We thank the editor and reviewers for their helpful reviews; the manuscript is stronger because of them. The comments are below, followed (in bold) our responses. All line numbers by us reference the marked-up copy of the revised manuscript.

EDITOR

Comments to the Author:
Thank you for addressing the reviewer comments. Both reviewers were positive about the paper and I also think it is an important contribution to the paleo-CO2 literature. The revisions proposed address most of their concerns. I note a few issues below that I would like to see addressed before final acceptance and ask that you submit a revised version of the paper that includes the revisions you proposed and addresses these further comments:

1) Line 98-101. Can you clarify this discussion of uncertainty? The issue is whether all studies using only stomatal density ignore the two sources of error you mention, or just some do. The wording here is not clear on that point.

We have clarified this point (lines 99-101). Very few studies propagate both sources of uncertainty (the Beerling 2009 study, mentioned at the end of the sentence, is one such study).

Line 102. Smaller than what? With fossil taxa? Please clarify.

We have clarified this point (smaller than with gas-exchange proxies; line 103).

Line 116-118. This sentence strikes me as awkward it refers to "elements" but then phrases them as questions. Can you reword?

Done (lines 116-117)

Line 121. Can you provide a more informative title for the table?

Done (line 122)

Line 170-171. How are the ambient CO2 values known?

References added to the Mauna Loa and Harvard Forests databases (lines 171-173).

Line 239. Regarding Milligan et al., please just provide the details of the method. Even if the paper is in press the reader would benefit.

Done (lines 242-243). Also, the Milligan paper is now published, with a doi.

Line 320-321. I think it would help here to explain that you do not have measurements of d13C-CO2.

We now note this (lines 324-325); we also mention this in the results (lines 508-510)

Line 342-346. I understand why you use the 2/3 range, but you should explain here why. Also, what does it mean to say that 28% overlap the target at 95% confidence if they are really noisy (that is, is this a

useful statement). And, how is the target value established? Assumed 400 ppm?

We now say that the 2/3 range is a close equivalent to +/- 1 standard deviation (lines 349-350). The target value is assumed to be 400 ppm (we now state this directly on line 171).

We agree that this is a noisy signal, and the reader can decide whether the signal:noise ratio is good enough for their purposes. We included this statement on the account of a direct request by one of the reviewers.

Line 526. Can you replace "tough" with a more specific term?

We replaced this word with "durable" (line 533) (toughness is a quantitative trait made by plant ecologists).

With best wishes, Ed Brook

REVIEWER #1

General comments: Fossil leaf gas-exchange based CO2 models are currently going through the "rigorous testing" phase and as the authors of this paper point out, this mechanistically, rather than empirically calibrated proxy, shows considerable promise. It is therefore of high relevance that studies, such as this one, are presented that provide quantification of potential confounding factors. In this case, the authors test three potential confounding factors (photorespiration, leaf temperature and canopy position) and provide quantifications on how these factors influence final CO2 estimates. They are capable of eliminating two of these factors as insignificantly affecting CO2 estimates (photorespiration and leaf temperature). The third factor, canopy position, is determined to strongly skew CO2 estimates, but the authors point out that it is possible to identify leaves that grew in lower canopy positions, based on leaf micromorphology and an uncharacteristically wide δ13C range. This paper is a relevant contribution towards quantification of the potential error in fossil leaf gas-exchange based CO2 models, and apart from minor suggested amendments, I have no problem with seeing this study being published.

Specific comments: In the materials and methods section, the authors lay out the specific ways that they are testing modern plants for potential bias in reconstructed CO2. In the appendix all the specific plants are listed with their input values and reconstructed CO2. However, from reading the methods section I get the impression that not each plant is being tested for the same potential confounding variable (photorespiration, leaf temperature and canopy position). It would be very helpful if there was a table that outlines specifically which plants were tested for what, or at least that this was made clear in the appendix, because in the main body of text it is hard to follow.

We now include this information in column E of the supplemental table.

In several places in the manuscript, including the abstract, it is mentioned that the random error propagation of the Franks et al. gas exchange model is better than uncertainty estimates of other leading paleo-CO2 proxies. It would be very helpful for the untrained reader to see some proof of this

statement in the form of a table that lists 1) the different CO2 proxies, 2) a method of error quantification, 3) the actual amount of uncertainty in those CO2 proxies and 4) the references to the case studies where this was tested. Such a table would lend credibility to the statement that gas-exchange models are quantifiably better than other CO2 proxies.

There are of course two elements of uncertainty: precision (spread of possible solutions) and accuracy (comparison to true answer; can only be quantified for times when CO2 has been measured). The abstract brings up the theme of accuracy (28% mean error rate). In the main text (section 3.1), the mean error rate is compared generally to that in other CO2 proxies by referencing the summary work of Franks et al. (2014).

The error propagation scheme noted by the reviewer is related to precision. We only mention precision in the Introduction by referencing what others have found (Franks et al., 2014). It is not a focal point of the current study.

The reviewer may (also) be referencing the paragraph in the Introduction where we argue that studies using other stomatal-based proxies probably overstate the accuracy and precision of their CO2 estimates (lines 98-107). Our arguments here are conceptual only—there are no data we can summarize in a table, unfortunately. The point we are trying to make is that the reported accuracies and precisions associated with these other methods—when applied to plants living today (not fossils)—are better than what we find with gas-exchange methods. But this is partly because these other methods are based on empirical calibrations with…present-day plants. So excellent accuracies and precisions are not particularly surprising. But when you apply these other methods to fossils that are millions of years old, the present-day empirical calibrations are likely less appropriate.

Final specific comment is on the title itself, for which I would like to suggest that the authors include what specifically is being tested. I.e. “Sensitivity of . . .. CO2 concentration to x, y & z”. There are other variables that the model is sensitive to and I believe the title would be more informative if the specifics were included.

The largest block of data (40 species) is “general” testing, that is, estimating CO_2 from field-grown trees without isolating any single confounding factor (summarized in Figure 2). Thus, it would not be fully representative to say that we were only testing the model for the influence of canopy position, temperature, and photorespiration.

Technical corrections: I could not find any spelling or styling errors in the manuscript. The paper is very well constructed and easy to follow.

REVIEWER #2

The authors present a sensitivity analysis of a mechanistic model (Franks model) to predict paleoatmospheric CO2. They explore several specific areas; the effect of gc(op)/gc(max), A0, temperature, photorespiration and leaf canopy position on the accuracy of CO2 estimates produced by the model. In doing so, the paper adds clarity, certainty or recommendations to the model for fossil application, all of which are important additions, especially as this model is being using in a growing number of research projects. Although the paper is an important contribution, it would benefit from clarity or expansion in certain areas:

1) Aims, methods and appendix: The aims and methods section is hard to follow. This may be due to the fact the aims and rationale are mixed in with the methods. It is unclear from the text or appendix data whether all or a subset of the data is being used for each of the analysis performed. A summarised table in the methods section containing the information on the analysis being performed, data source and parameters used or tested would be beneficial (i.e. a summary of the methods in tabular format). Similarly, in the appendix, additional information on the origin of the data, sample number per species, which data points/values are measured vs estimated/assumed and a direct comparison of measured vs model estimated CO2 would greatly improve clarity.

We now present a tabular summary of our study design (new Table 1).

In the Supplemental Table 1, we now give the sample size (column F), the target (i.e., correct) CO2 concentration (column G), and whether the input was measured or inferred (color coding of column headers). And column E gives what part of the study was addressed (general testing, temperature, or canopy position; reviewer #1 also asked for this information). We are not sure what is meant by "additional information on the origin of the data" beyond what is listed in column A and stated in the main-text Methods.

2) Statistical analysis: Accuracy was evaluated by the degree of error rate. These claims can be strengthened by using statistical analysis. How well the model predicts CO2 could be assessed by whether or not the estimates are statistically significant different (or hopefully not) from measured CO2 values.

We have added information about whether individual estimates depart from the target CO2 concentrations (lines 350-352 and 425-427).

3) gc(op)/gc(max) and A0 (section 3.1): This section gives details about when both gc(op)/gc(max) and A0 values are either known or values from Franks et al. 2014 are used, but it would be nice to see these two parameters evaluated separately i.e. how much does gc(op)/gc(max) alone improve estimates and the same for A0. Does one contribute more than the other for improving error rates?

We have added this information (lines 357-358).

Additional comments:
Line 86. Sensitivity saturates for some but not all taxa. See Haworth et al 2011.

We have added the qualifier "in many species".

Line 93. A Nearest living relative or equivalent approach also get around the issue of extinct taxa.

This is true for the stomatal ratio method, but these CO2 estimates are not meant to be quantitative in the same manner as estimates from the "full calibration" methods or the gas-exchange methods (as noted in the previous paragraph).

Line 156. Alternative approaches for fossils have been suggested such as estimating fossil A0 using scaling relationships between vein distance and assimilation rate however they are not discussed here (EG Montanez et al., 2016).

We have added a citation to the Montanez paper

Introduction – general comment. Critical published assessments of the Franks model are not cited (eg McElwain et al. 2016) yet they raise issues associated with parametrization of A0 and the insensitivity of CO2 estimates to variation in gamma star values which are both important discussion points in this manuscript in lines 454 -456 and 497-499.

As per a later comment, we have added a citation to McElwain et al. 2016 regarding gamma star on line 472.

Our study does not focus on the parameterization of A0, and so the associated literature does not seem relevant to the Introduction. Our study focuses on temperature, photorespiration, canopy position, as well as a general and broad test of the method.

Paragraph 201-217: A some information is missing here: chamber model/make, duration plants were grown in the chamber, light levels. What were measured vs set chamber conditions for temperature, light and CO2 (i.e. similar to how humidity is reported)

Chamber make/model (lines 216-217) and duration of experiment (line 233) are given. We have added information about light intensity as well as the standard deviations for temperature and CO2 concentrations in lines 217-221.

Lines 232: Stomatal density/stomatal measurements and leaf stable carbon isotopes were performed on the same leaves. Clarify how this was partitioned, e.g. was the leaf divided into 2 or was a whole punch used for carbon isotopes, etc.?

We now clarify our methodology in lines 241-242. We used either a hole punch or razor to remove two adjacent sections of leaf tissue near the leaf centers, avoiding major veins.

Lines 235: As Milligan et al is in review, I suggest adding more detail here on how δ13Ca of chamber CO2 was calculated. δ13Ca values of supplemented CO2 can be very negative and can vary between cylinders, unless the CO2 gas has a specific δ13Ca. What is the capacity of these cylinder, in L?

This paper is now published. In short, a mixing line was established based on direct d13C measurements of lab air, chamber air, and cylinder CO2 (= pure CO2). We were fortunate that the d13C of the cylinder was close to the well-mixed atmosphere (the d13C in most cylinders we have used in other experiments is much more depleted). We used only the single cylinder for the duration of the experiment. The target CO2 concentration (500 ppm) was not much higher than the CO2 concentration inside the lab (~440 ppm), so we did not use much CO2.

Figure 1: Does this need to be on a log scale? 1000 or 2000ppm are not very high values and the log scale visually skews data and error bars. A difference plot between measured and estimates plotted on a non-log scale would improve this figure.

We prefer a log scale because it is easier to differentiate estimates at the low-end of the CO2 scale, and because the uncertainties scale in a logarithmic fashion.

Line 351: Please provide supporting data for this statement in tabular form. What are the error rates of other proxies?

This information was summarized by Franks et al. (2014), so we prefer not to repeat it here.

Line 355: Might be helpful to report standard deviation of CO2 estimates, here and throughout the text.

We now report the range that encompasses two-thirds of all estimates (lines 350-352). (Because the individual estimates are not normally distributed (tail at the high end), reporting a standard deviation can be misleading.)

Line 411 to 413. Reporting of the difference between estimated and measured CO2 here is incomplete. Only means of all species investigated are provided rather than species-based diffeences or errors. For some species the error is substantial whereas other taxa show very small errors.

As per an earlier comment, we now report the species-level differences on lines 425-427; no individual species-level test was significant (line 414).

Line 454 to 456. This supports the findings of McElwain et al 2016 Paleo 3 but it is not cited. “This compensation point (Γ * in Eq. (2) is temperature, species and O2 dependent (Ethier and Livingston, 2004) but Franks et al. (2014) account only for the temperature dependency in the new paleo-CO2 proxy model. Allowing Γ * to vary in response to prevailing paleoatmospheric O2 concentration [O2] (Γ* = 1.78 × [O2]), which is known to have varied widely (10% to 30%) through the Phanerozoic (Bergman et al., 2004; Belcher and McElwain, 2008; Berner, 2009), would increase the precision of paleo-CO2 estimates but only fractionally.”

We have added a citation to McElwain et al. (2016 Palaeo3) (line 472).

Lines 500 to 506: A number of papers have suggested methods of estimating A0 to improve the accuracy of CO2 estimates using the Franks model but they are not discussed. This section would provide a good opportunity to discuss the proposed ideas and solutions.

This section deals with living leaves, where A could be measured directly. Measuring A wasn't part of our study design, unfortunately. In this section we are discussing possible reasons for noise in our mixing-model calculations. With regards to fossils, we are not recommending that our mixing model be used (line 527: "We note that our mixing-model strategy cannot be applied to fossils because…"), so the question of how to constrain A in fossils within the context of the mixing model is moot. Our take-home message for fossil applications is to avoid shade leaves (line 535), and we provide specific measurements that can be made on fossils to make this distinction, including vein density (lines 536-540).

Section 3.4: Have any values for δ13Ca been measured or are all calculated for this section? Is there any data set (from the literature or otherwise) this could be compared to? i.e. a dataset where known δ13Ca is compared to itself when calculated as per the manuscript? This would strengthen this section. If δ13Ca has only been calculated/inferred for this section without a comparison to measured δ13Ca I think claims on the effect of δ13Ca (or low canopy plants) on the model should be softened.

We made no direct measurements of understory d13Ca (multiple measurements over a growing season, and at different daytime hours, would be needed to calculate a representative mean value). As the reviewer correctly notes, we instead are assuming a well-behaved two end-member mixing model. We have added a note of caution related to this on lines 508-511.

Appendix: The authors used both known and general values for gc(op)/gc(max) and A0 to evaluate error rates but no measured values of either gc(op)/gc(max) or A0 are given in the appendix or text.

The Appendix summarizes all new data presented in the study (with the key graphics being Figures 2, 5, and 7). For these data, we *only* used "default" values of gop/gmax and Ao; that is, we did not measure these inputs on our leaves. As noted in the Introduction, this was a purposeful strategy because we wanted to test the CO2 model in a manner that would be similar to how most (but not all) folks will be applying the model to fossils. A "worst-case" test, if you will.

In the Introduction, we do summarize some of the already-published data (Figure 1). For these estimates, either gop/gmax or A0 were measured, and in most cases both were measured (lines 143-146). These data are not in the Appendix because they are already published and are not central to our study.

As the reviewer noted, we did additionally "degrade" these estimates by re-doing them assuming default values for gop/gmax and A0. We did this so we could compare them more directly to our estimates (lines 355-357).

Sensitivity of a leaf gas-exchange model for estimating paleoatmospheric CO_2 concentration

Dana L. Royer[1], Kylen M. Moynihan[1], Melissa L. McKee[1], Liliana Londoño[2], and Peter J. Franks[3]
[1]Department of Earth and Environmental Sciences, Wesleyan University, Middletown, Connecticut, USA
[2]Smithsonian Tropical Research Institute, Balboa, Ancón, Republic of Panamá
[3]Faculty of Agriculture and Environment, University of Sydney, Sydney, New South Wales, Australia
Correspondence: Dana L. Royer (droyer@wesleyan.edu)

Abstract. Leaf gas-exchange models show considerable promise as paleo-CO_2 proxies. They are largely mechanistic in nature, provide well-constrained estimates even when CO_2 is high, and can be applied to most subaerial, stomata-bearing fossil leaves from C_3 taxa, regardless of age or taxonomy. Here we place additional observational and theoretical constraints on one of these models, the "Franks" model. In order to gauge the model's general accuracy in a way that is appropriate for fossil studies, we estimated CO_2 from 40 species of extant angiosperms, conifers, and ferns based only on measurements that can be made directly from fossils (leaf $\delta^{13}C$ and stomatal density and size) and on a limited sample size (1-3 leaves per species). The mean error rate is 28%, which is similar to or better than the accuracy of other leading paleo-CO_2 proxies. We find that leaf temperature and photorespiration do not strongly affect estimated CO_2, although more work is warranted on the possible influence of O_2 concentration on photorespiration. Leaves from the lowermost 1-2 m of closed-canopy forests should not be used because the local air $\delta^{13}C$ value is lower than the global well-mixed value. Such leaves are not common in the fossil record, but can be identified by morphological and isotopic means.

1 Introduction

Leaves on terrestrial plants are well poised to record information about the concentration of atmospheric CO_2. They are in direct contact with the atmosphere and have large surface-area-to-volume ratios, so the leaf internal CO_2 concentration is tightly coupled to atmospheric CO_2 concentration. Also, leaves are specifically built for the purpose of fixing atmospheric carbon into structural tissue, and face constant selection pressure to optimize their carbon uptake relative to water loss. As a result, many components of the leaf system are sensitive to atmospheric CO_2, and these components feedback on one another to reach a new equilibrium when atmospheric CO_2 changes. In terms of carbon assimilation, Farquhar and Sharkey (1982) modeled this system in its simplest form as:

$$A_n = g_{c(tot)} \times (c_a - c_i), \quad (1)$$

where A_n is the leaf CO_2 assimilation rate (µmol m^{-2} s^{-1}), $g_{c(tot)}$ is the total operational conductance to CO_2 diffusion from the atmosphere to site of photosynthesis (mol m^{-2} s^{-1}), c_a is atmospheric CO_2 concentration (µmol mol^{-1} or ppm), and c_i is leaf intercellular CO_2 concentration (µmol mol^{-1} or ppm) (see also Von Caemmerer, 2000).

Rearranging Eq. (1) for atmospheric CO_2 yields:

$$c_a = \frac{A_n}{g_{c(tot)} \times (1 - \frac{c_i}{c_a})}. \quad (2)$$

Equation (2) forms the basis of two leaf gas-exchange approaches for estimating paleo-CO_2 from fossils (Konrad et al., 2008, 2017; Franks et al., 2014). In the Franks model, conductance is estimated in part from measurements of stomatal size and density, c_i/c_a from measurements of leaf $\delta^{13}C$ along with reconstructions of coeval air $\delta^{13}C$ (see also Eq. 9), and A_n from knowledge of living relatives and its dependency on c_a (Franks et al., 2014). Following Farquhar et al. (1980), the latter is modeled as (Franks et al., 2014; Kowalczyk et al., 2018):

$$A_n = A_0 \frac{[(\frac{c_i}{c_a})c_a - \Gamma^*][(\frac{c_{i0}}{c_{a0}})c_{a0} + 2\Gamma^*]}{[(\frac{c_i}{c_a})c_a + 2\Gamma^*][(\frac{c_{i0}}{c_{a0}})c_{a0} - \Gamma^*]}, \quad (3)$$

where Γ^* is the CO_2 compensation point in the absence of dark respiration (ppm) and the subscript "0" refers to conditions at a known CO_2 concentration (typically present-day). Equations (2) and (3) are then solved iteratively until the solution for c_a converges.

These gas-exchange approaches grew out of a group of paleo-CO_2 proxies based on the CO_2 sensitivity of stomatal density (D) or the similar metric stomatal index (Woodward, 1987; Royer, 2001). Here, the D-c_a sensitivity is calibrated in an extant species, allowing paleo-CO_2 inference from the same (or very similar) fossil species. These empirical relationships typically follow a power-law function (Wynn, 2003; Franks et al., 2014; Konrad et al., 2017):

$$c_a = \frac{1}{kD^{\alpha}}, \quad (4)$$

where k and α are species-specific constants.

The related stomatal ratio proxy is simplified: D is measured in an extant species (D_0, at present-day c_{a0}) and then the ratio of D_0 to D in a related fossil species is assumed to be linearly related to the ratio of paleo-c_a to present-day c_{a0} (Chaloner and McElwain, 1997; McElwain, 1998):

$$\frac{c_a}{c_{a0}} = k\frac{D_0}{D}. \quad (5)$$

Equation (5) can be rearranged to match Eq. (4) but with α fixed at 1. Thus, paleo-CO_2 estimates using the stomatal ratio proxy are based on a one-point calibration and an assumption that $\alpha = 1$; observations do not always support this assumption (e.g., α = 0.43 for *Ginkgo biloba*; Barclay and Wing, 2016). The scalar k was originally set at 2 for Paleozoic and Mesozoic reconstructions so that paleo-CO_2 estimates during the Carboniferous matched that from long-term carbon cycle models (Chaloner and McElwain, 1997). For younger reconstructions, k is probably closer to 1 (by definition, $k = 1$ for present-day plants). We note that the stomatal ratio proxy was originally conceived as providing qualitative information, only, about paleo-CO_2 (McElwain and Chaloner, 1995, 1996; Chaloner and McElwain, 1997; McElwain, 1998) and has not been tested with dated herbaria materials or with CO_2 manipulation experiments.

At high CO_2, the D-c_a sensitivity saturates in many species, leading to uncertain paleo-CO_2 estimates, often with unbounded upper limits (e.g., Smith et al., 2010; Doria et al., 2011). Stomatal density does not respond to CO_2 in all species (Woodward and Kelly, 1995; Royer, 2001), and because D-c_a relationships can be species-specific (that is, different species in the same genus with different responses; Beerling, 2005; Haworth et al., 2010), only fossil taxa that are still alive today should be used. The gas-exchange proxies partly address these limitations: 1) CO_2 estimates remain well-bounded—even at high CO_2—and their precision is similar to or better than other leading paleo-CO_2 proxies (~+35/-25% at 95% confidence; Franks et al., 2014); 2) the models are mostly mechanistic; that is, they are explicitly

driven by plant physiological principles, not just empirical relationships measured on living plants; 3) because the models retain sensitivity at high CO_2 and do not require that a fossil species still be alive today, much of the paleobotanical record is open for CO_2 inference, regardless of age or taxonomy; and 4) because the models are based on multiple inputs linked by feedbacks, they can still perform adequately even if one or more of the inputs in a particular taxon is not sensitive to CO_2, for example stomatal density (Milligan et al., 2019).

We note that the published uncertainties (= precision) associated with the stomatal density proxies are probably too small because they usually only reflect uncertainty in either the calibration regression or in the measured values of fossil stomatal density, but not both; when both sources are propagated~~this is done~~, errors often exceed ±30% at 95% confidence (Beerling et al., 2009). Also, error rates in estimates from extant taxa where CO_2 is known (= accuracy) are usually smaller with ~~the~~ stomatal density proxies than with gas-exchange proxies (e.g., Barclay and Wing, 2016), but this is expected because the same taxa have been calibrated in present-day (or near present-day) conditions. Because the gas-exchange proxies are largely built from physiological principles, they have less "recency" bias; that is, the gas-exchange proxies estimate present-day and paleo-CO_2 with similar certainty when the same methods are used to determine the inputs.

2 Study Aims and Methods

Leaf gas-exchange proxies for paleo-CO_2 are becoming popular (Konrad et al., 2008, 2017; Grein et al., 2011a, 2011b, 2013; Erdei et al., 2012; Roth-Nebelsick et al., 2012, 2014; Franks et al., 2014; Maxbauer et al., 2014; Montañez et al., 2016; Reichgelt et al., 2016; Tesfamichael et al., 2017; Kowalczyk et al., 2018; Lei et al., 2018; Londoño et al., 2018; Richey et al., 2018; Milligan et al., 2019). However, many elements in these models remain understudied. Here we ~~investigate~~ scrutinize four such elements of the Franks et al. (2014) model, and ask: how does the model perform across a large number of phylogenetically diverse taxa; and how is the model affected by temperature, photorespiration, and proximity to the forest floor? We describe next the motivation and details of the study design (see also Table 1 for summary).

Table 1. Attributes of data sets used to test the Franks et al. (2014) model.

Element of model tested	Number of species	Methods section	Notes
General testing in a phylogenetically diverse set of species and with a minimal number of leaves measured per species	40	2.1	Leaves come from Panama (published by Londoño et al., 2018), Connecticut, and Puerto Rico
Temperature	6	2.2	Theoretical calculations and growth chamber experiment
Photorespiration	NA	2.3	Theoretical calculations
Canopy position	6	2.4	Leaves come from Panama and Connecticut

2.1 General testing in living plants

Franks et al. (2014) tested the model on four species of field-grown trees (three gymnosperms and one angiosperm) and one conifer grown in chambers at 480 and 1270 ppm CO_2. The average error rate (absolute value of estimated CO_2 minus measured CO_2, divided by measured CO_2) was 5%. Follow-up

work with three field-grown tree species (Maxbauer et al., 2014; Kowalczyk et al., 2018), CO_2 experiments on seven tropical trees species (Londoño et al., 2018), and experiments on two fern and one conifer species (Milligan et al., 2019) indicate somewhat higher error rates (Fig. 1). Combined, the average error rate is 20% (median = 13%).

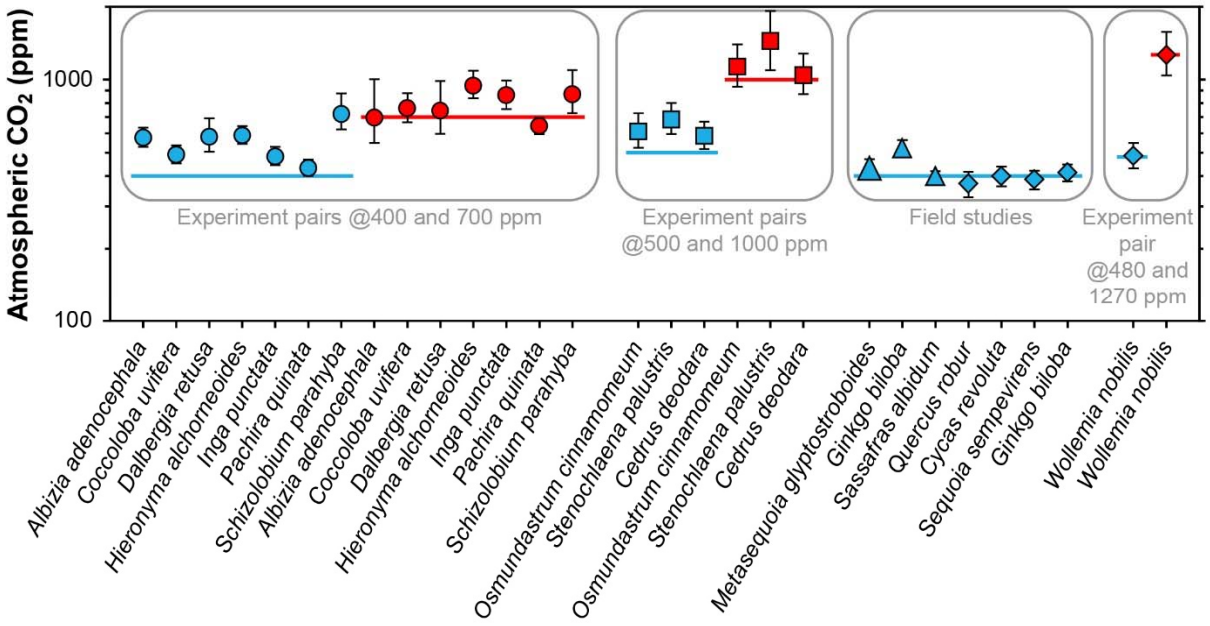

Figure 1. Published CO_2 estimates using the Franks model for extant plants where the physiological inputs A_0 (assimilation rate at a known CO_2 concentration) and/or $g_{c(op)}/g_{c(max)}$ (ratio of operational to maximum leaf conductance to CO_2) were measured directly. Horizontal lines are the correct CO_2 concentrations. Uncertainties in the estimates correspond to the 16^{th}-84^{th} percentile range. Circles are from Londoño et al. (2018), squares from Milligan et al. (2019), large triangle from Maxbauer et al. (2014), small triangles from Kowalczyk et al. (2018), and diamonds from Franks et al. (2014).

In these studies, two of the key physiological inputs were measured directly with an infrared gas analyzer: the assimilation rate at a known CO_2 concentration (A_0) and/or the ratio of operational to maximum stomatal conductance to CO_2 ($g_{c(op)}/g_{c(max)I}$, or ζ), the latter of which is important for calculating the total leaf conductance ($g_{c(tot)}$). These two inputs cannot be directly measured on fossils; thus, the error rates associated with Figure 1 may not be representative for fossil studies. Franks et al. (2014) argue that within plant functional types growing in their natural environment, mean A_0 is fairly conservative, leading to the recommended mean A_0 values in Franks et al. (2014) (12 μmol m^{-2} s^{-1} for angiosperms, 10 for conifers, and 6 for ferns and ginkgos). Along similar lines, the mean ratio $g_{c(op)}/g_{c(max)}$ tends to be conserved across plant functional types; Franks et al. (2014) recommend a value of 0.2, which may correspond to the most efficient setpoint for stomata to control conductance (Franks et al., 2012). This conservation of physiological function is one of the underlying principles in the Franks model.

Here we test this assumption by estimating CO_2 from 40 phylogenetically diverse species of field-grown trees. In making these estimates, we use the recommended mean values of A_0 and $g_{c(op)}/g_{c(max)}$ from Franks et al. (2014) instead of measuring them directly (see also Montañez et al., 2016 for other ways to infer assimilation rate from fossils). Thus, this dataset should be a more faithful gauge

for model accuracy as applied to fossils. Of the 40 species, 21 were previously published in Londoño et al. (2018), who collected sun-adapted canopy leaves of angiosperms using a crane in Parque Nacional San Lorenzo, Panama. To test the method in temperate forests, we collected leaves from eleven angiosperm and seven conifer species from Dinosaur State Park (Rocky Hill, Connecticut), Wesleyan University (Middletown, Connecticut), and Connecticut College (New London, Connecticut) during the summer of 2015. Here, all trees grew in open, park-like settings; one to three sun leaves were sampled from the lower outside crown of each tree. In January of 2015, we also sampled sun-exposed leaves from the tree fern *Cyathea arborea* in El Yunque National Forest, Puerto Rico (near the Yokahú Tower).

Stomatal size and density were measured either on untreated leaves using epifluorescence microscopy with a 420-490 nm filter, or on cleared leaves (using 50% household bleach or 5% NaOH) using transmitted-light microscopy. For most species, whole-leaf $\delta^{13}C$ comes from Royer and Hren (2017); the same leaves were measured for $\delta^{13}C$ and stomatal morphology. The UC Davis Stable Isotope Facility measured some additional leaf samples. Atmospheric CO_2 concentration (400 ppm) and $\delta^{13}C_{air}$ (-8.5‰) come from Mauna Loa, Hawaii (NOAA/ESRL, 2019), which we assume are representative of the local conditions where we sampled (e.g., Munger and Hadley, 2017). Table S1 summarizes for these 40 species all of the inputs needed to run the Franks model, along with the estimated CO_2 concentrations. Uncertainties in the estimates are based on error propagation using Monte Carlo simulations (Franks et al., 2014).

2.2 Temperature

The Franks model can be configured for any temperature. Franks et al. (2014) recommend that the photosynthesis parameters A_0 and Γ^*, and the air physical properties affecting diffusion of CO_2 into the leaf (the ratio of CO_2 diffusivity in air to the molar volume of air, or d/v) correspond with the mean daytime growing-season leaf temperature (more precisely, assimilation-weighted leaf temperature). The reasoning behind this is that (i) the assimilation-weighted leaf temperature ~~will~~ corresponds with the mean c_i/c_a derived from fossil leaf $\delta^{13}C$; and (ii) both theory (Michaletz et al., 2015, 2016) and observations (Helliker and Richter, 2008; Song et al., 2011) indicate that the control of leaf gas exchange leads to relatively stable assimilation-weighted leaf temperatures (~19-25 °C from temperate to tropical regions) despite large differences in air temperature. This is mostly due to the effects of transpiration on leaf energy balance. Franks et al. (2014) chose a fixed temperature of 25 °C because much of the Mesozoic and Cenozoic correspond to climates warmer than the present-day. When applying the Franks model to known cooler paleoenvironments, improved accuracy may be achieved with leaf-temperature-appropriate values for A_0, Γ^*, and d/v.

Bernacchi et al. (2003) proposed the following temperature sensitivity for Γ^* based on experiments:

$$\Gamma^* = e^{\left(19.02-\frac{37.83}{RT}\right)}, \quad (6)$$

where R is the molar gas constant (8.31446×10^{-3} kJ K^{-1} mol^{-1}) and T is leaf temperature (K). Marrero and Mason (1972) describe the sensitivity of water vapor diffusivity to temperature as:

$$d = 1.87 \times 10^{-10}\left(\frac{T^{2.072}}{P}\right), \quad (7)$$

where P is atmospheric pressure, which we fix at 1 atmosphere. Lastly, the temperature sensitivity of the molar volume of air follows ideal gas principles:

$$v = v_{STP}\left(\frac{T}{T_{STP}}\right)\left(\frac{P}{P_{STP}}\right), \quad (8)$$

where T_{STP} is 273.15 K, P_{STP} is 1 atmosphere, and v_{STP} is the air volume at T_{STP} and P_{STP} (0.022414 m^3 mol^{-1}).

Using Eqs. (6-8), we can describe how, conceptually, the sensitivities of *Γ** and *d/v* to leaf temperature affect estimates of CO_2 from the Franks model. We apply these relationships to a suite of 409 fossil and extant leaves from 62 species of angiosperms, gymnosperms, and ferns. These data come from the current study (see Sect. 2.1 and 2.4) and Londoño et al. (2018), Kowalczyk et al. (2018), and Milligan et al. (2019).

To experimentally test more generally how the Franks model is influenced by temperature, we grew six species of plants inside two growth chambers with contrasting temperatures (Conviron E7/2; Winnipeg, Canada). Air temperature was 28 ± 0.5 °C (1σ) and 20 ± 0.3 °C during the day, and 19 ± 0.7 °C and 11 ± 1.1 °C during the night. We note that the difference in leaf temperature was probably smaller than that in air temperature during the day (8 °C; see earlier discussion). We held fixed the day length (17 hours with a 30 minute simulated dawn and dusk) and CO_2 concentration (500 ± 10 ppm). Light intensity at the heights where we sampled leaves ranged from 100-400 µmol m^{-2} s^{-1}. Humidity differed moderately between chambers (76.5 ± 1.8% ~~1σ~~ and 90.0 ± 3.6%). To minimize any chamber effects, we alternated plants between chambers every two weeks.

Four of the species started as saplings purchased from commercial nurseries: bare-root, one-foot tall saplings of *Acer negundo* and *Carpinus caroliniana*, one-foot tall saplings of *Ostrya virginiana* with a soil ball, and bare-root, four-inch tall saplings of *Ilex opaca*. We grew the other two species from seed: *Betula lenta* from a commercial source, and *Quercus rubra* from a single tree on Wesleyan University's campus. All seeds were soaked in water for 24 hours and then cold stratified in a refrigerator for 30 and 60 days, respectively.

All seeds and saplings grew in the same potting soil (Promix Bx with Mycorise; Premier Horticulture; Quakertown, Pennsylvania, USA) and fertilizer (Scotts all-purpose flower and vegetable fertilizer; Maryville, Ohio, USA). They were watered to field capacity every other day, and we discarded any excess water passing through the pots. After three months of growth in the chambers, for each species-chamber pair we harvested the three newest fully expanded leaves whose buds developed during the experiment. In most cases, we harvested five plants per species-chamber pair; the one exception was *I. opaca*, where we were limited to three plants in the warm treatment and two in the cool treatment.

We measured stomatal size and density on cleared leaves (using 50% household bleach) with transmitted-light microscopy. Whole-leaf $\delta^{13}C$ comes from the UC Davis Stable Isotope Facility and the Light Stable Isotope Mass Spec Lab at the University of Florida; the same leaves were measured for $\delta^{13}C$ and stomatal morphology. We used either a hole punch or razor to remove two adjacent sections of leaf tissue near the leaf centers, avoiding major veins. Because we used the same CO_2 gas cylinder ($\delta^{13}C$ = -11.8‰) and laboratory space ($\delta^{13}C$ = -10.4‰) as Milligan et al. (2019), we used their two-end-member mixing model ($1/CO_2$ vs. $\delta^{13}C$; Keeling, 1958) to calculate the $\delta^{13}C$ of the chamber CO_2 at 500 ppm (-10.6 ‰). We used the recommended values from Franks et al. (2014) for the physiological inputs A_0 and $g_{c(op)}/g_{c(max)}$. Table S1 summarizes all of the inputs from this experiment needed to run the Franks model, along with the estimated CO_2 concentrations. The standard errors for the inputs are based on plant means.

To test if leaf $\delta^{13}C$ and stomatal morphology (stomatal density, stomatal pore length, and single guard cell width) differed between temperature treatments across species, we implemented a mixed model in R (R Core Team, 2016) using the lme4 (Bates et al., 2015) and lmerTest (Kuznetsova et al., 2017) packages, with temperature and species as the two fixed factors. To test if there was a significant

difference between CO_2 estimates from the two temperature treatments, we ran a Kolmogorov–Smirnov (KS) test in R. For each species, we first estimated CO_2 for each plant in the warm and cool treatments based on simulated inputs constrained by their means and variances. In the typical case with five plants per chamber, this produced five CO_2 estimates for the warm chamber and the same for the cool chamber. A KS test was then used to test for a significant temperature effect. We repeated this procedure 10,000 times, with 10,000 associated KS tests. The fraction of tests with a p-value < 0.05 was taken as the overall p value. An advantage of this approach is that it incorporates both within- and across-plant variation.

2.3 Photorespiration

c_i/c_a is estimated in the Franks model following Farquhar et al. (1982):

$$\Delta_{leaf} = a + (b - a) \times \frac{c_i}{c_a}, \quad (9)$$

where a is the carbon isotope fractionation due to diffusion of CO_2 in air (4.4‰; Farquhar et al., 1982), b is the fractionation associated with RuBP carboxylase (30‰; Roeske and O'Leary, 1984), and Δ_{leaf} is the net fractionation between air and assimilated carbon ($[\delta^{13}C_{air} - \delta^{13}C_{leaf}]/[1+\delta^{13}C_{leaf}/1000]$).

Equation (9) can be expanded to include other effects, including photorespiration (Farquhar et al., 1982):

$$\Delta_{leaf} = a + (b - a) \times \frac{c_i}{c_a} - \frac{f\Gamma*}{c_a}, \quad (10)$$

where f is the carbon isotope fractionation due to photorespiration. Photorespiration occurs when the enzyme rubisco fixes O_2, not CO_2 (i.e., RuBP oxygenase). One product of photorespiration is CO_2 (Jones, 1992), whose $\delta^{13}C$ is lower than the source substrate glycine. If this respired CO_2 escapes to the atmosphere, the $\delta^{13}C$ of the leaf carbon becomes more positive. Thus, if c_i/c_a is calculated using Eq. (9), as is common practice, the calculation may be falsely low, leading to an underprediction of atmospheric CO_2.

Measured values for f vary from ~9-15‰ (see compilation in Schubert and Jahren, 2018), which is in line with theoretical predictions (Tcherkez, 2006). At a 400 ppm atmospheric CO_2 and $\Gamma*$ of 40 ppm, Eq. (10) implies that ~1‰ of Δ_{leaf} is due to photorespiration, meaning that c_i/c_a should be ~0.04 higher relative to Eq. (9). Here, using the suite of fossil and extant leaves described in Sect. 2.2, we explore how the carbon isotopic fractionation associated with photorespiration affects CO_2 estimates with the Franks model. Because c_i/c_a is present in both of the fundamental equations (Eqs. 2 and 3), we solve them iteratively until c_i/c_a converges.

2.4 Leaves that grow close to the forest floor

The composition of air close to the forest floor can differ considerably from the well-mixed atmosphere. Of relevance to the Franks model, soil respiration can lead to a locally higher CO_2 concentration and lower $\delta^{13}C_{air}$ (Table 2~~1~~). This effect is strongest at night, when the forest boundary layer is thickest (e.g., Munger and Hadley, 2017), but we focus here on daylight hours because that is when most plants take up CO_2. In wet tropical forests, which can have very high soil respiration rates, CO_2 during the day near the forest floor can be elevated by tens-of-ppm, and the $\delta^{13}C_{air}$ can be 2-3‰ lower; in temperate forests,

the deviations are smaller (Table 2~~1~~). Above ~2 m, CO_2 concentrations and air $\delta^{13}C$ during the daytime largely match the well-mixed atmosphere.

Table 2~~1~~. Deviations in the $\delta^{13}C$ and concentration of CO_2 close to a forest floor relative to well-mixed air above the canopy. All measurements were made close to mid-day.

Study	$\delta^{13}C_{air}$ relative to well-mixed air (‰)	CO_2 relative to well-mixed air (ppm)	Height above forest floor (m)	Forest location
Tropical forest				
Broadmeadow et al. (1992)	-2	+20	0.15-1	Trinidad during dry season
Buchmann et al. (1997)	-2	+30	0.70-0.75	French Guiana during wet and dry seasons
Holtum and Winter (2001)	NA	+50	0.10	Panama during wet and dry seasons
Lloyd et al. (1996)	-3	+70	1	Brazil (Amazon Basin)
Quay et al. (1989)	-3	+20	2	Brazil (Amazon Basin)
Sternberg et al. (1989)	-2	+25	1	Panama during wet and dry seasons
Temperate forest				
Francey et al. (1985)	-1	+20	1	Tasmania
Munger and Hadley (2017)	NA	+15	1	Massachusetts (Harvard Forest)

As a result, leaves that grow close to the forest floor may cause the Franks model to produce CO_2 estimates higher than that of the mixed atmosphere for at least two reasons. First, the concentration of CO_2 near the forest floor is elevated; that is, the model may correctly estimate a CO_2 concentration that the user is not interested in. Second, because the $\delta^{13}C_{air}$ that a forest-floor plant experiences is lower than the global well-mixed value, if the user chooses the well-mixed value for model input (inferred, for example, from the $\delta^{13}C$ of marine carbonate; Tipple et al., 2010), then c_i/c_a and thus atmospheric CO_2 will be overestimated (see Eq. 2).

We sought to test how the Franks model is affected by the forest-floor microenvironment for five tropical angiosperm species and fifteen temperate angiosperm and fern species. The tropical leaves were sampled at ~1-2 m height from Parque Nacional San Lorenzo, Panama. In contrast to the canopy data set from San Lorenzo (Sect. 2.1), these CO_2 estimates have not been previously reported. In the summer of 2015, seven fern species were sampled at ~0.5 m height from Connecticut College and Wesleyan University. Also, we used leaf vouchers from Royer et al. (2010), who sampled eight herbaceous angiosperm species at ~0.1-0.2 m height from Reed Gap, Connecticut. For all 20 species, stomatal and carbon isotopic measurements follow the methods described in Sect. 2.1. Table S1 contains all of the inputs needed to run the Franks model, along with the estimated CO_2 concentrations.

We also investigated if we could include the forest-floor $\delta^{13}C_{air}$ effect in our estimates of atmospheric CO_2. We did not measure the CO_2 concentration and $\delta^{13}C_{air}$ around our plants, so we could not directly compare our values. But, ~~I~~if the only CO_2 inputs close to the forest floor are from the soil and well-mixed atmosphere, then the system can be modeled as a two-endmember mixing model

where $\delta^{13}C_{air}$ has a positive, linear relationship with $1/CO_2$ (Keeling, 1958). If the CO_2 concentration and $\delta^{13}C$ of both endmembers are known, the forest-floor microenvironment should fall somewhere on the modelled line. Importantly, the Franks model provides a second constraint on the system. Here, $\delta^{13}C_{air}$ has a negative, nonlinear relationship with $1/CO_2$ because $\delta^{13}C_{air}$ is positively related to c_i/c_a and CO_2. The Franks model thus provides a second calculation for the relationship between $\delta^{13}C_{air}$ and estimated CO_2 concentration. The intersection between the two curves should be the correct $\delta^{13}C_{air}$ and CO_2 concentration for the forest-floor microenvironment.

To estimate the soil CO_2 endmember, we measured the $\delta^{13}C$ of soil organic matter collected from the A horizons of 13 soil sites at San Lorenzo, and of five each at Reed Gap and Connecticut College. For all soils, we assume a 5000 ppm CO_2 concentration for a depth that is below the zone of CO_2 diffusion from the atmosphere (~0.3 m; Cerling, 1999; Breecker et al., 2009). The true value for wet temperate and tropical forest soils may be somewhat less or substantially more than 5000 ppm (Medina et al., 1986; Cerling, 1999; Hirano et al., 2003; Hashimoto et al., 2004; Sotta et al., 2004). Because the mixing model uses $1/CO_2$, a much higher CO_2 concentration (e.g., 10000 ppm) has little impact on our results.

3 Results and Discussion

3.1 General testing in living plants

Estimates of CO_2 across the 40 tree species sampled in the field range from 275 to 850 ppm, with a mean of 478 ppm and median of 472 ppm (Fig. 2); two-thirds of the estimates (a close equivalent to ±1 standard deviation) range between 353 and 585 ppm. In 28% of the tested species, the estimated CO_2 concentrations overlap with the target concentration (400 ppm) at 95% confidence; for these species, the CO_2 estimates do not differ significantly from the target. There are no strong differences across taxonomic orders, nor between leaves from tropical and temperate forests. The mean error rate across the estimates is 28% (median = 24%), which is higher than estimates that include direct measurements of the physiological inputs A_0 and $g_{c(op)}/g_{c(max)}$ (mean = 20%; median = 13%; Fig. 1). Along similar lines, if the estimates presented in Fig. 1 are re-estimated using the values for A_0 and $g_{c(op)}/g_{c(max)}$ recommended by Franks et al. (2014), the mean error rate increases to 37% (median = 33%). If only the default values of A_0 are used, the median error rate is 27%; and for only default values of $g_{c(op)}/g_{c(max)}$, 22%.

These results indicate that CO_2 accuracy is generally improved when A_0 and/or $g_{c(op)}/g_{c(max)}$ is measured. These measurements require expensive gas-exchange equipment and are not always easy or practical to make. Moreover, A_0 and $g_{c(op)}/g_{c(max)}$ cannot be measured on fossils. Some gains in accuracy are possible by measuring A_0 and $g_{c(op)}/g_{c(max)}$ on extant relatives of the fossil species (e.g., the same genus). Absent of this, our analysis using the recommended mean values of Franks et al. (2014) indicates an error rate, on average, of approximately 28%. This is comparable to or better than other leading paleo-CO_2 proxies (Franks et al., 2014).

One reliable way to improve accuracy is to estimate CO_2 with multiple species because the falsely-high and falsely-low estimates partly cancel each other out. The grand mean of estimates presented in Fig. 2 (478 ppm) is 20% from the 400 ppm target, which is less than the 28% mean error rate of individual estimates.

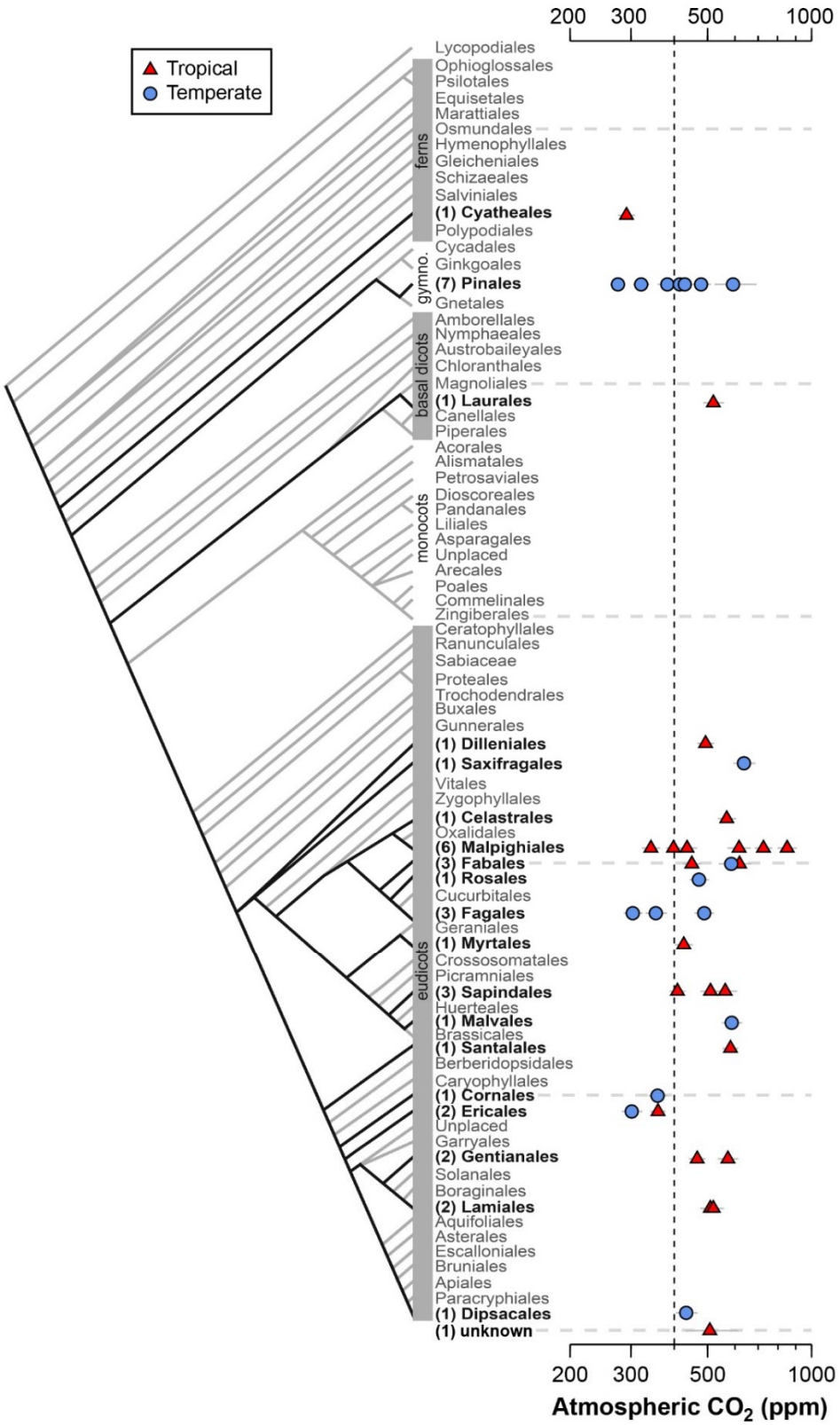

Figure 2. Estimates of CO_2 based on canopy leaves from 40 tree species. Uncertainties in the estimates correspond to the 16^{th}-84^{th} percentile range. Vertical line is the correct concentration (400 ppm). On the left is an order-level vascular plant phylogeny (APW v.13; Stevens, 2001 onwards).

Dow et al. (2014) observed that $g_{c(op)}/g_{c(max)}$ inversely varies with CO_2 in *Arabidopsis thaliana*, but primarily at subambient concentrations (yellow triangles in Fig. 3). At elevated CO_2, $g_{c(op)}/g_{c(max)}$ is close to 0.2, which is the value recommended by Franks et al. (2014). Data from eleven species of angiosperms, conifers, and ferns at present-day (or near present-day) and elevated CO_2 concentrations support the view of a limited effect at high CO_2 (Fig. 3; Franks et al., 2014; Londoño et al., 2018; Milligan et al., 2019). More data at subambient CO_2 are needed, but for CO_2 concentrations similar to or higher than the present-day, we see no strong reason to include a CO_2 sensitivity in $g_{c(op)}/g_{c(max)}$. The rather low values for *Cedrus deodara* and many of the tropical angiosperms (<0.1) are likely due to stress imposed by their growth chamber environment; these $g_{c(op)}/g_{c(max)}$ values are probably not representative of field-grown trees, which tend to be closer to 0.2 (Franks et al., 2014).

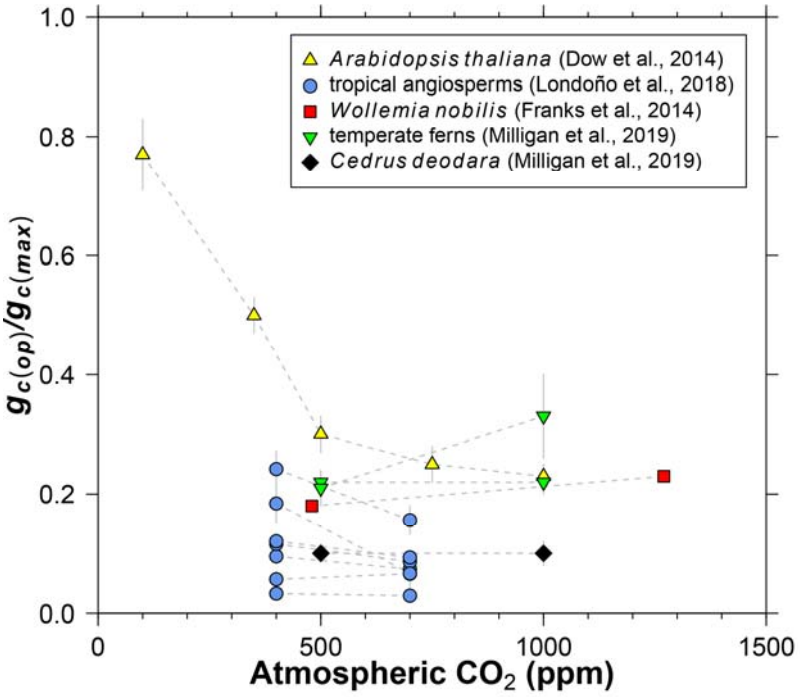

Figure 3. Literature compilation of the sensitivity of $g_{c(op)}/g_{c(max)}$ (ratio of operational to maximum leaf conductance to CO_2) to atmospheric CO_2 concentration.

3.2 Temperature

The temperature sensitivities of the ratio of diffusivity of CO_2 in air to the molar volume of air (d/v) and the CO_2 compensation point in the absence of dark respiration (Γ^*) have little effect on estimated CO_2 in the Franks model (Fig. 4). Given that assimilation-weighted leaf temperature only varies about 7 °C across plants today, the differences shown in Fig. 4—which are based on leaf temperatures of 25 °C and 15 °C—are likely a maximum effect. As such, we consider the use of a fixed leaf temperature (e.g., 25 °C) in the model to be a defensible simplification.

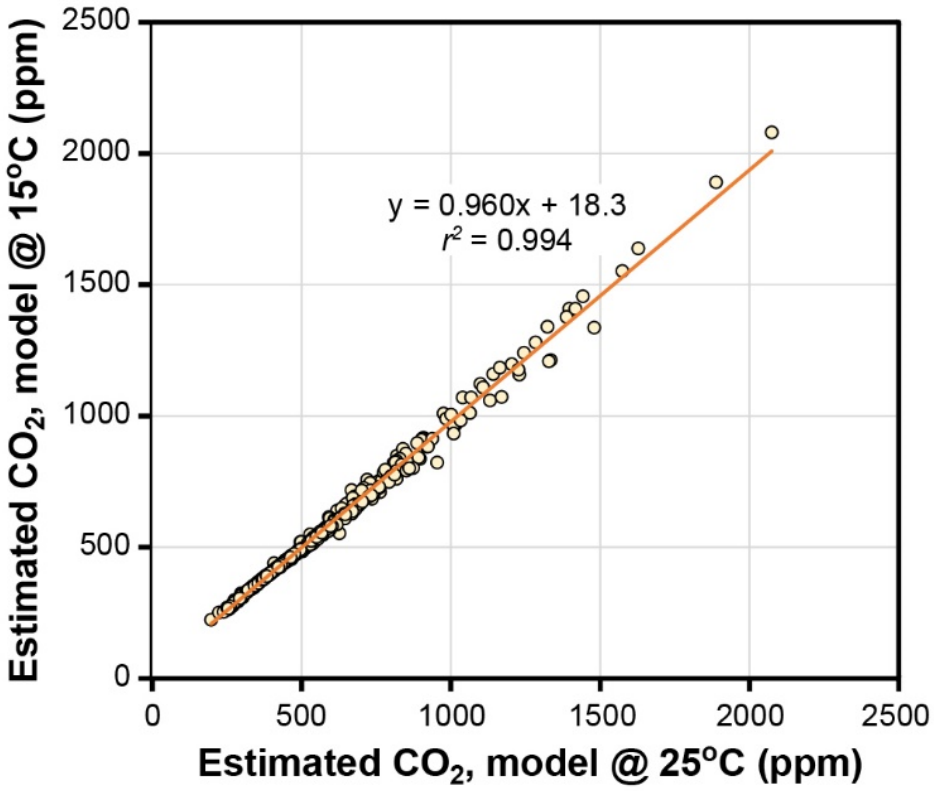

Figure 4. Estimates of CO_2 at leaf temperatures of 25 °C and 15 °C. Each symbol is an extant or fossil leaf. The difference in estimated CO_2 for any leaf is due to the theoretical effects of temperature on gas diffusion (d/v) and the CO_2 compensation point in the absence of dark respiration (Γ^*) (Eqs. 6-8).

Other inputs in the model may respond to temperature, though. In our growth chamber experiments where daytime air temperatures were 28 °C and 20 °C, the effect on estimated CO_2 was mixed (Fig. 5). In five out of six species, estimated CO_2 was higher in the warm treatment, but for all species these differences were not statistically significant ($P > 0.05$ based on a KS test; see Methods). Incorporating the temperature sensitivities in d/v and Γ^* had little effect ("adj" estimates in Fig. 5), as expected from Fig. 4.

None of the measured inputs—stomatal density, stomatal pore length, single guard cell width, and leaf $\delta^{13}C$—were significantly affected by temperature across all species ($P > 0.05$ for each of the four inputs based on a mixed model; see Methods). These small differences probably cannot account for the differences in estimated CO_2 between temperatures. It is more likely that some of the inputs that we did not directly measure, such as assimilation rate (A_0), the $g_{c(op)}/g_{c(max)}$ ratio, or mesophyll conductance (g_m), differ from the true mean value. In the cases for the five species where estimated CO_2 is higher in the warm treatment, our mean A_0 for the warm plants must be falsely high, or $g_{c(op)}/g_{c(max)}$ or g_m falsely low.

In summary, we see no strong reason to expand the parameterization of temperature in the model, though more growth-chamber experiments may be warranted. We note that in three out of the six species from the warm treatment, the estimated CO_2 cannot be distinguished from the 500 ppm target at 95% confidence; for the cool treatment, this is true for four of the species. Also, the across-species means of estimated CO_2 for the warm and cool treatments are reasonably close to the ~~500 ppm~~ target (590 and 502 ppm, respectively) and overall have a mean error rate of 25%. This level of uncertainty is similar to our field estimates where we did not measure A_0 or $g_{c(op)}/g_{c(max)}$ (28%; see Fig. 2).

This too provides support for our recommendation that it is not critical to include a broader treatment of temperature in the model.

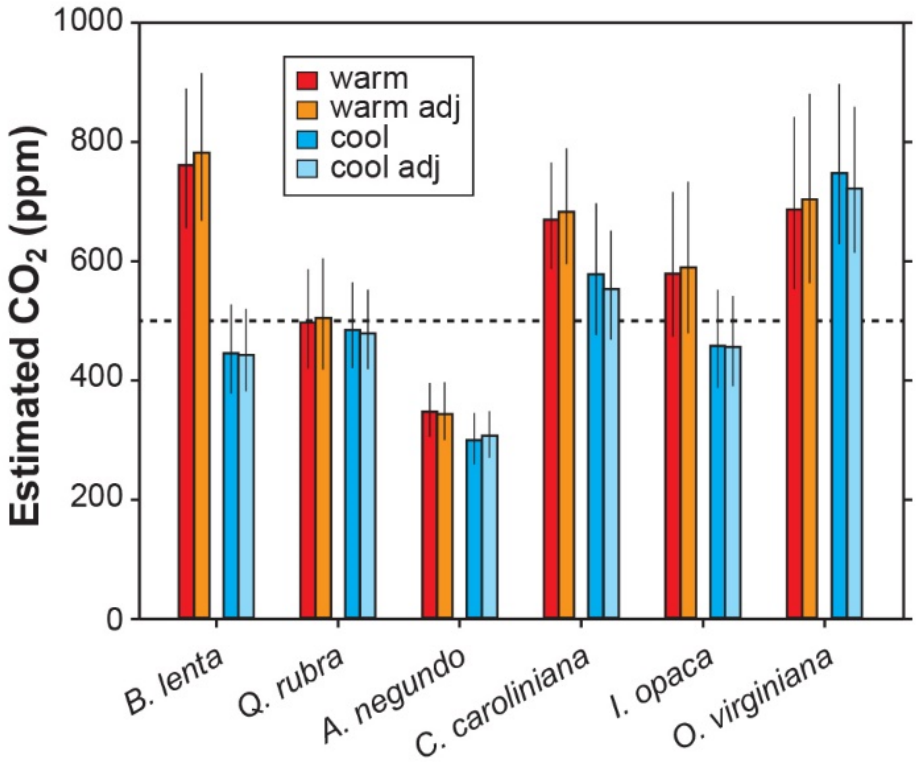

Figure 5. Estimates of CO_2 for plants grown inside growth chambers at daytime air temperatures of 28 °C and 20 °C. Also shown are estimates after taking into account the temperature sensitivity of gas diffusion (*d/v*) and the CO_2 compensation point in the absence of dark respiration (*Γ**) ("adj"; see also Fig. 4). Dashed line is the correct CO_2 concentration (500 ppm). Uncertainties in the estimates correspond to the 16^{th}-84^{th} percentile range.

3.3 Photorespiration

The theoretical effects of photorespiration do not strongly impact estimates of CO_2 in the Franks model. The average effect for our 409 extant and fossil leaves is to increase estimated CO_2 by 2.2% plus 38 ppm (Fig. 6). At 1000 ppm, for example, estimates would increase by 60 ppm. This calculation assumes a photorespiration fractionation (*f*) of 9.1‰, which is the value estimated for *Arabidopsis thaliana* (Schubert and Jahren, 2018). If a fractionation towards the upper bound of published estimates is used instead (15‰), estimated CO_2 increases on average by 3.8% plus 61 ppm. Across this range in *f*, the associated uncertainty in estimated CO_2 is well within the method's overall precision (~+35/-25% at 95% confidence; Franks et al., 2014). As such, CO_2 estimates made without these photorespiration effects (i.e. using Eq. 9 instead of Eq. 10) should be reliable, although some improvement is possible using Eq. 10 in cases where *f* is accurately known.

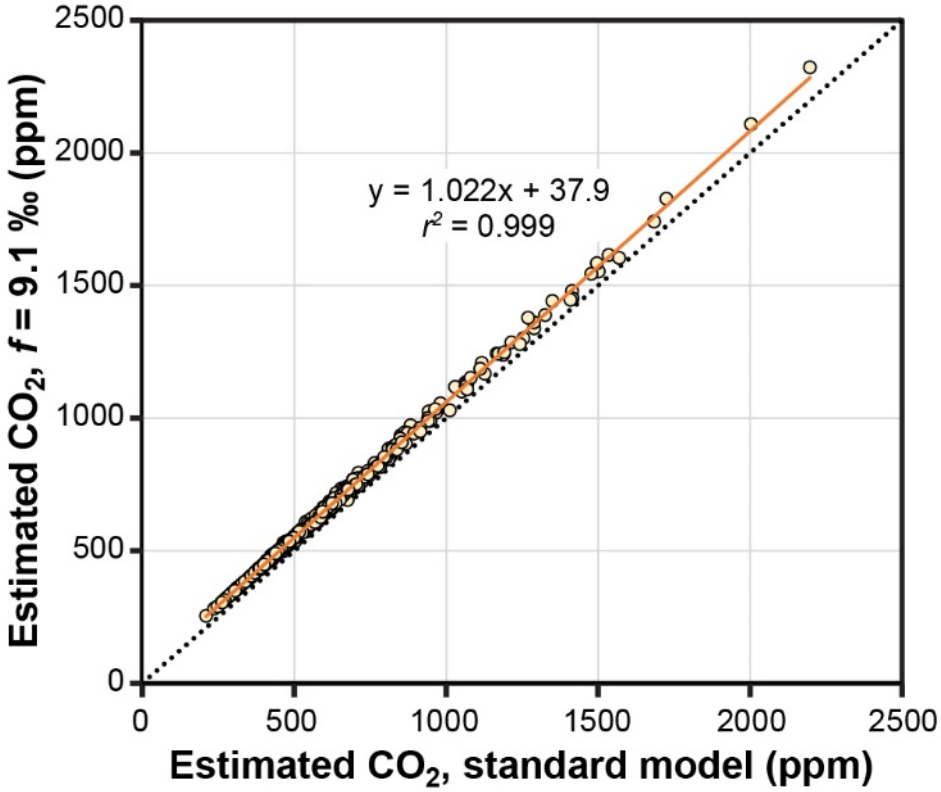

Figure 6. Estimates of CO_2 with and without a photorespiration effect (f = 9.1‰; see Eq. 10). Each symbol is an extant or fossil leaf. Dashed line is y=x.

We note that both f and Γ^* are also affected by atmospheric O_2 concentration. Because O_2 is directly responsible for photorespiration, f should scale with O_2 (or, more precisely, the O_2:CO_2 molar ratio). Unfortunately, this effect is poorly constrained (Beerling et al., 2002; Berner et al., 2003; Porter et al., 2017). In contrast, the theoretical effect of O_2 on Γ^* is known: it is linear with an approximate slope of 2 (Farquhar et al., 1982; see their Eq. B13). During the Phanerozoic, O_2 likely ranged from 10-30%, with lows during the early Paleozoic and early Triassic, and highs during the Carboniferous to early Permian and Cretaceous (Berner, 2009; Glasspool and Scott, 2010; Arvidson et al., 2013; Mills et al., 2016; Lenton et al., 2018). Assuming a present-day Γ^* of 40 ppm (at 21% O_2), Γ^* would be 60 ppm at 30% O_2 and 20 ppm at 10% O_2. Running the Franks model on our library of 409 extant and fossil leaves, we find little effect on estimated CO_2: estimates are 7.4% higher on average at 30% O_2 than at 10% O_2 (see also McElwain et al., 2016).

3.4 Leaves that grow close to the forest floor

CO_2 estimates for tropical understory leaves from five species at San Lorenzo, Panama, are very high, ranging from 1903 to 18863 ppm (species mean = 6837 ppm). For two of the species, Londoño et al. (2018) also analyzed canopy leaves from trees nearby, and these within-species comparisons highlight the vast discrepancy (*Ocotea* sp.: 541 vs. 5737 ppm; *Tontelea* sp.: 622 vs. 18863 ppm). The primary difference in the inputs between the canopy and understory leaves is the $\delta^{13}C_{leaf}$: Londoño et al. (2018) report a species-mean $\delta^{13}C_{leaf}$ of -30.0‰ for the 21 canopy species versus -35.6‰ for the five understory

species. This difference leads to very different mean estimates of c_i/c_a: 0.69 for canopy leaves versus a highly unrealistic (Diefendorf et al., 2010) 0.93 for understory leaves.

It is likely that the high CO_2 estimates from understory leaves are mostly driven by falsely high $\delta^{13}C_{air}$ inputs. Following the mixing model strategy outlined in Sect. 2.4 (and based on a soil organic matter $\delta^{13}C$ of -28.2‰ measured at San Lorenzo), we calculate a species-mean $\delta^{13}C_{air}$ of -16.7‰ (mean of intersection points in Fig. 7). When this $\delta^{13}C_{air}$ is used to estimate CO_2 with the Franks model (instead of -8.5‰), the species mean drops to 699 ppm. This is somewhat higher than the species mean from canopy leaves in the same forest (563 ppm; red triangles in Fig. 2; Londoño et al., 2018).

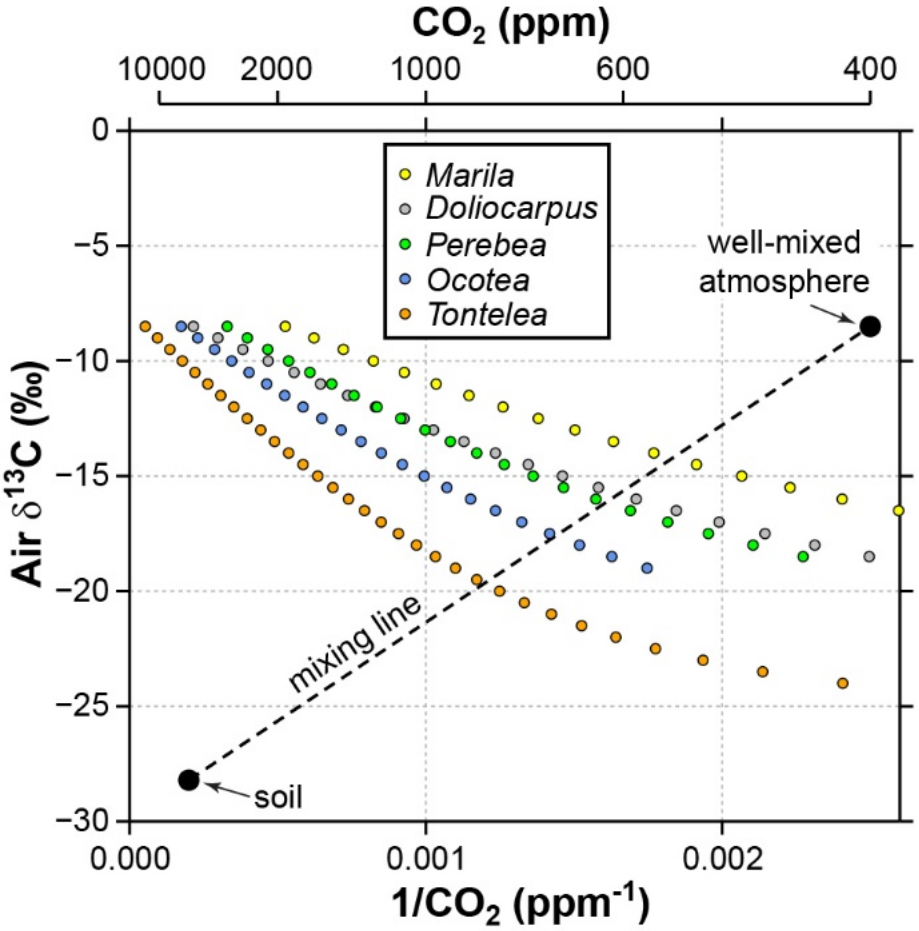

Figure 7. Sensitivity of estimated CO_2 in the Franks model to the $\delta^{13}C$ of atmospheric CO_2. Estimates come from leaves of five angiosperm species that grew close to the forest floor in Parque Nacional San Lorenzo, Panama. The step in $\delta^{13}C_{air}$ between estimates is 0.5‰. The dashed line is a two-endmember mixing model for CO_2 between the soil and well-mixed atmosphere. The intersection between the mixing model and the Franks model should correspond to the CO_2 concentration and $\delta^{13}C_{air}$ of the forest-floor microenvironment.

Understory leaves from Connecticut temperate forests show similar but less dramatic patterns, which we attribute to a more open canopy with stronger atmospheric mixing. CO_2 estimates for the 15 species range from 447 to 1567 ppm (mean = 794 ppm). Our intersection method identifies a mean

$\delta^{13}C_{air}$ of -11.2‰ for the Wesleyan and Connecticut College campuses (based on a soil $\delta^{13}C$ of -27.6‰ measured at Connecticut College) and -10.3‰ for Reed Gap (soil $\delta^{13}C$ = -26.4‰). Using these adjusted $\delta^{13}C_{air}$, the species mean of estimated CO_2 drops to 566 ppm, which is somewhat higher than the species mean from canopy leaves in the same areas (449 ppm; blue circles in Fig. 2).

We acknowledge that this analysis is too simple. First, we did not measure the understory CO_2 concentration and $\delta^{13}C_{air}$ (this would require repeated measurements during different daytime hours over a growing season to calculate a time-integrated value); instead, we assumed a simple two end-member mixing model between the soil and well-mixed atmosphere. Second, other factors probably contribute to the differences in estimated CO_2 between canopy and understory leaves. Our predicted $\delta^{13}C_{air}$ values are too low (~8‰ and 2‰ lower than the well-mixed atmosphere for the tropical and temperate forests) and our estimated CO_2 too high (~100 ppm higher than that from canopy leaves). In the lowermost 1-2 meters of the canopy, previous work suggests up to a -3‰ and +70 ppm deviation in tropical forests and -1‰ / +20 ppm in temperate forests (Table 1). One input that could help to resolve this discrepancy is the assimilation rate (A_0). We assumed a fixed A_0 of 12 µmol m^{-2} s^{-1} for all leaves, regardless of canopy position. Shade leaves often have lower assimilation rates than sun leaves (Givnish, 1988). Substituting lower A_0 values for understory leaves would lower estimated CO_2 roughly in proportion (Eqs. 2-3). Using lower A_0 values for shade leaves in the model is appropriate, but determining the best value is difficult. Typical A_0 values for leaves growing at the top of the canopy in full sun are far more consistent because photosynthesis in these leaves is usually at its maximum capacity (saturated at full sunlight) for the prevailing atmospheric CO_2 concentration. Because the degree of shadiness near the forest floor is highly variable, photosynthesis (A_0) in these leaves will be acclimated to some fraction of the full-sun maximum in a sun exposed leaf, but careful thought must go into determining what this fraction is.

We note that our mixing-model strategy cannot be applied to fossils because the global atmospheric CO_2 concentration is needed (one endpoint for dashed line in Fig. 7). Instead, our motivation for the analysis is to demonstrate that: 1) leaves growing in the lowermost 2 m of the canopy should be considered with caution in the context of the Franks model; and 2) the failure of the model is due to faulty inputs (mostly $\delta^{13}C_{air}$), not the model itself.

In most fossil leaf deposits, shade morphotypes are comparatively rare (e.g., Kürschner, 1997; Wang et al., 2018) because—relative to sun leaves—they are ~~not as~~less ~~tough~~durable, do not travel as far by wind, and are produced at a slower rate (Dilcher, 1973; Roth and Dilcher, 1978; Spicer, 1980; Ferguson, 1985; Burnham et al., 1992). Our recommendation is to exclude such leaves. There are several ways to differentiate sun vs. shade morphotypes: overall shape (Talbert and Holch, 1957; Givnish, 1978; Kürschner, 1997; Sack et al., 2006), shape of epidermal cells (larger and with a more undulated outline in shade leaves; Kürschner, 1997; Dunn et al., 2015), vein density (lower in shade leaves; Uhl and Mosbrugger, 1999; Sack and Scoffoni, 2013; Crifò et al., 2014; Londoño et al., 2018), and range in $\delta^{13}C_{leaf}$ (high when both sun and shade leaves are present, for example in our study; Graham et al., 2014). Not all shade leaves grow within 2 m of the forest floor, but excluding all such leaves would eliminate the forest-floor bias.

4 Conclusions

The Franks model is reasonably accurate (~28% error rate) even when the physiological inputs A_0 (assimilation rate at a known CO_2 concentration) and $g_{c(op)}/g_{c(max)}$ (ratio of operational to maximum leaf conductance to CO_2) are inferred, not measured. Accuracy does improve when these inputs are measured (~20% error rate), but such measurements are not possible with fossils and may not always

be feasible with nearest living relatives. A 28% error rate is broadly in line with (or better than) other leading paleo-CO_2 proxies.

Most of the possible confounding factors that we investigated appear minor. The temperature sensitivities of d/v (related to gas diffusion) and Γ^* (CO_2 compensation point in the absence of dark respiration) have a negligible impact on estimated CO_2. Our temperature experiments in growth chambers point to larger differences in some species, which must be related to incorrect values for inputs that were not directly measured, such as A_0, $g_{c(op)}/g_{c(max)}$, and g_m (mesophyll conductance). Overall, though, we find that the differences in estimated CO_2 imparted by temperature are generally smaller than the overall 28% error rate.

Incorporating the covariance between CO_2 concentration and photorespiration leads to only small changes in estimated CO_2. O_2 concentration affects photorespiration and thus may confound CO_2 estimates from the Franks model, but presently the effect is poorly quantified. The effect of O_2 on Γ^* is better known, and imparts only small changes in estimated CO_2 across a feasible range in Phanerozoic O_2 of 10-30%.

Leaves from the lowermost 1-2 m of the canopy experience slightly elevated CO_2 concentrations and lower air $\delta^{13}C$ during the daytime relative to the well-mixed atmosphere. We find that if we use the well-mixed air $\delta^{13}C$ to estimate CO_2 from leaves that grew near the forest floor, estimates are too high, especially in dense tropical canopies. When we use a two-endmember mixing model to calculate the correct local air $\delta^{13}C$, the falsely-high CO_2 estimates largely disappear. For fossil applications, shade leaves from the bottom of the canopy should be avoided. Shade leaves are typically rare in the fossil record (relative to sun leaves), and can be identified by their overall shape, the shape of their epidermal cells, their low leaf $\delta^{13}C$, and their low vein density.

Conceptually, the Franks model holds considerable promise for quantifying paleo-CO_2: it is mechanistically grounded and can be applied to most fossil leaves. Our tests of the model's accuracy and sensitivity to temperature and photorespiration largely uphold this promise.

Author contribution. DR, KM, MM, and LL designed and conducted the experiments; all authors interpreted the data; DR prepared the manuscript with contributions from all co-authors.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We thank G. Dreyer and P. Siver for logistical support at Connecticut College, S. Wang for lab assistance, ~~and~~ C. Crifò and A. Baresh for collecting the tropical samples, and J. McElwain, E. Brook, and an anonymous reviewer for helpful comments. Support for LL was provided by the Smithsonian Tropical Research Institute; the Mark Tupper Fellowship; National Science Foundation grants EAR 0824299 and OISE, EAR, DRL 0966884; the Anders Foundation; and the Gregory D. and Jennifer Walston Johnson and 1923 Fund.

References

Arvidson, R. S., Mackenzie, F. T., and Guidry, M. W.: Geologic history of seawater: A MAGic approach to carbon chemistry and ocean ventilation, Chemical Geology, 362, 287-304, https://doi.org/10.1016/j.chemgeo.2013.10.012, 2013.

Barclay, R. S. and Wing, S. L.: Improving the *Ginkgo* CO_2 barometer: implications for the early Cenozoic atmosphere, Earth and Planetary Science Letters, 439, 158-171, https://doi.org/10.1016/j.epsl.2016.01.012, 2016.

598 Bates, D., Mächler, M., Bolker, B., and Walker, S.: Fitting linear mixed-effects models using lme4, Journal
599 of Statistical Software, 67, https://doi.org/10.18637/jss.v067.i01, 2015.
600 Beerling, D. J.: Evolutionary responses of land plants to atmospheric CO_2, in: A History of Atmospheric
601 CO_2 and Its Effects on Plants, Animals, and Ecosystems, edited by: Ehleringer, J. R., Cerling, T. E.,
602 and Dearing, M. D., Springer, New York, 114-132, 2005.
603 Beerling, D. J., Fox, A., and Anderson, C. W.: Quantitative uncertainty analyses of ancient atmospheric
604 CO_2 estimates from fossil leaves, American Journal of Science, 309, 775-787,
605 https://doi.org/10.2475/09.2009.01, 2009.
606 Beerling, D. J., Lake, J. A., Berner, R. A., Hickey, L. J., Taylor, D. W., and Royer, D. L.: Carbon isotope
607 evidence implying high O_2/CO_2 ratios in the Permo-Carboniferous atmosphere, Geochimica et
608 Cosmochimica Acta, 66, 3757-3767, https://doi.org/10.1016/S0016-7037(02)00901-8, 2002.
609 Bernacchi, C. J., Pimentel, C., and Long, S. P.: *In vivo* temperature response functions of parameters
610 required to model RuBP-limited photosynthesis, Plant, Cell & Environment, 26, 1419-1430,
611 https://doi.org/10.1046/j.0016-8025.2003.01050.x, 2003.
612 Berner, R. A.: Phanerozoic atmospheric oxygen: new results using the GEOCARBSULF model, American
613 Journal of Science, 309, 603-606, https://doi.org/10.2475/07.2009.03, 2009.
614 Berner, R. A., Beerling, D. J., Dudley, R., Robinson, J. M., and Wildman, R. A.: Phanerozoic atmospheric
615 oxygen, Annual Review of Earth and Planetary Sciences, 31, 105-134,
616 https://doi.org/10.1146/annurev.earth.31.100901.141329, 2003.
617 Breecker, D. O., Sharp, Z. D., and McFadden, L. D.: Seasonal bias in the formation and stable isotopic
618 composition of pedogenic carbonate in modern soils from central New Mexico, USA, Geological
619 Society of America Bulletin, 121, 630-640, https://doi.org/10.1130/B26413.1, 2009.
620 Broadmeadow, M., Griffiths, H., Maxwell, C., and Borland, A.: The carbon isotope ratio of plant organic
621 material reflects temporal and spatial variations in CO_2 within tropical forest formations in
622 Trinidad, Oecologia, 89, 435-441, https://doi.org/10.1007/BF00317423, 1992.
623 Buchmann, N., Guehl, J.-M., Barigah, T., and Ehleringer, J. R.: Interseasonal comparison of CO_2
624 concentrations, isotopic composition, and carbon dynamics in an Amazonian rainforest (French
625 Guiana), Oecologia, 110, 120-131, https://doi.org/10.1007/s004420050140, 1997.
626 Burnham, R. J., Wing, S. L., and Parker, G. G.: The reflection of deciduous forest communities in leaf
627 litter: implications for autochthonous litter assemblages from the fossil record, Paleobiology, 18,
628 30-49, https://doi.org/10.1017/S0094837300012203, 1992.
629 Cerling, T. E.: Stable carbon isotopes in palaeosol carbonates, Special Publications of the International
630 Association of Sedimentologists, 27, 43-60, 1999.
631 Chaloner, W. G. and McElwain, J.: The fossil plant record and global climatic change, Review of
632 Palaeobotany and Palynology, 95, 73-82, https://doi.org/10.1016/S0034-6667(96)00028-0,
633 1997.
634 Crifò, C., Currano, E. D., Baresch, A., and Jaramillo, C.: Variations in angiosperm leaf vein density have
635 implications for interpreting life form in the fossil record, Geology, 42, 919-922,
636 https://doi.org/10.1130/g35828.1, 2014.
637 Diefendorf, A. F., Mueller, K. E., Wing, S. L., Koch, P. L., and Freeman, K. H.: Global patterns in leaf ^{13}C
638 discrimination and implications for studies of past and future climate, Proceedings of the
639 National Academy of Sciences, USA, 107, 5738-5743, https://doi.org/10.1073/pnas.0910513107,
640 2010.
641 Dilcher, D. L.: A paleoclimatic interpretation of the Eocene floras of southeastern North America, in:
642 Vegetation and Vegetational History of Northern Latin America, edited by: Graham, A., Elsevier,
643 Amsterdam, 39-53, 1973.

Doria, G., Royer, D. L., Wolfe, A. P., Fox, A., Westgate, J. A., and Beerling, D. J.: Declining atmospheric CO_2 during the late Middle Eocene climate transition, American Journal of Science, 311, 63-75, https://doi.org/10.2475/01.2011.03, 2011.

Dow, G. J., Bergmann, D. C., and Berry, J. A.: An integrated model of stomatal development and leaf physiology, New Phytologist, 201, 1218-1226, https://doi.org/10.1111/nph.12608, 2014.

Dunn, R. E., Strömberg, C. A. E., Madden, R. H., Kohn, M. J., and Carlini, A. A.: Linked canopy, climate, and faunal change in the Cenozoic of Patagonia, Science, 347, 258-261, https://doi.org/10.1126/science.1260947, 2015.

Erdei, B., Utescher, T., Hably, L., Tamás, J., Roth-Nebelsick, A., and Grein, M.: Early Oligocene continental climate of the Palaeogene Basin (Hungary and Slovenia) and the surrounding area, Turkish Journal of Earth Sciences, 21, 153-186, https://doi.org/10.3906/yer-1005-29, 2012.

Farquhar, G., von Caemmerer, S., and Berry, J.: A biochemical model of photosynthetic CO_2 assimilation in leaves of C_3 species, Planta, 149, 78-90, https://doi.org/10.1007/BF00386231, 1980.

Farquhar, G. D. and Sharkey, T. D.: Stomatal conductance and photosynthesis, Annual Review of Plant Physiology, 33, 317-345, https://doi.org/10.1146/annurev.pp.33.060182.001533, 1982.

Farquhar, G. D., O'Leary, M. H., and Berry, J. A.: On the relationship between carbon isotope discrimination and the intercellular carbon dioxide concentration in leaves, Australian Journal of Plant Physiology, 9, 121-137, https://doi.org/10.1071/PP9820121, 1982.

Ferguson, D. K.: The origin of leaf-assemblages—new light on an old problem, Review of Palaeobotany and Palynology, 46, 117-188, https://doi.org/10.1016/0034-6667(85)90041-7, 1985.

Francey, R., Gifford, R., Sharkey, T., and Weir, B.: Physiological influences on carbon isotope discrimination in huon pine (*Lagarostrobos franklinii*), Oecologia, 66, 211-218, https://doi.org/10.1007/BF00379857, 1985.

Franks, P. J., Leitch, I. J., Ruszala, E. M., Hetherington, A. M., and Beerling, D. J.: Physiological framework for adaptation of stomata to CO_2 from glacial to future concentrations, Philosophical Transactions of the Royal Society B, 367, 537-546, https://doi.org/10.1098/rstb.2011.0270, 2012.

Franks, P. J., Royer, D. L., Beerling, D. J., Van de Water, P. K., Cantrill, D. J., Barbour, M. M., and Berry, J. A.: New constraints on atmospheric CO_2 concentration for the Phanerozoic, Geophysical Research Letters, 41, 4685-4694, https://doi.org/10.1002/2014gl060457, 2014.

Givnish, T. J.: Ecological aspects of plant morphology: leaf form in relation to environment, Acta Biotheoretica, 27, 83-142, 1978.

Givnish, T. J.: Adaptation to sun and shade: a whole-plant perspective, Australian Journal of Plant Physiology, 15, 63-92, https://doi.org/10.1071/PP9880063, 1988.

Glasspool, I. J. and Scott, A. C.: Phanerozoic concentrations of atmospheric oxygen reconstructed from sedimentary charcoal, Nature Geoscience, 3, 627-630, https://doi.org/10.1038/ngeo923, 2010.

Graham, H. V., Patzkowsky, M. E., Wing, S. L., Parker, G. G., Fogel, M. L., and Freeman, K. H.: Isotopic characteristics of canopies in simulated leaf assemblages, Geochimica et Cosmochimica Acta, 144, 82-95, https://doi.org/10.1016/j.gca.2014.08.032, 2014.

Grein, M., Utescher, T., Wilde, V., and Roth-Nebelsick, A.: Reconstruction of the middle Eocene climate of Messel using palaeobotanical data, Neues Jahrbuch Für Geologie und Paläontologie Abhandlungen, 260, 305-318, https://doi.org/10.1127/0077-7749/2011/0139, 2011a.

Grein, M., Konrad, W., Wilde, V., Utescher, T., and Roth-Nebelsick, A.: Reconstruction of atmospheric CO_2 during the early Middle Eocene by application of a gas exchange model to fossil plants from the Messel Formation, Germany, Palaeogeography Palaeoclimatology Palaeoecology, 309, 383-391, https://doi.org/10.1016/j.palaeo.2011.07.008, 2011b.

Grein, M., Oehm, C., Konrad, W., Utescher, T., Kunzmann, L., and Roth-Nebelsick, A.: Atmospheric CO_2 from the late Oligocene to early Miocene based on photosynthesis data and fossil leaf

characteristics, Palaeogeography Palaeoclimatology Palaeoecology, 374, 41-51, https://doi.org/10.1016/j.palaeo.2012.12.025, 2013.
Hashimoto, S., Tanaka, N., Suzuki, M., Inoue, A., Takizawa, H., Kosaka, I., Tanaka, K., Tantasirin, C., and Tangtham, N.: Soil respiration and soil CO_2 concentration in a tropical forest, Thailand, Journal of Forest Research, 9, 75-79, https://doi.org/10.1007/s10310-003-0046-y, 2004.
Haworth, M., Heath, J., and McElwain, J. C.: Differences in the response sensitivity of stomatal index to atmospheric CO_2 among four genera of Cupressaceae conifers, Annals of Botany, 105, 411-418, https://doi.org/10.1093/aob/mcp309, 2010.
Helliker, B. R. and Richter, S. L.: Subtropical to boreal convergence of tree-leaf temperatures, Nature, 454, 511-514, https://doi.org/10.1038/nature07031, 2008.
Hirano, T., Kim, H., and Tanaka, Y.: Long-term half-hourly measurement of soil CO_2 concentration and soil respiration in a temperate deciduous forest, Journal of Geophysical Research, 108, 4631, https://doi.org/10.1029/2003JD003766, 2003.
Holtum, J. and Winter, K.: Are plants growing close to the floors of tropical forests exposed to markedly elevated concentrations of carbon dioxide?, Australian Journal of Botany, 49, 629-636, https://doi.org/10.1071/BT00054, 2001.
Jones, H. G.: Plants and Microclimate, Cambridge University Press, Cambridge, 1992.
Keeling, C. D.: The concentration and isotopic abundances of atmospheric carbon dioxide in rural areas, Geochimica et Cosmochimica Acta, 13, 322-334, https://doi.org/10.1016/0016-7037(58)90033-4, 1958.
Konrad, W., Roth-Nebelsick, A., and Grein, M.: Modelling of stomatal density response to atmospheric CO_2, Journal of Theoretical Biology, 253, 638-658, https://doi.org/10.1016/j.jtbi.2008.03.032, 2008.
Konrad, W., Katul, G., Roth-Nebelsick, A., and Grein, M.: A reduced order model to analytically infer atmospheric CO_2 concentration from stomatal and climate data, Advances in Water Resources, 104, 145-157, https://doi.org/10.1016/j.advwatres.2017.03.018, 2017.
Kowalczyk, J. B., Royer, D. L., Miller, I. M., Anderson, C. W., Beerling, D. J., Franks, P. J., Grein, M., Konrad, W., Roth-Nebelsick, A., Bowring, S. A., Johnson, K. R., and Ramezani, J.: Multiple proxy estimates of atmospheric CO_2 from an early Paleocene rainforest, Paleoceanography and Paleoclimatology, 33, 1427-1438, https://doi.org/10.1029/2018PA003356, 2018.
Kürschner, W. M.: The anatomical diversity of recent and fossil leaves of the durmast oak (*Quercus petraea* Lieblein/*Q. pseudocastanea* Goeppert)-implications for their use as biosensors of palaeoatmospheric CO_2 levels, Review of Palaeobotany and Palynology, 96, 1-30, https://doi.org/10.1016/S0034-6667(96)00051-6, 1997.
Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B.: lmerTest package: tests in linear mixed effects models, Journal of Statistical Software, 82, https://doi.org/10.18637/jss.v082.i13, 2017.
Lei, X., Du, Z., Du, B., Zhang, M., and Sun, B.: Middle Cretaceous pCO_2 variation in Yumen, Gansu Province and its response to the climate events, Acta Geologica Sinica, 92, 801-813, https://doi.org/doi:10.1111/1755-6724.13555, 2018.
Lenton, T. M., Daines, S. J., and Mills, B. J. W.: COPSE reloaded: an improved model of biogeochemical cycling over Phanerozoic time, Earth-Science Reviews, 178, 1-28, https://doi.org/10.1016/j.earscirev.2017.12.004, 2018.
Lloyd, J., Kruijt, B., Hollinger, D. Y., Grace, J., Francey, R. J., Wong, S.-C., Kelliher, F. M., Miranda, A. C., Farquhar, G. D., and Gash, J.: Vegetation effects on the isotopic composition of atmospheric CO_2 at local and regional scales: theoretical aspects and a comparison between rain forest in Amazonia and a boreal forest in Siberia, Australian Journal of Plant Physiology, 23, 371-399, https://doi.org/10.1071/PP9960371, 1996.

739 Londoño, L., Royer, D. L., Jaramillo, C., Escobar, J., Foster, D. A., Cárdenas-Rozo, A. L., and Wood, A.:
740 Early Miocene CO_2 estimates from a Neotropical fossil assemblage exceed 400 ppm, American
741 Journal of Botany, 105, 1929-1937, https://doi.org/10.1002/ajb2.1187, 2018.
742 Marrero, T. R. and Mason, E. A.: Gaseous diffusion coefficients, Journal of Physical and Chemical
743 Reference Data, 1, 3-118, https://doi.org/10.1063/1.3253094, 1972.
744 Maxbauer, D. P., Royer, D. L., and LePage, B. A.: High Arctic forests during the middle Eocene supported
745 by moderate levels of atmospheric CO_2, Geology, 42, 1027-1030,
746 https://doi.org/10.1130/g36014.1, 2014.
747 McElwain, J. C.: Do fossil plants signal palaeoatmospheric CO_2 concentration in the geological past?,
748 Philosophical Transactions of the Royal Society London B, 353, 83-96,
749 https://doi.org/10.1098/rstb.1998.0193, 1998.
750 McElwain, J. C. and Chaloner, W. G.: Stomatal density and index of fossil plants track atmospheric
751 carbon dioxide in the Palaeozoic, Annals of Botany, 76, 389-395,
752 https://doi.org/10.1006/anbo.1995.1112, 1995.
753 McElwain, J. C. and Chaloner, W. G.: The fossil cuticle as a skeletal record of environmental change,
754 Palaios, 11, 376-388, https://doi.org/10.2307/3515247, 1996.
755 McElwain, J. C., Montañez, I., White, J. D., Wilson, J. P., and Yiotis, C.: Was atmospheric CO_2 capped at
756 1000 ppm over the past 300 million years?, Palaeogeography Palaeoclimatology Palaeoecology,
757 441, 653-658, https://doi.org/http://dx.doi.org/10.1016/j.palaeo.2015.10.017, 2016.
758 Medina, E., Montes, G., Cuevas, E., and Rokzandic, Z.: Profiles of CO_2 concentration and $\delta^{13}C$ values in
759 tropical rain forests of the upper Rio Negro Basin, Venezuela, Journal of Tropical Ecology, 2, 207-
760 217, https://doi.org/10.1017/S0266467400000821, 1986.
761 Michaletz, S. T., Weiser, M. D., Zhou, J., Kaspari, M., Helliker, B. R., and Enquist, B. J.: Plant
762 thermoregulation: energetics, trait-environment interactions, and carbon economics, Trends in
763 Ecology & Evolution, 30, 714-724, https://doi.org/10.1016/j.tree.2015.09.006, 2015.
764 Michaletz, S. T., Weiser, M. D., McDowell, N. G., Zhou, J., Kaspari, M., Helliker, B. R., and Enquist, B. J.:
765 The energetic and carbon economic origins of leaf thermoregulation, Nature Plants, 2, 16129,
766 https://doi.org/10.1038/nplants.2016.129, 2016.
767 Milligan, J. N., Royer, D. L., Franks, P. J., Upchurch, G. R., and McKee, M. L.: No evidence for a large
768 atmospheric CO_2 spike across the Cretaceous-Paleogene boundary, Geophysical Research
769 Letters, 46, xxx-xxx, https://doi.org/10.1029/2018GL081215, 2019.
770 Mills, B. J. W., Belcher, C. M., Lenton, T. M., and Newton, R. J.: A modeling case for high atmospheric
771 oxygen concentrations during the Mesozoic and Cenozoic, Geology, 44, 1023-1026,
772 https://doi.org/10.1130/g38231.1, 2016.
773 Montañez, I. P., McElwain, J. C., Poulsen, C. J., White, J. D., DiMichele, W. A., Wilson, J. P., Griggs, G., and
774 Hren, M. T.: Climate, p_{CO2} and terrestrial carbon cycle linkages during late Palaeozoic glacial-
775 interglacial cycles, Nature Geoscience, 9, 824-828, https://doi.org/10.1038/ngeo2822, 2016.
776 Munger, W. and Hadley, J.: CO_2 profile at Harvard Forest HEM and LPH towers since 2009, Harvard
777 Forest Data Archive: HF197,
778 http://harvardforest.fas.harvard.edu:8080/exist/apps/datasets/showData.html?id=hf197, 2017.
779 NOAA/ESRL: https://www.esrl.noaa.gov/gmd/ccgg/trends/data.html, 2019.
780 Porter, A. S., Yiotis, C., Montañez, I. P., and McElwain, J. C.: Evolutionary differences in $\Delta^{13}C$ detected
781 between spore and seed bearing plants following exposure to a range of atmospheric O_2:CO_2
782 ratios: implications for paleoatmosphere reconstruction, Geochimica et Cosmochimica Acta,
783 213, 517-533, https://doi.org/10.1016/j.gca.2017.07.007, 2017.
784 Quay, P., King, S., Wilbur, D., Wofsy, S., and Rickey, J.: $^{13}C/^{12}C$ of atmospheric CO_2 in the Amazon Basin:
785 forest and river sources, Journal of Geophysical Research, 94, 18327-18336,
786 https://doi.org/10.1029/JD094iD15p18327, 1989.

R Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/, 2016.
Reichgelt, T., D'Andrea, W. J., and Fox, B. R. S.: Abrupt plant physiological changes in southern New Zealand at the termination of the Mi-1 event reflect shifts in hydroclimate and pCO_2, Earth and Planetary Science Letters, 455, 115-124, https://doi.org/10.1016/j.epsl.2016.09.026, 2016.
Richey, J. D., Upchurch, G. R., Montañez, I. P., Lomax, B. H., Suarez, M. B., Crout, N. M. J., Joeckel, R. M., Ludvigson, G. A., and Smith, J. J.: Changes in CO_2 during Ocean Anoxic Event 1d indicate similarities to other carbon cycle perturbations, Earth and Planetary Science Letters, 491, 172-182, https://doi.org/10.1016/j.epsl.2018.03.035, 2018.
Roeske, C. and O'Leary, M. H.: Carbon isotope effects on enzyme-catalyzed carboxylation of ribulose bisphosphate, Biochemistry, 23, 6275-6284, https://doi.org/10.1021/bi00320a058, 1984.
Roth-Nebelsick, A., Grein, M., Utescher, T., and Konrad, W.: Stomatal pore length change in leaves of *Eotrigonobalanus furcinervis* (Fagaceae) from the Late Eocene to the Latest Oligocene and its impact on gas exchange and CO_2 reconstruction, Review of Palaeobotany and Palynology, 174, 106-112, https://doi.org/10.1016/j.revpalbo.2012.01.001, 2012.
Roth-Nebelsick, A., Oehm, C., Grein, M., Utescher, T., Kunzmann, L., Friedrich, J.-P., and Konrad, W.: Stomatal density and index data of *Platanus neptuni* leaf fossils and their evaluation as a CO_2 proxy for the Oligocene, Review of Palaeobotany and Palynology, 206, 1-9, https://doi.org/10.1016/j.revpalbo.2014.03.001, 2014.
Roth, J. and Dilcher, D.: Some considerations in leaf size and leaf margin analysis of fossil leaves, Courier Forschungsinstitut Senckenberg, 30, 165-171, 1978.
Royer, D. L.: Stomatal density and stomatal index as indicators of paleoatmospheric CO_2 concentration, Review of Palaeobotany and Palynology, 114, 1-28, https://doi.org/10.1016/S0034-6667(00)00074-9, 2001.
Royer, D. L. and Hren, M. T.: Carbon isotopic fractionation between whole leaves and cuticle, Palaios, 32, 199-205, https://doi.org/10.2110/palo.2016.073, 2017.
Royer, D. L., Miller, I. M., Peppe, D. J., and Hickey, L. J.: Leaf economic traits from fossils support a weedy habit for early angiosperms, American Journal of Botany, 97, 438-445, https://doi.org/10.3732/ajb.0900290, 2010.
Sack, L. and Scoffoni, C.: Leaf venation: structure, function, development, evolution, ecology and applications in the past, present and future, New Phytologist, 198, 983-1000, https://doi.org/10.1111/nph.12253, 2013.
Sack, L., Melcher, P. J., Liu, W. H., Middleton, E., and Pardee, T.: How strong is intracanopy leaf plasticity in temperate deciduous trees?, American Journal of Botany, 93, 829-839, https://doi.org/10.3732/ajb.93.6.829, 2006.
Schubert, B. A. and Jahren, A. H.: Incorporating the effects of photorespiration into terrestrial paleoclimate reconstruction, Earth-Science Reviews, 177, 637-642, https://doi.org/10.1016/j.earscirev.2017.12.008, 2018.
Smith, R. Y., Greenwood, D. R., and Basinger, J. F.: Estimating paleoatmospheric pCO_2 during the Early Eocene Climatic Optimum from stomatal frequency of *Ginkgo*, Okanagan Highlands, British Columbia, Canada, Palaeogeography Palaeoclimatology Palaeoecology, 293, 120-131, https://doi.org/10.1016/j.palaeo.2010.05.006, 2010.
Song, X., Barbour, M. M., Saurer, M., and Helliker, B. R.: Examining the large-scale convergence of photosynthesis-weighted tree leaf temperatures through stable oxygen isotope analysis of multiple data sets, New Phytologist, 192, 912-924, https://doi.org/10.1111/j.1469-8137.2011.03851.x, 2011.

Sotta, E. D., Meir, P., Malhi, Y., Donato nobre, A., Hodnett, M., and Grace, J.: Soil CO_2 efflux in a tropical forest in the central Amazon, Global Change Biology, 10, 601-617, https://doi.org/10.1111/j.1529-8817.2003.00761.x, 2004.

Spicer, R. A.: The importance of depositional sorting to the biostratigraphy of plant megafossils, in: Biostratigraphy of Fossil Plants: Successional and Paleoecological Analyses, edited by: Dilcher, D. and Taylor, T., Dowden, Hutchinson, and Ross, Stroudsburg, PA, 171-183, 1980.

Sternberg, L., Mulkey, S. S., and Wright, S. J.: Ecological interpretation of leaf carbon isotope ratios: influence of respired carbon dioxide, Ecology, 70, 1317-1324, https://doi.org/10.2307/1938191, 1989.

Stevens, P. F.: Angiosperm Phylogeny Website. Version 13., www.mobot.org/MOBOT/research/APweb/, 2001 onwards.

Talbert, C. M. and Holch, A. E.: A study of the lobing of sun and shade leaves, Ecology, 38, 655-658, https://doi.org/10.2307/1943135, 1957.

Tcherkez, G.: How large is the carbon isotope fractionation of the photorespiratory enzyme glycine decarboxylase?, Functional Plant Biology, 33, 911-920, https://doi.org/10.1071/FP06098, 2006.

Tesfamichael, T., Jacobs, B., Tabor, N., Michel, L., Currano, E., Feseha, M., Barclay, R., Kappelman, J., and Schmitz, M.: Settling the issue of "decoupling" between atmospheric carbon dioxide and global temperature: $[CO_2]_{atm}$ reconstructions across the warming Paleogene-Neogene divide, Geology, 45, 999-1002, https://doi.org/10.1130/G39048.1, 2017.

Tipple, B. J., Meyers, S. R., and Pagani, M.: Carbon isotope ratio of Cenozoic CO_2: a comparative evaluation of available geochemical proxies, Paleoceanography, 25, PA3202, https://doi.org/10.1029/2009PA001851, 2010.

Uhl, D. and Mosbrugger, V.: Leaf venation density as a climate and environmental proxy: a critical review and new data, Palaeogeography Palaeoclimatology Palaeoecology, 149, 15-26, https://doi.org/10.1016/S0031-0182(98)00189-8, 1999.

Von Caemmerer, S.: Biochemical Models of Leaf Photosynthesis, CSIRO Publishing, Collingwood, Australia, 2000.

Wang, Y., Ito, A., Huang, Y., Fukushima, T., Wakamatsu, N., and Momohara, A.: Reconstruction of altitudinal transportation range of leaves based on stomatal evidence: an example of the Early Pleistocene *Fagus* leaf fossils from central Japan, Palaeogeography, Palaeoclimatology, Palaeoecology, 505, 317-325, https://doi.org/10.1016/j.palaeo.2018.06.011, 2018.

Woodward, F. I.: Stomatal numbers are sensitive to increases in CO_2 from pre-industrial levels, Nature, 327, 617-618, https://doi.org/10.1038/327617a0, 1987.

Woodward, F. I. and Kelly, C. K.: The influence of CO_2 concentration on stomatal density, New Phytologist, 131, 311-327, https://doi.org/10.1111/j.1469-8137.1995.tb03067.x, 1995.

Wynn, J. G.: Towards a physically based model of CO_2-induced stomatal frequency response, New Phytologist, 157, 391-398, https://doi.org/10.1046/j.1469-8137.2003.00702.x, 2003.