf spectroscopy as an accurate means for characterizing the seasonal variability in $V_{c,max25}$ across multiple tropical forests, leaf spectroscopy can thus likely be an accurate and scalable means for predicting $V_{c,max25}$ across both large spatial and temporal extents. Additionally, our results of the cross-site tight covariations between leaf spectra and the four traits (Fig. S9) also confirm the two very recent studies in which leaf spectroscopy accurately infers key leaf biochemical and morphological traits (e.g. $N_a$, LMA, leaf carbon content, and leaf total phenol content) across diverse forest types (Nakaji et al., 2019; Serbin et al., 2019). These together further suggest that leaf spectroscopy offers a promising and scalable means for monitoring leaf traits across various forest types, and that the measurement of leaf spectra alone together with a suitable PLSR model can accurately infer multiple leaf morphological, biochemical and physiological traits.
Our results further demonstrate that the cross-site scalable spectral-$V_{c,\text{max}25}$ relationship can be derived only when leaf samples covering sufficiently wide variability in both $V_{c,\text{max}25}$ and spectra are involved in the model development (Figs 4, 5; Schweiger, 2020). Thus, the prerequisite for developing a general and broadly applicable spectral model of $V_{c,\text{max}25}$ is to encompass the full trait space associated with both physiological and spectral variation in the training dataset. Many observational studies have suggested that both leaf spectra and $V_{c,\text{max}25}$ vary considerably among ecosystems (e.g. grasslands, shrublands, and forests), climate zones (e.g. boreal, temperate, subtropical and tropical zones), life forms (e.g. evergreen and deciduous trees, coniferous and broadleaved trees), leaf ages (e.g. young, mature and old leaves) and continents (e.g. China, Europe and North America) (Kattge et al., 2009; Ali et al., 2015; Dechant et al., 2017; Smith & Dukes, 2018; Albert et al., 2018; Wang et al., 2018; Serbin et al., 2019; Wu et al., 2019). However, most of these variabilities are currently either under or not sampled in our study. Therefore, further global sampling and analysis is still needed to explore and build a globally applicable spectral model of $V_{c,\text{max}25}$ as that has been developed for LMA (Serbin et al., 2019).

To elucidate the potential mechanism of spectral-$V_{c,\text{max}25}$ relationships, we explored the spectral regions that dominate the links between spectra and $V_{c,\text{max}25}$ by analyzing the patterns in the PLSR regression coefficients (Fig. 6a) and variable importance in projection (VIP; Fig. 6b). Our results show that the reflectance contribution to the estimation of $V_{c,\text{max}25}$ varies considerably across the full spectrum (Fig. 6a,b). Specifically, we found that the spectral domains of importance (i.e. VIP value>0.8; Wold et al., 2001) cover the visible range (450-700 nm), a strong response in the red edge range (700-800 nm), and multiple near infrared (NIR) (1030~1120 nm) and shortwave infrared (SWIR) bands (1350-1480 nm, 1640~1760 nm, and 1820~2180 nm). These identified $V_{c,\text{max}25}$-sensitive spectral domains not only agree closely with several previous similar efforts that analyze the spectral-$V_{c,\text{max}25}$ relationships (Serbin et al., 2012; Meacham-Hensold et al., 2019; Wang et al., 2020), but are also consistent with many other studies that broadly examine the relationship between reflectance spectra and leaf biochemistry and physiology (Curran, 1989; Elvidge, 1990; Kokaly et al., 2009; Ustin et al., 2009). For example, the visible range has been related to the concentration of leaf pigments (e.g. chlorophyll, carotenoids, and
anthocyanins; Wang et al., 2020; also see Fig. 6e,f), the red-edge range has been associated
with leaf Chl content (Fig. 6e,f), chlorophyll fluorescence emission, and photosystem II
function (Zarco-Tejada et al., 2000; Serbin et al., 2012; Wang et al., 2020), and the detected
NIR and SWIR bands have been found tightly connected with leaf N and protein contents
(Curran, 1989; Kokaly, 2009; also see Fig. 6c,d) as well as leaf water and dry matter
contents (Elvidge, 1990; Jacquemoud et al., 2009; also see Fig. 6g,h,i,j). These observed
\( V_{c,\text{max}25} \)-sensitive spectral regions, together with the previous understanding of trait-specific
sensitive spectral regions, thus support the trait-based hypothesis suggested by Wu et al.
(2019) that \( V_{c,\text{max}25} \) is tightly connected with multiple leaf traits retrievable from leaf spectra
as a candidate mechanism for the observed tight spectral-\( V_{c,\text{max}25} \) relationship (Fig. 4).

The four leaf traits measured here explain a much smaller fraction of the variation in
\( V_{c,\text{max}25} \) than the spectral model (15% vs. 77%; Fig. 2e vs. Fig. 4). Therefore, our results
suggest that there are likely other un-measured traits that aid in the determination of the
cross-site variability in \( V_{c,\text{max}25} \), and that information about those traits is present in the leaf
spectra. These un-measured traits or states might include leaf P content, leaf magnesium
content, leaf temperature, and leaf age (Walker et al., 2014; Asner et al., 2016; Wu et al.,
2019; Khan et al., 2020; Wang et al., 2020), but the direct quantitative evidence is currently
missing. In addition, we observed that accounting for site-specific trait-\( V_{c,\text{max}25} \)
relationships largely improved \( V_{c,\text{max}25} \) prediction \( (R^2=0.15 \text{ vs. } 0.61; \text{ Fig. 2e}) \) while there
was virtually no change of spectral-\( V_{c,\text{max}25} \) relationships between the cross-site model and
site-specific model \( (R^2=0.77 \text{ vs. } 0.76; \text{ Fig. 4 and Fig. 5a}) \). These results further suggest that
leaf spectra are able to capture the site-specific trait-\( V_{c,\text{max}25} \) relationships, and thus a single
spectral model consistently and accurately predicts \( V_{c,\text{max}25} \) across the three studied forest
types. Site-specific trait-\( V_{c,\text{max}25} \) relationships have often been attributed to environmental
acclimation of \( V_{c,\text{max}25} \) to growth temperature and other abiotic conditions (e.g. light, water,
and soil properties and nutrients) (Kattge et al., 2009; Prentice et al., 2014; Kumarathunge
et al., 2019; Smith et al., 2019), yet leaf spectra have also been shown to capture or rapidly
respond to changes in growth environment (Serbin et al., 2012; Khan et al., 2020). This
suggests that spectra can adapt to the underlying trait modifications in response to their
environmental acclimation over both short- and long-term environmental changes.
Therefore, it provides additional evidence that reflectance spectra can capture the site-specific variation in trait-$V_{c,max25}$ relationships associated with environmental acclimation.

Regardless, an in-depth exploration of the spectral-$V_{c,max25}$ relationships using experimental manipulations and measurements of more leaf traits, $V_{c,max25}$, and leaf spectra at both spatial and temporal (e.g. leaf age or seasonal) scales is still needed to fully elucidate the underlying mechanism that enables successful prediction of $V_{c,max25}$ from leaf spectra.

With these findings, our work also generates at least two implications as follows. First, our findings can complement current trait-based plant ecology studies. Empirical leaf trait relationships, regardless of their predictive strength, have been extensively used for reducing trait dimension and understanding plant adaptive strategies in functional ecology studies (Wright et al., 2004; Diaz et al., 2016). Our findings of the inconsistent leaf trait relationships across forest types suggest that the canonical leaf trait coordination theory (e.g. leaf economics spectrum that emphasizes the convergent leaf trait relationships) (Wright et al., 2004; Diaz et al., 2016) works at the global scale but breaks down at finer scales (Niinemets et al., 2015; Anderegg et al., 2018; Shiklomanov et al., 2018). Since leaf spectroscopy is able to accurately derive multiple leaf traits (Figs. 4, S9) and leaf trait relationships (Fig. S10), it thus offers a new way to characterize multi-trait variability and improve representations of leaf trait relationships (including those weak relationships) in TBMs and functional ecology studies.

Second, our findings also have important implications for future work that aims to characterize the $V_{c,max25}$ variability over large scales by leveraging leaf spectroscopy techniques. Our work suggests that leaf spectroscopic approaches can provide accurate, rapid, relatively low-cost and nondestructive estimates of $V_{c,max25}$ across diverse plant species and forest types, which can facilitate the broader characterization of $V_{c,max25}$ variability that is useful for ecological research and process modeling. Moving up in scale from leaves to landscapes, as well as whole-ecosystems, using remote sensing data depends on the effective extension of results like those presented here to vegetation canopies. Past research in agricultural landscapes suggests this scaling up is possible using similar approaches (Serbin et al., 2015), however additional work is needed to develop generalized,
robust methods. Importantly, canopy spectral variability is fundamentally tied to both leaf
spectra and canopy structural attributes (e.g. leaf area index and leaf angle distribution)
(Asner, 1998; Roberts et al., 2004; Ollinger, 2011). At the same time, other challenges to
spaceborne retrieval of $V_{c,\text{max}25}$ are associated with a multitude of issues, including sensor
design, uncertainties in the retrieval of surface reflectance, the sun-sensor geometry effect,
and the mixture effect associated with the spatial resolution issue (Roberts et al., 2004;
Thompson et al., 2019; Cavender-Bares et al., 2020). Therefore, additional research is
needed to understand the impacts of these issues on satellite retrievals of $V_{c,\text{max}25}$, yet new
opportunities in spaceborne image spectroscopy could yield new insights (Guanter et al.,
2015; Stavros et al., 2017; Schimel & Poulter, 2021). In addition, other novel platforms
including, unoccupied aerial systems, may also provide new opportunities for developing
large-scale maps of $V_{c,\text{max}25}$ (Singh et al., 2015; Asner et al., 2016).

Data availability statement
The data that support the findings of this study are available from the corresponding author
upon reasonable request.

Author contributions
J.W., Z.Y., S.P.S, A.R. and L.L. conceived the project idea; J.W. and Z.Y. designed the
study; Z.Y., Z.G., G. S., Y. Z., Y. C., X. W., J. L., B. W., Y. W., and J.W. collected the
field data; Z.Y., Z.G. and J.W. performed the synthetic data analysis; S.P.S., S.Y., W.H.,
A.R. and L.L. contributed to the result interpretation; Z.Y. and J.W. drafted the manuscript
and all authors contributed to the final version.

Acknowledgements
J.Wu, Z.Yan, Z.Guo, G.Song, Y.Zhao, Y.Chen, S.Wu, and J.Wang were supported by
National Natural Science Foundation of China (#31922090), National Natural Science
Foundation of China (#31901086), the Research Grants Council Early Career Scheme
(#27306020) and Seed Fund for Basic Research (#201905159005). Z.Yan was in part
supported by the Division of Ecology and Biodiversity PDF research award. A.Rogers and
S.P.Serbin were supported by the Next-Generation Ecosystem Experiments (NGEE
Tropics) project that is 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 (#DE-SC0012704) to Brookhaven National Laboratory. The authors would like to thank National Forest Ecosystem Research Station at Xishuangbanna, Dr. Guanhua Dai and Mr. Jinlong Dong for canopy crane access assistance, thank Mrs. Leung Kit Sum for providing help in laboratory chemical analyses of leaf samples, and thank the editor and the three anonymous reviewers for providing valuable suggestions and comments, which are greatly helpful in improving this manuscript.
Reference


Camino C, Gonzalez-Dugo V, Hernandez P, Zarco-Tejada PJ. 2019. Radiative transfer $V_{cmax}$ estimation from hyperspectral imagery and SIF retrievals to assess
photosynthetic performance in rainfed and irrigated plant phenotyping trials.


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

Fig. S1 Example of demonstration of the response of net CO₂ assimilation rate \( (A) \) to intracellular CO₂ concentration \( (C_i) \) for the typical species from the three forest types.

Fig. S2 The \( V_{c, \text{max}25} \) derived using the temperature response functions as Bernacchi et al. (2013) is highly consistent with that derived using the site-specific temperature response functions as Kumarathunge et al. (2019).

Fig. S3 Example of demonstration of leaf reflectance spectra for the typical species from the three forest types.

Fig. S4 Relationships of PROSPECT-5 inverted leaf chlorophyll (Chl) content with Chl SPAD and \( V_{c, \text{max}25} \).

Fig. S5 Histogram distributions of leaf traits under the controlled leaf age scenario (i.e. mature leaves only).

Fig. S6 Exploring the trait-\( V_{c, \text{max}25} \) relationships both within and across forest types under the mature leaf age scenario.

Fig. S7 The selection criterion for the optimal number of latent variables in the spectral model of \( V_{c, \text{max}25} \) under the cross-site scenario using partial least squares regression (PLSR).

Fig. S8 Histogram distribution of the coefficient of determination \( (R^2) \) for the PLSR spectral models over the 200 repetitions under the cross-site scenario.

Fig. S9 Accuracy assessments for the cross-site spectral models of leaf morphological and biochemical traits.

Fig. S10 Comparisons between the observed and spectral-modelled trait-\( V_{c, \text{max}25} \) relationships across the three forest types.

Table S1 The soil information of the three forest types in China.

Table S2 Summary of species, leaf traits and sample size of representative canopy trees across the three forest types in China.

Table S3 Statistical summary of the relationships between \( V_{c, \text{max}25} \) and the four leaf morphological and biochemical traits across the three forest types.

Supporting Method S1 The protocol for the \( A-C_i \) curve measurement.
Table 1 Evaluation of the spectral-$V_{c,\text{max}25}$ model under the three scenarios: the ‘site-specific model’ (trained and evaluated using the data from each forest site through the repeated double cross-validation (rdCV) method), ‘XSBN model’ (trained and evaluated using the data from the tropical forest site of XSBN through the rdCV method, and then applied the model to the independent sites at CB and DH), and ‘cross-site model’ (trained and evaluated using the data from all the three forest sites through the rdCV method). Leaf level measurements across all leaf ages were analyzed here. Abbreviations: $n$, sample size; $R^2$, the coefficient of determination; RMSE, the root mean square of error.

<table>
<thead>
<tr>
<th>Scenarios</th>
<th>Site</th>
<th>$n$</th>
<th>$R^2$</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site-specific model</td>
<td>CB</td>
<td>72</td>
<td>0.28</td>
<td>9.9</td>
</tr>
<tr>
<td></td>
<td>DH</td>
<td>91</td>
<td>0.55</td>
<td>9.5</td>
</tr>
<tr>
<td></td>
<td>XSBN</td>
<td>173</td>
<td>0.59</td>
<td>10.3</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>336</td>
<td>0.76</td>
<td>10.0</td>
</tr>
<tr>
<td>XSBN model</td>
<td>CB</td>
<td>72</td>
<td>0.20</td>
<td>10.4</td>
</tr>
<tr>
<td></td>
<td>DH</td>
<td>91</td>
<td>0.48</td>
<td>10.2</td>
</tr>
<tr>
<td></td>
<td>XSBN</td>
<td>173</td>
<td>0.59</td>
<td>10.3</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>336</td>
<td>0.66</td>
<td>11.9</td>
</tr>
<tr>
<td>Cross-site model</td>
<td>CB</td>
<td>72</td>
<td>0.35</td>
<td>9.4</td>
</tr>
<tr>
<td></td>
<td>DH</td>
<td>91</td>
<td>0.56</td>
<td>9.4</td>
</tr>
<tr>
<td></td>
<td>XSBN</td>
<td>173</td>
<td>0.63</td>
<td>9.8</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>336</td>
<td>0.77</td>
<td>9.7</td>
</tr>
</tbody>
</table>

Note: The three forest sites include the temperate forest in Mountain Changbai (CB), the subtropical forest in Mountain Dinghu (DH), and the tropical forest in Xishuangbanna (XSBN).
Fig. 1 Sites and histogram distributions of leaf traits across the three forest types in China. (a) Location of the three canopy crane sites that span a large latitudinal gradient, including a temperate forest in Mountain Changbai (CB), a subtropical forest in Mountain Dinghu (DH), and a tropical rainforest in Xishuangbanna (XSBN). The background shows a map of mean enhanced vegetation index (EVI) in the 2019 peak growing season (July-August) derived from the MOD13A1 products. (b-f) Site-specific histogram distributions of maximum carboxylation rate of Rubisco standardized to 25°C ($V_{c,max25}$) and the other four leaf traits (i.e. leaf nitrogen (N) content, leaf chlorophyll content (ChlSPAD), leaf mass per area (LMA), leaf water content (LWC)), with CB in red, DH in green, and XSBN in blue. The PDF shown in the figure legend refers to the probability distribution function (PDF) that was used to fit the site-specific histogram distribution. Numbers in panel (b-f) are the mean ± standard deviation of the trait values for each forest site. $V_{c,max25}$ and leaf traits were sampled in the 2019 peak growing season, and trait values of individual tree means from all leaf ages were presented here.
Fig. 2 Exploring trait-$V_{c,\text{max}25}$ relationships both within and across forest types. (a-d) Pairwise relationship between leaf $V_{c,\text{max}25}$ and the other four leaf traits (i.e. leaf N content, Chl$_{\text{SPAD}}$, LMA, LWC) within and across the three forest types; (e) performance (observed vs. predicted $V_{c,\text{max}25}$) of the multiple linear regression model using all four leaf traits as predictor variables. Three different color circles represent each of the three forest sites, with CB in red, DH in green, and XSBN in blue. Lines were fitted by ordinary least-squares regressions, with color lines for site-specific regression fitting and the gray line for all-sites’ regression fitting. In panel (e), dash line represents the 1:1 line, and the black line represents the regression fitting when combining the site-specific multiple linear regression model results. $R^2$ represents the coefficient of determination; ‘ns’ denotes the insignificant relationship with $p>0.05$. Leaf level trait measurements across all leaf ages were involved in this analysis.
Fig. 3 Exploring the cross-site trait-$V_{c,\text{max}25}$ relationships using principal component analysis. The trait loading on the plane below is defined by principal component 1 (PC1) and principal component 2 (PC2), with the green arrow for $V_{c,\text{max}25}$ and brown arrows for the other four leaf traits (leaf N content, Chl$_{SPAD}$, LMA, and LWC). PC1 and PC2 respectively explained 49% and 22% of total variation in all four traits and $V_{c,\text{max}25}$. The angle between $V_{c,\text{max}25}$ and each of the four traits indicates the extent of correlation between them, with the smaller angle representing the higher correlation, and the 90° indicating the orthogonal relationship. Based on this analysis, $V_{c,\text{max}25}$ shows higher relationship with leaf N content, followed by Chl$_{SPAD}$, but has nearly orthogonal relationships with LWC and LMA. Meanwhile, all leaf traits are located within the gray semicircle region of $V_{c,\text{max}25}$, suggesting that all these trait-$V_{c,\text{max}25}$ relationships are positive. Leaf level trait measurements across all ages and forest types were used in this analysis.
Fig. 4 Accuracy assessment for the cross-site spectral-$V_{c,max}^{25}$ relationship. The ‘cross-site’ spectral model of $V_{c,max}^{25}$ was trained and evaluated using the data from all the three forest sites through the repeated double cross-validation method. Error bars indicate the 95% confidence intervals for each predicted value derived from the ensemble partial least-squares regression (PLSR) models (i.e. each PLSR model is represented by a set of PLSR fitted spectral coefficients averaged over the 200 repetitions). The gray line represents the ordinary least-squares regression fit, and the black line indicates the 1:1 line. Three different color circles represent each of the three forest sites, with CB in red, DH in green, and XSBN in blue. Leaf level trait measurements across all leaf ages were involved in this analysis. Abbreviations: $n$, sample size; $R^2$, the coefficient of determination; RMSE, the root mean square of error.
Fig. 5 Accuracy assessments for the spectral models of $V_{c,max25}$ under the site-specific and XSBN scenarios. (a) ‘Site-specific model’ (trained and evaluated using the data from each forest site through the repeated double cross-validation (rdCV) method); (b) ‘XSBN model’ (trained and evaluated using the data from the tropical forest site of XSBN through the rdCV method, and then applied the model to the independent sites at CB and DH). Error bars indicate the 95% confidence intervals for each predicted value derived from the ensemble partial least-squares regression (PLSR) models. The gray line represents the ordinary least-squares regression fit, and the black line indicates the 1:1 line. Three different color circles represent each of the three forest sites, with CB in red, DH in green, and XSBN in blue. Leaf level trait measurements across all leaf ages were used here. Abbreviations: $n$, sample size; $R^2$, the coefficient of determination; RMSE, the root mean square of error.
**Fig. 6** Assessing the reflectance contributions to the spectral models of $V_{c,max25}$ and the four leaf traits (i.e. leaf N content, ChlSPAD, LMA, and LWC) under the cross-site scenario using the PLSR approach, including the left panel (a,c,e,g,i) for PLSR regression coefficients, and the right panel (b,d,f,h,j) for the variable importance in projection (VIP). The cross-site spectral models were trained and evaluated using the data from all the three forest sites through the repeated double cross-validation method. The central color lines indicate the mean values and the shaded regions indicate the 95% confidence interval of PLSR regression coefficients and VIP spectrum, respectively. On the right panel, VIP $\geq 0.8$ refers to the important spectral regions responsible for the spectral modelling of $V_{c,max25}$ and the four leaf traits (Wold et al., 2001). The shaded gray regions across all sub-panels identify the corresponding spectral bands for the $V_{c,max25}$ predictions. Abbreviations: VIR, visible range (450-700 nm); RE, red edge range (700-800 nm); NIR, near infrared range (800-1300 nm); SWIR, shortwave infrared range (1300~2500 nm).