# A GLOBAL DATABASE OF PHOTOSYNTHESIS MODEL PARAMETERS, AND PHYLOGENETICALLY CONTROLLED ANALYSIS OF PHOTOSYNTHETIC RESPONSES FROM EVERY MAJOR TERRESTRIAL PLANT CLADE

**RUNNING HEADLINE**: Global phylogenetic photosynthesis

Mina Rostamza[1,2,3], and Gordon G. McNickle[1,2,*]

[1]Department of Botany and Plant Pathology, Purdue University, West Lafayette, IN 47907-2054

[2]Center for Plant Biology, Purdue University, West Lafayette, IN 47907-2054

[3]Current address: Plant Molecular and Cellular Biology Laboratory, Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037

*Author for correspondence: gmcnickle@purdue.edu, 765-494-4645

**KEYWORDS**: Photosynthesis, Farquhar model, carbon dioxide, climate change, phylogeny, plant functional type

**ABSTRACT**

Plant photosynthesis is a major part of the global carbon cycle and climate system. Carbon capture by C3 plants is most often modelled using the Farquhar-von-Caemmerer-Berry (FvCB) equations. We undertook a global synthesis of all parameters required to solve the FvCB model. The publicly available dataset we assembled includes 3663 observations from 359 different plant species among 96 taxonomic families coming from every major vascular plant clade (lycophytes, ferns, gymnosperms, magnoliids, eudicots and monocots). Geographically, the species in the database have distributions that span the majority of the globe. We used the model to predict photosynthetic rates for the average plant in each major terrestrial plant clade and find that generally plants have dramatically increased their photosynthetic abilities through evolutionary time, with the average monocot (the youngest clade) achieving maximum rates of photosynthesis almost double that of the average lycophyte (the oldest clade). However, there was no evidence of niche conservatism with most variance occurring within, rather than among clades (K=0.357, p=0.001). We also solved the model for different average plant functional types (PFTs) and find that herbaceous species generally have much higher rates of photosynthesis compared to woody plants. Indeed, the maximum photosynthetic rate of graminoids is almost three times the rate of the average tree. The resulting functional responses to increasing $CO_2$ suggest that most groups are already at or near their maximum rate of photosynthesis. Unfortunately, only graminoids seem to have the capacity to continue to increase photosynthetic rates with increasing $CO_2$ concentrations in the atmosphere. We view this as version 1.0 of a database for global photosynthesis parameters and functional responses and hope that the publicly available dataset can improve models of photosynthesis.

## INTRODUCTION

Plant photosynthesis is a major factor in the global climate system. Indeed, the annual flux of atmospheric carbon (C) through the leaves of terrestrial plants is estimated to be $1 \times 10^{15}$ g $yr^{-1}$ (Beer *et al.*, 2010, Hetherington & Woodward, 2003). Carbon capture by C3 plants is most often modelled using models derived from the Farquhar-von-Caemmerer-Berry (FvCB) equations (Farquhar *et al.*, 1980, Farquhar & Wong, 1984, Sharkey *et al.*, 2007, Von Caemmerer, 2000). The FvCB model is a process based physiological model that accurately describes the rate of photosynthesis across light levels, and across both $CO_2$ and $O_2$ concentrations. In its modern form, the FvCB model also accounts for triose phosphate limitation (Lombardozzi *et al.*, 2018, Mcclain & Sharkey, 2019). Indeed, a version of the FvCB model forms the basis for most physiological, ecological, and earth system models that include plants (Rogers *et al.*, 2017).

Models that incorporate plant photosynthesis require accurate parameter estimates, estimates which are often spread across decades and journals and may be difficult to find for specific taxa. There have been several syntheses and meta-analyses that focus on two parameters of the FvCB model, $V_{c,max}$ and $J_{max}$ (E.g. Kattge & Knorr, 2007, Walker *et al.*, 2014, Wullschleger, 1993), but we are unaware of any attempt at a global synthesis of the full suite of at least 12 parameters needed to fully predict photosynthetic rates across the plant kingdom. In addition, the modern FvCB model of photosynthesis is well known to be over-parameterised (Qian *et al.*, 2012), and modern techniques for curve fitting and parameter estimation can benefit from better prior information. For example, Bayesian methods can work from a known prior distribution to enhance the ability to accurately estimate parameters (e.g. Patrick *et al.*, 2009). Thus, collecting all available parameter estimates into one database would enhance our ability to model global photosynthesis, as well as our ability to estimate parameters for new taxa.

Here, we describe a synthesis of all FvCB parameters where at least one parameter was estimated for a given species. The summary includes parameter estimates from 359 different plant species from 96 taxonomic families coming from every major vascular plant clade (lycophytes, ferns, gymnosperms, magnoliids, eudicots and monocots) whose distributions span the majority of the globe. The parameter estimates are presented using a number of summary statistics and probability density histograms. We also solve the FvCB model using the full range of parameter estimates to generate predictions about the breadth of plant photosynthetic responses across major vascular plant clades, plant functional types, and individual leaves. The full dataset is publicly available.

## MATERIALS AND METHODS

### *The FvCB photosynthesis model*

Here, we briefly describe the equations of the FvCB model (Farquhar *et al.*, 1980, Sharkey *et al.*, 2007, Von Caemmerer, 2000). The most basic modern FvCB photosynthesis approach for C3 plants assumes that the rate of carbon assimilation ($A$) in photosynthesis is co-limited by either carbon ($A_c$), light ($A_j$) or TPU ($A_p$) according to:

$$A = \min\left(A_c, A_j, A_p\right) - R_d, \qquad \text{Eqn 1}$$

where $R_d$ is the daytime respiration rate (See Table 1 for units).

The carbon-limited portion of assimilation by photosynthesis in Eqn 1 is given by:

$$A_c = \frac{V_{c,max} C_i}{C_i + K_c\left(1 + \frac{O_i}{K_o}\right)}, \qquad \text{Eqn 2}$$

where $C_i$ and $O_i$ are the intracellular concentrations of $CO_2$ and $O_2$ at the site of ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) activity respectively; $K_c$ and $K_o$ are the RuBisCO half saturation constants for $CO_2$ and $O_2$, respectively, and; $V_{c,max}$ is the maximum possible rate of photosynthesis. Half saturation constants are often called the enzyme “affinity” for the substrate,

but in reality they have limited biological meaning and simply describe the shape of the curvature of the functional response (Mcnickle & Brown, 2014).

The light limited portion of photosynthesis is given by:

$$A_j = \frac{JC_i}{4.5C_i + 10.5\Gamma^*}, \qquad \text{Eqn 3}$$

where $J$ is the realised electron chain transport determined by light; $\Gamma^*$ is the minimum partial pressure of $CO_2$ where carbon assimilation balances respiration (i.e. $A = R_d$) generally called the $CO_2$ compensation point; and $C_i$ is as above.

For the purposes of a global summary, the variable $J$ has been determined in several ways over the years. The most common approach following Farquhar and Wong (1984) $J$ was found by solving for the root of a simple quadratic equation:

$$\theta J^2 - J(J_{max} + \alpha Q) + J_{max}\alpha Q = 0, \qquad \text{Eqn 4a}$$

Where $Q$ is the light photon flux density striking the leaf (µmol photons $m^{-2}$ $s^{-1}$), $J_{max}$ is the light saturated maximum possible rate of electron transport; $\theta$ represents curvature of the light response; and; $\alpha$ is the efficiency of light conversion. Since this is a quadratic equation, and since negative values of $J$ have no biological meaning, we can use the quadratic formula to find the positive root of Eqn 4a where:

$$J = \frac{J_{max} + \alpha Q - \sqrt{(J_{max} + \alpha Q)^2 - 4\theta J_{max}\alpha Q}}{2\theta}. \qquad \text{Eqn 4b}$$

For some reason, most modern papers use different symbols from the original formulation. In addition, though significantly less common (Buckley & Diaz-Espejo, 2015), some authors use an approximation of eqns 4 where $J$ is approximated according to:

$$J = \frac{aQ}{\sqrt{1 + \frac{(aQ)^2}{{J_{max}}^2}}}, \qquad \text{Eqn 5}$$

where $a$ represents the efficiency of light conversion (similar to $\alpha$ in eqn 4), and $Q$ and $J_{max}$ have their usual meanings. Note that $a$ in Eqn 5 is more typically written in the literature as the Greek letter '*alpha*', but we altered this to avoid confusion with Eqn 4.

Finally, TPU limitation of photosynthesis in its most detailed form is given by:

$$A_p = \frac{3T_p(C_i - \Gamma^*)}{C_i - \Gamma^*(1 + 3r)}, \qquad \text{Eqn } 6a$$

where $r$ is a unitless scalar related to the proportion of glycolate recycled in chloroplasts (where $0 < r < 1$); $T_p$ is the rate of TPU, and the other parameters have their usual meanings given above (Table 1). However, it appears to be most common to assume that no glycolate is recycled (i.e. $r = 0$), and then Eqn 6a simplifies to:

$$A_p = 3T_p. \quad Eqn\ 6b$$

We note that there are even more complex versions than we have detailed here in Eqns 1-6. These more complex versions may include parameters for various resistances to the movement of CO2 into the leaf (Collatz *et al*., 1992) and transpiration (Farquhar & Sharkey, 1982, Farquhar & Wong, 1984). However, leaf resistance and transpiration could be considered their own rather large sub-fields independent from many attempts to estimate FvCB model parameters. Thus, we purposefully omitted resistances from our survey of the literature to make the scope of the literature synthesis more tractable. Also, models that include detailed temperature relationships via the Arrhenius equation are also somewhat common (Kattge & Knorr, 2007). Again however, to limit the scope of our synthesis we did not collect the parameters of these Arrhenius rate parameters for a more complex temperature dependent FvCB model.

*Literature search*

Briefly, we began with a search for "Farqhuar photosynthesis models" in the Web of Science, read each paper carefully, and extracted values for all parameters of the FvCB model (Table 1). The search was expanded by using the bibliographies of these papers. In total data was extracted from 202 papers. Units were standardized across studies for comparison. We also collected as many methodological details about the conditions in the cuvette as possible to attempt to understand sources of variation in the data that were not caused by taxonomy (E.g. humidity, temperature, vapour pressure deficit, light levels, $CO_2$ pressure). The full details of our systematic literature search and data extraction methods and criteria are given in the Supplementary Information.

We found that studies could generally be broadly organised into three types: temperature effects, leaf nitrogen effects, or mean plant effects. First, temperature is well known to affect the biochemical reactions in photosynthesis and we wanted to account for this variability (Kattge & Knorr, 2007). Thus, when temperature was the treatment parameter, values at each temperature were recorded separately. These rows are labelled as type '$T$' in the data file. Second, leaf nitrogen can also be related to photosynthesis parameters and we wanted to account for this variability as well (Kattge *et al.*, 2009, Walker *et al.*, 2014). Thus, when parameters were estimated separately by leaf nitrogen, each leaf nitrogen level was recorded separately. Nitrogen was sometimes reported as percent dry weight, and sometimes as mass per leaf area. Either were recorded, but in separate data columns. These rows are labelled as type '$N$' in the data file. Studies that reported just one value for each parameter and species (i.e. a species mean) are labelled as type '$M$' in the data file. Thus, there are three types of data: (i) mean values (type $M$); (ii) leaf temperature manipulations (type $T$), and; (iii) leaf nitrogen differences (type $N$). However, when available, temperature was recorded for type $M$ and type $N$ studies. Similarly, when available, leaf nitrogen was recorded for type $T$ and type $M$ studies.

*Taxonomy and biogeography*

Species names reported in the original papers were checked the National Center for Biotechnology Information's species taxonomy database using the brranching::phylomatic_names function in R. In general, this just updated any outdated species names to the most modern accepted name. In one case, we had to manually change a species names to a taxonomic synonym to match to the database: *Echinochloa crus-galli* (L.) Beauv was changed to an accepted synonym *Digitaria hispidula* (Retz.) Willd. The updated species names were then pruned from the plant megatree of Zanne *et al.* (2014) to visually represent the taxonomic coverage of the dataset using the brranching library in R (v. 0.6; Chamberlain, 2020). The search produced data for 21 C4 graminoid species, even though the basic FvCB model is meant to capture C3 photosynthesis (Collatz *et al.*, 1992). However, some parameters are common between C3 and C4 photosynthesis models, thus, we included these data in our database. Thus, species were also assigned to the following broad plant functional types (PFTs) based on growth form: C3 graminoid, C4 graminoid, Forb, Vine/Climbing, Shrub, or Tree. Finally, we recorded growth habit of each species in the following categories: annual, biennial, perennial, or crop for herbaceous species; deciduous, evergreen, or crop for woody species; and fern, tree, or club moss for ferns. Growth habit information was not available for three rare and exotic species and was recorded as NA.

Not all studies reported a location of plant material collection, measurement, or the cultivar examined. Thus, to obtain a sense of the geographic coverage of species in this dataset, we used the Botanical Information and Ecology Network (BIEN; Maitner *et al.*, 2018) to obtain museum occurrence records for each species in our global database. We then mapped these occurrence records. The resulting map shows the global distribution of all species in our dataset. Importantly this map is not the global distribution of measurement locations. However, the resulting map provides window into the range distributions of all species that have been studied for photosynthesis, and therefore details what regions of the world have good coverage of at least

178 approximate species level photosynthetic data. In addition, the resulting map also shows what
179 regions would benefit from increased empirical attention to improve global models.
180
181 *Data summary*
182 To create a global summary table, we treated the three data types differently. For studies
183 that report only mean values (type "M"), we used all values in the summary. For temperature
184 studies (type "T"), we only used the value nearest to 25°C. This was done because 25°C was the most
185 common temperature used in studies that did not manipulate leaf temperature and most
186 temperature studies measure the same leaf across many temperatures. Our approach of using only
187 one value per leaf avoids pseudoreplication. For leaf nitrogen studies (Type "N"), we averaged values
188 across leaf N amounts. Because different leaf nitrogen contents represented individual plants in
189 each study, this method creates a species mean and avoids pseudoreplication. To generate a global
190 summary, we calculated the mean, standard deviation, median, maximum, minimum, skew, and
191 kurtosis for each parameter from Eqns 1-6.
192
193 Data were also summarised by major taxonomic clade and PFT. For these summary rows,
194 there were generally too few categories in most groups to create density distributions, and we
195 report only the mean, standard deviation, and sample size. For sample sizes that were n<3, we
196 report the standard deviation as NA. For these summaries, we did not separate the *T*, *M*, and *N* data
197 types because many taxa were only studied once.
198
199 *Breadth of model outcomes*
200 It is useful to have a large database of FvCB model parameters, but the raw parameters do
201 not show the breadth of plant photosynthetic responses on their own, since 12 parameters (Table 1)
202 may combine in many different ways to produce the same rate of photosynthesis. To examine the
203 range of predicted photosynthetic responses, we used the parameter estimates to actually solve the

204 FvCB model for plant carbon assimilation rate by solving for model predicted $A - C_i$ and $A - Q$
205 curves. We did this in three ways. All three of these analysis uses all three types of data (M, T and N)
206 in order to show the breadth of responses.
207
208 First, we solved the model for the average hypothetical plant that describes each major
209 plant clade. To do this, each parameter value was set equal to the average of the clade, and the
210 predicted $A - C_i$ and $A - Q$ curves were solved. In addition, all parameters were set to the upper
211 and lower 95% confidence intervals around the mean, and the equations were solved again to give a
212 sense of variation within the clade. Missing values were replaced with the global mean in the clade
213 analysis. This let us compare the average photosynthetic functional response of lycophytes, ferns,
214 gymnosperms, magnoliids, eudicots, and monocots.
215
216 Second, we solved the model for the average hypothetical plant that describes each PFT.
217 This was done as above with the mean parameter values and the upper and lower confidence
218 intervals. Here, missing values were replaced with the mean of the appropriate clade (e.g. missing
219 graminoid parameters were replaced with the monocot mean, while missing vine or shrub
220 parameters with the eudicot mean). We included C4 species in this analysis.
221
222 Third, we went down each of the 3663 rows of our dataset and solved the model for every
223 individual leaf for which we had data. However, no study in our synthesis estimated all parameters
224 of the FvCB equation. Thus, to fill in gaps for any row, we used the global mean value for any missing
225 parameter values. In addition, the average $A_{max}$ for each species was calculated from these model
226 runs and then drawn onto the phylogeny. We used Bloomberg's K to examine phylogenetic signal in
227 the species level data using phytools in R (Revell, 2012).
228
229 **RESULTS**

*Taxonomy and biogeography*

In total, we obtained at least one parameter estimate from 359 species in 96 families (Fig 1A, Table S1). The data included all major vascular plant clades including: lycophytes (2 species, 1 family), ferns (33 species, 16 families), gymnosperms (23 species, 3 families), and angiosperms (303 species, 77 families). Angiosperms can be further separated into three more sub-clades comprised of monocots (62 species, 7 families), eudicots (235 species, 68 families), and magnoliids (6 species, 3 families). Tables S1 and S2 include more detailed summaries breaking the available data up among each of the 96 taxonomic families and within the six clades. In addition, Tables S3 and S4 contain more detailed summaries breaking the data up by PFT and growth habit.

The occurrence data from BIEN shows that the majority of data comes from North America, Europe, Australia, New Zealand, and Japan (Fig 1B). Coverage in Africa was patchy, with the best coverage in South Africa and parts of west Africa. However, excluding regions in Africa that mostly do not contain plants (i.e. the Sahara and Namib Deserts), there are entire countries in southern Africa for which very few species have been studied (e.g. Angola, Morocco, Nigeria). Similarly, coverage was spotty in northern Asia (E.g. Russia, Kazakhstan, Mongolia), Indochina, Indonesia and South America (e.g. Chile, Argentina, Uruguay, Paraguay). This is problematic, because many of these regions of Africa, Asia, and South America with patchy data are known diversity hotspots where a small number of taxonomic samples may not represent the average plant in those regions (Myers *et al.*, 2000). There was, however, a surprising amount of data from the southern foothills of the Himalayan mountains, and from Colombia and Ecuador. Regions of Brazil's Amazon rainforest are also sparsely measured, with much of the Brazilian data appearing to have come from the southern grassland regions.

Ignoring national borders, the geographic coverage suggests that most named ecosystem 'types' (e.g. sensu Whittaker, 1975) also have good coverage. Temperate and boreal forests and

256 grasslands have particularly good coverage (Fig 1B). However, given the high species diversity of
257 tropical ecosystems, there are likely important gaps in our understanding of the diversity of
258 photosynthetic responses in tropical regions across the globe. The arctic regions of Europe have very
259 good coverage, but there is little data from arctic Russia and large gaps in coverage for arctic North
260 America. Similarly, for grasslands, there is very good coverage in Australia and North America, but
261 relatively little for the grasslands of Asia, Africa, and South America.
262
263 *Summary statistics and parameter distributions*
264 By far, $V_{cmax}$ (n=1364), $J_{max}$ (n=961), and, to a lesser extent, $T_p$ (n=171) were the most
265 frequently estimated and reported parameter values (Table 2, 3). In general, most of the probability
266 density distributions of observed parameters were skewed or possibly multi-modal (Table 2, 3, Fig
267 S1, Fig S2). It is noteworthy that the minimum and maximum reported values of most parameters
268 differed by as much as four orders of magnitude across all vascular plants. Also, the coefficient of
269 variation in most cases was 0.5 or higher, suggesting high dispersion of the data among species.
270 However, the data show that the majority of this within- and among-species variation over orders of
271 magnitude was driven by methodological differences among research groups (Supplementary
272 information). For example, a few species were studied many times by different groups, and variation
273 in reported parameters for these species were largely explained by the nitrogen content of the
274 leaves each group measured, and the VPD tolerance they used when taking measurements (Table
275 S5, Fig S3). Similarly, among all species in our data base, methodological choices of different
276 research groups explained 60% of the variation in parameter estimates (Fig S4, Fig S5). Thus, once
277 methods are controlled, the variation is driven by species differences.
278
279 Among the six major clades, eudicots had the most data (n=2037 for $V_{c\,max}$) followed by
280 gymnosperms (n=461 for $V_{c\,max}$) and monocots (n=400 for $V_{c\,max}$). In general, for both $V_{cmax}$ and
281 $J_{max}$ there was a clear order to the means such that monocots > eudicots > magnoliids >

282 gymnosperms > lycophytes. This suggests that plants have become increasingly adept at
283 photosynthesis through evolutionary time. However, ferns do not seem to fit into this schema with
284 ferns in-between magnoliids and gymnosperms. The half saturation constants ($K_c$ and $K_o$) were only
285 estimated twice for gymnosperms and never for lycophytes, ferns, and magnoliids. Better sampling
286 of lycophyte, fern and magnoliid species may resolve phylogenetic differences in the future.
287
288 For PFTs, most researchers studied trees (n=2077 for $V_{c\ max}$) followed by C3 graminoids
289 (n=388 for $V_{c\ max}$) and forbs (n=222 for $V_{c\ max}$). There were relatively few vines, and shrubs. C4
290 graminoids were often presented alongside C3 plants, even though they do not quite conform to the
291 FvCB model. When they are included there was a clear order, for both $V_{cmax}$ and $J_{max}$ means such
292 that C4 graminoids > C3 graminoids > forbs > shrubs > trees > vines.
293
294 *Model outcomes*
295 We also used the parameters to solve the FvCB model to explore predicted photosynthesis
296 rates. For comparison among the six major clades represented in our data, we note that $A_{max}$ was
297 significantly different such that monocots > eudicots > gymnosperms (Fig 2A, B). There were too few
298 parameter estimates to draw confidence intervals for lycophyte, fern, and magnoliid clades, but
299 readers should assume they are very wide, and we hesitate to draw many conclusions without more
300 data. However, if the patterns stand, these differences among clades suggest that on average land
301 plants have substantially increased their photosynthetic ability over evolutionary time.
302
303 For PFT photosynthetic rates, $A_{max}$ was significantly different such that C3 graminoids >
304 shrubs > trees, but the confidence interval around forbs was so large that this group was not
305 significantly different from any of these groups (Fig 2C,D). Like magnoliids, there were too few
306 parameter estimates for vines to draw confidence intervals for these PFTs, and we hesitate to draw
307 conclusions about these PFTs.

We also went down our dataset row by row to generate the entire breadth of photosynthetic responses of every leaf and species for which we had data. When parameters were missing, we replaced them with the global mean (Table 1). This created 3663 $A - C_i$ and 3663 $A - Q$ curves (Fig 2E,F). When all 3663 curves were plotted on the same axis but with slightly transparent lines, the darker regions show the most common responses and the lighter regions show fewer common responses. Not surprisingly, the darkest lines largely appear to trace the average monocot and eudicot because these were best taxonomically sampled groups. The breadth of maximum photosynthesis values ranged from slightly negative rates of photosynthesis (-1.6 µmol C $m^{-2}$ $s^{-1}$) to 59.7 µmol C $m^{-2}$ $s^{-1}$.

When these data were averaged by species, there was no phylogenetic niche conservatism (Fig 1A, K=0.357, p=0.001) indicating that the majority of variation occurred within clades rather than among clades. Therefore, even though clades do appear to differ on average (Fig 2A,B), each clade is just as likely to contain some individual species with either low or high $A_{max}$ (Fig 1A).

**DISCUSSION**

Photosynthesis is an important part of the earth climate system. Enormous quantities of $CO_2$ pass through terrestrial plants each year across the globe. As such, good parameterisation of photosynthetic models is key to an understanding of phenomena ranging from simple leaf level physiology to global atmospheric dynamics. Here, we undertook a synthesis of parameter values in the literature. The data set described by this paper includes 3663 rows of data across 359 different plant species from 96 taxonomic families that spanned all major vascular plant clades including lycophytes, ferns, gymnosperms, magnoliids, monocots, and eudicots (Fig 1; Table 2). Our aim was that modellers with diverse interests and questions could make use of this dataset, either for some general plant summarised by clade or PFT (Table 2,3) or to generate prior probability distributions

for modern Bayesian model fitting approaches (Fig S1, S2). However, it is also possible with these data to get very detailed indeed, for example by family (Table S1, Table S2) or growth habit (Table S3, Table S4).

Generally, it is agreed that climate change has been somewhat slowed as plants have taken up more carbon that scientists originally expected (Beer *et al.*, 2010, Berendse *et al.*, 2001). Indeed, from our modelled $A - C_i$ curves, we can predict the average functional response of each major plant clade (Fig 1A) or PFT (Fig 1C) to rising $CO_2$ levels. Because our data go back to the 1970s, the average partial pressure of intracellular $CO_2$ across our data set was only 28.4 Pa (~280 ppm), while at the time of this writing the partial pressure of $CO_2$ in the atmosphere was approximately 42 Pa (~415 ppm). The model output shows that the average angiosperm has been able to take advantage of the increased CO2 in the atmosphere in this period from the 1970s to the present (Fig 2A). However, the PFT data show that the increased photosynthesis may have been largely by grassland graminoids and forbs, with most trees and shrubs already at their maximum photosynthetic rate with $CO_2$ at 28.4 Pa. Indeed, on average, our data suggest monocots may have increased their photosynthesis by more than 32% as intracellular $CO_2$ pressures rise from 28.4 to 42 Pa. Unfortunately, if we continue this trend and cast forward to the IPCC predictions for the year 2100 (Ipcc, 2012) of 74 – 103.4 Pa (~730-1020 ppm) $CO_2$, our model results suggest that most clades and PFTs are already near or at their maximum photosynthetic rates (Fig 1A, C). Our data suggest that only graminoids, have much capacity for increased $CO_2$ assimilation beyond their current assimilation rates as we move toward the expected $CO_2$ composition of the atmosphere by the year 2100. All that said, the lack of phylogenetic signal at the species level (Fig 1A) may mean that any given assemblage of species might contain species with larger half saturation constants, and therefore lower slopes along the $A - C_i$ curve. Unfortunately, we have far too few co-occuring species to examine this hypothesis currently, which would require much more comprehensive surveys of all species from a diversity of community types.

There are some holes in the global data that limit some conclusions that can be drawn. From a phylogenetic and biogeographic perspective, it seems that the lack of bryophyte and lichen data (Fig 1A), particularly in the arctic (Fig 1B), is a hole in our understanding of global photosynthesis. In some arctic and boreal systems, bryophytes and lichens can represent the bulk of net primary production (Limpens *et al.*, 2011). Most other taxa and ecosystems have quite good coverage coming from species that are probably representative of their commonness in different systems. For example, only 6 magnoliids are among the 303 angiosperms, but magnoliids only represent approximately 9000 of the known 300,000 angiosperm species (Joppa *et al.*, 2011). Thus, low taxonomic coverage of groups like magnoliids and lycophytes is at least roughly in proportion to their rarity in nature.

From a perspective of potential geographic bias, we cannot help but notice that – like many global datasets (E.g. Díaz *et al.*, 2016, Luyssaert *et al.*, 2007, Wright *et al.*, 2004) – the FvCB parameter data are very western-centric, primarily coming from North American, Australian, and European species. Many fields in science are well known to have a problem with diversity among members of the field (Swartz *et al.*, 2019). Setting aside moral concerns, a lack of diversity is thought to limit perspective and cost the field bright minds from underrepresented minorities. Indeed, more diverse collaborations have been shown to lead to higher impact research (Alshebli *et al.*, 2018). However, in a field like biology where diversity of life is also a part of the structure of our data, we suggest this western-centric data bias is also harming our global understanding of ecological systems. There could be much to be gained by increased coverage of species from the South American, African, and Asian continents. Particularly, since most terrestrial biodiversity hotspots are in these regions. Data exists for almost every species of tree in the boreal forest, and arguably the majority of the most common trees in temperate forests, yet we know comparatively little about the diversity of photosynthetic responses of the enormous diversity of plant species in the tropics.

However, we do not think the solution to this dual-diversity problem is for western scientists to move their research into regions of Africa, South America and Asia that are under sampled. These regions already have scientists who can become collaborators. Given that the cost of instruments required to estimate photosynthesis parameters is unusually high for ecological research, we recommend an international collaborative approach might be the most useful way to combine western access to expensive instruments, with local expertise in flora. Some have called such international collaborations the "fourth age of science" (Adams, 2013).

In terms of actual parameter estimates, we note that $C_i$ (n=112), $\Gamma^*$ (n=64), $K_c$ (n=43), $K_o$ (n=29), and all parameters from eqns 4-5 are rarely estimated (Table 1,2, Fig S1, S2). Most studies use previously published estimates of these parameters to remove the degrees of freedom problem caused by over-parameterization of the FvCB model, and then simply estimate $V_{c,max}$ (n=1294), $J_{max}$ (n=891), and increasingly $T_p$ (n=171). Parameters other than $V_{c,max}$ and $J_{max}$ are more challenging to estimate because of FvCB model over-parameterization. However, these more rarely estimated parameters are far from constant (Fig S1, S2) despite being required to solve the model. Thus, there is likely more variation in these more rarely estimated parameters (I.e. $\Gamma^*$, n=64; $K_c$, n=43; $K_o$, n=29) than is currently known, and we suggest that more attention to these parameter estimates would greatly improve our ability to accurately model photosynthesis. Indeed, all parameters show either a skewed distribution with high kurtosis, or perhaps even a multi-modal distribution (Fig S1, S2). Currently, there are too few data to know whether these multi-modal distributions are artefacts of low sample sizes, or if they represent important physiological trade-offs.

It should also be noted that though we recorded respiration rates ($R_d$, n=468) that were reported along with the main parameters of the FvCB equations, we did not seek out studies that report respiration rates of leaves. We are well aware that there are many more estimates of leaf respiration rates in the literature which are not associated with modelling exercises. However, like

various resistances to $CO_2$ movement into and through the leaf, seeking out leaf respiration rate measurements would require its own dedicated meta-analysis, which we view as a separate study. In the future, we seek to continue to update and expand this database of parameters.

The data also show that variation in parameter estimates within a species can be over an order of magnitude (Fig S3). Importantly, most of this variation can be explained by methodological differences among research groups and are driven by the leaf temperature and VPD inside the cuvette, as well as the nitrogen content of the leaves (i.e. the growing conditions of the plant). Variation among species can be over four orders of magnitude. Again however, this is largely explained by methods (Fig S4, S5) and clade (Fig 1,2, Table 1).

*Conclusion*

Photosynthesis is a key process in the global climate system and is often modelled with FvCB type models. However, considering that there are 300,000 estimated plant species on earth, there is the potential for a large diversity of photosynthetic ability among plants. Thus, accurate models should require large databases of parameter estimates. We have assembled such a database containing all parameter estimates required to solve the FvCB photosynthesis model. The publicly available database contains 3663 rows of data where at least one parameter was estimated for a given species. The summary includes parameter estimates from 359 different plant species from 96 taxonomic families representing all major vascular plant clades. The biogeographic coverage of species spans the majority of the globe, although there are some important gaps. Unfortunately, model outcomes using the parameter database show that most clades are approaching their maximum photosynthetic rate in response to elevated $CO_2$, with the notable exception graminoids. This means that most plants on Earth may already be at their maximum photosynthetic rate, and cannot increase any further with increasing $CO_2$ emissions. We hope that this database can improve our understanding of global carbon flux through the terrestrial biosphere.

**ACKNOWLEDGMENTS**

We thank Mike Mickelbart and Scott McAdam for many helpful discussions about photosynthesis, and we thank Laura Jessup and Abdel Halloway for comments on the manuscript. This work was funded in part by USDA NIFA Hatch funds to GGM (project number 1010722).

**DATA ACCESSABILITY**

All data will be made publicly available on dryad. All code used to make all figures and any data analysis are publicly available on GitHub (https://github.com/ggmcnickle/GlobalFvCB).

**AUTHOR CONTRIBUTIONS*:***

Both authors contributed equally to data collection, analysis and writing.

450 **REFERENCES**


451 Adams J (2013) The fourth age of research. *Nature,* **497**, 557-560.
452 Alshebli BK, Rahwan T, Woon WL (2018) The preeminence of ethnic diversity in scientific
453 collaboration. *Nature Communications,* **9**, 10.
454 Beer C, Reichstein M, Tomelleri E *et al.* (2010) Terrestrial gross carbon dioxide uptake: Global
455 distribution and covariation with climate. *Science,* **329**, 834-838.
456 Berendse F, Van Breemen N, Rydin H *et al.* (2001) Raised atmospheric co2 levels and increased n
457 deposition cause shifts in plant species composition and production in sphagnum bogs.
458 *Global Change Biology,* **7**, 591-598.
459 Buckley TN, Diaz-Espejo A (2015) Reporting estimates of maximum potential electron transport rate.
460 *New Phytologist,* **205**, 14-17.
461 Chamberlain S (2020) Brranching: Fetch 'phylogenies' from many sources. R package version 0.6.0.
462 pp Page.
463 Collatz G, Ribas-Carbo M, Berry J (1992) Coupled photosynthesis-stomatal conductance model for
464 leaves of $c_4$ plants. *Functional Plant Biology,* **19**, 519-538.
465 Díaz S, Kattge J, Cornelissen JHC *et al.* (2016) The global spectrum of plant form and function.
466 *Nature,* **529**, 167-171.
467 Farquhar GD, Caemmerer SV, Berry JA (1980) A biochemical-model of photosynthetic co2
468 assimilation in leaves of c-3 species. *Planta,* **149**, 78-90.
469 Farquhar GD, Sharkey TD (1982) Stomatal conductance and photosynthesis. *Annual Review of Plant*
470 *Physiology and Plant Molecular Biology,* **33**, 317-345.
471 Farquhar GD, Wong SC (1984) An empirical-model of stomatal conductance. *Australian Journal of*
472 *Plant Physiology,* **11**, 191-209.
473 Hetherington AM, Woodward FI (2003) The role of stomata in sensing and driving environmental
474 change. *Nature,* **424**, 901-908.
475 Ipcc (2012) *Managing the risks of extreme events and disasters to advance climate change*
476 *adaptation,* Cambridge, England, Cambridge University Press.
477 Joppa LN, Roberts DL, Pimm SL (2011) How many species of flowering plants are there? *Proceedings*
478 *of the Royal Society B: Biological Sciences,* **278**, 554-559.
479 Kattge J, Knorr W (2007) Temperature acclimation in a biochemical model of photosynthesis: A
480 reanalysis of data from 36 species. *Plant, Cell & Environment,* **30**, 1176-1190.
481 Kattge J, Knorr W, Raddatz T, Wirth C (2009) Quantifying photosynthetic capacity and its relationship
482 to leaf nitrogen content for global-scale terrestrial biosphere models. *Global Change Biology,*
483 **15**, 976-991.
484 Limpens J, Granath G, Gunnarsson U *et al.* (2011) Climatic modifiers of the response to nitrogen
485 deposition in peat-forming sphagnum mosses: A meta-analysis. *New Phytologist,* **191**, 496-
486 507.
487 Lombardozzi DL, Smith NG, Cheng SJ *et al.* (2018) Triose phosphate limitation in photosynthesis
488 models reduces leaf photosynthesis and global terrestrial carbon storage. *Environmental*
489 *Research Letters,* **13**, 074025.
490 Luyssaert S, Inglima I, Jung M *et al.* (2007) Co2 balance of boreal, temperate, and tropical forests
491 derived from a global database. *Global Change Biology,* **13**, 2509-2537.
492 Maitner BS, Boyle B, Casler N *et al.* (2018) The bien r package: A tool to access the botanical
493 information and ecology network (bien) database. *Methods in Ecology and Evolution,* **9**, 373-
494 379.
495 Mcclain AM, Sharkey TD (2019) Triose phosphate utilization and beyond: From photosynthesis to
496 end product synthesis. *Journal of Experimental Botany,* **70**, 1755-1766.
497 Mcnickle GG, Brown JS (2014) When michaelis and menten met holling: Towards a mechanistic
498 theory of plant nutrient foraging behaviour. *AoB Plants,* **6**, plu066.

Myers N, Mittermeier RA, Mittermeier CG, Da Fonseca GaB, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature,* **403**, 853-858.

Patrick LD, Ogle K, Tissue DT (2009) A hierarchical bayesian approach for estimation of photosynthetic parameters of c3 plants. *Plant, Cell & Environment,* **32**, 1695-1709.

Qian T, Elings A, Dieleman JA, Gort G, Marcelis LFM (2012) Estimation of photosynthesis parameters for a modified farquhar–von caemmerer–berry model using simultaneous estimation method and nonlinear mixed effects model. *Environmental and Experimental Botany,* **82**, 66-73.

Revell LJ (2012) Phytools: An r package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution,* **3**, 217-223.

Rogers A, Medlyn BE, Dukes JS *et al.* (2017) A roadmap for improving the representation of photosynthesis in earth system models. *New Phytologist,* **213**, 22-42.

Sharkey TD, Bernacchi CJ, Farquhar GD, Singsaas EL (2007) Fitting photosynthetic carbon dioxide response curves for c3 leaves. *Plant, Cell & Environment,* **30**, 1035-1040.

Swartz TH, Palermo A-GS, Masur SK, Aberg JA (2019) The science and value of diversity: Closing the gaps in our understanding of inclusion and diversity. *The Journal of Infectious Diseases,* **220**, S33-S41.

Von Caemmerer S (2000) *Biochemical models of leaf photosynthesis,* Collingwood, Vic, Australia, CSIRO Publishing.

Walker AP, Beckerman AP, Gu L *et al.* (2014) The relationship of leaf photosynthetic traits – vcmax and jmax – to leaf nitrogen, leaf phosphorus, and specific leaf area: A meta-analysis and modeling study. *Ecology and Evolution,* **4**, 3218-3235.

Whittaker RH (1975) *Communities and ecosystems,* New York, MacMillan.

Wright IJ, Reich PB, Westoby M *et al.* (2004) The worldwide leaf economics spectrum. *Nature,* **428**, 821-827.

Wullschleger SD (1993) Biochemical limitations to carbon assimilation in c3 plants—a retrospective analysis of the a/ci curves from 109 species. *Journal of Experimental Botany,* **44**, 907-920.

Zanne AE, Tank DC, Cornwell WK *et al.* (2014) Three keys to the radiation of angiosperms into freezing environments. *Nature,* **506**, 89-+.

**TABLE 1**: Summary of symbols, associated equation numbers in the text, plain descriptions and units for the parameters of the FvCB photosynthesis model that are included in this global summary. Note that those parameters in units of pressure (Pa), are reported differently by many authors. See text for unit conversions. Note that µmol $m^{-2}$ $s^{-1}$ represents carbon or photons depending on parameter.

| Parameter | Units | Eqn | Description |
|---|---|---|---|
| $V_{c,max}$ | µmol $m^{-2}$ $s^{-1}$ | 2 | The carbon saturated maximum rate of carbon fixation. |
| $J_{max}$ | µmol $m^{-2}$ $s^{-1}$ | 4, 5, 6 | The light saturated maximum possible rate of electron transport. |
| $C_i$ | Pa | 2, 3 | Intercellular partial pressure of $CO_2$ at the site of RUBISCO activity. (When constructing a $CO_2$ response curve, this is considered a variable, not a parameter.) |
| $\Gamma^*$ | Pa | 2, 3 | The partial pressure of the $CO_2$ compensation point. This is equivalent to a giving up density. It is where carbon assimilation (i.e. benefits) equal respiration costs. |
| $K_c$ | Pa | 2 | The half saturation constant for $CO_2$ of RUBISCO. This is sometimes called the enzyme "affinity" for the substrate, though in reality it has limited biological meaning. Rather, it describes the the curvature of the $CO_2$ functional response. It is simply the partial pressure that is half-way to the maximum rate. |
| $K_o$ | Pa | 2 | The half saturation constant for $O_2$ of RUBISCO. This is sometimes called the enzyme "affinity" for the substrate, though in reality it has limited biological meaning. Rather it describes the curvature of the oxygenation functional response of RUBISCO. It is simply the partial pressure of oxygen that is half-way to the maximum rate. |
| $R_d$ | µmol $m^{-2}$ $s^{-1}$ | 1 | The leaf respiration rate during the day of the leaf. |
| $T_p$ | µmol $m^{-2}$ $s^{-1}$ | 7a, b | The rate of TPU production. |
| $r$ | unitless | 7a | The rate of TPU recycling. Often assumed to be 0. |
| $\alpha$ | unitless | 4 | The efficiency of light conversion. |
| $\theta$ | unitless | 4 | A curvature factor of the light response. |
| $a$ | unitless | 5 | The efficiency of light conversion. |
| **Variable** | | | |
| $Q$ | µmol photons $m^{-2}$ $s^{-1}$ | 4, 5 | The light photon flux density striking the leaf. (When constructing a $CO_2$ response curve, this can be considered a parameter not a variable) |
| $J$ | µmol $m^{-2}$ $s^{-1}$ | 3, 4, 5 | The actual rate of electron transport going to support $NADP^+$ reduction for RuBP regeneration. |
| $A_c$ | µmol $m^{-2}$ $s^{-1}$ | 1 | Rate of carbon assimilation limited by carbon reactions. |
| $A_j$ | µmol $m^{-2}$ $s^{-1}$ | 1 | Rate of carbon assimilation limited by light and electron transport. |
| $A_p$ | µmol $m^{-2}$ $s^{-1}$ | 1 | Rate of carbon assimilation limited by TPU. |
| $A$ | µmol $m^{-2}$ $s^{-1}$ | 1 | Actual rate of carbon assimilation |

**TABLE 2**: Summary statistics and distribution parameters for each parameter of the Farquhar photosynthesis equations (Eqn 1-4). The global density distribution histogram for each parameter is shown in Fig S1. The mean, standard deviation (SD) and sample size is also shown for each major clade. SD is reported as NA for n<3, and empty cells represent no data. C4 grasses that came up in our literature search are included even though different models than the one described here are more appropriate.

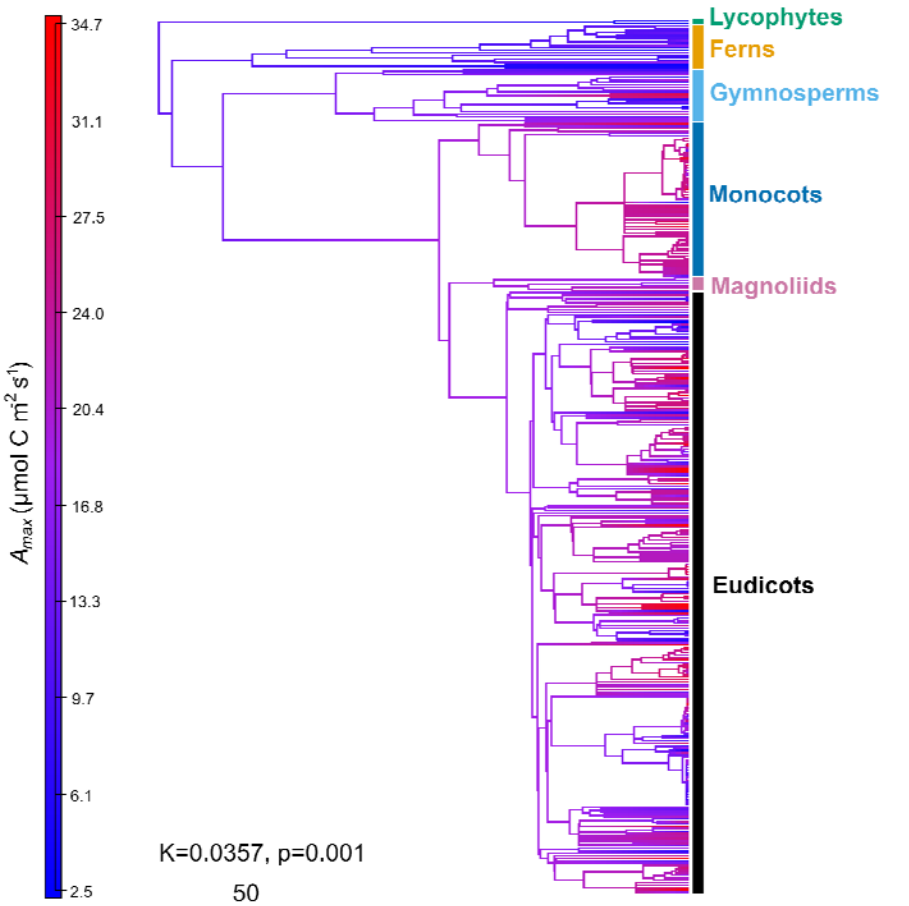
| Clade | Stat | $V_{c\,max}$ (µmol m$^{-2}$ s$^{-1}$) | $J_{max}$ (µmol m$^{-2}$ s$^{-1}$) | $C_i$ (Pa) | $\Gamma^*$ (Pa) | $K_c$ (Pa) | $K_o$ (Pa) | $R_d$ (µmol m$^{-2}$ s$^{-1}$) | $T_p$ (µmol m$^{-2}$ s$^{-1}$) |
|---|---|---|---|---|---|---|---|---|---|
| All | n | 1364 | 961 | 112 | 64 | 43 | 29 | 468 | 171 |
| | Mean | 61.0 | 135.6 | 30.5 | 4.1 | 35.8 | 29169.1 | 1.2 | 12.0 |
| | SD | 39.1 | 82.0 | 6.8 | 2.8 | 12.6 | 9829.1 | 1.7 | 5.4 |
| | CV | 0.6 | 0.6 | 0.22 | 0.68 | 0.35 | 0.33 | 1.4 | 0.45 |
| | Median | 52.9 | 118.6 | 32.7 | 3.9 | 30.5 | 28168.4 | 0.8 | 12 |
| | Min | 0.3 | 10.4 | 15.9 | 0.01 | 17.0 | 16500 | 0.02 | 1.7 |
| | Max | 267.1 | 498.4 | 39.8 | 9.2 | 70.7 | 47926.7 | 20.3 | 29.2 |
| | Skew. | 1.09 | 1.12 | -0.73 | 0.19 | 0.82 | 0.43 | 7.62 | 0.64 |
| | Kurtosis | 4.60 | 4.47 | 2.28 | 2.25 | 3.08 | 2.04 | 76.31 | 3.59 |
| Lycoph. | n | 2 | 2 | | | | | 2 | |
| | Mean | 17.2 | 25.5 | | | | | 0.5 | |
| | SD | NA | NA | | | | | NA | |
| Trache. | n | 33 | 33 | | | | | 33 | |
| | Mean | 51.2 | 59.7 | | | | | 0.5 | |
| | SD | 27.9 | 24.1 | | | | | 0.3 | |
| Gymno. | n | 461 | 301 | 37 | 27 | 2 | 2 | 41 | 49 |
| | Mean | 34.7 | 65.2 | 19.4 | 7.7 | 27.4 | 41543.3 | 0.9 | 2.8 |
| | SD | 32.25 | 42.93 | 5.14 | 3.04 | NA | NA | 0.45 | 2.62 |
| Magno. | n | 18 | 3 | | | | | 1 | |
| | Mean | 21 | 119.3 | | | | | 0.7 | |
| | SD | 12.85 | 29.16 | | | | | NA | |
| Eudic. | n | 2037 | 1060 | 43 | 63 | 42 | 31 | 283 | 48 |
| | Mean | 76.3 | 135.9 | 28.8 | 4.3 | 47.5 | 26442.6 | 1.5 | 6.8 |
| | SD | 56.88 | 64.8 | 5.67 | 1.87 | 31.63 | 11382.05 | 1.76 | 2.6 |
| Monoc. | n | 400 | 352 | 27 | 8 | 22 | 20 | 228 | 236 |
| | Mean | 85.9 | 197.4 | 31.7 | 1.9 | 51.9 | 37944 | 0.8 | 13 |
| | SD | 41.06 | 93.85 | 6.99 | 1.97 | 32.59 | 17120.35 | 0.52 | 5.35 |
| **PFT** | | | | | | | | | |
| C3 gramin. | n | 388 | 342 | 12 | 5 | 21 | 20 | 222 | 236 |
| | Mean | 87.2 | 199.6 | 29.7 | 3.0 | 53.6 | 37944 | 0.8 | 13 |
| | SD | 41.1 | 92.2 | 7.1 | 1.7 | 32.8 | 17120.3 | 0.5 | 5.4 |
| C4 gramin. | n | 6 | 4 | 12 | | | | 3 | |
| | Mean | 65.2 | 254.9 | 35.4 | | | | 1.7 | |
| | SD | 60.8 | 149.3 | 5.2 | | | | 0.7 | |
| Forb | n | 222 | 123 | 22 | 43 | 23 | 21 | 40 | 6 |
| | Mean | 83.5 | 167.5 | 31.2 | 4.3 | 41.1 | 24574.4 | 1.5 | 9.4 |
| | SD | 50.3 | 75.4 | 6.0 | 1.6 | 32.2 | 10685.8 | 2.2 | 3.4 |
| Vine | n | 55 | 54 | 1 | 1 | | | 5 | 2 |
| | Mean | 55.9 | 95.4 | 20 | 4.1 | | | 0.5 | 9.1 |
| | SD | 23.1 | 37.2 | NA | NA | | | 0.3 | NA |
| Shrub | n | 166 | 143 | 3 | 4 | 3 | | 23 | 21 |
| | Mean | 72.1 | 125.7 | 18.7 | 2.8 | 26.8 | | 2.1 | 6.6 |
| | SD | 58.4 | 63.3 | 0.8 | 3.5 | 5.0 | | 1.7 | 2.5 |
| Tree | n | 2077 | 1048 | 57 | 42 | 19 | 12 | 259 | 68 |
| | Mean | 66.6 | 114.4 | 22.4 | 6.6 | 54.9 | 32073 | 1.3 | 3.7 |
| | SD | 56.1 | 65.4 | 6.0 | 3.0 | 30.1 | 12035.2 | 1.6 | 2.7 |

* only $R_d$ values reported with estimates of FvCB model parameters were sought. This will not be a complete literature review of plant respiration rates.

**TABLE 3**: Summary statistics across all measurements for 334 species for each parameter of the Farquhar photosynthesis equations 4-5. The global density distribution histogram for each parameter is shown in Fig S2. The mean, standard deviation (SD) and sample size is also shown for each major clade. SD is reported as NA for n<3, and empty cells represent no data. Clades are summarised by family in Table S1 and by functional type growth habit in Table S4. C4 grasses that came up in our literature search are included even though different models than the one described here are more appropriate.

| Clade | Stat | $\alpha$ (Eqn 4) | $\theta$ (Eqn 4) | $a$ (Eqn 5) |
|---|---|---|---|---|
| ALL | n | 41 | 204 | 6 |
| | Mean | 0.22 | 0.58 | 0.17 |
| | SD | 0.21 | 0.29 | 0.06 |
| | CV | 0.95 | 0.50 | 0.35 |
| | Median | 0.16 | 0.66 | 0.19 |
| | Min | 0.02 | 0.04 | 0.06 |
| | Max | 0.95 | 1.0 | 0.23 |
| | Skewness | 1.55 | -0.73 | -1.81 |
| | Kurtosis | 5.86 | 2.35 | 6.87 |
| Gymnosperm | n | 14 | 13 | 3 |
| | Mean | 0.23 | 0.53 | 0.19 |
| | SD | 0.05 | 0.17 | 0.02 |
| Magnoliid | n | | 1 | |
| | Mean | | 0.06 | |
| | SD | | NA | |
| Eudicot | n | 26 | 44 | 3 |
| | Mean | 0.27 | 0.51 | 0.16 |
| | SD | 0.22 | 0.41 | 0.09 |
| Monocot | n | 2 | 230 | |
| | Mean | 0.27 | 0.67 | |
| | SD | NA | 0.18 | |
| **PFT** | | | | |
| C3 graminoid | n | | 220 | |
| | Mean | | 0.68 | |
| | SD | | 0.20 | |
| C4 graminoid | n | | 7 | |
| | Mean | | 0.44 | |
| | SD | | 0.40 | |
| Forb | n | 17 | 23 | |
| | Mean | 0.32 | 0.77 | |
| | SD | 0.20 | 0.30 | |
| Vine | n | | 3 | |
| | Mean | | 0.65 | |
| | SD | | 0.50 | |
| Shrub | n | 4 | 4 | |
| | Mean | 0.02 | 0.33 | |
| | SD | 0.006 | 0.5 | |
| Tree | n | 20 | 30 | 5 |
| | Mean | 0.24 | 0.33 | 0.17 |
| | SD | 0.20 | 0.30 | 0.10 |

Note: Lycophytes and ferns are absent from this table because no data were available for these parameters.



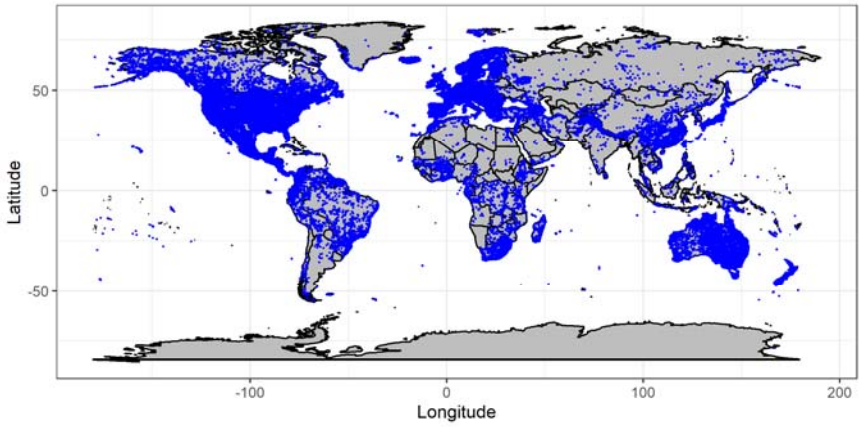
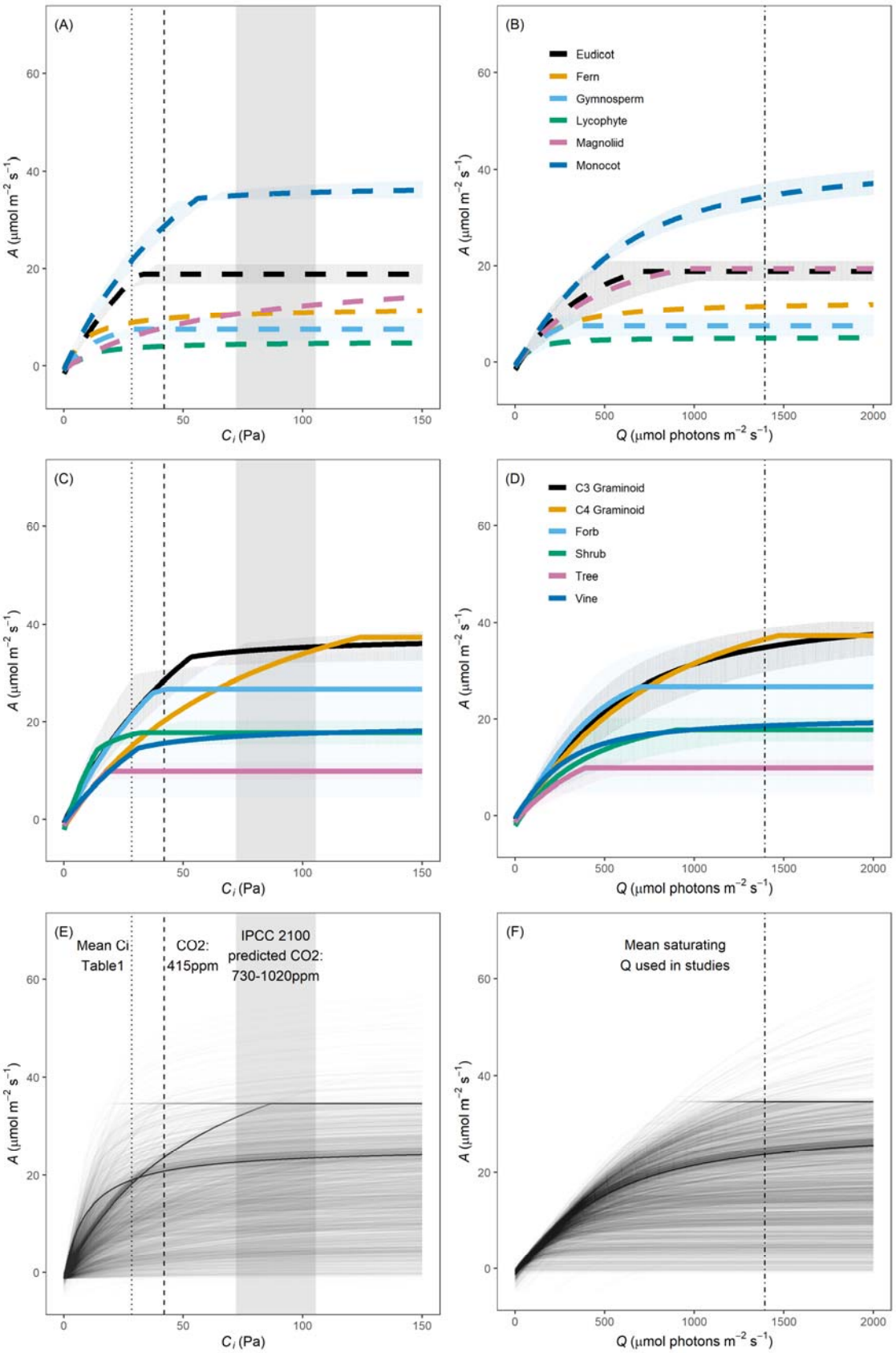




**FIG 1**: (A) The evolutionary relationships among all 359 species among 6 major terrestrial plant clades for which photosynthesis parameters were recovered from the literature and included in our database. Scale bar represents 50 million years of evolution, and colours represent modelled $A_{max}$. Unresolved taxonomy are drawn as polytomies at the genus level. The raw phylogeny in newick format is given in the supplementary information to allow digital visualisation of such a large phylogeny. There was no evidence of phylogenetic niche conservatism across the entire dataset, meaning species with low or high $A_{max}$ are equally likely in any clade. (B) Global distribution data for all 334 species included in the data set (blue points). Points represent occurrence distribution data for all species retrieved from BIEN, not the locations of the measurements.


(A)
(B)
(C)
(D)
(E)
(F)
Eudicot
Fern
Gymnosperm
Lycophyte
Magnoliid
Monocot
C3 Graminoid
C4 Graminoid
Forb
Shrub
Tree
Vine
Mean Ci
Table1
CO2:
415ppm
IPCC 2100
predicted CO2:
730-1020ppm
Mean saturating
Q used in studies
A (μmol m⁻² s⁻¹)
Ci (Pa)
Q (μmol photons m⁻² s⁻¹)

561

**Fig 2:** (**A**) Mean $A - C_i$ and (**B**) $A - Q$ curves among each of the six major plant clades for which we have parameter estimates (solid lines). We also show (**C**) Mean $A - C_i$ and (**D**) $A - Q$ curves with species grouped by broad functional type (dashed lines). Shading around the mean curve in panels A-D represents solving the model with the 95% confidence interval around all parameters. These 95% confidence intervals cross on the $A - C_i$ curve because of the oxygenation behaviour of RUBISCO at low $CO_2$ partial pressures. For, lycophytes, ferns, magnoliids, C4 graminoids, and vines the large number of missing values were replaced with the global mean, and we did not draw confidence intervals for these groups (readers should assume they are very wide due to low taxonomic sampling). Finally, each line in the lower panels represent 3663 separate (**E**) $A - C_i$ or (**F**) $A - Q$ curves generated for every single row in our database and show the breadth of individual species responses across diverse conditions. For these 3663 lines, missing values were replaced with the global mean (Table 2). In all panels, for $A - C_i$ curves, the black vertical dotted line marks the estimated mean partial pressure of $CO_2$ inside the leaf in reported in the literature, while the dashed line represents the approximate current partial pressure of $CO_2$ in the atmosphere at the time of writing. The shaded rectangle represents the predicted range of $CO_2$ Intergovernmental panel on climate change (IPCC) predicted partial pressures of $CO_2$ by the year 2100. For $A - Q$ curves, the black dot-dash line represents the mean saturating Q used among studies in our database.