Impact of mesophyll diffusion on estimated global land CO_2 fertilization

Ying Sun[a], Lianhong Gu[b,1], Robert E. Dickinson[a,1], Richard J. Norby[b], Stephen G. Pallardy[c], and Forrest M. Hoffman[d]

[a]Department of Geological Sciences, University of Texas at Austin, Austin, TX 78712; [b]Environmental Sciences Division and Climate Change Science Institute, Oak Ridge National Laboratory, Oak Ridge, TN 37831; [c]Department of Forestry, University of Missouri, Columbia, MO 65211; and [d]Climate Change Science Institute and Computational Earth Sciences Group, Oak Ridge National Laboratory, Oak Ridge, TN 37831

Contributed by Robert E. Dickinson, September 19, 2014 (sent for review April 9, 2014; reviewed by Thomas D. Sharkey and Ying-Ping Wang)

In C_3 plants, CO_2 concentrations drop considerably along mesophyll diffusion pathways from substomatal cavities to chloroplasts where CO_2 assimilation occurs. Global carbon cycle models have not explicitly represented this internal drawdown and therefore overestimate CO_2 available for carboxylation and underestimate photosynthetic responsiveness to atmospheric CO_2. An explicit consideration of mesophyll diffusion increases the modeled cumulative CO_2 fertilization effect (CFE) for global gross primary production (GPP) from 915 to 1,057 PgC for the period of 1901–2010. This increase represents a 16% correction, which is large enough to explain the persistent overestimation of growth rates of historical atmospheric CO_2 by Earth system models. Without this correction, the CFE for global GPP is underestimated by 0.05 PgC/y/ppm. This finding implies that the contemporary terrestrial biosphere is more CO_2 limited than previously thought.

mesophyll conductance | CO_2 fertilization | carbon cycle | gross primary production | photosynthetic model

To reach Rubisco, the carboxylating enzyme of the Calvin cycle, CO_2 molecules must diffuse through two consecutive segments of a continuous pathway in leaves of C_3 plant species. The first segment connects leaf intercellular air space with ambient air and is controlled by stomata; the second one consists of mesophyll layers from intercellular air space to stroma of chloroplasts where Rubisco resides (1, 2). These two stages differ in the media through which CO_2 moves. Diffusion in the first segment (stomatal diffusion) is through gases only; that in the second segment (mesophyll diffusion) occurs in a variety of media including liquids and lipids, i.e., cell walls, plasmalemma, cytosol, chloroplast envelope membranes, and stroma. The path length of this mesophyll diffusion is generally shorter than that of stomatal diffusion (2). However, diffusion of CO_2 through liquids is several orders of magnitude slower than it is through gases; diffusion through lipids in membranes is even slower than it is through liquid water (3), although it may be facilitated by aquaporin-like channels (4). Consequently, mesophyll layers constitute a major barrier for CO_2 movement inside leaves (5–9).

However, the importance of this mesophyll diffusion limitation for photosynthesis has yet to be reflected in carbon cycle modeling. Current large-scale carbon cycle models (10, 11) have explicitly considered stomatal diffusion but not mesophyll diffusion. Most carbon cycle models use some form of the biochemical model of Farquhar, von Caemmerer, and Berry (FvCB) for modeling photosynthesis (12). In theory, the FvCB model should use the CO_2 concentration at the site of carboxylation inside the chloroplast (C_c). Nevertheless, most modelers have knowingly or unknowingly applied it directly to the CO_2 concentration inside the substomatal cavity (C_i). Since C_c can be much smaller than C_i, because of the mesophyll resistance to CO_2 diffusion, a compensating adjustment is needed to correct for this overestimate of CO_2 available for carboxylation, an adjustment that has been provided by the use of phenomenological, rather than actual, values of fundamental photosynthetic parameters. These parameters include the maximum carboxylation rate (V_{cmax}), maximum electron transport rate (J_{max}), and triose phosphate utilization rate (TPU) and have typically been estimated from leaf gas exchange measurements commonly known as A/C_i curves obtained under carefully controlled environmental conditions (13, 14). The parameter estimation procedures used in such efforts have treated mesophyll conductance (g_m) as if it were infinitely large, even though laboratory studies indicate that it is finite and that the mesosphyll diffusion limitation on photosynthesis can be substantial (1–9, 15, 16).

Without explicit consideration of mesophyll diffusion, fundamental photosynthetic parameters inferred from A/C_i curves are significantly underestimated (7, 15, 17). V_{cmax} is particularly sensitive to g_m and is underestimated by as much as 75% if g_m is assumed infinite (15). Therefore, the phenomenological parameters used in current carbon cycle models substantially undervalue the actual biochemical capacities of the photosynthetic machinery.

Will this underestimation of actual biochemical capacities of photosynthetic apparatus compensate for the overestimation of CO_2 available for carboxylation in determining the long-term terrestrial fertilization effect of anthropogenic CO_2 emissions estimated by carbon cycle models? Photosynthesis is more sensitive to changes in CO_2 at low CO_2 than at high CO_2 concentrations because the photosynthetic response to CO_2 as an enzyme-catalyzed reaction is a saturating curve. Consequently an overestimation of CO_2 at the carboxylation site leads to an underestimation of photosynthetic sensitivity to variations in ambient CO_2 concentration. The degree of this underestimation is

Significance

Understanding and accurately predicting how global terrestrial primary production responds to rising atmospheric CO_2 concentrations is a prerequisite for reliably assessing the long-term climate impact of anthropogenic fossil CO_2 emissions. Here we demonstrate that current carbon cycle models underestimate the long-term responsiveness of global terrestrial productivity to CO_2 fertilization. This underestimation of CO_2 fertilization is caused by an inherent model structural deficiency related to lack of explicit representation of CO_2 diffusion inside leaves, which results in an overestimation of CO_2 available at the carboxylation site. The magnitude of CO_2 fertilization underestimation matches the long-term positive growth bias in the historical atmospheric CO_2 predicted by Earth system models. Our study will lead to improved understanding and modeling of carbon–climate feedbacks.

Author contributions: Y.S., L.G., and R.E.D. designed research; Y.S., L.G., and R.E.D. performed research; R.J.N. and F.M.H. contributed new reagents/analytic tools; Y.S. analyzed data; and Y.S., L.G., R.E.D., R.J.N., S.G.P., and F.M.H. wrote the paper.

Reviewers: T.D.S., Michigan State University; and Y.-P.W., The Commonwealth Scientific and Industrial Research Organization.

The authors declare no conflict of interest.

[1]To whom correspondence may be addressed. Email: lianhong-gu@ornl.gov or robted@jsg.utexas.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1418075111/-/DCSupplemental.

not constant; rather, it will dynamically vary with all environmental factors that affect photosynthesis. As a result, it may be difficult for this dynamically varying bias to be compensated for by the use of a few phenomenological parameters tuned to a limited number of measurements made under narrow environmental conditions. If so, lacking explicit consideration of g_m represents an inherent structual deficiency that may prevent carbon cycle models from adequately simulating the long-term responses of global photosynthesis to historial and future changes in atmospheric CO_2 concentration due to anthropogenic emissions.

To evaluate the consequence of this deficiency, we examine the simulated responses of global annual gross primary production (GPP) to the increase in atmospheric CO_2 concentration since the beginning of the last century. We focus on GPP because it is the first step of the terrestrial carbon cycle and is affected directly by mesophyll diffusion of CO_2. Our interest is in the long-term trend of global GPP, rather than in its absolute magnitude for a particular year or the mean GPP over a period. For short-term applications, the model structural deficiency can be easily compensated for by a tuning of model parameters, i.e., V_{cmax} and/or J_{max} can be adjusted so that a carbon cycle model lacking g_m will give the same GPP as estimated by a model including g_m, e.g., for the first year of a 100-year period alone, or for that matter, for the last year alone; however, it is considerably more difficult to match these two models all the way from the first to last year for atmospheric CO_2 that keeps rising during this 100-year period. Thus, focusing on the long-term trend is an effective way to quantify the effects of model structural deficiencies and their potential consequences.

We developed an empirical global g_m model for C_3 plant species based on a synthesis of data in the literature (*SI Text*) and implemented it into the state-of-the-art Community Land Model 4.5 (CLM4.5) (18, 19). This implementation allows us to contrast simulations that either consider or omit the mesophyll diffusion limitation. We refer to these simulations as the g_m-including and g_m-lacking simulations, respectively.

To enable a correct comparison between the g_m-including and g_m-lacking simulations, a matching correspondence must be established between the original phenomenological photosynthetic parameters in CLM4.5 (denoted thereafter as the g_m-lacking parameters) and the fundamental photosynthetic parameters that reflect the actual capacities of the photosynthetic machinery (denoted thereafter as the g_m-including parameters). This matching correspondence was achieved via a parameter conversion function that was developed from a global leaf gas exchange dataset collected by LeafWeb (leafweb.ornl.gov) (15, 20). The development of this conversion function was based on the CLM4.5 formulation of the FvCB model (*SI Text*). Other measures have also been taken to ensure that any difference in the trend of GPP between simulations can be attributed unambiguously to the mesophyll diffusion treatments (*SI Text*).

We ran CLM4.5 including or lacking g_m from 1901 to 2010 with historical climate in conjunction with either observed or constant atmospheric CO_2 concentrations (*SI Text*). In a given year t, the CO_2 fertilization effect [CFE, in units of petagram (10^{15}g) carbon (PgC) per year] on GPP of the historical anthropogenic carbon emissions was quantified relative to a baseline reference (GPP_{ref}), set to be the average of annual GPP of 1901–1910 from simulations with constant CO_2 (296 ppm)

$$CFE(t) = GPP(t) - GPP_{ref}. \quad [1]$$

We examined the impact of accounting for mesophyll diffusion via the difference in CFE (ΔCFE) between the g_m-including and g_m-lacking simulations. If both models were to predict the CO_2 fertilization effect equally well, there would be no long-term trend in ΔCFE.

We also applied the so-called Keeling's β factor to measure divergence in the degree of CO_2 fertilization between the g_m-including and g_m-lacking simulations. The β factors for these two types of simulations were compared through their ratio R

$$R(t) = \frac{\beta_I(t)}{\beta_L(t)} = \frac{[GPP_I(t) - GPP_{I,ref}]GPP_{L,ref}}{[GPP_L(t) - GPP_{L,ref}]GPP_{I,ref}}, \quad [2]$$

where the subscripts I and L denote the g_m-including and g_m-lacking simulations, respectively. The denominators of β_I and β_L share a common logarithmic CO_2 term, which cancels in the expression of R. Because R is the ratio of two β factors, its dynamic behavior (e.g., in response to changes in ambient CO_2) will be different from those of the β factors themselves. If models including and lacking g_m were to predict the same effect of CO_2 fertilization on GPP, then R should be close to 1. A value $R > 1$ indicates that the g_m-lacking model underestimates the CO_2 fertilization effect compared with the g_m-including model; $R < 1$ indicates the opposite.

Results and Discussion

The ΔCFE for global GPP between the g_m-including and g_m-lacking simulations increases from 1901 to 2010 (Fig. 1*A*). A relatively gentle pre-1950 increase in ΔCFE is followed by an upsurge after 1950. This pattern closely matches that of the rising atmospheric CO_2 over the same period (Fig. 1*A*) such that there is a strong positive linear relationship between ΔCFE and atmospheric CO_2 (Fig. 1*B*). These results indicate that the global GPP modeled without explicit consideration of mesophyll diffusion substantially underestimates the long-term fertilization effect of anthropogenic CO_2 emissions on global photosynthesis. Globally, this underestimation is 0.05 PgC/y/ppm (Fig. 1*B*).

From 1901 to 2010, the fertilization of the anthropogenic fossil CO_2 emissions stimulates a cumulative total of 915 PgC (the time integration of Eq. **1**) to the global GPP of the entire period in the g_m-lacking simulations. In the g_m-including simulations, however, the cumulative anthropogenic stimulation is estimated to be 1,057 PgC. Thus, with the baseline reference CO_2 of 296 ppm, an explicit consideration of mesophyll diffusion increases the modeled cumulative CFE on global GPP by 16% by 2010.

Climate variability affects the magnitude of underestimation of the CO_2 fertilization effect by the g_m-lacking simulations, which is seen as interannual variations in ΔCFE (Fig. 1 *A* and *B*). The relationship between ΔCFE and CO_2 also varies spatially (Fig. 1 *B* and *C*). The global trend is mostly contributed by the tropics (15°S to 15°N) and the boreal and arctic regions (>45°N). However, essentially all regions with vegetation activity have a positive relationship between ΔCFE and CO_2, suggesting that the need for representing mesophyll diffusion to correctly model the CO_2 fertilization effect is universal.

The results based on the analysis of R likewise indicate that models lacking g_m underestimate the degree of CO_2 fertilization. R is consistently larger than 1 for the globe as well as for all latitudinal bands in northern and southern hemispheres (Fig. 2). It is larger for the boreal and arctic regions, indicating that in a relative sense, the simulated CO_2 fertilization effect in high latitudinal regions is more sensitive to the consideration of mesophyll diffusion than in lower latitudinal regions. This latitudinal trend in R can be explained by the generally lower temperatures and increased presence of needleleaf evergreen trees in higher latitudes, both factors leading to smaller mesophyll conductance (*SI Text* and Fig. S1) and therefore larger sensitivity of modeled CO_2 fertilization to the representation of mesophyll diffusion.

This identification of elevated importance of mesophyll diffusion in the tropics and the boreal and arctic regions over the midlatitudes highlights the necessity of using both ΔCFE and R in our evaluation. ΔCFE provides an absolute measure of the

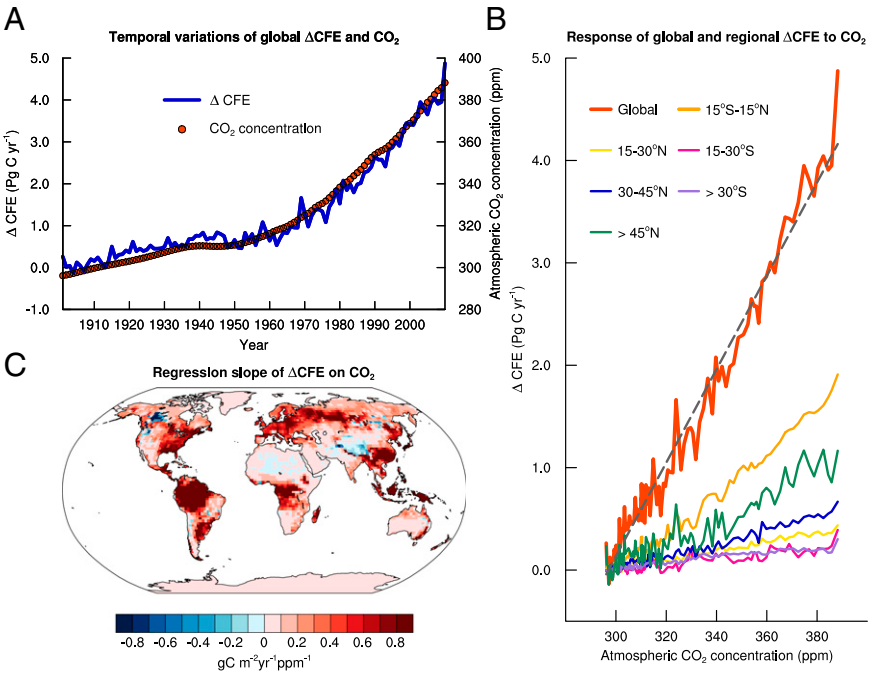

Fig. 1. Temporal and spatial variations of the difference in the CO_2 fertilization effect (ΔCFE, PgC/y) on annual gross primary production simulated with CLM4.5 between including and lacking explicit consideration of mesophyll conductance (g_m). (*A*) Historical trends in ΔCFE (blue curve, left ordinate) and in atmospheric CO_2 concentration (ppm, red dots, right ordinate) from 1901 to 2010. (*B*) The variation of the global and latitudinal ΔCFE with atmospheric CO_2 concentration. The global curve (red) is fitted with a line ($y = -13.55 + 0.05x$, $r^2 = 0.98$). (*C*) The spatial variation in the slope (gC/m^2/y/ppm) of the linear regression of the grid-based ΔCFE as a function of atmospheric CO_2 concentration. The increase of ΔCFE with time and atmospheric CO_2 concentration demonstrates that carbon cycle models without explicit representation of mesophyll diffusion underestimate CO_2 fertilization effect.

impact of mesophyll diffusion on estimated land CO_2 fertilization in carbon flux units. Consequently, it scales with vegetation productivity and total baseline GPP. In contrast, *R* is calculated from relative changes in GPP. As a result, it is suited for sensitivity comparison across climate regions and vegetation types which may differ in productivity. Thus, Δ*CFE* and *R* complement each other by revealing different aspects of the importance of mesophyll diffusion for modeling long-term global and regional CO_2 fertilization effects. The importance of representing mesophyll diffusion for modeling tropical photosynthesis lies in the region's high productivity and large contribution to global GPP, whereas for the boreal and arctic regions, the importance comes from the increased photosynthetic sensitivity associated with their relatively high mesophyll diffusion limitation.

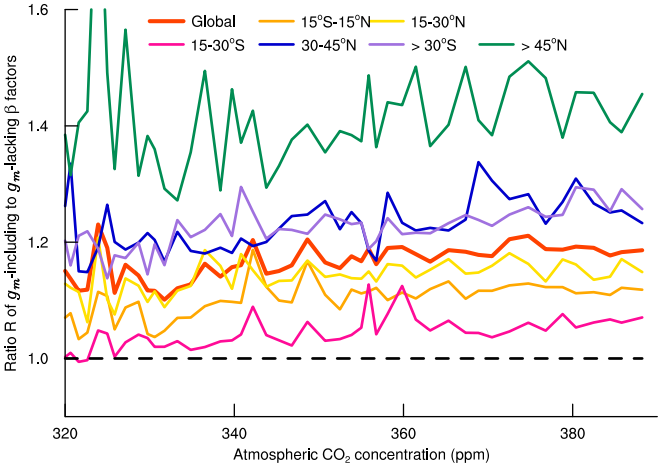

Fig. 2. Changes with atmospheric CO_2 concentration of the global and latitudinal ratio (*R*) of the β factors calculated with CLM4.5 between including and lacking explicit consideration of mesophyll conductance (g_m). *R* fluctuates erratically for the first half of the 20th century when the atmospheric CO_2 is close to the baseline reference of 296 ppm in 1901, causing the denominator of *R* to vary around zero. Therefore, values of *R* when the atmospheric CO_2 is less than 320 ppm are not shown. The dashed line at $R = 1$ separates the underestimation ($R > 1$) from the overestimation ($R < 1$) of the CO_2 fertilization effect on gross primary production by CLM4.5 lacking g_m compared with CLM4.5 including g_m.

Our results imply that Earth system models (ESMs) will overpredict the long-term growth rate of atmospheric CO_2 concentrations due to anthropogenic carbon emissions when their terrestrial carbon cycle modeling components do not consider mesophyll diffusion explicitly. Most ESMs investigated by the fifth phase of the Coupled Model Intercomparison Project (CMIP5) show persistent high bias in their predictions of historical atmospheric CO_2, and it has been suggested that weak ocean uptake could contribute to this high bias (21, 22). We analyzed the outputs of CMIP5 for our simulation period of 1901–2010 and found that in 14 of 17 ESMs, the prognostically computed atmospheric CO_2 grows too fast compared with the observations (Fig. 3). Relative to the 1901 value, the overpredictions of these 14 ESMs range from 10 to 25 ppm by 2010.

By how much can their lack of explicit representation of mesophyll diffusion help explain the overpredictions of historical atmospheric CO_2 growth by ESMs? To answer this question simply, we assume that net primary production is half of GPP (23) and half of the CO_2 released into the atmosphere stays in the atmosphere (24). With these two empirically supported assumptions, we estimate that a quarter of the underestimated cumulative CO_2 fertilization effect by models lacking explicit representation of mesophyll diffusion will be reflected in an overpredicted growth of the atmospheric CO_2 concentration. For the period from 1901 to 2010, the underestimated cumulative *CFE* is 1,057 − 915 = 142 PgC, which corresponds to an overprediction of 17 ppm in atmospheric CO_2 by 2010 [$142/(4 \times 2.123) \sim 17$ ppm; 1 ppm = 2.123 PgC], a value right in the middle of the range of overpredictions by current ESMs. This 17-ppm bias is significant because it almost equals the observed increase in recent decades from intensive fossil CO_2 emissions and occurs over a period with a 100-ppm increase in atmospheric CO_2 (i.e., a 17% overestimation). The lack in the ESMs of an explicit representation of mesophyll diffusion provides a plausible explanation to their atmospheric CO_2 prediction bias, as an alternative to weak ocean uptake.

A rigorous evaluation of the question posted above would be much more complicated and require detailed global carbon budget accounting and expensive simulations in a costly ESM

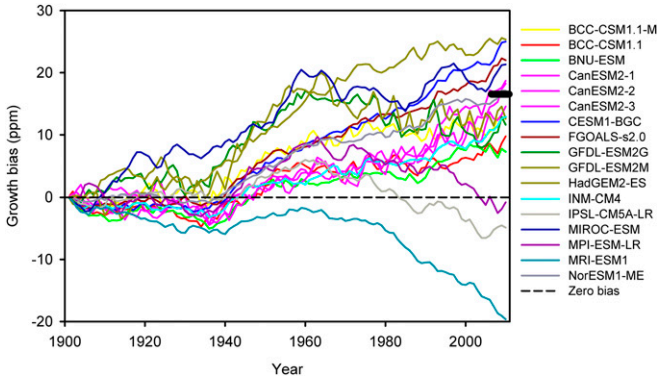

Fig. 3. Growth biases of atmospheric CO_2 concentrations prognostically computed by emission-driven ESMs in the fifth phase of CMIP5. The growth bias in a given year t is calculated as $C_m(t) - C_m(1901) - [C_o(t) - C_o(1901)]$, where C is atmospheric CO_2 concentration and the subscripts m and o denote model and observation, respectively. The reference baseline year is 1901 for which the growth bias is forced to be zero, allowing a focus on the long-term trend. The thick black bar indicates our estimated bias (~17 ppm) caused by lacking explicit representation of mesophyll diffusion. Details about these ESMs and CMIP5 can be found elsewhere (21, 22).

framework for carbon-climate feedbacks. Biomass carbon residence time will need to be considered as not all net primary production once fixed remains with vegetation for very long (25). Also, recent studies found that plant respiratory CO_2 may be transported upward via a transpiration stream in xylem and then be refixed (26), reducing direct dependence of carboxylation on ambient CO_2. However, as all carbon in plant organs must ultimately originate from CO_2 moving through mesophyll, the refixation of xylem-transported respiratory CO_2 will unlikely substantially diminish the impact of the mesophyll diffusion limitation on photosynthesis.

Other model deficiencies besides lack of explicit representation of mesophyll diffusion, for example, inadequate consideration of nitrogen deposition and land use and land cover change (27), could also explain much of the overprediction of the growth rate of historical atmospheric CO_2 by current ESMs, assuming the cause of their overprediction does reside over land. However, because mesophyll diffusion affects the first step of terrestrial carbon cycle, it is important for its effect to be accounted for so that impacts of other factors can be evaluated more reliably.

Although our g_m model is based on empirical data from a large number of species (*SI Text*), its global application, as in any such effort, is bound to have uncertainties. We therefore conducted a leaf-scale uncertainty analysis with the global LeafWeb database of leaf gas exchange measurements. In this leaf-scale analysis, no g_m model was needed; instead, actual parameters optimized from leaf gas exchange measurements were used directly. We calculated leaf-scale R (Eq. **2**) for about 130 C_3 species covering all major plant functional types of the world from herbaceous temperate plants to woody tropical species (15). This obtained leaf scale R is consistent with our global analyses with CLM4.5 (Fig. 4). The results show that under nonsaturating levels of photosynthetic photon flux density (typically PPFD < 1,000 μmol/m^2/s; Fig. 4 *A*–*C*), the leaf scale R, which is averaged across all species, is significantly larger than 1 across a wide range of intercellular CO_2 for all temperature levels examined (intercellular instead of ambient CO_2 was used in this leaf-scale calculation to avoid uncertainties associated with modeling stomatal conductance). For a given nonsaturating PPFD level, the averaged leaf scale R tends to increase with decrease in temperatures, consistent with the trend of latitudinal R in the northern hemisphere simulated by CLM4.5. Only under conditions similar to the measurement conditions of the original A/C_i curves (i.e., saturating light levels and high temperatures) is the averaged leaf-scale R close to 1 (Fig. 4*D*).

Why do carbon cycle models without explicit representation of mesophyll diffusion underestimate the degree of CO_2 fertilization, even when their key photosynthetic parameters have been fitted against A/C_i curves? Why does R vary so much with environmental conditions? The answers are not straightforward but can be most clearly understood through demonstrations with the Excel spreadsheet-based Tool for Evaluating Mesophyll Impact on Predicting Photosynthesis (TEMIPP), which is provided as part of *SI Text*.

The key to understanding these two questions lies in a combination of two factors: (*i*) the way A/C_i curves are measured and (*ii*) the unique structure of the highly nonlinear FvCB model. A/C_i curves are generally made under a saturating level of PPFD (typically >1,000 μmol/m^2/s) and fixed temperature (e.g., 25 °C) and thus represent measurements in one dimension (A vs. CO_2 concentrations). This strict control of measurement conditions at saturating PPFD and fixed temperature is necessary for obtaining data about key biochemical processes of photosynthesis (13, 14). In contrast, carbon cycle models have to run for natural environmental conditions in a 3D space with PPFD, temperature, and CO_2 concentrations all varying simultaneously. Thus, model applications fall outside the ranges of conditions used for parameter calibration. This mismatch is of concern because the highly nonlinear FvCB model consists of three distinct submodels (Rubisco-, RuBP regeneration-, and TPU-limited carboxylation rates) whose respective applicable domains vary dynamically with environmental

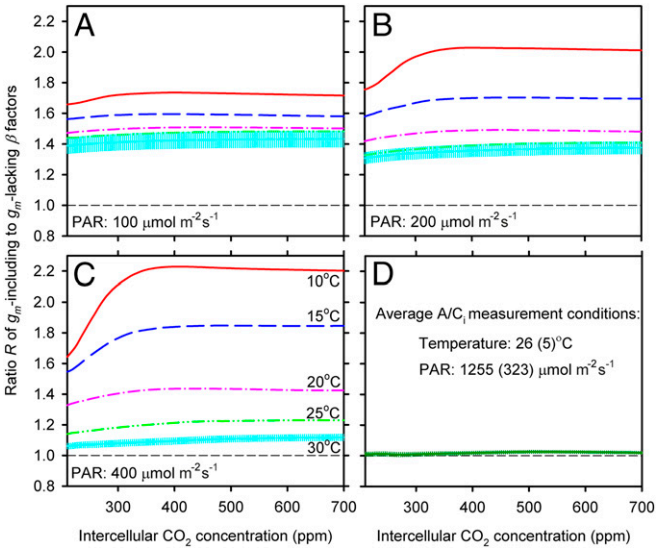

Fig. 4. Changes with intercellular CO_2 concentration of the averaged ratio (R) of the β factors for leaf net photosynthetic rates calculated with a leaf photosynthetic model between including and lacking explicit consideration of mesophyll conductance (g_m). Each curve in each plot represents an average of >1,000 R ratio curves. For each nonsaturating level of photosynthetic photon flux density (PPFD = 100, 200, and 400 μmol/m^2/s, respectively, for *A*, *B*, and *C*), five levels of temperature (10, 15, 20, 25, and 30 °C) are used. Calculations are also done at the original leaf gas exchange (A/C_i) measurement conditions which vary somewhat from measurement to measurement (*D*). The averaged PPFD for the leaf gas exchange measurements is 1,255 ± 323 μmol/m^2/s and the averaged temperature is 26 ± 5 °C. The 95% confidence interval is shown for the mean R ratio curve at 30 °C (*A*–*C*) and at the measurement conditions (*D*; although barely seen due to a large volume of samples). This figure demonstrates that, although photosynthetic models with and without g_m show similar leaf-level CO_2 fertilization effects for the conditions used in measurements, this similarity degrades with increasing distances from the measurement conditions.

conditions that affect photosynthesis. When mesophyll diffusion is not considered explicitly, the condition mismatch discussed above will cause the submodel domains to be incorrectly identified, which results in carboxylation rates being determined by wrong submodels. Hence, even if A/C_i curves were fit apparently well, predictions under naturally varying conditions will be still problematic.

Fig. S7 demonstrates the points made above with two measured and one simulated examples of leaf photosynthetic response to changes in intercellular CO_2 concentration. An unlimited number of cases can be generated with TEMIPP. In Fig. S7, the data were obtained over a range of CO_2 concentrations but at a constant saturating PPFD and a fixed temperature, following the common practices in A/C_i curve measurements that are used for tuning model photosynthetic parameters. On a first look, both the g_m-lacking and g_m-including models fit the original data very well. Such an apparent good fit to data even by g_m-lacking models is very common in the literature of A/C_i curve analyses and probably has contributed to the underappreciation by carbon cycle modelers of the importance of mesophyll diffusion for modeling photosynthesis. However, the residual plots (the *Insets* in Fig. S7) reveal biases easy to miss in visual examination: there are always systematic differences in the predicted photosynthetic rate between the two models, depending on CO_2 concentrations. More importantly, when the fitted parameters are used to predict photosynthetic rates at other values of PPFD and temperatures, the difference between them enlarges. The magnitude of this enlargement depends on levels of CO_2 concentration. At extremely high values of CO_2 that saturate photosynthesis, the limiting effect of mesophyll diffusion diminishes and the g_m-lacking and g_m-including curves tend to merge, regardless of specific values of PPFD and temperature. However, within the intermediate range of CO_2, the g_m-lacking curves are always above the corresponding g_m-including curves, indicating that the g_m-lacking model approaches photosynthetic saturation faster and at a lower CO_2 than does the g_m-including model.

The reason for this faster approach to photosynthetic saturation by the g_m-lacking model is that it overestimates CO_2 concentrations available at the site of carboxylation and therefore underestimates photosynthetic sensitivity to variations in ambient CO_2 concentration. Although this underestimation of sensitivity to CO_2 is minimized when model parameters are tuned against measurements, the effectiveness of the tuned parameters is limited to the narrow PPFD and temperature conditions under which the measurements for parameter tuning are made. When the natural environmental conditions deviate from these parameter tuning conditions, the tuned parameters become less effective as the applicable domains of the highly nonlinear FvCB submodels are misidentified and wrong submodels are applied to calculate carboxylation rates. Thus, a lack of explicit consideration of mesophyll diffusion represents an inherent structural deficiency for carbon cycle models, a structural deficiency that cannot be compensated for by the use of phenomenologically obtained photosynthetic parameters.

In addition to providing a potential explanation for the cause of the overprediction of historical atmospheric CO_2 growth by ESMs, our study has identified a common mechanism that could help resolve several other important issues. A recent inventory and field observation-based report showed that, contrary to expectation, the world's forests have continued to serve as a large persistent carbon sink (28), consistent with our finding that terrestrial ecosystems may have responded to historical anthropogenic CO_2 emissions more strongly than models have indicated. Also, carbon cycle models generally underestimate the long-term trends in the seasonal amplitude of atmospheric CO_2 (29) and in forest ecosystem water use efficiencies (30). These underestimations could be explained by the underestimation by carbon cycle models of the long-term response of net ecosystem productivity to the increase in atmospheric CO_2 concentration, which in turn could be explained by our finding that models lacking explicit consideration of mesophyll diffusion underestimate the CO_2 fertilization effect. Furthermore, in the northern hemisphere, the increase in the seasonal amplitude of atmospheric CO_2 has been larger in high latitudes than in low latitudes and substantially larger than simulated by carbon cycle models (31), which agrees with our finding that the estimated CO_2 fertilization effect in the regions of >45°N is particularly sensitive to the consideration of mesophyll diffusion (Fig. 2).

Our results at the global scale are a logical extension of what has long been known at the leaf scale by plant physiologists. Numerous studies have reported that mesophyll and stomatal conductances have similar magnitudes and are equally important in controlling CO_2 concentrations available for photosynthesis (1–9, 15–17, 20). The drawdown of CO_2 from substomatal cavities to chloroplasts may reduce photosynthesis by 25–75%, depending on species (32). Differences in g_m may even have the potential to alter the balance in species competitiveness as plant communities respond to rising atmospheric CO_2 because an atmosphere enriched in CO_2 may favor species with lower g_m such as needleleaf evergreen trees and others (2, 6). Thus, mesophyll diffusion can play a crucial role in our understanding and predicting photosynthetic responses to the increase in atmospheric CO_2 concentration from leaf to global scales.

An uncertainty in estimating the long-term impact of mesophyll diffusion on global land CO_2 fertilization is the acclimation of biochemical capacities of photosynthetic machinery to elevated atmospheric CO_2 concentrations. Photosynthetic acclimation has been observed in many free-air CO_2 enrichment (FACE) experiments of C_3 plant species (33). Typical acclimation involves down-regulation of V_{cmax} and J_{max} to balance resource allocation to reactions controlling photosynthesis. Our leaf-level analysis indicates that mesophyll diffusion plays similar roles in controlling CO_2 availability at the site of carboxylation regardless of leaf productivity, which can be seen by comparing Fig. S7 *A* and *B* with *C*. Thus, it is unlikely that photosynthetic acclimation to elevated CO_2 will qualitatively change the findings reported here.

A related issue is how limited availability of nutrients, particularly nitrogen and phosphorous, interacts with the impact of mesophyll diffusion on the CO_2 fertilization effect. Model simulations have shown that explicit representations of nitrogen and phosphorus limitations generally result in reduced response of the terrestrial carbon sink to historical increases in atmospheric CO_2 concentrations (27, 34, 35). Thus, nutrient limitations may reduce the impact of mesophyll diffusion on estimated CO_2 fertilization effect. However, if the main effect of nutrient limitations is to reduce V_{cmax} and J_{max}, then our argument made above on photosynthetic acclimation also applies to nutrient limitation. Furthermore, it is possible that an explicit consideration of mesophyll diffusion would influence model evaluations of nutrient limitations on terrestrial carbon cycle. Given the unequivocal fact that mesophyll diffusion limits CO_2 availability at the site of carboxylation, an interesting question is as follows: will explicit representation of mesophyll diffusion delay the development of nutrient limitation in the terrestrial carbon cycle in response to increase in atmospheric CO_2 concentrations? With improved representations of nutrient limitation and mesophyll diffusion, this question may be answerable in the near future.

In summary, the terrestrial biosphere may be more CO_2 limited and therefore absorb more carbon per unit increase of atmospheric CO_2 than previously thought. Over the period investigated in this study (1901–2010), atmospheric CO_2 only started to rise rapidly after the 1950s. If the current trend of increasing atmospheric CO_2 continues, the underestimation of the CO_2 fertilization effect by models lacking explicit representation of mesophyll diffusion should grow to well above the 16% determined for the period of 1901–2010. Although mesophyll conductance increases

with warming, this increase does not keep pace with the increased carboxylation capacity of Rubisco and thus mesophyll diffusion may become even more limiting for photosynthesis at high temperatures (i.e., as measured by ΔCFE) (4). To adequately predict long-term effects of anthropogenic emissions and carbon–climate interactions, carbon cycle models should explicitly consider mesophyll resistance to CO_2 diffusion. As we demonstrated with our global mesophyll conductance model (*SI Text*), this consideration does not add substantial computational burden or excessive new parameterization. Carbon cycle models that lack explicit representation of mesophyll diffusion will underestimate historical and future terrestrial carbon uptake. Consequently, they will overestimate historical and future growth rates of atmospheric CO_2 concentration due to fossil fuel emissions, with ramifications for predicted climate change.

Methods

The global g_m model was developed by synthesizing the latest advances in plant physiological literature. Large-scale carbon cycle models generally use the concept of plant functional types (PFTs) to simulate carbon, water, and energy fluxes in terrestrial ecosystems. Our global g_m model is likewise based on the PFT concept so that it is consistent with current large-scale modeling philosophy and applicable broadly to different vegetation types, rather than to particular ecosystems. In the g_m model, leaf structures, which differ among PFTs, determine the maximum attainable g_m, whereas temperature and moisture stress factors and within-canopy environmental gradients modify this maximum value (*SI Text*). The global g_m model and the conversion function that establishes the g_m-lacking to g_m-including parameter correspondence are all of the structural elements that we have added to CLM4.5. A series of global simulations were conducted to verify the internal consistency and validity of the modified CLM4.5 (*SI Text*). All simulations are offline experiments driven by the CRU/NCEP/NCAR reanalysis (www.cru.uea.ac.uk/cru/data/ncep/) and transient land use. The historical atmospheric CO_2 concentrations, which are used in all simulations except in constant CO_2 runs, are derived from ice cores and atmospheric observations. A detailed description on methods is available in *SI Text*.

ACKNOWLEDGMENTS. We thank Drs. Paul Hanson, Stan Wullschleger, Anthony Walker, and David Weston for comments and suggestions. This material is based upon work supported by the US Department of Energy, Office of Science, Biological and Environmental Research Program Grants DE-FG02-01ER64746 to University of Texas at Austin and DE-AC05-00OR22725 to Oak Ridge National Laboratory (ORNL). The ORNL's Laboratory Directed Research and Development (LDRD) program also partially supported the research. ORNL is managed by University of Tennessee-Battelle, LLC.

1. Evans JR, Sharkey TD, Berry JA, Farquhar GD (1986) Carbon isotope discrimination measured concurrently with gas-exchange to investigate CO_2 diffusion in leaves of higher-plants. *Aust J Plant Physiol* 13(2):281–292.
2. Niinemets U, Flexas J, Peñuelas J (2011) Evergreens favored by higher responsiveness to increased CO_2. *Trends Ecol Evol* 26(3):136–142.
3. Evans JR, Kaldenhoff R, Genty B, Terashima I (2009) Resistances along the CO_2 diffusion pathway inside leaves. *J Exp Bot* 60(8):2235–2248.
4. Bernacchi CJ, Portis AR, Nakano H, von Caemmerer S, Long SP (2002) Temperature response of mesophyll conductance. Implications for the determination of Rubisco enzyme kinetics and for limitations to photosynthesis in vivo. *Plant Physiol* 130(4):1992–1998.
5. Warren CR (2008) Stand aside stomata, another actor deserves centre stage: The forgotten role of the internal conductance to CO_2 transfer. *J Exp Bot* 59(7):1475–1487.
6. Niinemets U, Díaz-Espejo A, Flexas J, Galmés J, Warren CR (2009a) Role of mesophyll diffusion conductance in constraining potential photosynthetic productivity in the field. *J Exp Bot* 60(8):2249–2270.
7. Niinemets U, Díaz-Espejo A, Flexas J, Galmés J, Warren CR (2009b) Importance of mesophyll diffusion conductance in estimation of plant photosynthesis in the field. *J Exp Bot* 60(8):2271–2282.
8. Flexas J, et al. (2012) Mesophyll diffusion conductance to CO_2: An unappreciated central player in photosynthesis. *Plant Sci* 193-194:70–84.
9. Sharkey TD (2012) Virtual special issue on [corrected] mesophyll conductance: Constraint on carbon acquisition by C3 plants. *Plant Cell Environ* 35(11):1881–1883.
10. Taylor KE, Meehl GA, Stouffer RJ (2012) An overview of CMIP5 and the experiment design. *Bull Am Meteorol Soc* 93(4):485–498.
11. Piao S, et al. (2013) Evaluation of terrestrial carbon cycle models for their response to climate variability and to CO_2 trends. *Glob Change Biol* 19(7):2117–2132.
12. Farquhar GD, von Caemmerer S, Berry JA (1980) A biochemical model of photosynthetic CO_2 assimilation in leaves of C $_3$ species. *Planta* 149(1):78–90.
13. Long SP, Bernacchi CJ (2003) Gas exchange measurements, what can they tell us about the underlying limitations to photosynthesis? Procedures and sources of error. *J Exp Bot* 54(392):2393–2401.
14. Sharkey TD, Bernacchi CJ, Farquhar GD, Singsaas EL (2007) Fitting photosynthetic carbon dioxide response curves for C($_3$) leaves. *Plant Cell Environ* 30(9):1035–1040.
15. Sun Y, et al. (2014) Asymmetrical effects of mesophyll conductance on fundamental photosynthetic parameters and their relationships estimated from leaf gas exchange measurements. *Plant Cell Environ* 37(4):978–994.
16. Evans JR, von Caemmerer S (2013) Temperature response of carbon isotope discrimination and mesophyll conductance in tobacco. *Plant Cell Environ* 36(4):745–756.
17. Ethier GJ, Livingston NJ (2004) On the need to incorporate sensitivity to CO_2 transfer conductance into the Farquhar-von Caemmerer-Berry leaf photosynthesis model. *Plant Cell Environ* 27(2):137–153.
18. Bonan GB, et al. (2011) Improving canopy processes in the Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data. *J Geophys Res* 116(G2):G02014.
19. Oleson KW, et al. (2013) Technical description of version 4.5 of the community land model (CLM). *NCAR Technical Note*. Available at www.cesm.ucar.edu/models/cesm1.2/clm/. Accessed October 2, 2013.
20. Gu L, Pallardy SG, Tu K, Law BE, Wullschleger SD (2010) Reliable estimation of biochemical parameters from C_3 leaf photosynthesis-intercellular carbon dioxide response curves. *Plant Cell Environ* 33(11):1852–1874.
21. Friedlingstein P, et al. (2014) Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks. *J Clim* 27(2):511–526.
22. Hoffman FM, et al. (2014) Causes and implications of persistent atmospheric carbon dioxide biases in Earth System Models. *J Geophys Res Biogeosci* 119(2):141–162.
23. Waring RH, Landsberg JJ, Williams M (1998) Net primary production of forests: A constant fraction of gross primary production? *Tree Physiol* 18(2):129–134.
24. Keeling CD, Whorf TP, Wahlen M, van der Plichtt J (1995) Interannual extremes in the rate of rise of atmospheric carbon dioxide since 1980. *Nature* 375(6533):666–670.
25. Friend AD, et al. (2014) Carbon residence time dominates uncertainty in terrestrial vegetation responses to future climate and atmospheric CO2. *Proc Natl Acad Sci USA* 111(9):3280–3285.
26. Trumbore SE, Angert A, Kunert N, Muhr J, Chambers JQ (2013) What's the flux? Unraveling how CO_2 fluxes from trees reflect underlying physiological processes. *New Phytol* 197(2):353–355.
27. Zhang Q, Pitman AJ, Wang YP, Dai YJ, Lawrence PJ (2013) The impact of nitrogen and phosphorous limitation on the estimated terrestrial carbon balance and warming of land use change over the last 156 yr. *Earth System Dynamics* 4(2):333–345.
28. Pan Y, et al. (2011) A large and persistent carbon sink in the world's forests. *Science* 333(6045):988–993.
29. McGuire AD, et al. (2001) Carbon balance of the terrestrial biosphere in the twentieth century: Analyses of CO_2, climate and land use effects with four process-based models. *Global Biogeochem Cycles* 15(1):183–206.
30. Keenan TF, et al. (2013) Increase in forest water-use efficiency as atmospheric carbon dioxide concentrations rise. *Nature* 499(7458):324–327.
31. Graven HD, et al. (2013) Enhanced seasonal exchange of CO_2 by northern ecosystems since 1960. *Science* 341(6150):1085–1089.
32. Terashima I, Hanba YT, Tazoe Y, Vyas P, Yano S (2006) Irradiance and phenotype: Comparative eco-development of sun and shade leaves in relation to photosynthetic CO_2 diffusion. *J Exp Bot* 57(2):343–354.
33. Ainsworth EA, Long SP (2005) What have we learned from 15 years of free-air CO_2 enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO_2. *New Phytol* 165(2):351–371.
34. Wang YP, Polglase PJ (1995) Carbon balance in the tundra, boreal forest and humid tropical forest during climate change: Scaling up from leaf physiology and soil carbon dynamics. *Plant Cell Environ* 18(10):1226–1244.
35. Zhang Q, Wang YP, Pitman AJ, Dai YJ (2011) Limitations of nitrogen and phosphorous on the terrestrial carbon uptake in the 20th century. *Geophys Res Lett* 38(22):L22701.

Supporting Information

Sun et al. 10.1073/pnas.1418075111

SI Text

Development of a Global Mesophyll Conductance (g_m) Model. A model of mesophyll conductance (g_m) suitable for global applications has not previously been available. We develop such an empirical g_m model by synthesizing the latest advances in field plant physiological studies. Large-scale carbon cycle models generally use the concept of plant functional types (PFTs) to simulate carbon, water, and energy fluxes of terrestrial ecosystems. We use a similar strategy to develop a global g_m model so that it is consistent with large-scale modeling philosophy and applicable broadly to different vegetation types, rather than to particular ecosystems. Field measurements have shown that g_m varies with leaf structures and environmental conditions (1–4). Leaf structures determine the maximum attainable g_m with external environmental forcings modifying this maximum value. This consensus reflects a recent understanding that environmental stress factors (e.g., temperature and water) can induce rapid physiological changes (e.g., hardening of cell walls and aquaporin-mediated alteration of membrane permeability) that cause g_m to vary on time scales of minutes to hours (5–8). Accordingly, we model g_m as

$$g_m = g_{\max 0} \cdot f_l(x) \cdot f_T(T_l) \cdot f_w(\theta), \quad \textbf{[S1]}$$

where $g_{\max 0}$ is the maximum g_m (i.e., a value of g_m under nonstress conditions, here referring to the presence of ample soil water and a temperature of 25 °C) of a leaf at the canopy top of a PFT; $f_l(x)$ represents the vertical variation of g_m as a function of cumulative leaf area index x from canopy top, driven by light gradient within the canopy; and $f_T(T_l)$ and $f_w(\theta)$ are the response functions of g_m to leaf temperature T_l and to soil moisture θ, respectively.

The present study makes no attempt to represent other potential environmental effects, e.g., salinity, O_3, nutrient availability (4), on g_m because these effects are much less well understood and seldom quantified in field studies. Also we do not consider the potential rapid, direct responses of g_m to changes in ambient CO_2 concentration and irradiance reported in some previous studies as a recent analysis shows that such responses may be due to methodological artifacts in experiments (9). There is also the possibility that the spatial separation between Rubisco and the releasing site of CO_2 from dark respiration and photorespiration (mitochondria) may render g_m to be a composite variable, rather than a stable parameter (10). This possibility can be addressed with a two-component model of mesophyll conductance (10). However, currently there are no data available to parameterize this two-component model. Fortunately, observations often show that in C_3 plant species, mitochondria occupy the center of the cell and are surrounded by chloroplasts that are positioned just under the plasmalemma (11). In such a spatial configuration, CO_2 molecules evolved from mitochondria and released into cytosol must first diffuse through chloroplasts to reach the intercellular air space. From a modeling point of view, this arrangement has the same effect as if Rubisco and mitochondria shared the same compartment (12). Sun et al. (13) and Gu and Sun (9) used simulations to demonstrate that a single g_m model is sufficient for the purpose of modeling photosynthesis. Therefore, in this study, we adopt the framework of a single g_m model.

The $g_{\max 0}$ varies significantly across plant species. A synthesis of measurements for ∼100 plant species showed that this term is related to the leaf dry mass per unit area M_a through an empirical power law (2)

$$g_{\max 0} = a \cdot M_{a0}^{b}, \quad \textbf{[S2]}$$

where M_{a0} represents M_a at canopy top. The constants a = 24.240338 and b = −0.6509 are two empirical parameters that have been determined by fitting compiled data to Eq. **S2** [r^2 = 0.79; $P < 0.001$; see figure 2.1 in Niinemets et al. (2)]. All empirical constants used in our global mesophyll conductance model are listed in Table S1. Similar patterns were also reported in other studies (4, 14) but with fewer species samples. The values of a and b depend on the units of data used for the nonlinear regression of Eq. **S2**. Our study uses μmol/m^2/s/Pa for $g_{\max 0}$ and g_m, and g/m^2 for M_{a0} and M_a. The coefficient b is negative, indicating $g_{\max 0}$ decreases with M_{a0} across PFTs under nonstress conditions. The $g_{\max 0}$ in Eq. **S2** is area based, whereas Niinemets et al. (2) used a mass-based unit. The area- and mass-based $g_{\max 0}$ differ by a factor of M_{a0}. Accordingly, the value of b in our study is a notation on an area basis. In Eq. **S2**, $g_{\max 0}$ is represented as a leaf trait associated with PFTs because it is determined by M_{a0}, which is the product of two important leaf traits: leaf thickness and foliar mass density (2).

Multiple steps are needed to derive an expression for $f_l(x)$, which describes the variation of g_m associated with the prevailing light regime within a plant canopy. The light intensity shapes the mesophyll cell morphology, the number of mesophyll cell layers, and the leaf thickness. These factors are components that determine $M_a(x)$, the leaf dry mass per area at a cumulative leaf area index x (15). They also affect the total surface area of mesophyll cells exposed to intercellular air space per unit leaf area and hence g_m (2, 11). Consequently, g_m tends to change systematically from top to bottom of a canopy (16–22) and scales well with the variation of $M_a(x)$ with canopy depth (18, 21). Therefore, $M_a(x)$ is an important link between g_m and the prevailing light gradient within a canopy. Here this vertical variation of g_m is modeled as follows:

$$g_{\max}(x) = g_{\max 0} \cdot [M_a(x)/M_{a0}]^{d}, \quad \textbf{[S3]}$$

where d = 0.8109, an empirical parameter fitted from the data (r^2 = 0.67) in Montpied et al. (21). Rearranging Eq. **S3**, we have

$$f_l(x) = g_{\max}(x)/g_{\max 0} = [M_a(x)/M_{a0}]^{d}. \quad \textbf{[S4]}$$

According to Niinemets (15), $M_a(x)$ is related to the seasonally integrated photosynthetic active radiation $I(x)$ via

$$M_a(x) = M_{a0} \cdot [I(x)/I_0]^{f}, \quad \textbf{[S5]}$$

where I_0 is the value of $I(x)$ at canopy top; f = 0.221897, a parameter fitted (r^2 = 0.57) from Niinemets (15). Applying the Beer's law

$$I(x) = I_0 \cdot \exp(-k_I \cdot x), \quad \textbf{[S6]}$$

where k_I = 0.50, a commonly used value for the seasonally averaged light extinction coefficient (23, 24), and substituting Eqs. **S6** and **S5** to Eq. **S4**, we obtain

$$f_l(x) = \exp(-k_I \cdot d \cdot f \cdot x) = \exp(-k_g \cdot x). \quad \textbf{[S7]}$$

Here, $k_g = k_I \cdot d \cdot f$ = 0.08997, a composite parameter that is the product of three empirical coefficients k_I, d, and f. Eq. **S7** shows

that the vertical variation of g_m within a canopy can be modeled as an exponentially decreasing function of cumulative leaf area index x from the top of canopy, characterized with a single decay coefficient k_g. The advantage of creating a single composite parameter k_g is that it facilitates the sensitivity test with this parameter, which may guide process-based measurements. The joint control of M_{a0} and x on g_m is illustrated in Fig. S1, showing that g_m decreases with M_{a0} and x.

It is important to clarify that the dependence of $g_{\max 0}$ on M_{a0} in Eq. **S2** is fundamentally different from the relationship between g_m and $M_a(x)$ in Eq. **S3**. Eq. **S2** applies across PFTs, whereas Eq. **S3** is used to formulate $f_I(x)$ and applies to the depth of a canopy within a PFT. $M_a(x)$ is a composite leaf structural variable and is modified by the prevailing vertical gradient in light regime within a plant canopy. The variations of leaf structures along canopy depth and thus $M_a(x)$ are different from the variations of leaf structures across leaf forms of PFTs and thus the composite leaf trait M_{a0}. As stated earlier, the light intensity gradient along the canopy depth affects $M_a(x)$ through its effects on the mesophyll cell morphology, the number of mesophyll cell layers, and the leaf thickness. In contrast, the variations of M_{a0} across leaf forms of PFTs reflect the changes in leaf robustness, e.g., the compactness of mesophyll cells, the thickness of cell walls, and the foliar mass density. Detailed discussion on this issue is beyond the scope of this paper but can be found elsewhere (2, 11).

Several functions have been proposed to describe the temperature response of g_m (5, 25–27). Here the formulation of Bernacchi et al. (5), which was based on detailed measurements and showed agreement with data from independent researchers (28), is used

$$f_T(T_l) = \exp[c - \Delta H_a/(R \cdot T_l)]/\{1 + \exp[(\Delta S \cdot T_l - \Delta H_d)/(R \cdot T_l)]\}, \quad \textbf{[S8]}$$

where $c = 20.0$, a scaling constant; $\Delta H_a = 49.6 \times 10^3$ J/mol, the activation energy; $\Delta H_d = 437.4 \times 10^3$ J/mol, the deactivation energy; $\Delta S = 1.4 \times 10^3$ J/mol/K, an entropy term; and $R = 8.314$ J/mol/K, the universal gas constant. Eq. **S8** is normalized to 25 °C; hence, $f_T(25\ ^\circ\text{C}) = 1$. It features an initial increase of g_m with T_l (10–35 °C) and decline at high T_l, thus allowing simulation of high temperature inhibition.

Different forms of the water stress term have been applied to g_m in canopy models (29–32). For convenience, we use the CLM4.5 water stress function (33), which has already been applied to V_{cmax} and stomatal conductance g_s (34). In the CLM4.5 formulation

$$f_w(\theta) = \sum_i^n f_{\text{root},i} \cdot w_i(\theta), \quad \textbf{[S9]}$$

where n is the total number of soil layers; $f_{\text{root},i}$ is the root fraction within soil layer i; and $w_i(\theta)$ is the plant wilting factor, derived from the soil water content θ for each layer. The calculation of $f_{\text{root},i}$ and $w_i(\theta)$ follows CLM4.5 procedures (33). The term $f_w(\theta)$ ranges from 1 (wet soil) to ~0 (dry soil), depending on the soil water potential of each layer, and root distribution of PFTs.

Implementation of a Global g_m Model in CLM4.5. CLM4.5 divides the canopy into sunlit and shaded fractions and calculates photosynthesis separately for these two groups of leaves. To be consistent with the CLM4.5 canopy integration scheme, the mean, weighted g_m for sunlit and shaded fractions is calculated, respectively, as

$$\overline{g_{\max\_sun}} = g_{\max 0} \frac{\int_0^L \exp(-k_g \cdot x) \cdot f_{sun}(x) dx}{\int_0^L f_{sun}(x) dx} = g_{\max 0} \cdot \frac{k_b}{k_g + k_b} \cdot \frac{1 - \exp[-(k_g + k_b) \cdot L]}{1 - \exp(-k_b \cdot L)}, \quad \textbf{[S10a]}$$

$$\overline{g_{\max\_sha}} = g_{\max 0} \frac{\int_0^L \exp(-k_g \cdot x) \cdot [1 - f_{sun}(x)] dx}{\int_0^L [1 - f_{sun}(x)] dx} = g_{\max 0} \cdot \frac{k_b}{k_g (k_g + k_b)} \cdot \frac{k_b - (k_g + k_b) \cdot \exp(-k_g \cdot L) + k_g \cdot \exp[-(k_g + k_b) \cdot L]}{\exp(-k_b \cdot L) - 1 + k_b \cdot L}. \quad \textbf{[S10b]}$$

Here, L is the leaf area index; and k_b is the direct beam extinction coefficient and is used to calculate $f_{sun}(x)$ and $f_{sha}(x)$, the fractions of sunlit and shaded leaves respectively. k_I in Eqs. **S6** and **S7** differs from k_b in that k_I is a seasonal mean, whereas k_b varies with solar zenith angle and thus the course of a day and throughout a year and hence is typically updated each time step in model simulations.

g_m links the CO_2 concentration inside leaf chloroplast (C_c) and that at intercellular air space (C_i) with the net carbon assimilation rate A through $C_c = C_i - A/g_m$. Here C_c and C_i are in units of Pa because A and g_m are in units of μmol/m^2/s and μmol/m^2/s/Pa, respectively; elsewhere in the paper, CO_2 concentrations are expressed in units of ppm. The original CLM4.5 used a numerical scheme to solve for photosynthesis by iterating over C_i. In the g_m-enabled CLM4.5, we iterate over C_c and A is calculated with the photosynthesis model from C_c.

$g_{\max 0}$ is obtained from M_{a0}, which in turn is calculated from the inverse of the canopy-top specific leaf area SLA_0 [m^2(gC)], a parameter already specified in CLM4.5 (Fig. S1). M_{a0} (g/m^2) in Eq. **S2** differs from $1/SLA_0$ (gC/m^2) by a factor of two because the former refers to total leaf dry mass, whereas the latter includes only the carbon fraction. Our g_m model is applicable only to C_3 plants; the calculations for C_4 photosynthesis are unchanged in the present study.

Conversion of the g_m-Lacking to g_m-Including Photosynthetic Parameters for CLM4.5. The development and implementation of a global g_m model allows photosynthesis to be calculated at the correct CO_2 concentration; that is, the CO_2 concentration at the site of carboxylation inside the chloroplast. Accordingly, the fundamental biochemical photosynthetic parameters (i.e., V_{cmax}, $J_{\max}$, and TPU) of the FvCB model used in conjunction with the g_m model must reflect the actual photosynthetic capacities of the chloroplast. Current global carbon cycle models use phenomenological photosynthetic parameters (i.e., the g_m-lacking parameters) that are smaller than the true photosynthetic capacities (i.e., the g_m-including parameters) of the chloroplast because their derivation did not represent mesophyll diffusion of CO_2 explicitly (13, 35–37). A matching correspondence between the g_m-including and g_m-lacking photosynthetic parameters must be established so that the dynamic behaviors of carbon cycle models with and without explicit representation of mesophyll diffusion can be compared.

A complication in establishing the matching correspondence between the g_m-including and g_m-lacking fundamental photosynthetic parameters is that at least two different forms of the FvCB model have been used to model leaf photosynthesis. There

is no a priori argument that the correspondence between the g_m-including and g_m-lacking parameters is invariant between the different forms of the FvCB model. In one form, which we call the monolimiting form [see, for example, the implementation in Gu et al. (38)], the net photosynthetic rate A is given by the following expression:

$$A = \min\{W_c, W_j, W_p\}(1-\Gamma^*/C_c) - R_d, \qquad \textbf{[S11]}$$

where W_c, W_j, and W_p represent the carbon carboxylation rate limited by Rubisco, RuBp regeneration, and TPU, respectively, and Γ^* and R_d denote the chloroplastic CO_2 photosynthetic compensation point and day respiration, respectively. In the monolimiting FvCB model, the smallest of three limitations exclusively defines the photosynthetic rate for any particular set of forcing conditions, whereas the other two limitations play no role (except that they help to determine which one is the smallest). It is the most frequently used form. In another form, which is called the colimiting form (39), A is given by the smaller root of the following equations:

$$\begin{aligned} &\Theta_{cj}A_i^2 - (A_c + A_j)A_i + A_cA_j = 0 \\ &\Theta_{ip}(A+R_d)^2 - (A_i + A_p)(A+R_d) + A_iA_p = 0, \end{aligned} \qquad \textbf{[S12]}$$

where A_i in the first equation is a transient variable for calculating A in the second equation; Θ_{cj} and Θ_{ip} are two fixed empirical curvature parameters (0.98 and 0.95, respectively, in CLM4.5); and A_c, A_j, and A_p are, respectively, given by

$$\begin{aligned} A_c &= W_c(1-\Gamma^*/C_c) \\ A_j &= W_j(1-\Gamma^*/C_c) \\ A_p &= W_p(1-\Gamma^*/C_c). \end{aligned} \qquad \textbf{[S13]}$$

In contrast to the monolimiting FvCB model, the colimiting FvCB model calculates the photosynthetic rate for any particular set of forcing conditions with inputs from all three limitations, although its value is dominated by the most limiting of the three limitations (when Θ_{cj} and Θ_{ip} are set to 1, the mono- and colimiting FvCB models are identical).

Sun et al. (13) showed that, for the monolimiting FvCB model, the following relationship can be used to estimate the g_m-including parameters accurately from the corresponding g_m-lacking parameters if g_m is known:

$$y = w\exp\left(p\frac{w^u}{g_m^q + v}\right). \qquad \textbf{[S14]}$$

Here (w, y) represents the pairs of g_m-lacking and g_m-including V_{cmax}, J_{max}, and TPU at a reference temperature of 25 °C; and p, q, u, and v are empirical constants and differ among these pairs (Table S2). To estimate p, q, u, and v in Eq. **S14**, Sun et al. (13) used a worldwide database of leaf gas exchange measurements collected by LeafWeb (leafweb.ornl.gov). The measurements used in Sun et al. (13) contained more than 1,000 A/C_i curves from nearly 130 C_3 plant species covering all major plant functional types of the world, which include grasses, herbs, crops, shrubs, and trees (deciduous and evergreen broadleaf and conifers). These curves were analyzed for g_m-lacking and g_m-including V_{cmax}, J_{max}, and TPU, as well as g_m with an optimization approach (36, 38). The obtained parameters are then used to estimate the empirical constants in Eq. **S14**.

However, CLM4.5 in contrast to its earlier versions applies the colimiting FvCB model and also somewhat different temperature response functions (33, 34). To ensure consistency and comparability between the g_m-lacking and g_m-including simulations, we reanalyzed the LeafWeb A/C_i curves for g_m-lacking and g_m-including V_{cmax}, J_{max}, and TPU, as well as g_m with exactly the same formulations of the FvCB model and temperature response functions used by CLM4.5. The obtained parameters are then used to reestimate p, q, u, and v in Eq. **S14**. The reestimated p, q, u, and v are given in Table S2. For convenience, we use monolimiting and colimiting conversion functions to denote Eq. **S14** with the empirical constants estimated for the monolimiting and colimiting FvCB models, respectively.

Fig. S2 shows the relationship between the g_m-including parameters estimated directly from the A/C_i curves for the colimiting FvCB model and those calculated from the colimiting conversion function. The agreement is reasonable for V_{cmax}, J_{max}, and TPU, although the monolimiting conversion function appears to fit better overall [compare Fig. S2 with figure 5 in Sun et al. (13)]. Fig. S3 compares the difference between the mono- and colimiting conversion functions for a few selected values of g_m. Both conversion functions show that V_{cmax} is the most sensitive, TPU is the least sensitive, and J_{max} has an intermediate sensitivity to g_m. However, the colimiting V_{cmax} and J_{max} appear to be more sensitive to g_m than their monolimiting counterparts, whereas the opposite is true for TPU. In fact, the colimiting TPU differs little between with and without explicit consideration of mesophyll diffusion such that all empirical constants are effectively zero for the colimiting TPU conversion function (Table S2).

Because the empirical constants in the colimiting conversion function are estimated in accordance with CLM4.5 photosynthetic model formulations, we report global simulation results based on the colimiting conversion function. However, the colimiting conversion function may be more uncertain than the monolimiting conversion function as the fitting of an A/C_i curve with the colimiting FvCB model is considerably more difficult than with the monolimiting FvCB model. In any A/C_i curve measurement, one or two carboxylation limitation states (Rubisco, RuBP regeneration, and TPU) may be missing (40, 41). When the monolimiting FvCB model is used to fit the data, any missing limitation state can be detected and its associated parameters be removed from the parameter estimation process with the enumeration strategy developed by Gu et al. (38). In contrast, the application of the colimiting FvCB model has to simply assume all three limitation states are always present in the data, which may or may not be true for any particular curve. Consequently, more unreasonable parameter values are produced with the colimiting FvCB model and have to be disqualified according to the objective criteria proposed in Gu et al. (38). Even after the application of this data quality control procedure, there are still some unusually large parameter values compared with those obtained for the monolimiting FvCB model [compare Fig. S2 with figure 5 in Sun et al. (13); the same original dataset are used in both analyses], indicating not all curves used in the final fitting contain sufficient constraining power for parameter estimation for the colimiting conversion function. As a precaution, the monolimiting conversion function is also applied in additional global simulations so that impacts of uncertainty in the conversion function can be evaluated.

CLM4.5 accounts for the influence of day length $f(D)$ on (g_m-lacking) V_{cmax} (adjusted to 25 °C), i.e., $V_{cmax} = V^0_{cmax}\cdot f(D)$. Here V^0_{cmax} is the value of V_{cmax} unattenuated by shortening in day length. The same $f(D)$ is propagated to J_{max} and TPU (at 25 °C) via their linear dependence on V_{cmax}. We perform parameter conversion on the unattenuated (g_m-lacking) parameters (PFT-specific) and then apply the factor $f(D)$ to the corresponding, converted g_m-including values. In this way, we preserve the functional relationships among the fundamental photosynthetic parameters both before and after the conversion (13). Finally, the conversions of photosynthetic parameters are all performed on values at canopy top so that the vertical profiles of photosynthetic parameters are not altered. In short, the canopy integration strategy of the g_m-enabled CLM4.5 is not altered from

that of the original CLM4.5, which helps ensure that any difference between the g_m-lacking and g_m-including simulations is caused entirely by the explicit consideration of mesophyll diffusion. Table S3 compares g_m-lacking and g_m-including photosynthetic parameters in CLM4.5 with either monolimiting or colimiting conversion functions.

Global Simulations and Consistency Checks. Because mesophyll diffusion of CO_2 affects gross primary production (GPP) directly and because GPP is the first step of the terrestrial carbon cycle, we focus on GPP in this current study. The cascade effects of mesophyll diffusion through GPP on downstream carbon processes are also important to recognize. However, models of terrestrial carbon cycling are complicated and many processes are represented without precise knowledge. As the first study (to our knowledge) of mesophyll diffusion effects at the global scale, focusing on GPP provides a clear demonstration of the impact of mesophyll diffusion and a direct explanation of the cause of its impact.

We also decided to conduct simulations only on a historical period (1901–2010) for which observational forcing data are available. Thus, our estimation of impact of mesophyll diffusion is likely to be conservative because this impact is expected to be more important on long term than on short term and for the historical period, models, including the g_m-lacking models, are presumably already well calibrated. An extension of the simulations to future climate, for example, to 2100 using representative concentration pathways (RCPs) of the Intergovernmental Panel on Climate Change (IPCC) should show larger impacts of mesophyll diffusion. However, focusing on a relatively short historical period removes many uncertainties inherently involved in making future prediction and allows us to concentrate on illuminating processes and mechanisms rather than on showing magnitudes.

As stated earlier, for the objective of this study, it is important to ensure that any difference between the g_m-lacking and g_m-including simulations is caused entirely by an explicit consideration of mesophyll diffusion. That is why strict correspondence between g_m-lacking and g_m-including fundamental photosynthetic parameters must be established. However, ensuring this strict parameter correspondence is not sufficient to obtain comparable simulations. We must also make sure that the g_m-lacking and g_m-including simulations are done for the same model terrestrial biosphere. This is achieved by using the same state variables (e.g., CLM4.5 specified leaf area indices) in all our simulations. This strategy also avoids potentially large biases in prognostically evolving state variables of the biosphere which can be caused by, for example, uncertainties in the representations of limitations of nutrients (particularly nitrogen and phosphorous) on photosynthesis.

Multiple simulations are carried out to clarify and quantify the dynamic behaviors of CLM4.5 with or without the developed g_m model implemented (Table S4): the CTRL run, a control simulation with the original CLM4.5 that has no explicit representation of mesophyll diffusion; the IMED run, a simulation that incorporates the developed g_m model only but retains the original g_m-lacking photosynthetic parameters in CLM4.5; and the MESO run, a simulation with both the explicit consideration of mesophyll diffusion and the corresponding g_m-including photosynthetic parameters. The MESO run is repeated with the colimiting and monolimiting conversion functions as a test for the effect of uncertainty in the conversion function. The IMED run serves as an intermediate step to illustrate the effect of incorporating a g_m model and also as a consistency check. Because the g_m-lacking photosynthetic parameters are smaller than the corresponding g_m-including photosynthetic parameters (13, 35–37) and because the implementation of a g_m model will lead to reduced CO_2 concentrations at the site of carboxylation, we expect the annual GPP of the IMED run will be smaller than either that of the CTRL run or that of the MESO run if the g_m model and the conversion function have been applied correctly in CLM4.5. We also carry out two additional experiments to provide baseline references for quantifying the historical CO_2 fertilization effect on GPP: CTRL_cCO_2 and MESO_cCO_2, both of which are similar to CTRL and MESO except that the atmospheric CO_2 concentration is fixed at the level of year 1901, i.e., 296 ppm.

All simulations are offline experiments driven by the CRU/NCEP/NCAR reanalysis (www.cru.uea.ac.uk/cru/data/ncep/). The historical atmospheric CO_2 concentrations, which are used in all simulations except in constant CO_2 runs, are derived from ice cores and atmospheric observations (42). All simulations have a spatial resolution of 1.9° × 2.5° with a prescribed transient land cover (33, 43).

We compare the mean global annual GPP of 1985–2004 from different runs with the reanalysis climate and observed atmospheric CO_2 concentrations to ensure that the g_m-enabled CLM4.5 behaves as expected. From 1985 to 2004, the atmospheric CO_2 concentration increased from 345 to 377 ppm. The CTRL run estimates the global annual GPP to be 145.17 PgC/y averaged over the period of 1985–2004 (Fig. S4*A*). The IMED run reduces the mean global annual GPP by ∼15 PgC/y (Fig. S4 *B* and *D*) compared with CTRL. This reduction is consistent with our expectation as mentioned above. The GPP decrease is throughout the globe, but most prevalent in the tropics. The use of a g_m model together with the g_m-including photosynthetic parameters in the MESO run with the colimiting conversion function leads to a mean annual GPP of 146.60 PgC/y, a slight increase of ∼1 PgC/y compared with CTRL (Fig. S4*C*) and about 17 PgC/y higher than the IMED run, again consistent with expectation. The mean annual GPP for the MESO run with the monoconversion function is 144.60 PgC/y, a less than 1 PgC/y decrease from CTRL and more than 15 PgC/y higher than IMED, again consistent with expectations.

Currently there are no direct global GPP measurements to verify our simulation results. Published estimates of contemporary global GPP are highly uncertain. One estimate based on scaled-up eddy flux measurements coupled with diagnostic models placed global GPP at 123 PgC/y (44). A different approach using atmospheric oxygen isotopes suggested 150–175 PgC/y (45). Thus, our simulated GPP values for the MESO runs for the contemporary time (1985–2004) are in the middle of these published estimates.

The comparisons among the CTRL, IMED, and MESO runs indicate that our implementation of the global g_m model in CLM4.5 is internally consistent. This internal consistency is an important check on representation of mesophyll diffusion in CLM4.5 because our g_m model and the parameter conversion function are independently derived. These comparisons also demonstrate that it is important to update the g_m-lacking photosynthetic parameters to the corresponding g_m-including photosynthetic parameters in models that represent mesophyll diffusion explicitly; without doing so, the models would not be self-consistent.

Fig. S5 shows the difference in the CO_2 fertilization effect (ΔCFE) on GPP between the g_m-including and g_m-lacking CLM4.5 when the monolimiting conversion function is used in the simulations. The results are similar to those with the colimiting conversion function (compare Fig. S5 with Fig. 1), even though the two conversion functions differ somewhat (Fig. S3). This similarity suggests that any imperfection in establishing the matching correspondence between the g_m-including and g_m-lacking biochemical parameters through the parameter conversion functions does not fundamentally alter our findings.

Under What Conditions Do the g_m-Lacking and g_m-Including Models Match? Comparison of absolute values of GPP predicted by g_m-lacking and g_m-including models is not an appropriate way of

evaluating the impact of explicit representation of mesophyll diffusion on modeling the long-term trend of the CO_2 fertilization effect because the two models may have different baseline references against which the fertilization effects are quantified. However, a direct comparison of absolute values of GPP can help identify conditions under which the phenomenological photosynthetic parameters can be used to compensate for the effect of the model structural deficiency due to a lack of explicit representation of mesophyll diffusion. Fig. S6 shows the temporal variation of the difference in the annual global GPP simulated with CLM4.5 between with and without explicit consideration of mesophyll diffusion. Before the early1970s (atmospheric $CO_2 < 320$ ppm), the g_m-lacking model predicts a higher GPP, and after the middle 1980s (atmospheric $CO_2 > 350$ ppm), it predicts a lower GPP compared with the g_m-including model. From the early 1970s to the middle 1980s (320–350 ppm), the two models have similar GPP. Thus, 320–350 ppm is the range of ambient CO_2 for which the g_m-lacking model is adequate for simulating GPP.

The minimal bias of the g_m-lacking model for a range of ambient CO_2 in some intermediate past probably reflects the history of model development and calibration. It is interesting to note that 320–350 ppm is generally within the range of ambient CO_2 emphasized in leaf gas exchange measurements and thus is well constrained in curve fitting for photosynthetic parameters (40, 41). Therefore, Fig. S6 may be viewed as a vindication of a well-known modeling principle: a structurally deficient model may work well for the conditions to which its parameters are carefully tuned, but once the conditions deviate from the tuning conditions, its reliability becomes questionable. This general principle evidently applies to carbon cycle models.

Analyses of Leaf-Scale CO_2 Fertilization Effect Based on Leaf Gas Exchange Measurements

Leaf gas exchange measurements (A/C_i curves) collected by LeafWeb (leafweb.ornl.gov) were used in this study for two purposes. One was to determine the empirical constants in the conversion function (Eq. **S14**). The other was to analyze the difference in leaf-scale CO_2 fertilization effect between models with and without explicit consideration of mesophyll diffusion. Both tasks depended on the paired g_m-including and g_m-lacking V_{cmax}, J_{max}, and TPU, as well as g_m estimated from these A/C_i curves (13, 38).

The leaf-scale analyses (Fig. 4) used directly in the calculations more than 1,000 pairs of g_m-including and g_m-lacking photosynthetic parameters and g_m values, removing any uncertainty that may be related to the parameter conversion function or to the global g_m model. The leaf-scale simulations are run at multiple levels of PPFD and temperature and also at all measurement conditions of the original A/C_i curves from which the g_m-including and g_m-lacking photosynthetic parameters and g_m are estimated. The average measurement conditions of the original A/C_i curves are $1{,}255 \pm 323$ μmol/m^2/s for PAR and 26 ± 5 °C for temperature. From these leaf-scale runs, the ratio R of the g_m-including to g_m-lacking β factors for each PPFD and temperature combinations was calculated and averaged across all curves.

An Excel Spreadsheet-Based Tool for Evaluating Mesophyll Impact on Predicting Photosynthesis.

Purpose. Tool for Evaluating Mesophyll Impact on Predicting Photosynthesis (TEMIPP) is a Microsoft Excel spreadsheet-based, leaf-scale photosynthetic modeling tool. It is used for demonstrating the impact of lacking an explicit representation of mesophyll diffusion in a photosynthetic model on the predicted response of photosynthesis to the increase in atmospheric CO_2 concentrations.

Approach. TEMIPP simulates the measurement, analysis and application of curves of photosynthesis A against intercellular CO_2 concentrations C_i (i.e., the so-called A/C_i curves). A/C_i curves are typically measured at a saturating level of photosynthetic photon flux density (PPFD) and a fixed temperature. TEMIPP generates an A/C_i curve at a set of measuring environmental conditions (PPFD, temperature, atmospheric pressure, and oxygen) and a set of actual fundamental photosynthetic parameters (e.g., V_{cmax}, J_{max}, TPU, dark respiration R_d, mesophyll conductance g_m), all specified by the user. The photosynthetic rate is then calculated by applying the Farquhar–von Caemmerer–Berry (FvCB) model (46) extended with a finite g_m (36, 38). A g_m-lacking model, which is the FvCB model applied with an assumption of an infinite g_m, is fit to the generated A/C_i curve. The obtained key photosynthetic parameters are then used in the g_m-lacking model as in current carbon cycle models to predict photosynthesis at a new set of conditions that is different from the original set of measuring conditions under which the A/C_i curve for fitting was produced.

Instead of using simulated A/C_i curves, users have the option to apply real A/C_i measurements to TEMIPP. When real A/C_i curves are used, users will need to provide TEMIPP independently estimated photosynthetic parameters including g_m. Users can examine the impact of lacking an explicit representation of g_m by comparing model performance between the fitting to the original A/C_i curve and the prediction at new conditions. It is useful to check the residuals between the actual value and the value calculated by the g_m-lacking model as the residuals can reveal model performance more clearly than a simple direct comparison which can be misleading. Also it is important to compare the limitation states determined by the g_m-lacking model with the actual limitation states. This comparison will provide insight as to why an apparently well-calibrated g_m-lacking model can perform poorly in predictions.

The fitting uses the evolutionary method in Solver provided by Microsoft Excel. The Evolutionary algorithm is selected because the FvCB model is a classic change-point model, and its optimization is not smooth (38). If users do not wish to use the Microsoft Solver, they can use any optimization software they might have or LeafWeb (leafweb.ornl.gov) to estimate the parameters and then input their own parameters directly into TEMIPP.

The temperature response functions used in TEMIPP are from Sharkey et al. (41). If users wish to use different temperature response functions, they can input their own temperature response funcitons as well.

Detailed instructions.

i) Generate a new A/C_i curve. A new A/C_i curve can be generated in any of the following ways:

- *a*) Change the values of the standardized fundamental parameters Vcmax25, Jmax25, TPU25, gm25, Rd25 (cells E10–I10). Users can also change the Rubisco kinetic parameters (J10–L10) or the leaf absorptance parameter (M10) if they wish.
- *b*) Change the A/C_i curve measuring conditions of temperature, PPFD, atmospheric barometric pressure, and oxygen partial pressure (E20–H20).
- *c*) If they wish, users can provide their own coefficients in the temperature response functions in the section from E15 to L17.
- *d*) The A/C_i data for fitting are automatically computed from cells B36 to B53, depending on the values of C_i from A36 to A53. Users can adjust the C_i values from A36 to A53 as they wish. Leave any unused cell blank.

ii) Fit the g_m-lacking model:

- *a*) Click the cursor at Data in the top of Excel Spreadsheet.
- *b*) Click Solver. You may have to install the Excel Solver first.
- *c*) This brings up the Solver Parameters menu.
- *d*) The settings should have been already specified.
- *e*) Click Solve to minimize the value in the objective cell F54.
- *f*) Wait for the Solver to complete its job. This may take a while.

g) When Solver results menu appears, choose "Keep Solver Solution" and click OK.

iii) Provide a new set of environmental conditions for which the g_m-lacking model will make predictions. Put these values in E21–H21. Try different conditions to see how the performance of the g_m-lacking model vary as the conditions for prediction deviate from the conditions for which the original A/Ci curve for fitting was produced.

iv) Examine the two plots around row 70. Also check to see if the g_m-lacking model has identified the limitation states correctly (the limitation states are displayed in the section AD34–AE56 and AU34–AV136).

v) The default setting in the Solver Parameters menu is for optimizing Vcmax25, Jmax25, and Rd25 for the g_m-lacking model. TPU25 for the g_m-lacking model is set to be equal to that users provide in cell G10 to take advantage of the fact that TPU-limited photosynthesis is 3*TPU – R_d, which does not depend on CO_2 concentrations and therefore g_m. This avoids potential overfitting and unreasonable parameter values. However, if users wish to estimate TPU25 for the g_m-lacking model as well, go to the Solver Parameters menu, add ",G11" (without the quotation marks) after "I11" in the box under "By Changing Variable Cells."

vi) If they wish, users can also optimize for the Rubisco kinetic parameters for the g_m-lacking model by adding ",J11,K11,L11" under "By Changing Variable Cells" in the Solver Parameters menu. However, A/C_i data generally do not contain enough information to constrain all these parameters.

vii) Use real A/C_i measurements with independently estimated parameters. To use real A/C_i measurements with parameters estimated by other means for TEMIPP, do the following steps:

a) Save a copy of TEMIPP.
b) Manually input the real A/C_i data in the section A36–B53 and leave any unused cells blank (do not cut and paste as this will cause disabling of the auto-computing functions).
c) Input the standardized fundamental parameters estimated with explicit consideration of g_m into E10–L10 (TEMIPP can be modified to estimate g_m for the purpose of testing).
d) Input the A/C_i measuring conditions in E20–H20.
e) If in their A/C_i curve analysis, users used a set of coefficients for the temperature response functions different from those listed in TEMIPP, input the users' coefficients into the section E15–L17.
f) If users have independent estimates of the corresponding parameters for the g_m-lacking model, input them to E11–I11 and skip the Microsoft Solver; otherwise, invoke the Solver.
g) Check the plots and limitation states.

viii) Modify TEMIPP to estimate g_m and associated fundamental photosynthetic parameters.

a) Save a copy of TEMIPP.
b) Manually input the real C_i data in the section A36–A53 and leave any unused cells empty (do not cut and paste).
c) Manually input the A (net photosynthesis) data in the section M36–M53 and leave any unused cells empty (do not cut and paste).
d) If users wish to use a different set of coefficients for the temperature response functions, input these different coefficients into the section E15–L17.
e) Bring up the Solver.
f) Replace the content in "Set Objective:" with "O54" (without the quotation marks).
g) Replace the content in "By Changing Variable Cells:" with "E10,F10,G10,H10,I10".
h) Replace all "$11"s in the box under "Subject to the Constraints" with "$10".
i) Click Solve.
j) Wait for the Solver to complete its job.
k) The optimized parameters will be displayed in the cells from E10–I10.

TEMIPP is meant to be a demonstration tool only. For actual A/C_i curve analyses, methods such as LeafWeb (leafweb.ornl.gov) are more appropriate (38).

1. Warren CR (2008) Stand aside stomata, another actor deserves centre stage: The forgotten role of the internal conductance to CO_2 transfer. *J Exp Bot* 59(7):1475–1487.
2. Niinemets U, Díaz-Espejo A, Flexas J, Galmés J, Warren CR (2009a) Role of mesophyll diffusion conductance in constraining potential photosynthetic productivity in the field. *J Exp Bot* 60(8):2249–2270.
3. Nobel PS (1977) Internal leaf area and cellular CO_2 resistance: Photosynthetic implications of variations with growth conditions and plant species. *Physiol Plant* 40(2): 137–144.
4. Flexas J, Ribas-Carbó M, Diaz-Espejo A, Galmés J, Medrano H (2008) Mesophyll conductance to CO_2: Current knowledge and future prospects. *Plant Cell Environ* 31(5): 602–621.
5. Bernacchi CJ, Portis AR, Nakano H, von Caemmerer S, Long SP (2002) Temperature response of mesophyll conductance. Implications for the determination of Rubisco enzyme kinetics and for limitations to photosynthesis in vivo. *Plant Physiol* 130(4): 1992–1998.
6. Chazen O, Neumann PM (1994) Hydraulic signals from the roots and rapid cell-wall hardening in growing maize (Zea mays L.) leaves are primary responses to polyethylene glycol-induced water deficits. *Plant Physiol* 104(4):1385–1392.
7. Miyazawa SI, Yoshimura S, Shinzaki Y, Maeshima M, Miyake C (2008) Deactivation of aquaporins decreases internal conductance to CO_2 diffusion in tobacco leaves grown under long-term drought. *Funct Plant Biol* 35(7):553–564.
8. Kaldenhoff R (2012) Mechanisms underlying CO_2 diffusion in leaves. *Curr Opin Plant Biol* 15(3):276–281.
9. Gu L, Sun Y (2014) Artefactual responses of mesophyll conductance to CO2 and irradiance estimated with the variable J and online isotope discrimination methods. *Plant Cell Environ* 37(5):1231–1249.
10. Tholen D, Ethier G, Genty B, Pepin S, Zhu XG (2012) Variable mesophyll conductance revisited: Theoretical background and experimental implications. *Plant Cell Environ* 35(12):2087–2103.
11. Nobel PS (2009) *Physicochemical and Environmental Plant Physiology* (Academic Press, Oxford, UK), 4th Ed.
12. Cernusak LA, et al. (2013) Environmental and physiological determinants of carbon isotope discrimination in terrestrial plants. *New Phytol* 200(4):950–965.
13. Sun Y, et al. (2014) Asymmetrical effects of mesophyll conductance on fundamental photosynthetic parameters and their relationships estimated from leaf gas exchange measurements. *Plant Cell Environ* 37(4):978–994.
14. Niinemets U, Wright IJ, Evans JR (2009c) Leaf mesophyll diffusion conductance in 35 Australian sclerophylls covering a broad range of foliage structural and physiological variation. *J Exp Bot* 60(8):2433–2449.
15. Niinemets U (2007) Photosynthesis and resource distribution through plant canopies. *Plant Cell Environ* 30(9):1052–1071.
16. Terashima I, Hanba YT, Tazoe Y, Vyas P, Yano S (2006) Irradiance and phenotype: Comparative eco-development of sun and shade leaves in relation to photosynthetic CO_2 diffusion. *J Exp Bot* 57(2):343–354.
17. Hanba YT, Kogami H, Terashima I (2002) The effect of growth irradiance on leaf anatomy and photosynthesis in Acer species differing in light demand. *Plant Cell Environ* 25(8):1021–1030.
18. Piel C, Frak E, Le Roux X, Genty B (2002) Effect of local irradiance on $CO(_2)$ transfer conductance of mesophyll in walnut. *J Exp Bot* 53(379):2423–2430.
19. Laisk A, et al. (2005) Adjustment of leaf photosynthesis to shade in a natural canopy: Rate parameters. *Plant Cell Environ* 28(3):375–388.
20. Warren CR, Low M, Matyssek R, Tausz M (2007) Internal conductance to CO_2 transfer of adult Fagus sylvatica: Variation between sun and shade leaves and due to free-air ozone fumigation. *Environ Exp Bot* 59(2):130–138.
21. Montpied P, Granier A, Dreyer E (2009) Seasonal time-course of gradients of photosynthetic capacity and mesophyll conductance to CO_2 across a beech (Fagus sylvatica L.) canopy. *J Exp Bot* 60(8):2407–2418.
22. Han Q, Iio A, Naramoto M, Kakubari Y (2010) Response of internal conductance to soil drought in sun and shade leaves of adult. *Acta Silv. Lign. Hung.* 6:123–134.
23. Campbell GS, Norman JM (1998) *An Introduction to Environmental Biophysics* (Springer, New York).
24. Richardson JJ, Kim SH, Moskal LM (2009) Modeling approaches to estimate effective leaf area index from aerial discrete-return LIDAR. *Agric For Meteorol* 149(6-7):1152–1160.
25. Warren CR, Dreyer E (2006) Temperature response of photosynthesis and internal conductance to CO_2: Results from two independent approaches. *J Exp Bot* 57(12): 3057–3067.

26. Yamori W, Noguchi K, Hanba YT, Terashima I (2006) Effects of internal conductance on the temperature dependence of the photosynthetic rate in spinach leaves from contrasting growth temperatures. *Plant Cell Physiol* 47(8):1069–1080.
27. Scafaro AP, Von Caemmerer S, Evans JR, Atwell BJ (2011) Temperature response of mesophyll conductance in cultivated and wild Oryza species with contrasting mesophyll cell wall thickness. *Plant Cell Environ* 34(11):1999–2008.
28. Evans JR, von Caemmerer S (2013) Temperature response of carbon isotope discrimination and mesophyll conductance in tobacco. *Plant Cell Environ* 36(4):745–756.
29. Keenan T, Sabate S, Gracia C (2010a) The importance of mesophyll conductance in regulating forest ecosystem productivity during drought periods. *Glob Change Biol* 16(3):1019–1034.
30. Keenan T, Sabate S, Gracia C (2010b) Soil water stress and coupled photosynthesis-conductance models: Bridging the gap between conflicting reports on the relative roles of stomatal, mesophyll conductance and biochemical limitations to photosynthesis. *Agric For Meteorol* 150(3):443–453.
31. Egea G, Verhoef A, Vidale PL (2011) Towards an improved and more flexible representation of water stress in coupled photosynthesis-stomatal conductance models. *Agric For Meteorol* 151(10):1370–1384.
32. Oliver RJ, Taylor G, Finch JW (2012) Assessing the impact of internal conductance to CO2 in a land-surface scheme:Measurement and modelling of photosynthesis in Populus nigra. *Agric For Meteorol* 152:240–251.
33. Oleson KW, et al. (2013) Technical description of version 4.5 of the community land model (CLM). *NCAR Technical Note.* www.cesm.ucar.edu/models/cesm1.2/clm/. Accessed October 2, 2013.
34. Bonan GB, et al. (2011) Improving canopy processes in the Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data. *J Geophys Res* 116(G2):G02014.
35. Niinemets U, Díaz-Espejo A, Flexas J, Galmés J, Warren CR (2009b) Importance of mesophyll diffusion conductance in estimation of plant photosynthesis in the field. *J Exp Bot* 60(8):2271–2282.
36. Ethier GJ, Livingston NJ (2004) On the need to incorporate sensitivity to CO_2 transfer conductance into the Farquhar-von Caemmerer-Berry leaf photosynthesis model. *Plant Cell Environ* 27(2):137–153.
37. Zeng W, Zhou GS, Jia BR, Jiang YL, Wang Y (2010) Comparison of parameters estimated from A/C_i and A/C_c curve analysis. *Photosynthetica* 48(3):323–331.
38. Gu L, Pallardy SG, Tu K, Law BE, Wullschleger SD (2010) Reliable estimation of biochemical parameters from C_3 leaf photosynthesis-intercellular carbon dioxide response curves. *Plant Cell Environ* 33(11):1852–1874.
39. Collatz GJ, Ball JT, Grivet C, Berry JA (1991) Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: A model that includes a laminar boundary layer. *Agric For Meteorol* 54(2-4):107–136.
40. Long SP, Bernacchi CJ (2003) Gas exchange measurements, what can they tell us about the underlying limitations to photosynthesis? Procedures and sources of error. *J Exp Bot* 54(392):2393–2401.
41. Sharkey TD, Bernacchi CJ, Farquhar GD, Singsaas EL (2007) Fitting photosynthetic carbon dioxide response curves for C_3 leaves. *Plant Cell Environ* 30(9):1035–1040.
42. Piao S, et al. (2013) Evaluation of terrestrial carbon cycle models for their response to climate variability and to CO_2 trends. *Glob Change Biol* 19(7):2117–2132.
43. Hurtt GC, et al. (2011) Harmonization of land-use scenarios for the period 1500-2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. *Clim Change* 109(1-2):117–161.
44. Beer C, et al. (2010) Terrestrial gross carbon dioxide uptake: Global distribution and covariation with climate. *Science* 329(5993):834–838.
45. Welp LR, et al. (2011) Interannual variability in the oxygen isotopes of atmospheric CO2 driven by El Niño. *Nature* 477(7366):579–582.
46. Farquhar GD, von Caemmerer S, Berry JA (1980) A biochemical model of photosynthetic CO_2 assimilation in leaves of C $_3$ species. *Planta* 149(1):78–90.

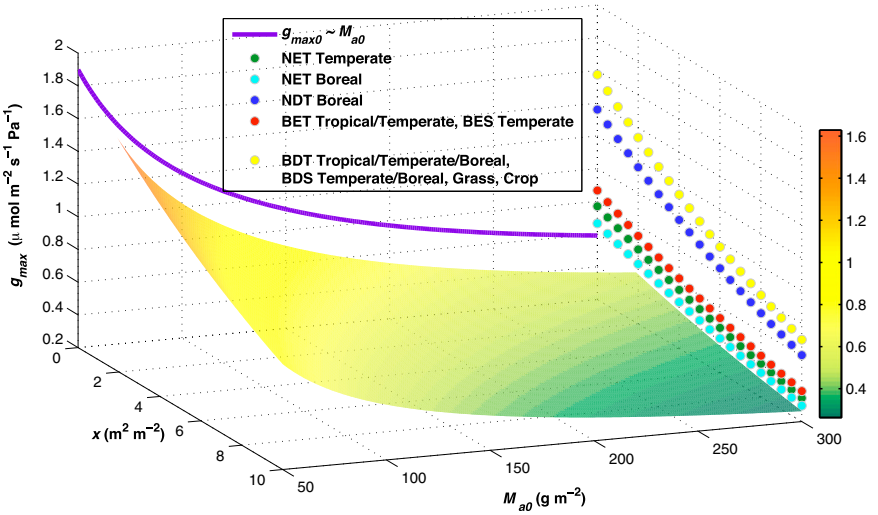

Fig. S1. The g_m model (scaled to the leaf temperature of 25 °C and saturating soil moisture). The 3D surface shows the maximum mesophyll conductance g_{max} [$=a \cdot M_{a0}^{b} \cdot \exp(-k_g \cdot x)$, the product of Eqs. **S2** and **S7**] as a function of canopy-top M_a (M_{a0}) and cumulative leaf area index (x). The purple curve in the $x = 0$ plane shows the relationship between g_{max0} and M_{a0} (Eq. **S2**). The dotted curves show the g_m profile within canopy (Eq. **S7**) for the specified PFTs in CLM4.5. The values of M_{a0} are PFT specific, calculated from the canopy top-specific leaf area SLA_0 in CLM4.5. For clarity, we display these dotted curves on the $M_{a0} = 300$ (g/m^2) plane. Here, NET, NDT, BET, BDT, and BDS denote needleleaf evergreen tree, needleleaf deciduous tree, broadleaf evergreen tree, broadleaf deciduous tree, and broadleaf deciduous shrub, respectively.

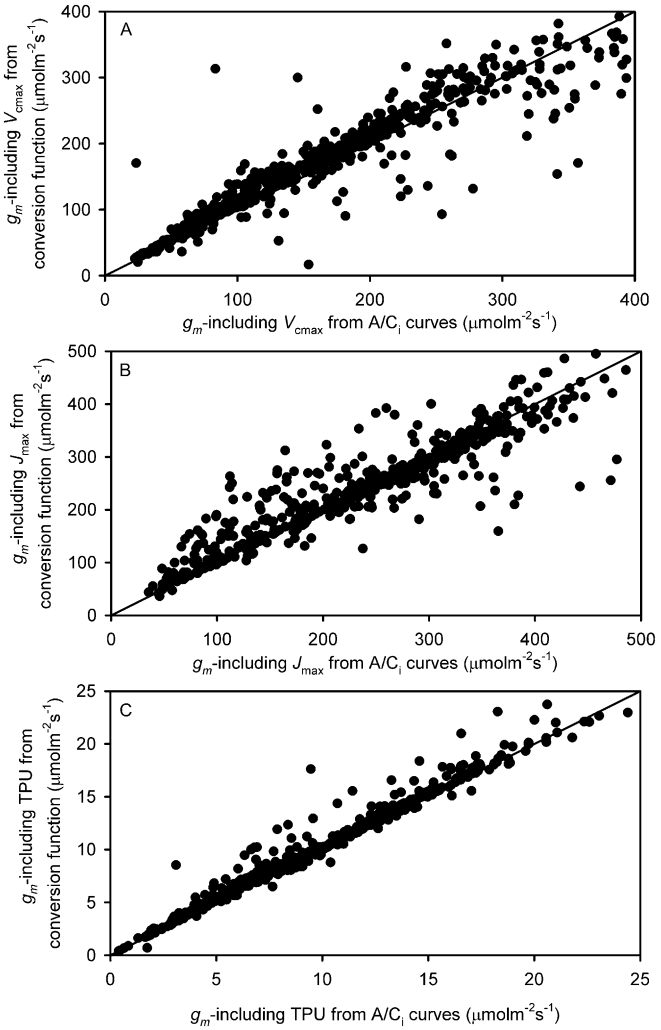

Fig. S2. Comparison of the g_m-including (*A*) V_{cmax}, (*B*) J_{max}, and (*C*) TPU predicted with the colimiting conversion function to corresponding values optimized directly from A/C_i curves with the colimiting FvCB model. In each plot, the 1:1 line is included as a reference. The correlation coefficient is 0.91, 0.87, and 0.98 for V_{cmax}, J_{max}, and TPU, respectively.

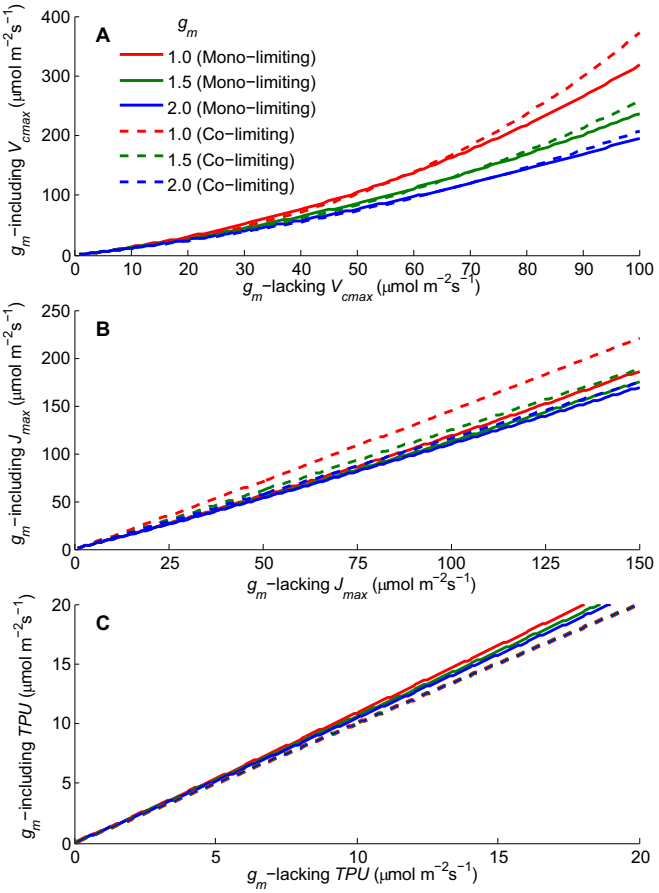

Fig. S3. The comparison of the monolimiting (solid curves) and colimiting (dashed curves) conversion functions at different values of mesophyll conductance g_m (μmol/m^2/s/Pa) for (*A*) V_{cmax}, (*B*) J_{max}, and (*C*) TPU.

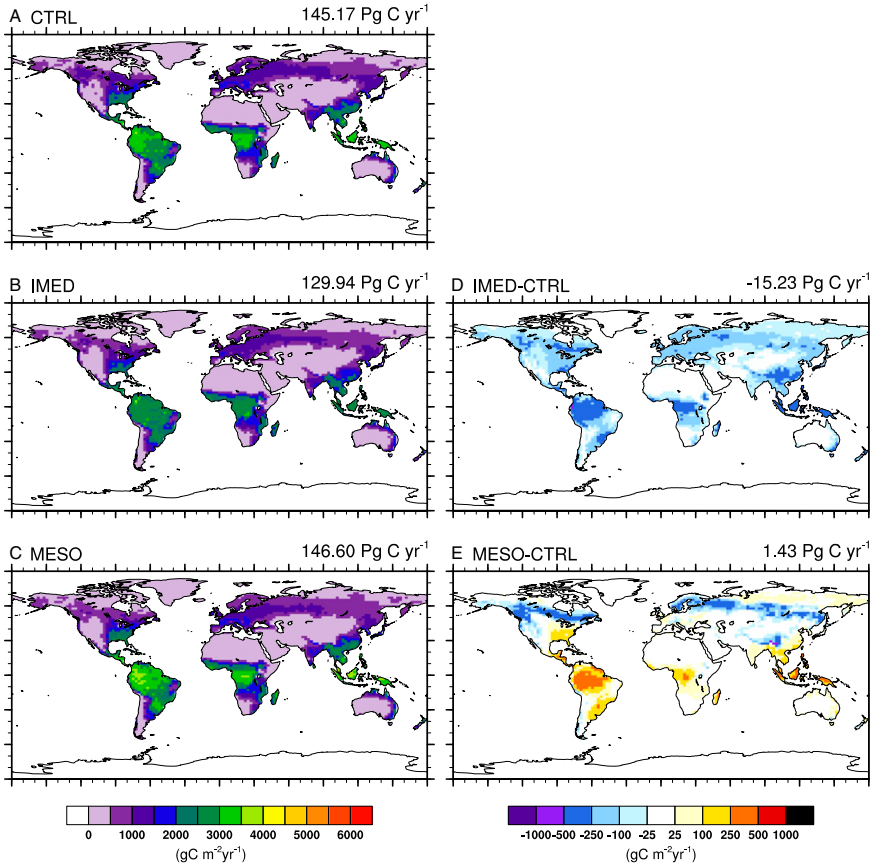

Fig. S4. Mean annual GPP (gC/m^2/y) of the period 1985–2004 for the simulation of (*A*) CTRL, the control simulation with CLM4.5; (*B*) IMED, the intermediate simulation with mesophyll conductance model and g_m-lacking photosynthetic parameters; and (*C*) MESO, the final, fully updated simulation with mesophyll conductance model and g_m-including photosynthetic parameters. (*D* and *E*) Differences from the CTRL simulation for the IMED and MESO simulations, respectively. The value shown on the *Upper Right* of each panel is the estimated mean annual global GPP (PgC/y). All simulations are driven by reanalysis climate and observed atmospheric CO_2 concentrations. The colimiting conversion function is used in these simulations.

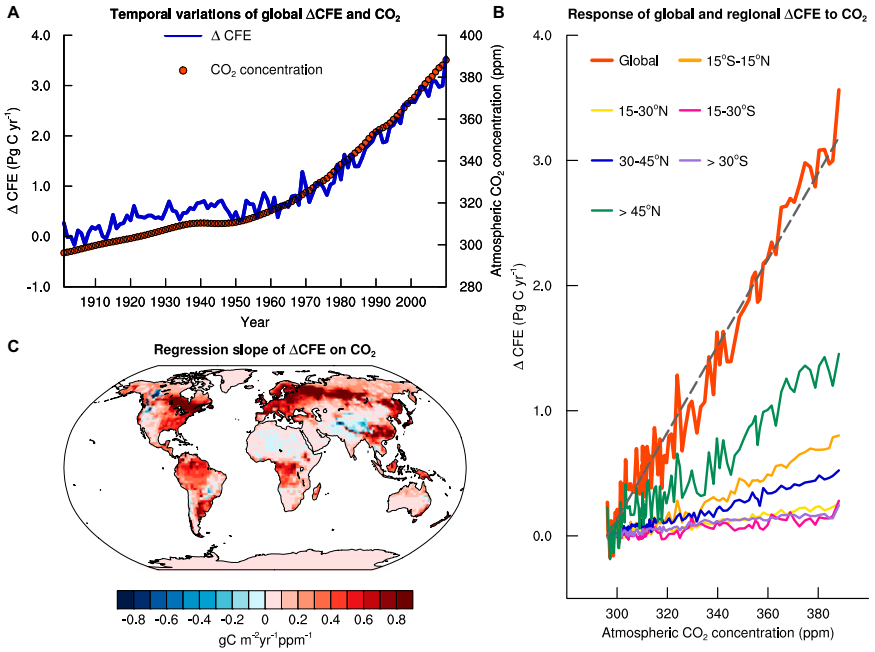

Fig. S5. Temporal and spatial variations of the difference in the CO_2 fertilization effect (ΔCFE, PgC/y) on annual gross primary production simulated with CLM4.5 between with and without explicit consideration of mesophyll conductance (g_m). The monolimiting conversion function is used to establish the g_m-lacking to g_m-including photosynthetic parameter correspondence. (*A*) Historical trends in ΔCFE (blue curve, left ordinate) and in atmospheric CO_2 concentration (ppm, red dots, right ordinate) from 1901 to 2010. (*B*) The variation of the global and latitudinal ΔCFE with atmospheric CO_2 concentration. The global curve (red) is fitted with a line ($y = -10.233 + 0.035x$, $r^2 = 0.96$). (*C*) The global variation in the slope (gC/m^2/y/ppm) of the linear regression of the grid-based ΔCFE as a function of atmospheric CO_2 concentration.

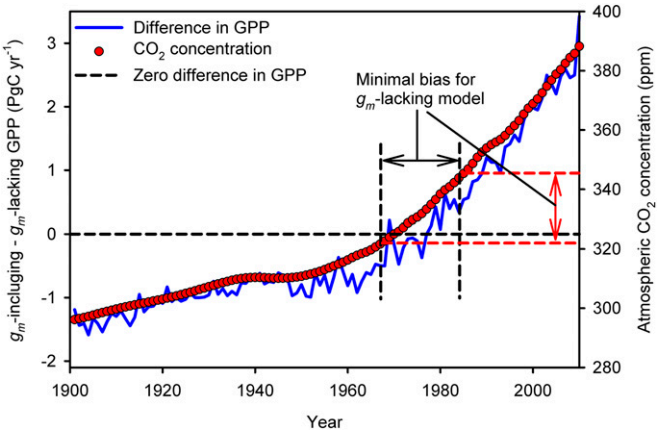

Fig. S6. Temporal variation of the difference in the annual global GPP simulated with CLM4.5 between with and without explicit consideration of mesophyll conductance (g_m). This figure is similar to Fig. 1*A* except that there is no offset in the baseline reference GPP between the two models. The colimiting conversion function is used to establish the g_m-lacking to g_m-including photosynthetic parameter correspondence. The historical atmospheric CO_2 concentration is also shown. The two dash vertical (black) and horizontal (red) lines identify the years and the range of atmospheric CO_2, respectively, in which the g_m-lacking model has minimal bias.

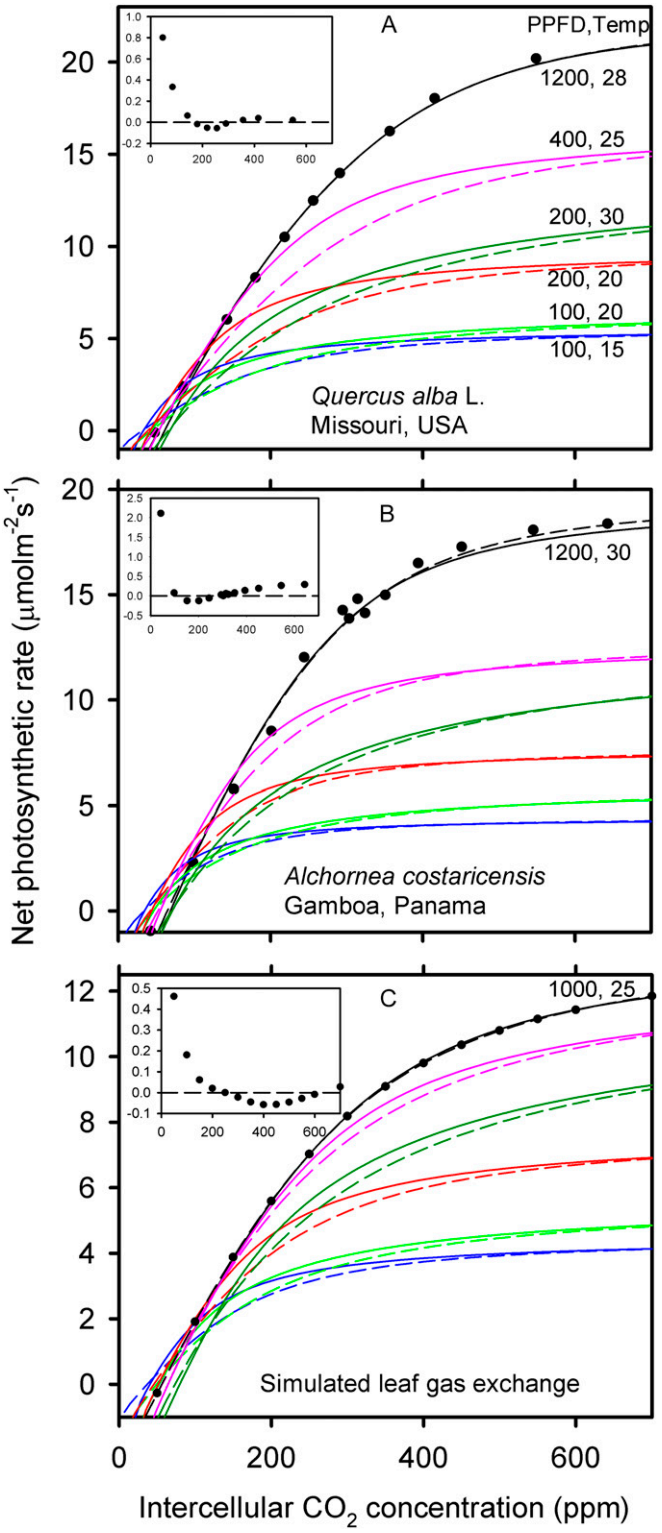

Fig. S7. Typical examples of the contrast in leaf net photosynthetic rate as a function of intercellular CO_2 concentration calculated with a leaf photosynthetic model between including (dashed curves) and lacking (solid curves) explicit consideration of mesophyll conductance (g_m). In each panel, the uppermost dashed and solid curves represent direct fitting of the g_m-including and g_m-lacking models, respectively, to the data (dots) measured (*A* and *B*) or simulated (*C*) at indicated PPFD (μmol/m^2/s) and temperatures (°C). The difference in fit between the g_m-including and g_m-lacking models is more clearly seen in the *Inset* of each panel, which shows the difference in calculated leaf photosynthetic rate (*y* axis, μmol/m^2/s) as a function of intercellular CO_2 concentration (ppm). The photosynthetic parameters obtained in the fittings are then used to calculate photosynthesis at other PPFD and temperatures, which are indicated in *A* only but are the same in *B* and *C*.

Table S1. Empirical constants used in the g_m model

Symbol	Equation	Value	Source
Modeling mesophyll conductance at canopy top			
a	Eq. **S2**	24.240338	Fit from Niinemets et al. (2)
b		−0.6509	
Canopy integration of mesophyll conductance			
d	Eqs. **S3, S4, S7**	0.8109	Fit from Montpied et al. (21)
f	Eqs. **S5, S7**	0.221897	Fit from Niinemets (15)
k_l	Eqs. **S6, S7**	0.5	Theoretical value (23)
k_g	Eq. **S7**	0.08997	$=k_l \cdot d \cdot f$
Temperature response function			
c	Eq. **S8**	20	Bernacchi et al. (5)
ΔH_a		49.6×10^3 J/mol	
ΔH_d		437.4×10^3 J/mol	
ΔS		1.4×10^3 J/mol/K	

Table S2. Empirical constants in the conversion function that relates the g_m-lacking to g_m-including parameters at a reference temperature of 25 °C

Photosynthetic parameter (μmol/m^2/s)	p	q	u	v	r^2	RMS (μmol/m^2/s)	Source
Monolimiting conversion function							
V_{cmax}	0.1164	1.2643	0.6429	0.9431	0.83	18.0437	Sun et al. (13)
J_{max}	0.0084	0.7552	0.6230	−0.1166	0.97	7.5290	
TPU	0.1249	1.8059	0.2525	1.5905	0.99	0.3597	
Colimiting conversion function							
V_{cmax}	0.0340	1.1253	0.8787	0.4801	0.83	42.9184	This study
J_{max}	0.2935	1.4838	0.0858	0.1726	0.75	62.7638	
TPU	0	0	0	0	0.97	0.9254	

Two sets of constants are given: one is for the monolimiting FvCB model and reproduced from Sun et al. (13) and the other is for the colimiting FvCB model and estimated in this study. TPU for the colimiting FvCB model differs little between the g_m-including and -lacking considerations and as such all its associated constants are effectively zero (Eq. **S14**).

Table S3. Values of key photosynthetic parameters for each PFT in CLM4.5

PFTs	M_{a0} (gC/m^2)*	g_{max0} (μmol/m^2/s/Pa)	g_m-lacking V_{cmax} (μmol/m^2/s)	g_m-lacking J_{max} (μmol/m^2/s)	g_m-including V_{cmax} (μmol/m^2/s)†	g_m-including J_{max} (μmol/m^2/s)
NET temperate	100.00	1.21	62.50	107.19	132.16 (132.31)	143.59 (124.39)
NET boreal	125.00	1.05	62.60	107.36	145.20 (143.68)	152.81 (127.06)
NDT boreal	41.67	2.14	39.10	67.06	52.83 (55.23)	76.29 (71.89)
BET tropical	83.33	1.36	55.00	94.33	100.90 (103.48)	120.77 (106.83)
BET temperate	83.33	1.36	61.50	105.47	120.11 (121.29)	135.36 (120.53)
BDT tropical	33.33	2.47	41.00	70.32	53.89 (55.91)	78.14 (74.94)
BDT temperate	33.33	2.47	57.70	98.96	83.47 (84.91)	110.32 (107.07)
BDT boreal	33.33	2.47	57.70	98.96	83.47 (84.91)	110.32 (107.07)
BES temperate	83.33	1.36	61.70	105.82	120.73 (121.85)	135.81 (120.96)
BDS temperate	33.33	2.47	54.00	92.61	76.50 (78.20)	103.18 (99.89)
BDS boreal	33.33	2.47	54.00	92.61	76.50 (78.20)	103.18 (99.89)
C3 arctic grass	33.33	2.47	78.20	134.11	126.66 (125.09)	149.94 (147.52)
C3 grass	33.33	2.47	78.20	134.11	126.66 (125.09)	149.94 (147.52)

Leaf temperature and mean growth temperature are assumed to be 25 °C. Here the mean growth temperature refers to the 10-d mean air temperature, which is used to account for the acclimation effect of J_{max}/V_{cmax} ratio in CLM (33). BDS, broadleaf deciduous shrub; BDT, broadleaf deciduous tree; BES, broadleaf evergreen shrub; BET, broadleaf evergreen tree; NDT, needleleaf deciduous tree; NET, needleleaf evergreen tree.

*Some PFTs share the same values of SLA_0 and therefore M_{a0} (=$1/SLA_0$) in CLM, which leads to the same values of M_a and g_{max} at canopy top for these PFTs.

†Two sets of V_{cmax} and J_{max} values are shown here when g_m is included, corresponding to colimiting and monolimiting conversion functions, respectively. Values in parentheses are those from the application of the monolimiting conversion function.

Table S4. Summary of global simulations

Simulation	Description
Transient CO_2 simulations	
CTRL	Control simulations with the default CLM4.5
IMED	Intermediate simulations that use CLM4.5 with the mesophyll conductance (g_m) model only, but retain the original phenomenological (g_m-lacking) photosynthetic parameters
MESO	Fully updated simulations that use CLM4.5 with the g_m model and the g_m-including photosynthetic parameters
Constant CO_2 simulations	
CTRL_cCO_2	Same as CTRL, but with a constant atmospheric CO_2 concentration (296 ppm at 1901)
MESO_cCO_2	Same as MESO, but with a constant atmospheric CO_2 concentration (296 ppm at 1901)

The MESO simulations are repeated with the colimiting and monolimiting conversion functions.

Other Supporting Information Files

Dataset S1 (XLSX)