Fundamental Research xxx (xxxx) xxx

Contents lists available at ScienceDirect

Fundamental Research

journal homepage: http://www.keaipublishing.com/en/journals/fundamental-research/

Article

Global photosynthetic capacity jointly determined by enzyme kinetics and eco-evo-environmental drivers

Zhengbing Yan[a,b,c], Matteo Detto[d], Zhengfei Guo[b], Nicholas G. Smith[e], Han Wang[f,g], Loren P. Albert[h], Xiangtao Xu[i], Ziyu Lin[b], Shuwen Liu[b], Yingyi Zhao[b], Shuli Chen[j], Timothy C. Bonebrake[b], Jin Wu[b,k,l,*]

[a] State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Xiangshan, Beijing 100093, China
[b] School of Biological Sciences, The University of Hong Kong, Hong Kong Special Administrative Region 999077, China
[c] China National Botanical Garden, Beijing 100093, China
[d] Department of Ecology and Evolutionary Biology, Princeton University, Princeton NJ 08544, USA
[e] Department of Biological Sciences, Texas Tech University, Lubbock TX 79409, USA
[f] Ministry of Education Key Laboratory for Earth System Modelling, Department of Earth System Science, Tsinghua University, Beijing 100084, China
[g] Joint Centre for Global Change Studies, Tsinghua University, Beijing 100084, China
[h] Department of Biology, West Virginia University, Morgantown WV 26506, USA
[i] Department of Ecology and Evolutionary Biology, Cornell University, Ithaca NY 14853, USA
[j] Department of Ecology and Evolutionary Biology, University of Arizona, Tucson AZ 85721, USA
[k] Institute for Climate and Carbon Neutrality, The University of Hong Kong, Hong Kong 999077, China
[l] State Key Laboratory of Agrobiotechnology, The Chinese University of Hong Kong, Shatin, Hong Kong 999077, China

ARTICLE INFO

Article history:
Received 27 February 2023
Received in revised form 7 October 2023
Accepted 27 December 2023
Available online xxx

Keywords:
Global carbon cycling
Leaf photosynthetic capacity
Enzyme kinetics
Eco-evolutionary optimality
Ecophysiology
Climate
Leaf traits
Belowground resource constraint

ABSTRACT

Accurate understanding of global photosynthetic capacity (i.e. maximum RuBisCO carboxylation rate, $V_{c,\,max}$) variability is critical for improved simulations of terrestrial ecosystem photosynthesis metabolisms and carbon cycles with climate change, but a holistic understanding and assessment remains lacking. Here we hypothesized that $V_{c,\,max}$ was dictated by both factors of temperature-associated enzyme kinetics (capturing instantaneous eco-physiological responses) and the amount of activated RuBisCO (indexed by $V_{c,\,max}$ standardized at 25 °C, $V_{c,\,max25}$), and compiled a comprehensive global dataset (n = 7339 observations from 428 sites) for hypothesis testing. The photosynthesis data were derived from leaf gas exchange measurements using portable gas exchange systems. We found that a semi-empirical statistical model considering both factors explained 78% of global $V_{c,\,max}$ variability, followed by 55% explained by enzyme kinetics alone. This statistical model outperformed the current theoretical optimality model for predicting global $V_{c,\,max}$ variability (67%), primarily due to its poor characterization on global $V_{c,\,max25}$ variability (3%). Further, we demonstrated that, in addition to climatic variables, belowground resource constraint on photosynthetic machinery built-up that directly structures the biogeography of $V_{c,\,max25}$ was a key missing mechanism for improving the theoretical modelling of global $V_{c,\,max}$ variability. These findings improve the mechanistic understanding of global $V_{c,\,max}$ variability and provide an important basis to benchmark process-based models of terrestrial photosynthesis and carbon cycling under climate change.

1. Introduction

Terrestrial photosynthesis is the largest carbon flux exchange between biosphere and atmosphere and exerts an important role in buffering atmospheric CO_2 growth [1,2]. Under many real world conditions, photosynthesis is expected to be limited by the maximum carboxylation rate of RuBisCO in the chloroplasts ($V_{c,\,max}$) [3–5]. $V_{c,\,max}$ shows a large variability globally and is affected by multiple environmental and biotic variables, such as climate, edaphic properties and leaf traits [6–9], all of which will be altered by future climate and land use changes. Although predicting $V_{c,\,max}$ variability has received much scientific attention [8,10-12], a holistic understanding of what drives global $V_{c,\,max}$ variability remains lacking.

Traditionally, empirically derived trait coordinated relationships and theory-based optimality models are used to explain $V_{c,\,max}$ variability [3,8-10]. The trait-based approach suggests that $V_{c,\,max}$ can be estimated using its empirical relationships with other leaf biochemical traits (e.g., leaf N content per unit leaf area (N_a) and leaf

* Corresponding author.
E-mail address: jinwu@hku.hk (J. Wu).

https://doi.org/10.1016/j.fmre.2023.12.011

Please cite this article as: Z. Yan, M. Detto, Z. Guo et al., Global photosynthetic capacity jointly determined by enzyme kinetics and eco-evo-environmental drivers, Fundamental Research, https://doi.org/10.1016/j.fmre.2023.12.011

chlorophyll content) that tightly correlate with photosynthetic biochemistry [3,7,10]. However, mounting evidence suggests that trait-$V_{c,\,max}$ relationships vary with plant growing environments, seasons and other biotic variables [13–16], which further induces large uncertainties in representing $V_{c,\,max}$ in terrestrial biosphere models (TBMs). Additionally, the trait-based approach is ultimately based on empirical relationships, lacking mechanistic foundation that is critical for long-term predictions under global change.

Recently, theoretical optimality models have been increasingly used for explaining and modelling large-scale $V_{c,\,max}$ variability [8,17,18]. The models were built based on eco-evolutionary first-principles [13,19-21], hypothesizing that plants can adjust their photosynthetic biochemistry for maximizing their net carbon gain, which is the difference between photosynthetic carbon gain and all carbon costs of building and maintaining the photosynthetic machinery [13,22]. Because optimality models are mechanistic and prognostic, they have been suggested as a novel means to infer $V_{c,\,max}$ variability [18,23]. To support its validity is a recent theoretical study [8], demonstrating that optimality models driven by aboveground environmental variables (i.e., temperature, incoming photosynthetically active radiation (PAR), vapor pressure deficit (VPD), atmospheric CO_2 concentration (C_a), and elevation (a proxy of atmospheric pressure)) alone can accurately capture global $V_{c,\,max}$ variability. However, it remains contentious whether the optimality models could accurately predict $V_{c,\,max}$ at a standardized temperature (e.g. $V_{c,\,max25}$, $V_{c,\,max}$ standardized at 25 °C) [8,12]. Some studies suggested that the optimality model could be more appropriate to predict $V_{c,\,max}$ during the growing-season temperature ($V_{c,\,maxTg}$) rather than the $V_{c,\,max25}$, because growing-season temperature is the temperature regularly experienced by plants rather than a standard temperature, which may be atypical for that environment [8,24]. In contrast, some other studies hold the opposite viewpoint that the optimality model could also accurately predict $V_{c,\,max25}$, but with the use of limited dataset and the growing-season temperature as the proxy of leaf measurement temperature when calculating $V_{c,\,max25}$ from the measured $V_{c,\,max}$ [12,25]. Therefore, despite its promising application in plant ecophysiology and modelling studies, a more rigorous and holistic assessment remains needed. Meanwhile, several empirical studies demonstrate the critical role of belowground resource availability on regulating large-scale $V_{c,\,max}$ variability [26,27]. For example, a large fraction of leaf nitrogen is allocated to RuBisCO enzyme, which is ultimately related to soil nutrient availability and associated plant evolved strategies for root nitrogen uptake [13,27]. With these, it remains unclear whether and how belowground environmental conditions exert a role in constraining photosynthetic machinery built-up in the theoretical optimality modelling framework.

On the other hand, $V_{c,\,max}$ is the product of catalytic rate and amount of RuBisCO enzyme (indicated by $V_{c,\,max25}$) in the chloroplasts [17,28]. This can be interpreted as two key factors governing $V_{c,\,max}$ variability, namely the temperature-associated enzyme kinetics of a given $V_{c,\,max25}$ and the eco-evo-environmental drivers of $V_{c,\,max25}$. The second category, the eco-evo-environmental drivers of $V_{c,\,max25}$, stands in contrast with pure enzyme kinetics. Within this second category, $V_{c,\,max25}$ variability is associated with environmental factors and leaf traits [3,6,10,29], which can influence plant photosynthetic processes, the construction and maintenance of photosynthetic apparatus, carbon uptake capacity and then the subsequent responses to changing ambient conditions [30,31]. Thus, incorporating the relationships between $V_{c,\,max25}$ variability and these environmental factors and leaf traits could further enhance the characterization of global $V_{c,\,max}$ variability.

The goal of this study is to improve the process understanding of global $V_{c,\,max}$ variability by first assessing the relative role of enzyme kinetics and eco-evo-environmental drivers of $V_{c,\,max25}$, and then exploring the way to improve current optimality models for characterizing global $V_{c,\,max}$ variability. Specifically, we test the hypothesis that global patterns of $V_{c,\,max}$ are driven by (1) enzyme kinetics, (2) eco-evo-environmental drivers of $V_{c,\,max25}$, or (3) both enzyme kinetics and the eco-evo-environmental drivers of $V_{c,\,max25}$. Additionally, we test whether current optimality model explains global pattern of $V_{c,\,max}$, once $V_{c,\,max}$ has been standardized to 25 °C (to remove the temperature response of $V_{c,\,max}$). Last, we test whether including belowground resource constraints through edaphic properties on photosynthetic machinery would improve current optimality modelling of global variability in both $V_{c,\,max}$ and $V_{c,\,max25}$. To test these hypotheses, we collated a comprehensive global dataset of $V_{c,\,max}$ for C_3 plants with concurrent measurements of environmental variables and key leaf traits, and then integrated this unique global dataset with both statistical and optimality modelling analyses.

2. Materials and methods

2.1. Field dataset of $V_{c,\,max}$, climate and edaphic variables, and leaf traits

To compile the global $V_{c,\,max}$ dataset of C_3 plants (Fig. S1), we turned to the following three data sources: two global data sources assembled by Smith et al. [8] and Peng et al. [12], respectively, and one data source from three contrasting forest ecosystems in China [16]. The two global data sources were mainly derived from earlier compilations by Meir et al. [32], Domingues et al. [33,34], Cernusak et al. [35], Walker et al. [3], Atkin et al. [36], Maire et al. [37], Bahar et al. [26], Smith & Dukes [38], Dong et al. [39], Wang et al. [40], Bloomfield et al. [41], Xu et al. [42] and the TRY plant trait database. It is worthy to note, when compiling this dataset, we only retained the records with concurrent measurement of leaf temperature. This also represents one major difference between our compiled data and those compiled by Peng et al. [12], in which Peng et al. [12] used the growing-season temperature as the proxy of leaf measurement temperature for those $V_{c,\,max}$ records lack of the information on leaf measurement temperature (details about the cross-comparison between ours and Peng et al. [12] are shown in Fig. S2 and in discussion Section 2). In our data, all $V_{c,\,max}$ records came from natural vegetation, including 7339 measurements from 2250 species and 428 sites covering all major biomes worldwide. In order to quantify the separate and joint effects of enzyme kinetics and eco-evo-environmental drivers of $V_{c,\,max25}$ on global $V_{c,\,max}$ variability, only those $V_{c,\,max}$ records with concurrent measurements of environmental variables and leaf traits (i.e., LMA and N_a) were chosen. Totally, 5748 measurements from 281 sites met these selection criteria.

Besides, we extracted six climate-related variables (i.e. temperature, VPD, PAR, C_a, elevation, and precipitation). For each climate variable, the average value across the full growing season (defined as all months with mean monthly air temperature higher than 0 °C) was calculated for each site [8]. These include (1) the mean growing-season temperature (T_g) and precipitation extracted using the corresponding latitude and longitude from monthly, 0.5 degree resolution data of 1901-2015 provided by Climatic Research Unit [43]; (2) VPD and PAR calculated from the CRU data using SPLASH model; (3) C_a derived primarily from original records in earlier compilations, but when there was no C_a record, it was estimated using corresponding value from global average estimates by NASA GISS model (https://data.giss.nasa.gov/modelforce/ghgases/); and (4) elevation was partly derived from original records, but when there was no record, it was extracted from 0.5 degree resolution data from WFDEI meteorological forcing dataset [44]. Notably, T_g and precipitation were three-dimensionally interpolated to the actual site locations using Geographically Weighted Regression (GWR) following Peng et al. [12], while VPD and PAR were adjusted to the actual elevation following Smith et al. [8].

Moreover, we extracted ten edaphic variables, including pH, carbon (C) content, nitrogen (N) content, C:N ratio, Priestley-Taylor coefficient (α; this variable indicated plant-available surface moisture, and was calculated as the ratio of actual evapotranspiration to equilibrium evapotranspiration. Equilibrium evapotranspiration refers to theoretical value of evaporation from a wet surface to a saturated atmosphere), cation

exchange capacity (CEC), silt content, clay content, sand content, and bulk density. α of each 0.5 degree resolution was calculated using the SPLASH model run at a monthly timescale [45]. Other edaphic variables, comprehensively reflecting soil physical and chemical conditions, were extracted from a 250-m resolution global data at the top 30 cm depth provided by ISRIC SoilGrids database (www.soilgrids.org) based on site-specific latitude and longitude.

2.2. Field dataset of $V_{c,\ maxTmeas}$ and $V_{c,\ max25}$

With $V_{c,\ max}$ derived at its measurement temperature (T_{meas}, °C), or $V_{c,\ maxTmeas}$, we calculated $V_{c,\ max}$ respectively at T_g ($V_{c,\ maxTg}$, °C) and 25 °C ($V_{c,\ max25}$), using a modified Arrhenius function (see Eqs. (1)-(2)) that described the instantaneous response of enzyme kinetics to any given temperature [28]. Given a reference temperature T_0 (°C), $V_{c,\ max}$ at temperature T_1 (°C) can be computed as

$$V_{c,\ maxT_1} = V_{c,\ maxT_0} \times f(T_0, T_1) \tag{1}$$

where

$$f(T_0, T_1) = e^{\frac{H_a(T_1 - T_0)}{R(T_0 + 273.15)(T_1 + 273.15)}} \times \frac{1 + e^{\frac{(T_0 + 273.15)(\Delta S) - H_d}{R(T_0 + 273.15)}}}{1 + e^{\frac{(T_1 + 273.15)(\Delta S) - H_d}{R(T_1 + 273.15)}}} \tag{2}$$

where H_a is the activation energy (71,513 J mol^{-1}), R is the universal gas constant (8.314 J mol^{-1} K^{-1}), H_d is the deactivation energy (200,000 J mol^{-1}), and ΔS is an entropy term (J mol^{-1} K^{-1}) calculated following Kattge and Knorr (2007) [28].

$$\Delta S = -1.07 \times T_g + 668.39 \tag{3}$$

Spurious correlations may arise because the same temperature response function was applied to both observed and modeled $V_{c,\ max}$. Smith et al. [8] has examined this issue and found that the temperature scalar did yield some but low spurious correlations. We also compared models against $V_{c,\ maxTmeas}$ that did not subject to the temperature scalar issue, and found similar results as those for $V_{c,\ maxTg}$ (Fig. S3), suggesting that spurious correlation effects induced by the temperature scalar function would not affect our result interpretations.

2.3. Statistical and optimality modelling approaches for predicting global $V_{c,\ maxTg}$ variability

To improve the process understanding of global $V_{c,\ max}$ variability, we cross-compared a statistical approach with an optimality modelling approach for predicting $V_{c,\ maxTg}$ and $V_{c,\ max25}$, respectively. The statistical model examined the separate and joint effects of the two factors in explaining $V_{c,\ maxTg}$ variability, with two levels of analyses. In the first level, we used the temperature-associated enzyme kinetics response assuming no intra- or inter-specific variability in $V_{c,\ max25}$ across global C_3 plants. In the second level, we added temperature-associated enzyme kinetics to site-specific $V_{c,\ max25}$ derived from its empirical relationship with both environmental variables and leaf traits, assuming that large geographical variability in $V_{c,\ max25}$ was correlated with their living environments [6,29] and leaf traits [3,10]. The optimality modelling relies on an established optimality theory [8], and is able to infer $V_{c,\ max,Tg}$ from five aboveground environmental variables (i.e., temperature, PAR, VPD, elevation and C_a). The current optimality model by default assumes that the cost of building photosynthetic machinery is independent of species and the belowground environmental condition. To account for variable costs, particularly the cost of nutrient acquisition [27,37], we built and evaluated a modified optimality model that considered the belowground resource constraints through edaphic variables on the unit cost of building and maintaining the photosynthetic machinery.

All analyses were performed at site-mean level by averaging all measurements of each site. Details of these modelling analyses are shown below. For simplicity, we primarily focused on the data analyses for the subset dataset with concurrent measurements of environmental variables and leaf traits in the main text, while the results from the data analyses for the entire dataset (solely including the aboveground environmental measurements) are shown in the supplementary materials with similar findings (Figs. S4-S6).

2.3.1. Statistical modelling approach for predicting global $V_{c,\ maxTg}$ variability

There are two levels of statistical modelling analyses:

(1) Quantifying the effect of enzyme kinetics on global $V_{c,\ maxTg}$ variability. To assess the effect of enzyme kinetics alone on explaining global $V_{c,\ maxTg}$ variability, we first derived the site-mean $V_{c,\ max25}$ based on field measurements, and then calculated the average $V_{c,\ max25}$ across all sites, assuming that all the C_3 plants shared the same $V_{c,\ max25}$ as this average value. Afterwards, for each site, we calculated the modelled $V_{c,\ maxTg}$ using Eqs. (1)-(3) that multiplied this global-average $V_{c,\ max25}$ with $f(25,\ T_g)$.

(2) Quantifying the joint effects of enzyme kinetics and eco-evo-environmental drivers of $V_{c,\ max25}$ on global $V_{c,\ maxTg}$ variability. We first built a multiple linear regression model of $V_{c,\ max25}$ with site-specific environmental variables and leaf traits as the input (Table S1), and then added this modelled $V_{c,\ max25}$ with temperature-associated enzyme kinetics (Eqns. 1-3) for deriving $V_{c,\ maxTg}$ of each site. The coefficients of predictors in the multiple linear regression were determined using the Ordinary Least-Squares method by minimizing the root mean square error (RMSE) between the observed and predicted $V_{c,\ max25}$ values. Specifically, $V_{c,\ max25}$ and $V_{c,\ maxTg}$ modelling were built and evaluated using a repeated cross-validation method, including four steps: (1) the full dataset was split into calibration and independent validation subsets using a 5-fold cross-validation with 100 repetitions; (2) for each repetition, the calibration subset was used to build the model, with the validation subset for model evaluation; (3) the modelled $V_{c,\ max25}$ for each validation subset was calculated and averaged across all 100 repetitions to obtain the ensemble predicted value for each record; and (4) the ensemble predicted $V_{c,\ max25}$ and their associated $f(25,\ T_g)$, were used together to estimate $V_{c,\ maxTg}$ of each record.

2.3.2. The optimality modelling approach for $V_{c,\ maxTg}$

(1) *The optimality theory*

The optimality model shown in Smith et al. [8] is based on two separate optimization processes. The first process involves the leaf photosynthetic process, and the total carbon gain is derived from leaf assimilation rate (A), whereas the respiration carbon cost is required to build and maintain the photosynthetic machinery, including the light harvesting component (via J_{max}) and the carbon reduction component (via $V_{c,\ max}$) [13]. Thus, the first optimization is expressed as maximizing the net carbon gain, which is the difference between A and the carbon costs associated with maintaining $V_{c,\ max}$ and J_{max} [8]. The second optimization process applies the least-cost theory to find the optimal leaf internal to external CO_2 ratio (χ) at the lowest carbon cost to maintain photosynthetic machinery ($V_{c,\ max}$) and water transpiration (E) at a given assimilation rate (A) [13]. These two optimality conditions can be expressed as:

$$\max_{V_{c,\ max},\ J_{max}} \left(A - b \times V_{c,\ max} - d \times J_{max}\right) \tag{4}$$

$$\min_{\chi} \left(a \times \frac{E}{A} + b \times \frac{V_{c,\ max}}{A}\right) \tag{5}$$

where a, b, and d are the dimensionless carbon cost factors for E, $V_{c,\ max}$ and the maximum electron transport rate (J_{max}), respectively.

Based on an assumption that $V_{c,\,max}$ is linearly proportional to J_{max} (i.e. $V_{c,\,max} = e{\times}J_{max,}$ where e is the ratio of $V_{c,\,max}$ to J_{max}), the total carbon costs for both $V_{c,\,max}$ and J_{max} could then be expressed by the carbon cost for J_{max} only. Thus, the first optimality condition can further be reduced to

$$\max_{J_{max}} \left(A - c \times J_{max}\right) \tag{6}$$

where c is defined as the total unit carbon cost of building and maintaining the two components of photosynthetic machinery. By building upon these two independent optimizations, $V_{c,\,maxTg}$ can further be inferred with solely environmental inputs from T_g, PAR, VPD, C_a, and elevation. The key mathematical formulations have been introduced in Smith et al. (2019) [8]. Here we provided the details about the optimality model in the supporting method S1.

(2) Evaluating the optimality model with a constant or dynamic cost

The marginal cost c (i.e., $\frac{\partial A_j}{\partial J_{max}}$) reflects the fundamental tradeoff between photosynthetic carbon gain (i.e., A_j) and the relevant cost for photosynthetic machinery built-up and maintenance (i.e., J_{max} and $V_{c,\,max}$, which in this study are assumed to be linearly related), and is an important parameter for predicting $V_{c,\,max}$ in the optimality model (Eqns 6 and S10). In the optimality model of Smith et al., [8], c was parameterized with a default globally constant at 0.053. However, as several empirical and theoretical studies demonstrated, the costs in constructing and maintaining the photosynthetic machinery could also depend on environmental factors and leaf traits [13,22,26,27,46]. To explore whether incorporating the constraints from environmental variables and leaf traits on the cost factor c can improve the current optimality model performance, we parameterized c under four scenarios (Table S2 and S3): (1) a globally constant c at 0.053 across all sites as used in the current optimality model [8], and a site-specific variable c respectively constrained by (2) the edaphic variables alone, (3) both environmental (i.e. climatic and edaphic) variables and leaf traits, and (4) the five most important variables (i.e. N_a, VPD, soil pH, precipitation and elevation) identified from the statistical modelling (Fig. 4). To differentiate these four optimality models, we called (1) as optimality-constant (with a fixed c) and (2-4) as optimality-dynamic models (with a variable c).

For the three optimality-dynamic models, the cost factor c was fitted with a four-step approach illustrated below. First, the full dataset was split into calibration and validation subsets using a 5-fold cross-validation with 100 repetitions. Different choices of folds and repetitions yielded similar results. Second, for each repetition, the calibration subset was used to build a multiple linear regression between the cost factor c_i at site i, and their corresponding site-specific environmental variables and leaf traits, x_{ij}, depending on the scenario:

$$c_i = \beta_0 + \sum_j \beta_j x_{ij} \tag{7}$$

where β_j are fitted parameters for the specific environmental variable or leaf trait (see the example demonstration in Table S4 and S5). The multiple linear regression coefficients of predictors were fitted using a genetic algorithm (GA) technique through minimizing the RMSE between observed and optimality-dynamic modelled $V_{c,\,maxTg}$. GA is a heuristic global optimization method [47] that avoids the dependence on the initial parameter values and efficiently identifies the global parameter optimization solution when there are many acceptable local solutions [48]. Third, for each repetition, we used the multiple linear regression coefficients estimated from the calibration sub-dataset to predict $V_{c,\,max}$ of the independent validation sub-dataset. Fourth, the predicted $V_{c,\,max}$ was then averaged across all 100 repetitions to obtain the ensemble predicted value for each measurement record, which was used to evaluate the optimality model performance under each scenario.

2.4. Cross-model performance comparisons and attributions of important variables of the two modelling approaches

With the above model-derived $V_{c,\,maxTg}$ and $V_{c,\,max25}$, we evaluated the model performance with reference to field-derived $V_{c,\,maxTg}$ and $V_{c,\,max25}$. Four statistical metrics for assessing model performance include (1) r^2–the square of correlation coefficient, (2) Bias–the residual bias, (3) RMSE, and (4) AIC–Akaike Information Criterion. The AIC was calculated with the formula AIC = $2 \times k + n \times \log(\mathrm{RMSE}^2)$, where n is the number of observations, and k is the number of model parameters.

For the statistical modelling approach, we assessed the relative importance of each environmental or biotic variable on $V_{c,\,max25}$ modelling. To avoid the strong multicollinearity amongst predictor variables (Fig. S7), we first computed the variance inflation factor (VIF) and iteratively removed variables with very high VIF [49–50], until the remaining variables with VIF less than 10 and the absolute values of correlation coefficients amongst these variables were less than 0.7. Afterwards, we conducted a model selection for $V_{c,\,max25}$ based on corrected AIC using 'glmulti' R package [51]. We estimated the relative importance of each variable as the sum of the Akaike weights for the models in which the variable appeared. We set relative importance value of 0.8 as a cut-off to differentiate between important and unimportant variables [52]. Finally, we illustrated the sign of the effects from each selected variable through conducting partial regression plots while holding all the other variables constant [51,53].

For the optimality-constant model, we used a sensitivity analysis to assess the variable importance. Specifically, we partitioned the variance of the key model output into the variation in each model input (i.e., T_g, PAR, VPD, elevation, and C_a), allowing these inputs to cover the full range of environmental variables. The analysis includes three steps. First was the set-up of the five environmental variables. We built a one-time random sampling for each of the five environmental variables using Sobol' quasi-random sequences [54]. Second were the model simulations of $V_{c,\,max25}$ using the optimality-constant model, with the five environmental variables as input. Third was the variance partitioning. We employed a global variance-based sensitivity analysis algorithm developed by Saltelli et al. [55], and partitioned the variance of modelled $V_{c,\,max25}$ driven by each environmental variable, by which we quantified the variables' relative importance.

3. Results

3.1. Relative importance of enzyme kinetics and eco-evo-environmental drivers of $V_{c,\,max25}$ in explaining global $V_{c,\,maxTg}$ variability

To investigate the relative importance of enzyme kinetics and eco-evo-environmental drivers of $V_{c,\,max25}$ in explaining global $V_{c,\,maxTg}$ variability, we analyzed the relationships between field-derived and model-predicted $V_{c,\,maxTg}$ using the statistical modelling approach. We found that the statistical approach considering both factors explained 78% of global $V_{c,\,maxTg}$ variability (RMSE = 11.87 μmol CO_2 m^{-2} s^{-1}, AIC = 1423; Fig. 1b), followed by 55% explained by enzyme kinetics alone (RMSE = 17.17 μmol CO_2 m^{-2} s^{-1}; AIC = 1600; Fig. 1a). This result demonstrates that temperature-associated enzyme kinetics dominate the explanation of global $V_{c,\,maxTg}$ variability, with eco-evo-environmental drivers of $V_{c,\,max25}$ also exerting a considerable role.

3.2. The performance of optimality-constant model and the reasons underlying its poorer performance

To investigate whether optimality-constant model explains global pattern of $V_{c,\,max}$, we cross-compared the optimality-constant model with the statistical model. We found that the optimality-constant model explained 67% of global $V_{c,\,maxTg}$ variability but with comparable model

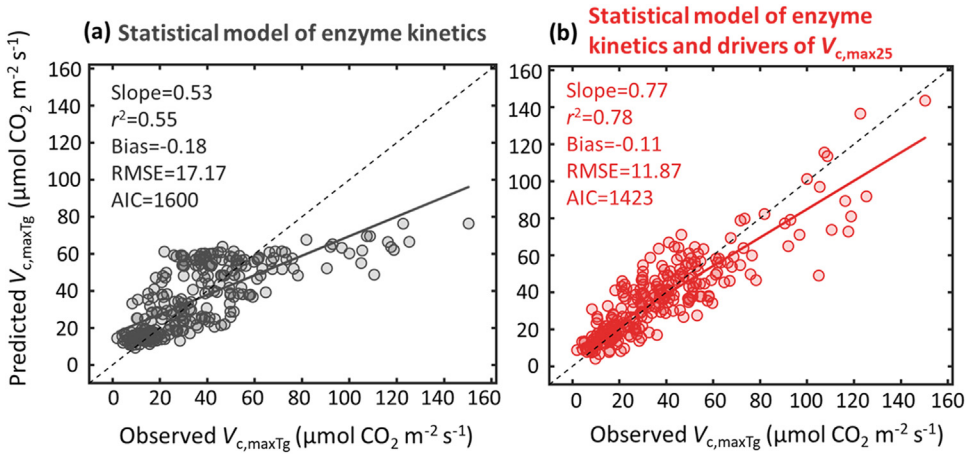

Fig. 1. The relative importance of enzyme kinetics and eco-evo-environmental drivers of $V_{c,\,max25}$ in controlling global $V_{c,\,maxTg}$ variability. The statistical modelling approach is analyzed at two levels: global $V_{c,\,maxTg}$ variability is described using (a) the temperature-associated enzyme kinetics together with a globally averaged $V_{c,\,max25}$, or (b) by adding temperature-associated enzyme kinetics to site-specific $V_{c,\,max25}$ derived from its empirical relationship with both environmental variables and leaf traits. Details of these two models with the fitted equations are presented in Tables S5 and S6. Four statistical metrics for assessing model performance include (1) r^2–the square of correlation coefficient, (2) Bias–the residual bias, (3) RMSE–the root mean square of error, and (4) AIC–Akaike Information Criterion. Lines are fitted by the ordinary least-square regressions.

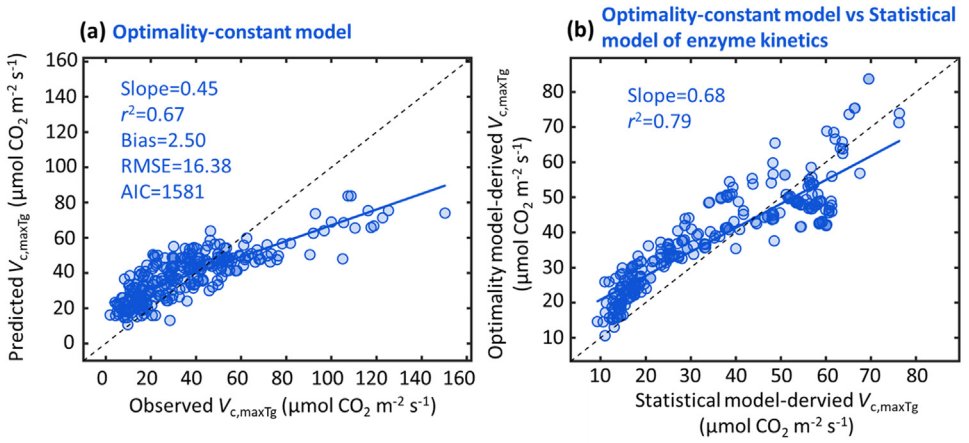

Fig. 2. The performance of optimality-constant model in predicting global $V_{c,\,maxTg}$ variability. The optimality model is used for global $V_{c,\,maxTg}$ prediction with a default globally constant c (i.e., the total unit carbon cost of building and maintaining the photosynthetic machinery), or the optimality-constant model (a), which displays high correlation as the statistical model of enzyme kinetics (b). The statistical model of enzyme kinetics here is the same as Fig. 1a.

error (RMSE = 16.38 µmol CO_2 m^{-2} s^{-1}; AIC = 1581; Fig. 2a) as the statistical model considering enzyme kinetics alone (RMSE = 17.17 µmol CO_2 m^{-2} s^{-1}; AIC = 1600; Fig. 1a). We further found that the predicted $V_{c,\,maxTg}$ by optimality-constant model was highly correlated with that predicted using the statistical model considering enzyme kinetics alone (r^2 = 0.79; Fig. 2b), suggesting that the optimality-constant model likely captured similar ecophysiological processes as the enzyme kinetics' response, and thus showed comparable model performance in predicting $V_{c,\,maxTg}$ (Fig. 1a vs Fig. 2a). This also agreed with the $V_{c,\,max25}$ modelling result, as environmental variables and leaf traits altogether explained 36% of $V_{c,\,max25}$ variability (Fig. 3a; Table S1), but optimality-constant model explained only 3% of $V_{c,\,max25}$ variability (Fig. 3b).

To explore the reasons underlying the poorer performance of optimality-constant model in predicting $V_{c,\,max25}$ variability further, we cross-compared the relative importance of each variable that was represented in the optimality-constant and statistical modelling of $V_{c,\,max25}$. Based on the statistical modelling approach, we identified five most important biotic and environmental variables for explaining $V_{c,\,max25}$ variability, i.e., N_a, VPD, soil pH, precipitation and elevation, following a descending order of their relative importance (Fig. 4a). This result was different from the optimality-constant model, in which PAR and temperature were identified as the two most important variables (with relative contributions of 50.0% and 48.5%, respectively) for $V_{c,\,max25}$ (Fig. 4a). Collectively, these results demonstrate that the important drivers of global $V_{c,\,max25}$ (and thus $V_{c,\,maxTg}$) variability were not well represented in optimality-constant model.

3.3. The improvement of optimality model through considering belowground resource constraints

To investigate whether including belowground resource constraints on photosynthetic machinery would improve current optimality modelling of global $V_{c,\,max}$ variability, we revised the current optimality model by including a dynamic representation of c that described belowground resource constraints through edaphic variables. We found that optimality-dynamic model had a much-improved predictive power and reduced model error for $V_{c,\,maxTg}$ variability (r^2 = 0.74; RMSE = 13.80 µmol CO_2 m^{-2} s^{-1}; AIC = 1501; Fig. 5a), but still showed a slightly subordinate performance compared with the statistical model of both enzyme kinetics and eco-evo-environmental drivers of $V_{c,\,max25}$ (Fig. 1b).

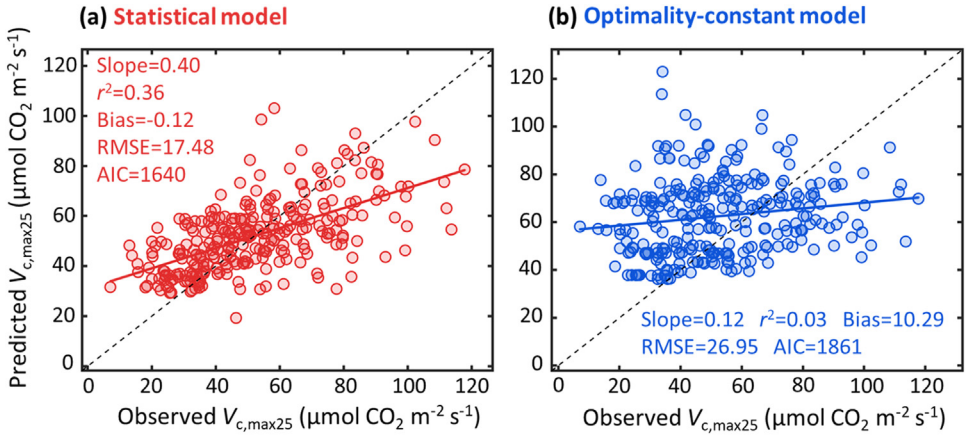

Fig. 3. The statistical model of environmental variables and leaf traits outperforms the optimality-constant model in predicting $V_{c,\,max25}$. The comparison of field-derived $V_{c,\,max25}$ with (a) the statistical model derived $V_{c,\,max25}$ that relies on its multiple linear regression relationship with both environmental variables and leaf traits, and (b) the $V_{c,\,max25}$ derived from the optimality-constant model. Details of the statistical model with the fitted equations are presented in Tables S4 and S5.

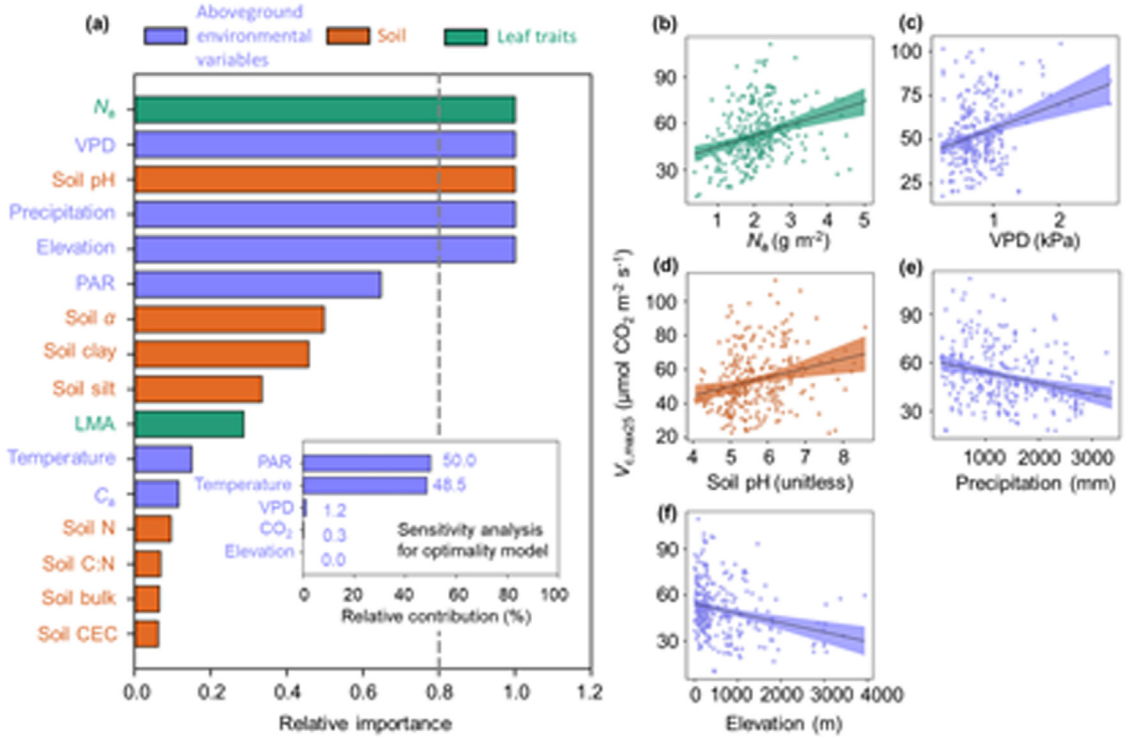

Fig. 4. Relative importance of environmental variables and leaf traits in predicting global $V_{c,\,max25}$ variability. (a-main) The relative importance of each variable based on the sum of the Akaike weights derived from a model selection using corrected AIC; (a-inset) the relative contribution of the five aboveground environmental variables to $V_{c,\,max25}$ based on the sensitivity analysis of the optimality-constant model; (b-f) partial regression plots of $V_{c,\,max25}$ with the predictor of area-based leaf nitrogen content (N_a), vapor pressure deficit (VPD), soil pH, precipitation, and elevation, respectively. The cutoff (dashed line) of panel (a) is set at 0.8 for identifying the most important predictor variables; the shade areas in (b-f) are 95% confidential intervals around the predicted relationships.

We also observed that optimality-dynamic model explained 19% of $V_{c,\,max25}$ variability (Fig. 5b), which was larger than that (3%) explained by optimality-constant model (Fig. 3b). Such a cross-comparison further demonstrates that the improvement in modelling $V_{c,\,max25}$ through considering belowground resource constraints on c indeed represents a key avenue for improved modelling of global $V_{c,\,max}$ variability.

4. Discussion

This study has three main findings. First, the statistical modelling approach considering enzyme kinetics and eco-evo-environmental drivers of $V_{c,\,max25}$ explained a large portion (78%) of global $V_{c,\,maxTg}$ variability, with enzyme kinetics as the dominant factor explaining 55% of global $V_{c,\,maxTg}$ variability. Second, treating the statistical model as a benchmark for the optimality approach to predict $V_{c,\,maxTg}$, we found that the optimality-constant model effectively captured the role of enzyme kinetics in controlling global $V_{c,\,maxTg}$ variability, but was ineffective in capturing the second control on global $V_{c,\,maxTg}$ variability through eco-evo-environmental drivers of $V_{c,\,max25}$. Third, we found that belowground resources added constraints on building and maintaining photosynthetic machinery, and improved the modelling of $V_{c,\,max25}$ variability.

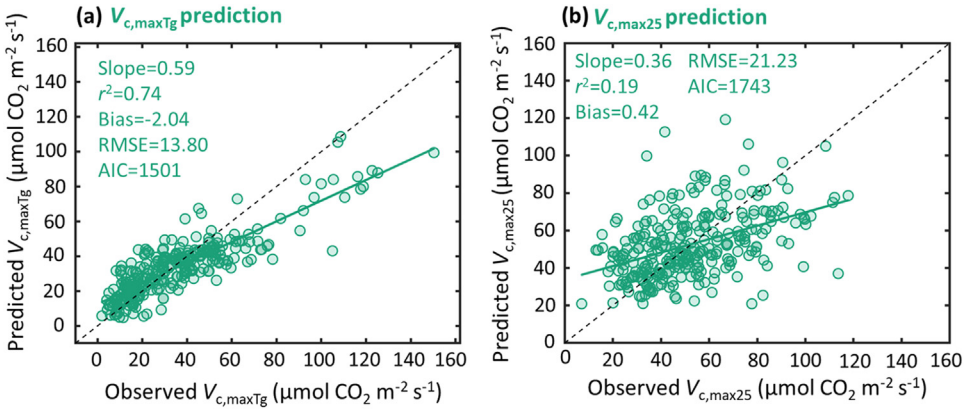

Fig. 5. The performance of optimality-dynamic model in predicting global variability of $V_{c,\,maxTg}$ (a) and $V_{c,\,max25}$ (b). The optimality model is used for global $V_{c,\,maxTg}$ and $V_{c,\,max25}$ predictions with a site-specific dynamic *c* (i.e., the total unit carbon cost of building and maintaining the photosynthetic machinery) constrained by edaphic variables, or the optimality-dynamic model. Details of the optimality-dynamic model with the fitted equations for cost factor *c* parameterization are presented in Tables S4 and S5.

4.1. Enzyme kinetics and eco-evo-environmental drivers of $V_{c,\,max25}$ jointly determine global $V_{c,\,maxTg}$ variability

Our results first showed that enzyme kinetics' response was the dominant regulator of global $V_{c,\,maxTg}$ variability (Fig. 1). One possible explanation for this is the thermal sensitivity of the biochemical components involved in the RuBisCO enzyme kinetics [56–57], and thus the temperature-associated enzyme kinetics exert a direct, positive control on global $V_{c,\,maxTg}$ variability. To further elucidate the role of enzyme kinetics, we examined the empirical relationship between $V_{c,\,maxTg}$ and both environmental variables and leaf traits. Our analysis revealed that these predictors together accounted for 71% of $V_{c,\,maxTg}$ variability (Fig. S8a), with growing-season temperature being the dominant predictor (55%) (Fig. S8b), which is equivalent to the explanatory capacity to the enzyme kinetics statistical model (Fig. 1a). We also investigated the relative importance of each variable in predicting $V_{c,\,maxTg}$ variability, and identified the four most biotic and environmental variables, namely T_g, VPD, N_a, and elevation, in descending order of their relative importance (Fig. S9). These results collectively consolidate the dominant role of enzyme kinetics in global $V_{c,\,maxTg}$ variability.

Interestingly, several previous studies have reported weak, yet significant, negative relationships between $V_{c,\,max25}$ and site-mean growing temperature across global vegetative landscapes [6,29,58]. However, utilizing the largest global $V_{c,\,max}$ dataset compiled in this study, we found that the bivariate ordinary least-squares regression relationship between T_g and $V_{c,\,max25}$ was insignificant (Fig. S7). The key factors for explaining $V_{c,\,max25}$ variability, based on the sum of the Akaike weights, included leaf N content, VPD, soil pH, precipitation, and elevation, but not the T_g factor (Fig. 4). The minimal effect of T_g on $V_{c,\,max25}$ variability is also supported by another similar study that explored the best model of global $V_{c,\,max25}$ variability [62]. The potential cause for T_g's minimal effect might be related to the interactive effects among multiple interrelated environmental variables at a global scale (Fig. S7), with the apparent temperature effect being hidden by other associated variables, such as VPD, soil pH and precipitation (Fig. 4c-e). These results together suggest that such a temperature-induced $V_{c,\,max25}$ down-regulation is not sufficient to compensate the direct, positive effect of temperature on $V_{c,\,maxTg}$ through enzyme kinetics' response, and thus the positive temperature-$V_{c,\,maxTg}$ relationship is observed on a global scale. Collectively, our results imply that temperature-associated enzyme kinetics represent the first-order control on the variability of $V_{c,\,maxTg}$, which would increase with warmer growing season temperatures.

Second, we found that eco-evo-environmental drivers of $V_{c,\,max25}$ also importantly regulated global $V_{c,\,maxTg}$ variability (Fig. 1b). Thus, when we added the site-specific $V_{c,\,max25}$ to the temperature-associated enzyme kinetics, the explained variance of global $V_{c,\,maxTg}$ variability increased from 55% to 78%, even though the environmental variables and leaf traits altogether only explained 36% of $V_{c,\,max25}$ variability. These results altogether further generate two important implications: (1) improved characterization of $V_{c,\,max25}$ variability represents a very effective avenue for modelling global $V_{c,\,maxTg}$ variability, and (2) there remains a large proportion of $V_{c,\,max25}$ variability unexplained. Some of the unexplained $V_{c,\,max25}$ variability may be attributed to random measuring and sampling error in the evaluation of site-mean level $V_{c,\,max25}$ [41]. Alternatively, other important environmental factors (e.g., day length, soil moisture, and soil phosphorus (P) content) [6,29,37], unmeasured leaf traits (e.g., leaf P content, and leaf age) [3,59], species properties [29,30], and evolutionary history [60] might have also played important roles in driving $V_{c,\,max25}$ variability.

4.2. Limitations of optimality-constant model in capturing global $V_{c,\,maxTg}$ variability

There is an increasing interest in using the optimality-constant model for characterizing $V_{c,\,max25}$ variability for different applications, such as exploring the patterns of $V_{c,\,max25}$ variation across environmental gradients and parameterizing TBMs [12,18]. However, our study revealed that optimality-constant model performed relatively worse than the optimality-dynamic model in predicting global $V_{c,\,maxTg}$ variability, which was primarily attributed to the low predictive power (r^2 = 0.03) of the global $V_{c,\,max25}$ (Fig. 3b). Our findings thus highlight that optimality-constant model development should focus on improvements in capturing large-scale $V_{c,\,max25}$ variability. Notably, Peng et al. (2021) using the similar global dataset, reported a much higher $V_{c,\,max25}$ prediction (i.e. r^2 = 0.36) based on the optimality model [12], which is different from ours (r^2 = 0.03). Our further analysis suggested that this discrepancy is primarily linked to the inaccurate dataset (ultilizing the mean growing season temperature to approximate $V_{c,\,max}$ records without direct leaf temperature measurements) employed in Peng et al. (2021)'s data analysis [12]. Upon removing this inaccurate dataset (i.e. the $V_{c,\,max}$ records transformed from the mean growing season temperature) from the data analysis, the result comparable to ours (r^2 = 0.03) was then found (Fig. S2b). Such cross-comparisons also suggest that our analysis is rigorous and the generated result is trustworthy.

Our results also help identify that the poorer performance of optimality-constant model was mainly because of its failure to capture the dominant environmental and biotic regulations of global $V_{c,\,max25}$ variability, including N_a, VPD, soil pH, precipitation, and elevation (Fig. 4). In contrast, optimality-constant model only included five above-

ground environmental variables (i.e., T_g, VPD, PAR, C_a and elevation), with PAR and T_g alone capturing 98.5% of $V_{c,\,max25}$ variability in model predictions (Fig. 4a). Our results suggested that other variables, including belowground resource availability, also importantly regulated $V_{c,\,max25}$ variability (Fig. 4b-f). For example, N_a may reflect soil N availability [61] and has been proven to be an essential element for the construction of RuBisCO and other essential enzymes of the photosynthetic machinery and thus, N_a positively correlates with $V_{c,\,max25}$ [3,10]. Soil pH is a positive indicator of soil fertility, and higher soil pH usually means a lower nutrient cost for the construction of RuBisCO, thus favoring plant investment in greater $V_{c,\,max25}$ [27,37]. VPD and precipitation reflect the water stress status that importantly mediates plant stomatal behaviors, and high VPD and low precipitation generally reduce stomatal conductance and then increase the investment in $V_{c,\,max25}$ required to achieve optimal photosynthesis [13,62]. However, the relationship between soil pH and soil fertility is non-linear and dependent on plant species types [66], leading to a peak in $V_{c,\,max25}$ at intermediate soil pH levels (Fig. 4d). Similarly, the relationship between VPD and $V_{c,\,max25}$ exhibits non-linear and monotonic trends, rather than linear ones (Fig. 4c). These findings indicate that the use of linear regressions to derive explanatory variables for $V_{c,\,max25}$ has certain limitations, and future studies should consider non-linear explanations for key factors influencing $V_{c,\,max25}$ variability..

We also observed a negative effect of elevation alone on $V_{c,\,max25}$ (Fig. 4f), which contradicts previous findings that higher $V_{c,\,max25}$ is necessary to maximize carbon assimilation at high elevations when examining specific elevation transects [19,25,42]. This discrepancy may be attributed to complex and interactive effects of numerous interrelated environmental variables and leaf traits across large geographical extents (Fig. S5). As a result, the direct effect of elevation on $V_{c,\,max25}$ differs from the apparent effects often confounded by other interrelated variables, including N_a, VPD, soil pH and precipitation (Fig. 4b-f). Collectively, these explored dominant environmental and biotic variables further consolidate important drivers associated with global $V_{c,\,max25}$ (and thus $V_{c,\,maxTg}$) variability, which are not comprehensively represented in current optimality-constant model.

It is important to note that, besides soil nutrient availability, atmospheric N deposition also serves as a significant source of plant N uptake and demand, potentially affecting plant N concentration and subsequently $V_{c,\,max25}$ [10,67,68]. The magnitude of N deposition has fluctuated significantly over the time across the past decades [67,68]. However, our study lacks the necessary sampling year information to match global $V_{c,\,max}$ observations at a temporal scale, making it difficult to incorporate atmospheric N deposition into the analysis of factors driving global $V_{c,\,max25}$ variability. Further manipulative experiments and field monitoring are warranted to disentangle the relative importance of atmospheric N deposition and soil N availability in regulating the amount of RuBisCO enzyme and, subsequently, $V_{c,\,max25}$ variability. Besides, our study primarily examined the individual role of each environmental variable and leaf trait in driving global $V_{c,\,max25}$ variability, as previously mentioned. However, our findings also revealed the combined influence of climate variables, soil properties, and leaf traits on $V_{c,\,max25}$ variability (Fig. S10), suggesting that belowground properties can regulate $V_{c,\,max25}$ variability by indirectly affecting aboveground plant performance, while aboveground climate properties and plant traits can also influence $V_{c,\,max25}$ variability through indirect changes to belowground chemical and physical properties. Thus, to fully understand the intrinsic mechanisms of global $V_{c,\,max25}$ variability, both independent and interactive effects of explanatory variables must be thoroughly considered.

4.3. Considering the belowground resource constraints can improve the performance of optimality model

Our results showed that optimality-dynamic model considering the belowground resource constraints outperformed optimality-constant model in predicting both $V_{c,\,maxTg}$ and $V_{c,\,max25}$. This finding is consistent with recent studies, which suggest that edaphic variables can modify leaf economics [27,37]. The explanation is that edaphic variables influence the availability of soil nutrients and water and, consequently, the uptake costs for the construction and maintenance of photosynthetic machinery [27,37]. For example, a large fraction of leaf N (~50%) is allocated to the photosynthetic apparatus [63]. The cost of root N uptake varies with soil N availability and N acquisition strategies [37,64]. Similarly, soil water content influences the water uptake and transport costs necessary to deliver water to the leaves, which can affect the investments to support assimilation [13,27]. These considerations thus imply that the cost of belowground resource uptake and use could be a first-order priority for improving theoretical modelling of global $V_{c,\,max}$ variability.

In addition, we tested the environmental constraints on parameter c through two other scenarios, i.e., the five most important variables revealed by the statistical model (Fig. 4), and all the environmental and biotic variables. Overall, these three scenarios achieved very comparable results in predicting both $V_{c,\,maxTg}$ (Table S2) and $V_{c,\,max25}$ (Table S3), with the r^2 of 0.74-0.75 and RMSE of 13.51-13.80 for $V_{c,\,maxTg}$, and r^2 of 0.19-0.22 and RMSE of 19.64-21.63 for $V_{c,\,max25}$, respectively. These further lend us confidence in the following three interpretations: (1) our empirically identified five most important variables for $V_{c,\,max25}$ (Fig. 4) are ecologically meaningful, (2) these environmental and biotic variables could shape the $V_{c,\,max25}$ biogeography through regulating the site-specific variability in c (Fig. 6b-f), and (3) the optimality model when considering a variable c indeed improves the modelling of both $V_{c,\,max25}$ and $V_{c,\,maxTg}$.

The trends of c with environmental and biotic variables also agree with the idea that allocation costs vary with environmental gradients (Fig. 6b-f). Parameter c represents the total unit cost of constructing and maintaining the photosynthetic apparatus. Previous empirical studies have demonstrated that the construction cost of photosynthetic apparatus is influenced by belowground resource availability [26,27]. The proportion of leaf N allocated to the RuBisCO enzyme and the cost of root N uptake are influenced by soil nutrient and water availability, as well as associated N acquisition strategies [13,27,37,64]. Our results reveal that the cost factor c exhibits considerable variability and is influenced by edaphic properties. We identified a significantly negative relationship between parameter c and soil pH, indicating decreasing carbon costs at high soil pH levels with greater nutrient availability. This observation is supported by increased N allocation to the RuBisCO enzyme at higher pH levels [62] and lower leaf respiration and nutrient acquisition costs on alkaline, more fertile soils [27,65]. Furthermore, we discovered a significantly positive relationship between parameter c and soil α (a positive indicator of plant-available surface moisture) (results not shown), suggesting the decreasing carbon costs under drier conditions, accompanied by higher N allocation to the RuBisCO enzyme and subsequently, $V_{c,\,max25}$ [13,27].The parameterization of the cost functions allows including important environmental variables in a more mechanistic framework than the statistical regression modelling. However, our current parametrization of c is purely phenomenological. More research will be needed to understand how the cost functions relate to the biological processes of building and maintaining the photosynthetic machinery.

4.4. Implication and future directions

Our work generates two implications and future directions for understanding controls of global $V_{c,\,maxTg}$ variability and terrestrial biosphere modelling. First, the proposed statistical modelling approach captures the synthetic effects of enzyme kinetics, environmental variables and leaf traits on $V_{c,\,maxTg}$, providing an improved understanding of global $V_{c,\,maxTg}$ variability. The accurate prediction of $V_{c,\,maxTg}$ by considering both factors of enzyme kinetics and eco-evo-environmental drivers of $V_{c,\,max25}$ highlights that there are timescale-dependent

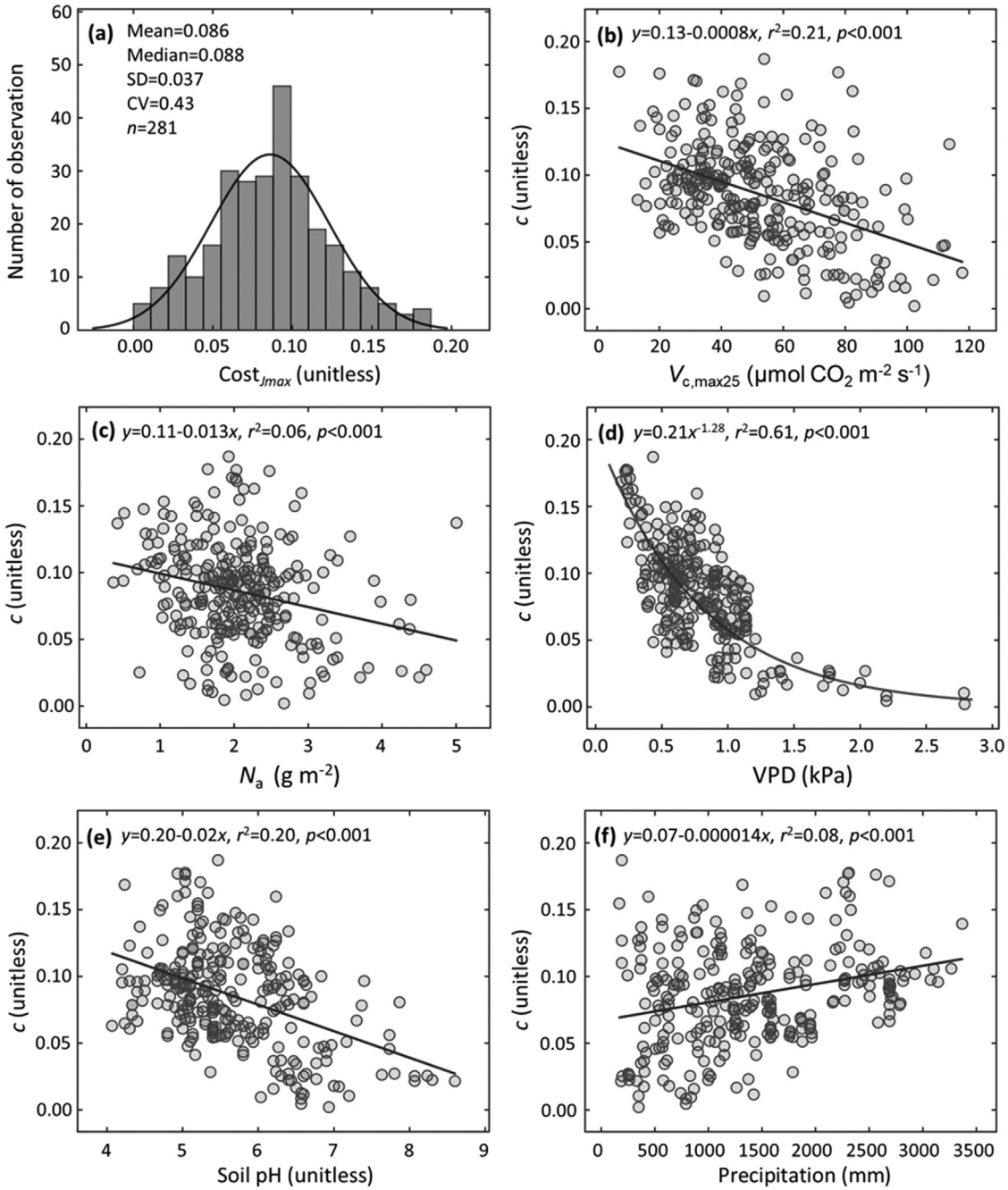

Fig. 6. The cost factor *c* constrained by edaphic variables shows large variability and is highly related to environmental variables and leaf traits. (a) The histogram distribution of the *c* across the global 281 sites. The bolded black line indicates the probability distribution function used to fit the histogram distribution; four statistical metrics are used to indicate the characteristics of *c* distribution, including mean, median, standard deviation (SD) and coefficient of variation (CV). (b-f) The ordinary least-square regression plots of *c* with the predictor of $V_{c,\,max25}$, area-based leaf nitrogen content (N_a), vapor pressure deficit (VPD), soil pH, and precipitation, respectively. r^2 and *p*-value represent the square of correlation coefficient and significance level, respectively.

mechanisms in regulating global $V_{c,\,maxTg}$ variability, with enzyme kinetics capturing instantaneous ecophysiological responses and explaining the dominant variance of global $V_{c,\,maxTg}$. Meanwhile, the second factor related to environmental variables and leaf traits likely represents the eco-evolutionary control of $V_{c,\,maxTg}$ through structuring $V_{c,\,max25}$ biogeography. Although the second factor is empirical, we further hypothesize that this might be associated with several key eco-evolutionary processes, e.g., abiotic filtering of species pools, competition for limited resources and the resultant trait-trait relationships subject to fundamental evolutionary principles [23,30,31,60]. However, further rigorous hypothesis testing related to eco-evolutionary processes of $V_{c,\,max25}$ through experimental manipulation and field observation approaches across large environmental gradients is still needed.

Second, the two factors that both regulate global $V_{c,\,maxTg}$ variability also provide an important benchmark and theoretical basis for evaluating current optimality models [8,19] while inspiring future improvements in the model representation of $V_{c,\,maxTg}$. For example, our results highlight that optimality-constant model may struggle primarily to capture $V_{c,\,max25}$ variability (Fig. 3b), probably due to an incomplete representation of environmental and biotic regulations on the cost functions. These findings suggest that these deficiencies should be addressed to more reliably model photosynthetic processes in TBMs [5,18]. To

achieve this, sophisticated observational and experimental studies are still needed to help elucidate the reasons for the near-zero predictability of $V_{c,\ max25}$ in current optimality-constant model, as well as exploring potential ways to mechanistically improve optimality modelling, and thus better constrain TBMs for improved representation of terrestrial photosynthesis, carbon cycling and climate change [1–2].

Data availability statement

All data used in our study will be uploaded to the Dryad Digital Repository after the manuscript is accepted.

Declaration of competing interest

The authors declare that they have no conflicts of interest in this work.

CRediT authorship contribution statement

Zhengbing Yan: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Matteo Detto:** Methodology, Supervision, Writing – review & editing. **Zhengfei Guo:** Methodology, Writing – review & editing. **Nicholas G. Smith:** Writing – review & editing. **Han Wang:** Writing – review & editing. **Loren P. Albert:** Writing – review & editing. **Xiangtao Xu:** Writing – review & editing. **Ziyu Lin:** Methodology. **Shuwen Liu:** Methodology. **Yingyi Zhao:** Methodology. **Shuli Chen:** Methodology. **Timothy C. Bonebrake:** Writing – review & editing. **Jin Wu:** Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing.

Acknowledgments

This work was supported by National Natural Science Foundation of China (31922090 and 31901086), Hong Kong Research Grant Council Early Career Scheme (27306020), and HKU Seed Fund for Basic Research (201905159005 and 202011159154). J. Wu was in part supported by the Innovation and Technology Fund (funding support to State Key Laboratories in Hong Kong of Agorobiotechnology) of the HKSAR, China. M.D was supported by the Carbon Mitigation Initiative of the Princeton University. NGS acknowledges support from the National Science Foundation (DEB-2045968) and Texas Tech University. The authors would like to thank Prof. Iain Colin Prentice for the constructive suggestions on the earlier draft of this manuscript, Mr. Yunke Peng for the sharing of globally compiled V_{cmax} data, and Dr. Calvin Ka Fai Lee for the proofread of this manuscript.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.fmre.2023.12.011.

References

[1] G.B. Bonan, S.C. Doney, Climate, ecosystems, and planetary futures: The challenge to predict life in Earth system models, Science 359 (2018) eaam8328.
[2] Y. Ryu, J.A. Berry, D.D. Baldocchi, What is global photosynthesis? History, uncertainties and opportunities, Remote Sens. Environ. 223 (2019) 95–114.
[3] A.P. Walker, A.P. Beckerman, L.H. Gu, et al., The relationship of leaf photosynthetic traits-V_{cmax} and J_{max} to leaf nitrogen, leaf phosphorus, and specific leaf area: A meta–analysis and modeling study, Ecol. Evol. 4 (2014) 3218–3235.
[4] G.V. Farquhar, S.V. von Caemmerer, J.A. Berry, A biochemical model of photosynthetic CO_2 assimilation in leaves of C_3 species, Planta 149 (1980) 78–90.
[5] A. Rogers, B.E. Medlyn, J.S. Dukes, et al., A roadmap for improving the representation of photosynthesis in Earth system models, New Phytol. 213 (2017) 22–42.
[6] A.A. Ali, C.G. Xu, A. Rogers, et al., Global-scale environmental control of plant photosynthetic capacity, Ecol. Appl. 25 (2015) 2349–2365.
[7] H. Croft, J.M. Chen, X.Z. Luo, et al., Leaf chlorophyll content as a proxy for leaf photosynthetic capacity, Glob. Change Biol. 23 (2017) 3513–3524.
[8] N.G. Smith, T.F. Keenan, I.C. Prentice, et al., Global photosynthetic capacity is optimized to the environment, Ecol. Lett. 22 (2019) 506–517.
[9] X. Qian, L. Liu, H. Croft, et al., Relationship between leaf maximum carboxylation rate and chlorophyll content preserved across 13 species, J. Geophys. Res.-Biogeosci. 126 (2021) e2020JG006076.
[10] J. Kattge, W. Knorr, T. Raddatz, et al., Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global-scale terrestrial biosphere models, Glob. Change Biol. 15 (2009) 976–991.
[11] A.A. Ali, C.G. Xu, A. Rogers, et al., A global scale mechanistic model of photosynthetic capacity (LUNA V1. 0), Geosci. Model Dev. 9 (2016) 587–606.
[12] Y.K. Peng, K.J. Bloomfield, L.A. Cernusak, et al., Global climate and nutrient controls of photosynthetic capacity, Commun. Biol. 4 (2021) 462.
[13] I.C. Prentice, N. Dong, S.M. Gleason, et al., Balancing the costs of carbon gain and water transport: Testing a new theoretical framework for plant functional ecology, Ecol. Lett. 17 (2014) 82–91.
[14] R.J. Norby, L.H. Gu, I.C. Haworth, et al., Informing models through empirical relationships between foliar phosphorus, nitrogen and photosynthesis across diverse woody species in tropical forests of Panama, New Phytol. 215 (2017) 1425–1437.
[15] M. Detto, X. Xu, Optimal leaf life strategies determine $V_{c,max}$ dynamic during ontogeny, New Phytol. 228 (2020) 361–375.
[16] Z.B. Yan, Z.F. Guo, S.P. Serbin, et al., Spectroscopy outperforms leaf trait relationships for predicting photosynthetic capacity across different forest types, New Phytol. 232 (2021) 134–147.
[17] H. Wang, O.K. Atkin, T.F. Keenan, et al., Acclimation of leaf respiration consistent with optimal photosynthetic capacity, Glob. Change Biol. 26 (2020) 2573–2583.
[18] X.Z. Luo, T.F. Keenan, Global evidence for the acclimation of ecosystem photosynthesis to light, Nat. Ecol. Evol. 4 (2020) 1351–1357.
[19] H. Wang, I.C. Prentice, T.F. Keenan, et al., Towards a universal model for carbon dioxide uptake by plants, Nat. Plants 3 (2017) 734–741.
[20] H. Wang, I.C. Prentice, I.J. Wright, et al., Leaf economics fundamentals explained by optimality principles, Sci. Adv. 9 (2023) eadd5667.
[21] N. Dong, I.C. Prentice, I.J. Wright, et al., Leaf nitrogen from the perspective of optimal plant function, J. Ecol. 110 (2022) 2585–2602.
[22] I.J. Wright, P.B. Reich, M. Westoby, Least-cost input mixtures of water and nitrogen for photosynthesis, Am. Nat. 161 (2003) 98–111.
[23] S.P. Harrison, W. Cramer, O. Franklin, et al., Eco-evolutionary optimality as a means to improve vegetation and land-surface models, New Phytol. 231 (2021) 2125–2141.
[24] N.G. Smith, T.F. Keenan, Mechanisms underlying leaf photosynthetic acclimation to warming and elevated CO_2 as inferred from least-cost optimality theory, Glob. Change Biol. 26 (2020) 5202–5216.
[25] Y.K. Peng, K.J. Bloomfield, I.C. Prentice, A theory of plant function helps to explain leaf-trait and productivity responses to elevation, New Phytol. 226 (2020) 1274–1284.
[26] N.H.A. Bahar, F.Y. Ishida, L.K. Weerasinghe, et al., Leaf-level photosynthetic capacity in lowland Amazonian and high-elevation Andean tropical moist forests of Peru, New Phytol. 214 (2017) 1002–1018.
[27] J. Paillassa, I.J. Wright, I.C. Prentice, et al., When and where soil is important to modify the carbon and water economy of leaves, New Phytol. 228 (2020) 121–135.
[28] J. Kattge, W. Knorr, Temperature acclimation in a biochemical model of photosynthesis: A reanalysis of data from 36 species, Plant Cell Environ. 30 (2007) 1176–1190.
[29] N.G. Smith, J.S. Dukes, Drivers of leaf carbon exchange capacity across biomes at the continental scale, Ecology 99 (2018) 1610–1620.
[30] I.J. Wright, P.B. Reich, M. Westoby, et al., The worldwide leaf economics spectrum, Nature 428 (2004) 821–827.
[31] P.B. Reich, The world-wide 'fast–slow' plant economics spectrum: A traits manifesto, J. Ecol. 102 (2014) 275–301.
[32] P. Meir, A. Shenkin, M. Disney, et al., in: Plant structure-Function Relationships and Woody Tissue Respiration: Upscaling to Forests from Laser-Derived Measurements in Advances in Photosynthesis and Respiration, Springer International Publishing, 2017, pp. 89–105.
[33] T.F. Domingues, P. Meir, T.R. Feldpausch, et al., Co-limitation of photosynthetic capacity by nitrogen and phosphorus in West Africa woodlands, Plant Cell Environ. 33 (2010) 959–980.
[34] T.F. Domingues, F.Y. Ishida, T.R. Feldpausch, et al., Biome-specific effects of nitrogen and phosphorus on the photosynthetic characteristics of trees at a forest-savanna boundary in Cameroon, Oecologia 178 (2015) 659–672.
[35] L.A. Cernusak, L.B. Hutley, J. Beringer, et al., Photosynthetic physiology of eucalypts along a sub-continental rainfall gradient in northern Australia, Agric. For. Meteorol. 151 (2011) 1462–1470.
[36] O.K. Atkin, K.J. Bloomfield, P.B. Reich, et al., Global variability in leaf respiration in relation to climate, plant functional types and leaf traits, New Phytol. 206 (2015) 614–636.
[37] V. Maire, I.J. Wright, I.C. Prentice, et al., Global effects of soil and climate on leaf photosynthetic traits and rates, Glob. Ecol. Biogeogr. 24 (2015) 706–717.
[38] N.G. Smith, J.S. Dukes, LCE: Leaf carbon exchange dataset for tropical, temperate, and boreal species of North and Central America, Ecology 98 (2017) 2978.
[39] N. Dong, I.C. Prentice, B.J. Evans, et al., Leaf nitrogen from first principles: Field evidence for adaptive variation with climate, Biogeosciences 14 (2017) 481–495.
[40] H. Wang, S.P. Harrison, I.C. Prentice, et al., The China plant trait database: Towards a comprehensive regional compilation of functional traits for land plants, Ecology 99 (2018) 500.

[41] K.J. Bloomfield, L.A. Cernusak, D. Eamus, et al., A continental-scale assessment of variability in leaf traits: Within species, across sites and between seasons, Funct. Ecol. 32 (2018) 1492–1506.
[42] H.Y. Xu, H. Wang, I.C. Prentice, et al., Predictability of leaf traits with climate and elevation: A case study in Gongga Mountain, China, Tree Physiol. 41 (2021) 1336–1352.
[43] I. Harris, P.D. Jones, T.J. Osborn, et al., Updated high resolution grids of monthly climatic observations – the CRU TS3.10 dataset, Int. J. Climatol. 34 (2014) 623–642.
[44] G.P. Weedon, G. Balsamo, N. Bellouin, et al., The WFDEI meteorological forcing data set: WATCH forcing data methodology applied to ERA-interim reanalysis data, Water Resour. Res. 50 (2014) 7505–7514.
[45] T.W. Davis, I.C. Prentice, B.D. Stocker, et al., Simple process-led algorithms for simulating habitats (SPLASH v.1.0): Robust indices of radiation, evapotranspiration and plant-available moisture, Geosci. Model Dev. 10 (2017) 689–708.
[46] J.A. Quebbeman, J.A. Ramirez, Optimal allocation of leaf-level nitrogen: Implications for covariation of V_{cmax} and J_{max} and photosynthetic downregulation, J. Geophys. Res.-Biogeosci. 121 (2016) 2464–2475.
[47] D.E. Goldberg, Genetic Algorithms in Search, Optimization, and Machine Learning, Addison-Wesley, Reading, MA, 1989.
[48] S. Hamblin, On the practical usage of genetic algorithms in ecology and evolution, 4 (2013) 184-194.
[49] C.F. Dormann, J. Elith, S. Bacher, et al., Collinearity: A review of methods to deal with it and a simulation study evaluating their performance, Ecography 36 (2013) 27–46.
[50] S. Doetterl, A. Stevens, J. Six, et al., Soil carbon storage controlled by interactions between geochemistry and climate, Nat. Geosci. 8 (2015) 780–783.
[51] V. Calcagno, C. de Mazancourt, glmulti: An R package for easy automated model selection with (generalized) linear models, J. Stat. Softw. 34 (2010) 1–29.
[52] C. Terrer, S. Vicca, B.A. Hungate, et al., Mycorrhizal association as a primary control of the CO_2 fertilization effect, Science 353 (2016) 72–74.
[53] P. Breheny, W. Burchett W, Visualization of regression models using visreg, R. J. 9 (2017) 56–71.
[54] I.M. Sobol', D. Asotsky, A. Kreinin, et al., Construction and comparison of high-dimensional Sobol' generators, Wilmott 2011 (2011) 64–79.
[55] A. Saltelli, P. Annoni, I. Azzini I, et al., Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index, Comput. Phys. Commun. 181 (2010) 259–270.
[56] W. Yamori, K. Hikosaka, D.A. Way, Temperature response of photosynthesis in C_3, C_4, and CAM plants: Temperature acclimation and temperature adaptation, Photosynth. Res. 119 (2014) 101–117.
[57] J. Galmes, M.V. Kapralov, L.O. Copolovici, et al., Temperature responses of the Rubisco maximum carboxylase activity across domains of life: Phylogenetic signals, trade-offs, and importance for carbon gain, Photosynth. Res. 123 (2015) 183–201.
[58] D.P. Kumarathunge, B.E. Medlyn, J.E. Drake, et al., Acclimation and adaptation components of the temperature dependence of plant photosynthesis at the global scale, New Phytol. 222 (2019) 768–784.
[59] J. Wu, A. Rogers, L.P. Albert, et al., Leaf reflectance spectroscopy captures variation in carboxylation capacity across species, canopy environment and leaf age in lowland moist tropical forests, New Phytol. 224 (2019) 663–674.
[60] Z.B. Yan, J. Sardans, J. Peñuelas, et al., Global patterns and drivers of leaf photosynthetic capacity: The relative importance of environmental factors and evolutionary history, Glob. Ecol. Biogeogr. (2023), doi:10.1111/geb.13660.
[61] J. Firn, J.M. McGree, E. Harvey, et al., Leaf nutrients, not specific leaf area, are consistent indicators of elevated nutrient inputs, Nat. Ecol. Evol. 3 (2019) 400–406.
[62] X.Z. Luo, T.F. Keenan, J.M. Chen, et al., Global variation in the fraction of leaf nitrogen allocated to photosynthesis, Nat. Commun. 12 (2021) 1–10.
[63] J.R. Evans, J.R. Seemann, The allocation of protein nitrogen in the photosynthetic apparatus: Costs, consequences, and control, Photosynthesis 8 (1989) 183–205.
[64] J.B. Fisher, S. Sitch, Y. Malhi, et al., Carbon cost of plant nitrogen acquisition: A mechanistic, globally applicable model of plant nitrogen uptake, retranslocation, and fixation, Glob. Biogeochem. Cycle 24 (2010) GB1014.
[65] O.K. Atkin, M.H. Turnbull, J. Zaragoza-Castells, et al., Light inhibition of leaf respiration as soil fertility declines along a post-glacial chronosequence in New Zealand: An analysis using the Kok method, Plant Soil 367 (2013) 163–182.
[66] H. Lambers, R.S. Oliveira, in: Plant Physiological Ecology, Springer Nature Switzerland, Switzerland, 2019, pp. 301–331.
[67] X.J. Liu, Y. Zhang, W.X. Han, et al., Enhanced nitrogen deposition over China, Nature 494 (2013) 459–463.
[68] J. Peñuelas J, I.A. Janssens, P. Ciais, et al., Anthropogenic global shifts in biospheric N and P concentrations and ratios and their impacts on biodiversity, ecosystem productivity, food security, and human health, Glob. Change Biol. 26 (2020) 1962–1985.

Zhengbing Yan is a staff scientist at the State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Science, with broad research interest in plant functional ecology, biogeochemical cycles, vegetation spectroscopy and global ecology.

Jin Wu is an assistant professor at the University of Hong Kong, with an interest in ecosystem ecology, plant physiological ecology, climate science, vegetation spectroscopy and earth system science. He was funded by National Outstanding Youth Science Fund Project of National Natural Science Foundation of China (2020).