

# *Gaps in knowledge and data driving uncertainty in models of photosynthesis*

**Michael C. Dietze**

## **Photosynthesis Research**

Official Journal of the International  
Society of Photosynthesis Research

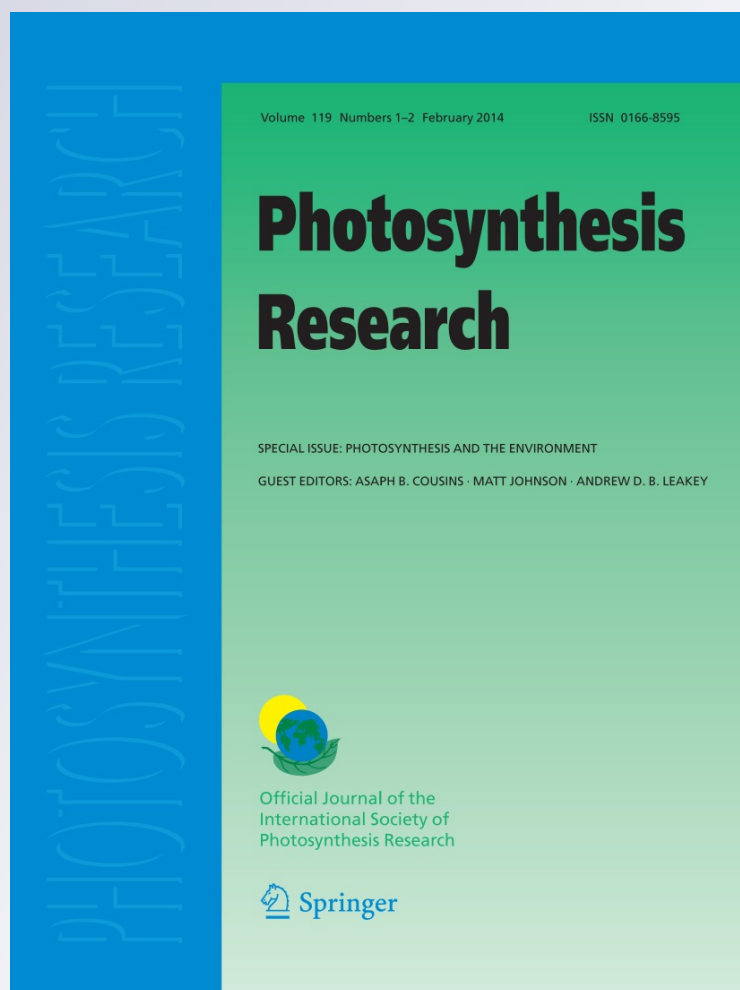
ISSN 0166-8595

Volume 119

Combined 1-2

Photosynth Res (2014) 119:3-14

DOI 10.1007/s11120-013-9836-z



**Your article is protected by copyright and all rights are held exclusively by Springer Science +Business Media Dordrecht. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at [link.springer.com](http://link.springer.com)".**

# Gaps in knowledge and data driving uncertainty in models of photosynthesis

Michael C. Dietze

Received: 10 October 2012 / Accepted: 23 April 2013 / Published online: 5 May 2013  
© Springer Science+Business Media Dordrecht 2013

**Abstract** Regional and global models of the terrestrial biosphere depend critically on models of photosynthesis when predicting impacts of global change. This paper focuses on identifying the primary data needs of these models, what scales drive uncertainty, and how to improve measurements. Overall, there is a need for an open, cross-discipline database on leaf-level photosynthesis in general, and response curves in particular. The parameters in photosynthetic models are not constant through time, space, or canopy position but there is a need for a better understanding of whether relationships with drivers, such as leaf nitrogen, are themselves scale dependent. Across time scales, as ecosystem models become more sophisticated in their representations of succession they need to be able to approximate sunfleck responses to capture understory growth and survival. At both high and low latitudes, photosynthetic data are inadequate in general and there is a particular need to better understand thermal acclimation. Simple models of acclimation suggest that shifts in optimal temperature are important. However, there is little advantage to synoptic-scale responses and circadian rhythms may be more beneficial than acclimation over shorter timescales. At high latitudes, there is a need for a better understanding of low-temperature photosynthetic limits, while at low latitudes the need is for a better understanding of phosphorus limitations on photosynthesis. In terms of sampling, measuring multivariate photosynthetic response surfaces are potentially more efficient and more accurate than traditional univariate response curves. Finally, there is

a need for greater community involvement in model validation and model-data synthesis.

**Keywords** Acclimation · Database · Ecosystem model · Statistical design

## Introduction

Global change is having numerous impacts on the global carbon cycle. This is not just limited to the impacts of elevated CO<sub>2</sub> and climate change, but involves the complex interactive effects of numerous simultaneous drivers, such as land use change, fragmentation, invasive competitors and pathogens, and atmospheric pollutants such as ozone and acid deposition. Projections of global change impacts on the carbon cycle at a regional to global scale are currently made using process-based ecosystem models, since disentangling and generalizing the multitude of interactions is beyond direct measurement and experimentation. Photosynthesis is the engine behind the global carbon cycle, and thus it is critical that we understand the impacts of these drivers on gross primary productivity (GPP) as this will feed forward to impacts on carbon storage. However, one could easily take the position that we already know everything we need to know about photosynthesis and we should worry about other parts of the carbon cycle. This is a testament to decades of hard work that has not only taken photosynthesis apart through reductionist experimentation, but has put it back together again to generate simple mathematical models with an amazing predictive capacity. Enzyme kinetic models of photosynthesis, such as the Farquhar–von Caemmerer–Berry model (FvCB) (Farquhar et al. 1980, 2001), have been around for over thirty years. While physiologists continue to refine our understanding of

M. C. Dietze (✉)  
Department of Earth and Environment, Boston University,  
675 Commonwealth Ave, Rm 130, Boston, MA 02215, USA  
e-mail: dietze@bu.edu

photosynthesis, and related processes such as stomatal and mesophyll conductance, these models are without doubt the most accurate and mechanistic part of a modern ecosystem model. Model inter-comparisons have shown that ecosystem models that include an enzyme kinetic approach to photosynthesis provide more accurate predictions of net ecosystem exchange (NEE) than other approaches (Schwalm et al. 2010; Dietze et al. 2011; Schaefer et al. 2012). Furthermore, when coupled to models of stomatal conductance, enzyme kinetic models explain how key global change drivers, such as temperature, CO<sub>2</sub>, and humidity, affect GPP. Extensions of these models also exist that account for the effects of other global change drivers, such as ozone, on photosynthesis (Reich 1987; Sitch et al. 2007). Other extensions also account for isotopic partitioning (Riley et al. 2002), providing another tracer of ecosystem carbon and water fluxes.

Despite the apparent success of photosynthesis models, in this paper, written from the perspective of field ecologist and ecosystem modeler, I want to touch on a few specific points that I feel need greater attention from the research community. This includes in part an assessment of which avenues of photosynthesis research are most pressing from an ecosystem modeling perspective. It is also an assessment of where we need a better exchange of knowledge and data, in both directions, between modelers and empiricists, and a greater engagement in the model-data synthesis process. Specifically, my objective is to address the following questions:

1. What are the remaining data needs required to parameterize photosynthetic models at regional to global scales?
2. What scales are driving uncertainty?
3. What's the most efficient way to get this information?

### What are the data needs for ecosystem models?

In discussing the data needs of ecosystem models, I will focus on models that employ an enzyme kinetic approach to modeling photosynthesis as it is both more mechanistic and more accurate (Schwalm et al. 2010; Dietze et al. 2011). First off, uncertainty analyses demonstrate that when starting from scratch, modeling a new species or functional group, that photosynthetic parameters such as  $V_{\text{cmax}}$  (maximum velocity of carboxylase) are among the most important parameters in an ecosystem model to constrain (LeBauer et al. 2012). However, despite the high sensitivity to these parameters (Booth et al. 2012), the volume and quality of data available frequently results in low uncertainties leading these parameters to drop in importance relative to other parameters in ecosystem

models (LeBauer et al. 2012). This low importance of photosynthetic parameters should be viewed as a resounding success of photosynthesis research in generating accurate models and high quality measurements. That said, beyond  $V_{\text{cmax}}$ ,  $J_{\text{max}}$  (rate of RuBP regeneration), and quantum efficiency, many of the other parameters in FvCB have substantially lower data coverage (Ziehn et al. 2011) and are frequently assumed by modelers to be constant. In addition, model inter-comparison has demonstrated that there are nonetheless significant errors and biases in current predictions of GPP (Schaefer et al. 2012). The simplest of these errors to correct is that models tend to underestimate maximum GPP (Schaefer et al. 2012), which is related to the fact that models do not take full advantage of the wealth of data that has been collected due to the lack of databases, an issue that will be discussed in detail below. Furthermore, the under-prediction results from the fact that, for reasons of scaling, models generally use canopy-level values of  $V_{\text{cmax}}$  that are lower than those observed at the leaf-level (Bonan et al. 2012).

While models underestimate peak GPP, they tend to overestimate photosynthetic rates in the fall, winter, and spring (Schaefer et al. 2012). These errors currently appear to have less to do with an incorrect temperature scaling (Bernacchi et al. 2009), and are more driven by a failure to shutdown photosynthesis at temperatures around freezing. Indeed, there is far less data available on photosynthetic responses near freezing (Starr and Oberbauer 2003) than at ambient or high temperatures, though there also remains a good bit of uncertainty about high temperature photosynthesis (Kattge and Knorr 2007; Smith and Dukes 2013). Low-temperature photosynthetic responses are particularly important for evergreen vegetation and for biomes, such as the tundra, where the mean annual temperature is below freezing and low temperatures are not uncommon in the growing season (Davidson 2012). Finally, the other consistent GPP bias across ecosystem models is an overprediction under dry conditions in response to drought and humidity stress (Schaefer et al. 2012). These responses are not driven by photosynthesis itself, but rather by direct stomatal responses to the atmosphere and indirect responses due to hydraulic limitation. I will not focus on them further other than to note that while volumes have been written on both stomatal regulation and plant hydraulics there remains a strong need for additional research and better models (De Kauwe et al. 2013). These responses are of particular importance in climate change feedbacks across the full moisture gradient from rainforests to deserts (Berry et al. 2010).

In addition to the need to understand low-temperature photosynthesis at high latitudes, it almost goes without saying that we simply need more data at low latitudes. Sampling across tropical biomes is disproportionately

small relative to both the magnitude of carbon fluxes and the diversity of species. These issues are exacerbated even further as we begin to utilize models that account for successional dynamics, as these models currently use a larger number of plant functional types to describe northern hardwood forests (Medvigy et al. 2009) than they do to describe tropical rainforests (Moorcroft et al. 2001). This is purely a reflection of data limitation. In addition, as will be discussed below, issues of multiple nutrient limitations on photosynthesis are greater in these biomes and currently not dealt with by most ecosystem models. Our understanding of responses to elevated CO<sub>2</sub> are likewise inadequate in both the tropics and at high latitudes (Leakey et al. 2012).

So far we have treated photosynthetic parameters in ecosystem models as if they are fixed constants with a temperature scaling, as indeed this is how many ecosystem models represent photosynthesis. However, there is ample evidence that these parameters are not fixed but vary across a multitude of spatial, temporal, and phylogenetic scales. One of the strongest factors driving variability in photosynthesis is leaf nitrogen, as Rubisco alone makes up around half of a leaf's protein (Ellis 1979). Relationships between leaf N and  $V_{\text{cmax}}$  exist in most models, either explicitly as a term that dynamically varies  $V_{\text{cmax}}$  with leaf N (Bonan et al. 2011), or implicitly due to across species trade-offs (Wright et al. 2004). Global syntheses not surprisingly demonstrate lower uncertainty in the N– $V_{\text{cmax}}$  relationship across the temperate biomes, due to significantly higher sample sizes, while the tropics, grasslands, and shrublands are more uncertain (Kattge et al. 2009). More interesting, they demonstrate a much shallower relationship on tropical oxisols, suggesting P limitation (Kattge et al. 2009) and a need to improve models that consider multiple nutrient limitation (Harpole et al. 2011; Wang et al. 2010; Goll et al. 2012). While there has not yet been a formal partitioning of uncertainty in these relationships, the other thing that is noteworthy is that despite a notable amount of scatter it appears that within-PFT variability in nitrogen dominates the variability in PFT-level estimates of  $V_{\text{cmax}}$  (Kattge et al. 2009, Fig. 2, Feng and Dietze in review). However, it is important to note that this statement does not tell us on what *scale* this variability occurs. For example, most ecosystem models with dynamic N feedbacks will use relationships such as these to allow for within-PFT variability in response to differences in N-uptake across space and time. However, it is also true that much of this within-PFT variability in foliar traits may be due to differences among the species that make up a PFT (Townsend et al. 2008), which are ignored in most ecosystem models. Within a site, there is also considerable variability in leaf traits associated with canopy position, leaf age, and responses to stress, damage, and thermal

acclimation. However, it is not a priori obvious that across-species relationships can be used to drive variability across space, time, age, canopy position, etc., or that the relationships are the same for each of these effects. It would obviously be a great boon to modelers if these relationships were scale invariant, but this has yet to be demonstrated rigorously. These ideas are revisited in the second section “What scales are driving uncertainty?”.

Addressing the data needs of models requires the collection of new data, but even more than this it requires that we do better job at making existing data open and organized. I believe that the plant biology research community as a whole would benefit considerably from an open database project to archive data on leaf-level gas exchange and photosynthesis. Even if we ignore the needs of ecosystem models, there are many clear examples, from GenBank to FLUXNET, where the construction of databases have enabled new and creative syntheses that were previously unimaginable and well beyond the initial concept of the original architects of these databases. Leaf-level data are uniquely suited for such a database as it is relevant across the full range of the scales of organization in plant biology, from the gene to the global biosphere. Gas exchange is one of the few places in plant biology where a global scale modeler and an Arabidopsis geneticist can find overlap in the data of interest and the methods used for data collection and analysis. Both parties can, for example, find common cause in the task of keeping a LI-6400 running smoothly. There is growing interest from the community in connecting data across this full range of scales of organization (Neale et al. 2010), but leaf-level physiology represents a gap in cyberinfrastructure between the molecular and ecological. For example, the iPlant Tree Biology project recently launched CartograTree (<http://dendrome.ucdavis.edu/cartogratree/>), a meta-database that allows spatial queries of databases on trees across scales from the genome to the ecosystem (Vasquez-Gross *in review*). However, the only leaf-level data currently in CartograTree are summary statistics from the TRY-DB (<http://try-db.org>) trait database (Kattge et al. 2011). Work with summary statistics on photosynthetic traits has admittedly been very useful for parameterizing ecosystem models (Kattge et al. 2009; Bonan et al. 2012), but this approach is inherently limited when new analytical methods are developed (Patrick et al. 2009; Feng and Dietze in review) or when models are refined (Bernacchi et al. 2009; Medlyn et al. 2011). Only by archiving the raw measurements can the data be re-analyzed using consistent approaches as models evolve. Furthermore, summary data and the literature meta-analyses are tedious to compile and generally lack information on important covariates, such as vapor pressure deficit (VPD). Decades of raw data are in danger of being lost if archives are not developed.



Another key point in favor of database construction is that leaf-level gas exchange measurements currently benefit from a remarkable degree of uniformity in data formats and measurement technologies. We are long past the days when labs had to construct their own measurement systems and current research is overwhelmingly done using instrumentation from a handful of companies (LI-COR, PP Systems, ADC BioScientific, CDI Bio-Science). The benefit of this uniformity is that cyberinfrastructure need only deal with a small number of file formats and the required meta-data is simplified. The creation of such databases will allow this data to be mined and manipulated in novel ways (Hey et al. 2009). In addition, the diversity of possible covariates and applications suggests that database interoperability be a key goal, so that gas exchange data can be easily linked to data on genomes, gene expression, plant traits, spectroscopy, flux towers, etc. Furthermore, funding agencies are increasingly requiring that data be publicly archived, but existing database options do not meet these criteria. The aforementioned TRY-DB database is capable of storing gas exchange data and parameters estimated from these data, but is not set up to handle raw gas exchange files and only a subset of the data is public. The only gas exchange database I am aware of, the LeafWeb project (<http://leafweb.ornl.gov>), is designed to promote a specific statistical method for fitting  $A-C_i$  curves (Gu et al. 2010) and likewise does not currently make data publicly available. While LeafWeb is set up to ingest raw data, it places unnecessary and complex formatting requirements on the data, while at the same time requiring meta-data that is biased toward observational ecological measurements. As such, there is no easy way to record experimental designs or connect these data to ancillary data on the multitude of other things we have been known to measure in conjunction with gas exchange, such as microarrays or spectroscopy. At present, there is not even a requirement for including the scientific name of the species measured or a way to archive measurements other than  $A-C_i$  curves. It may be that LeafWeb proves to be an important first start, and greater interoperability with projects like TRY-DB could join the strengths of both approaches. However, regardless of whether it builds upon existing efforts or is an entirely new project, there is a need for an open database that the full span of research communities find useful and can support.

### What scales drive uncertainties?

As alluded to in the last section, photosynthetic parameters can vary across a multitude of scales and in response to a range of factors. In this section, I provide a perspective on which scales are likely to be most important for ecosystem

models, and sketch out what these mathematical models might look like. First, let me reiterate from the previous sections that no other process in an ecosystem model is as mechanistic or predictive as enzyme kinetic photosynthesis. Therefore, it is *not* the calculation of GPP at each model time step, in response to changes in temperature, light, and humidity, which drives ecosystem model uncertainty. Second, there are a number of processes that affect how photosynthetic models scale in space, first from the leaf to the canopy, and then from the canopy to the larger landscape or model grid cell. I will not focus on this scaling because it has been discussed extensively elsewhere, including in this issue (Desai *in review*), and because spatial scaling is the motivation behind elegant models like the ecosystem demography (ED) model (Moorcroft et al. 2001; Medvigy et al. 2009). I instead want to focus on other aspects of the scaling problem, and will instead focus on addressing the question “*What are the important time scales on which plants adjust their photosynthetic parameters?*” First, at the instantaneous scale, it bears reminding that enzyme kinetic models are *equilibrium* models not instantaneous models of photosynthesis. Given the 30–60 min time step of most models, this equilibrium assumption appears valid; however, this is only true if the environmental drivers are approximately constant across the time step. Given that the time step was chosen so that there is little trend in the drivers, this is probably a safe assumption for the upper canopy. However this assumption is unlikely to be true in the understory where the light environment is driven by sunflecks (Naumburg and Ellsworth 2002). Responses to sunflecks involve dynamic responses and feedbacks that differ non-trivially from the square wave response one would predict from an equilibrium model, and enormously from the response under the average understory light level (Way and Pearcy 2012). These dynamics can have a large impact on understory growth and survival, and can interact with responses to global change drivers such as  $\text{CO}_2$  (Leakey et al. 2002). In many ecosystem models these effects will have little impact, as the understory contributes little to total GPP. However for models that include successional dynamics the growth and survival of the understory plays an important role in determining the rate and trajectory of species turn over (e.g., Pacala et al. 1996). The modeling challenges that remain are in how, within a computationally-efficient canopy radiation model, to approximate the distribution of sunflecks, and how to approximate a dynamic photosynthesis model based on the sunfleck distribution (e.g., Gross et al. 1991) without having to simulate the dynamic responses second-by-second. One option for the latter might be a mixture model between shade and sunlight conditions with an exponential rate parameter controlling the transition from one state to the other. Such a

model may have an approximate solution in the time-spectral domain using Fourier or Laplace transforms. From a physiological perspective, there is a need for simple generalizations of the enzyme kinetic models that account for these dynamics (e.g., Gross et al. 1991) and data with which to calibrate these models across a wide range of biomes.

In considering time scales of variability in photosynthesis, let us next jump to the process of acclimation. By acclimation, we refer to the change in biochemical and physiological processes associated with responses to environmental conditions that shift instantaneous response functions. For example, a sustained increase in temperature will trigger genetic signals that lead to a shift in  $V_{\text{cmax}}$  toward a higher photosynthetic optimum. However, this longer-term shift in  $V_{\text{cmax}}$  falls along a very different response curve than the instantaneous Arrhenius response to short-term temperature fluctuations. Smith and Dukes (2013) recently published a detailed review of photosynthetic acclimation to temperature and  $\text{CO}_2$ , and the representation of these processes in ecosystem models, so I will not repeat that discussion here but will instead touch on a few points that bear greater scrutiny. First, it is important to separate the acclimation process into one part addressing what the new equilibrium parameter will be, and another part describing the rate at which that equilibrium is reached. We know more about the equilibrium (Kattge and Knorr 2007), and ultimately this is the more important of the two parts. It is also worth mentioning that most data on temperature optima document changes in  $A_{\text{max}}$  (“photosynthetic assimilation rates measured under high light, ample soil moisture, and ambient  $\text{CO}_2$ ,” censu Wright et al. 2004) not  $V_{\text{cmax}}$ , which is not adequate for the purpose of parameterizing ecosystem models. It is also noteworthy that the temperature optima for  $A_{\text{max}}$  and  $V_{\text{cmax}}$  qualitatively seems to follow a different shape, with the former being more symmetric and possibly lower in the peak temperature. This difference may be due the fact that  $A_{\text{max}}$  is more responsive to high temperatures because of the difference in VPD, which increases rapidly with temperature, as well as the differences in temperature dependence among other model parameters, such as  $K_c$ ,  $K_o$ , and  $I^*$ , and the temperature dependence in leaf respiration (Lin et al. 2012).

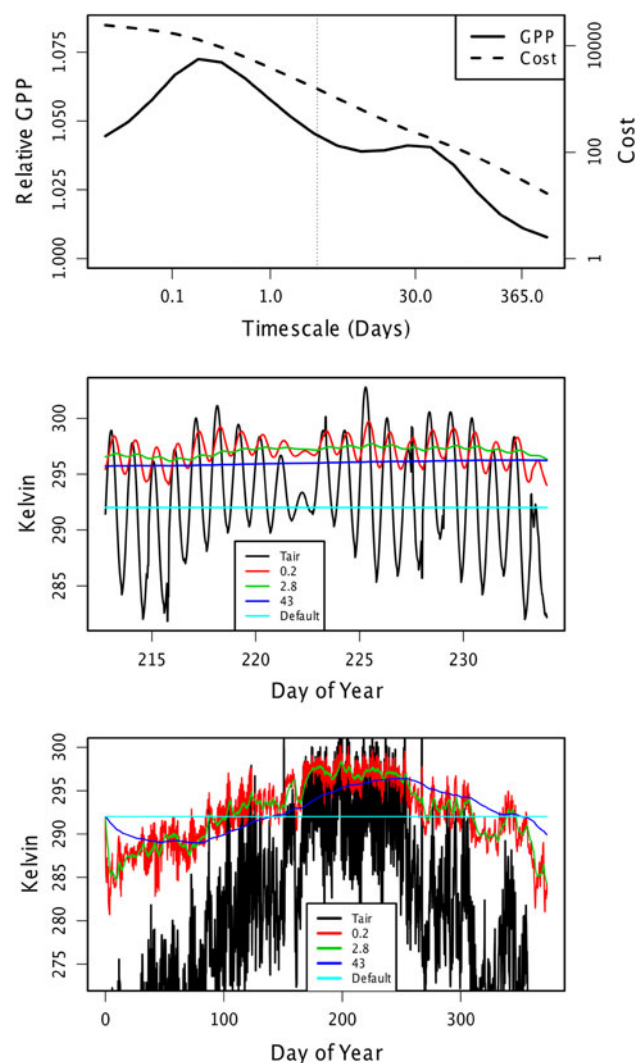
Global models that include acclimation find that there is only a modest overall response from including these dynamics, but that this is the result of compensating responses between high and low latitudes, which are each in the 20 % range and thus represent important processes (Friend 2010). While Smith and Dukes (2013) claim that “The absence of long-term plant carbon exchange responses in most process-based models can no longer be attributed to a lack of suitable algorithms” I feel this view is a bit optimistic. Of the two models, they present both are, in essence, just

regressions between average temperature and the optimum temperature for  $V_{\text{cmax}}$  and  $J_{\text{max}}$ , and thus provide little mechanistic understanding of acclimation. One is based on a single site and an optimization assumption (Friend 2010). The other is much more robust, based on experimental data for 36 species, but the species are predominantly temperate (Kattge and Knorr 2007). There remains a pressing need to determine if the assumption of linearity holds in polar and tropical regions, especially since this is where the largest impact on GPP occurs. Also, the linear model only explains approximately 40 % of the variability, which means that there are clearly other processes contributing to variability in optimum temperature. Indeed, parametric uncertainty associated with acclimation was a major driver of uncertainty in a level-level model of net photosynthesis (Ziehn et al. 2011). Acclimation is an area ripe for greater exploration across the scales of organization, as a better understanding of the genetic and biochemical drivers of acclimation may open the door for a more mechanistic and general model.

Returning to the question of the rate of acclimation, different models currently appear to use different time-scales (e.g., 3, 15, and 30 days) and, to the best of my knowledge, there is little data behind these choices. In order to explore these choices let us consider a simple model of acclimation along the lines of those presented above. I start from the variant of the FvCB model as described by Medlyn et al. (2002) using  $V_{\text{cmax}} = 100$  and the weighted mean value for Ha (64.4). Next, we will apply the Friend (2010) relationship between leaf temperature ( $T_L$ , °C) and “equilibrium” optimal temperature,  $T_e = 17 + 0.35 T_L$ . Finally, we need a model to relate the optimum temperature at any point in time to the current leaf temperature and the equilibrium temperature. As a simple starting point consider the simple exponential decay function  $dT_{\text{opt}}/dt = 1/\lambda(T_e - T_{\text{opt}})$ . Here  $\lambda$  is the time-scale (units = days) at which the instantaneous optimum temperature approaches the equilibrium optimum temperature. GPP is then simulated from 1998 to 2006 using 30 min meteorological data from the Willow Creek Ameriflux tower in a mature hardwood forest in northern Wisconsin (45° 48'N, 90° 05'W), that was gap-filled by the North American Carbon Program (Ricciuto et al. 2009). In addition to just evaluating the impacts of acclimation, let us also estimate the costs, as the processes of up- or down-regulating photosynthesis is undoubtedly associated with additional metabolic costs of protein synthesis (Amthor 2000). Here, the actual carbon costs of acclimation are unknown, but are assumed to be proportional to the cumulative number of degrees that  $T_{\text{opt}}$  is shifted up and down. For example, if  $T_{\text{opt}}$  is shifted one degree up every morning and one degree down every afternoon the annual “cost” would be  $2 \times 365 = 730$  degrees, while if  $T_{\text{opt}}$  is

shifted up 5 degrees in the spring and down 5 degrees in the fall then the annual “cost” would be 10 degrees. Figure 1a shows the impact of choosing different timescales (presented on a log axis) on the GPP of the site compared to the GPP calculated without acclimation. The absolute units for the cost curve in Fig. 1a are not directly comparable to the carbon benefits of acclimation, thus we cannot perform a cost-benefit analysis, but this still provides a relative metric of the costs of acclimation. From this simple theoretical exercise, we learn a number of interesting things. First is that GPP is maximized when acclimation occurs at a sub-daily timescale. However, if we look at the estimate of  $T_{\text{opt}}$  when  $\lambda = 0.2$  days in Fig. 1b we see that, compared to air temperature, the optimum temperature lags the diurnal cycle. GPP would thus actually be maximized if plants could anticipate the diurnal cycle ahead of time rather than responding to it using an acclimation response. In other words, from a theoretical perspective, circadian rhythms could have a larger impact on GPP than acclimation. There is evidence for circadian cycles in plants at the molecular levels (Harmer 2009), at the leaf photosynthetic level (Hennessey and Field 1991), as well as at the ecosystem level (Dios et al. 2012). However, it remains unclear how important circadian rhythms are to photosynthesis under field conditions (Williams and Gorton 1998), nor are these effects incorporated into current ecosystem models. The second thing we learn from this simple model is that the GPP response is fairly flat over a span of timescales from  $\sim 3$  to  $\sim 45$  days. This means that there is little benefit gained from faster acclimation within this range, and indeed GPP is slightly lower at the faster timescale. In practical terms, this means there is little benefit to try to acclimate to synoptic-scale weather patterns, as a plant is just as likely to find itself, at the end of a weather front, disacclimated to the new conditions as it is to benefit from such acclimation (Fig. 1b, c). It is also possible that failing to include acclimation responses plays a role in the progressively larger errors in ecosystem models at timescales beyond the synoptic (Dietze et al. 2011). From this simple model, we also learn that the apparently arbitrary choice of timescale in previous studies is unlikely to have had a large impact on results, as the values in use are in this flat region. That said, it is still important to make measurements that document the temporal dynamics of acclimation responses as this simple model may not hold in practice. For example, one could impose a large step change in temperature and then record the daily changes in photosynthetic parameters and  $T_{\text{opt}}$ .

Moving to the next time scale, feedbacks through resources occur across a range of scales. As discussed earlier, relationships between photosynthetic parameters and



**Fig. 1** Timescales of photosynthetic acclimation, Willow Creek Ameriflux Tower, WI. *Top* allowing for the acclimation of the optimal temperature for photosynthesis increases GPP (unitless, expressed relative to a model without acclimation) regardless of the timescale at which acclimation occurs, but GPP is maximized at a sub-daily timescale. GPP is relatively insensitive to acclimation over a range from 45 days down to 3 days (vertical dotted line, value from Friend 2010). The “cost” of acclimation, expressed in terms of the cumulative degrees of acclimation applied to  $T_{\text{opt}}$ , declines rapidly as timescale increases (values on log scale) suggesting that there is little benefit gained from acclimating to synoptic-scale variation. *Middle* comparison between air temperature ( $T_{\text{air}}$ , black) and the optimal photosynthetic temperature ( $T_{\text{opt}}$ ) at different timescales (red = 0.2 - days, green = 2.8 days, blue = 43 days) during peak summer demonstrates the strong benefit of acclimation at the intermediate time scales. The  $T_{\text{opt}}$  in the model without acclimation is denoted “Default” (cyan). Sub-daily acclimation lags observed temperatures, suggesting that processes such as the circadian rhythm that anticipate the diurnal cycle are likely to provide a greater net benefit than acclimation processes that respond to the environment. *Bottom* same comparison between  $T_{\text{air}}$  and optimal temperature as middle panel but over a full year



resources, especially nitrogen, play an important role in many ecosystem models. One of the important empirical questions that need be resolved is whether there is a degree of “universality” to these relationships or whether they are scale dependent. For example, are the across-species relationships between  $V_{\text{cmax}}$  and N discussed earlier (Kattge et al. 2009), equivalent to the within-species relationships? There is evidence to suggest not, but rather that species vary in their inherent nitrogen use efficiency based on life history and depending upon environment, meaning that the within-species slope of this line differs from the slope of the across-species relationship (Feng and Dietze *in review*). This difference is important because it means that if a model employs the across-species relationship to estimate  $V_{\text{cmax}}$  from N then the response to nitrogen deposition or fertilization would be incorrect because the response is driven by the within-species relationship. However, even here the response depends upon time scale, as on a longer time scale the change in resources will drive a community shift in species composition rather than a phenotypic shift in foliar chemistry. Even when addressing the within-species relationship, it is not yet clear whether a single relationship drives different responses on different scales. For example, is the relationship between  $V_{\text{cmax}}$  and N associated with a change in light environment along a canopy profile the same as the relationship associated with leaf aging, and is either the same as the relationship governing the response to nitrogen additions? Also, what is the relative importance of each of these processes? In our own work in grasslands we find that, across many species and functional groups, variations in foliar N account for most of the variation in  $V_{\text{cmax}}$  throughout the growing season (Feng and Dietze, *in review*). We also show that this seasonal variability associated with leaf aging has a larger impact on estimated GPP than the effects of canopy position (Wang et al. 2012). This is noteworthy because far more ecosystem models include the effects of canopy position than include leaf aging. Finally, induced foliar responses to damage and herbivory are likely to drive yet another different  $V_{\text{cmax}}$ –N relationship than age, canopy position, or fertilization, because nitrogen may be allocated away from photosynthesis and into defense compounds. The challenge is that models require generality, and if we end up with scores of different  $V_{\text{cmax}}$ –N relationships then large-scale models will be paralyzed unless there are general mechanisms that explain these responses. In fact, one of the most critical benefits that molecular, genetic, and physiological plant biologists can provide to global ecology is this type of quantitative understanding of general mechanisms that allows modelers to see above the Sisyphean task of developing statistical relationships for a quarter billion different plant species. Also related is a need to improve models to predict foliar nitrogen, as a flawless  $V_{\text{cmax}}$ –N relationship is of limited value if N is unconstrained.

The final time scale that I want to touch on is that of succession. This time scale is often overlooked by modelers and physiologists alike; however, it is important to remember that the responses to climate change will play out over centuries and that on these time scales it is the community processes of growth, mortality, dispersal, recruitment, and disturbance that shape ecosystems. Despite many textbook examples of successional processes being orderly and predictable, reality tends to be more complex and dependent upon both current and historical conditions (Foster et al. 2003; Dietze and Clark 2008). On these time scales, the factors that affect understory survival, such as photosynthesis at low light, sunfleck utilization, and the coupling between photosynthate production, respiration and allocation, can play a role in determining long-term dynamics that is disproportionate to the amount of carbon fixed by these plants (e.g., Pacala et al. 1996). It is also important to remember that because succession plays out over long time scales, land use history plays an important role in determining the current state and trajectory (Foster and Aber 2006; Albani et al. 2006). For example, many of the forests in the eastern U.S. are still recovering from logging and agriculture that occurred in the 1800s and early 1900s, which predicts a continuing shift to later successional species that have lower maximum photosynthetic rates and deeper canopies (Albani et al. 2006). However, past FACE experiments in this region have all looked at early successional species and less is known about how mid and late successional species will respond to elevated  $\text{CO}_2$ , or how shifts in photosynthetic responses might accelerate or retard successional trajectories (Mohan et al. 2007). Ironically, photosynthesis research remains relevant to the issue of land use legacies, as hyperspectral remote sensing provides a way of connecting photosynthetic chemistry to species-level mapping (Asner and Martin 2009).

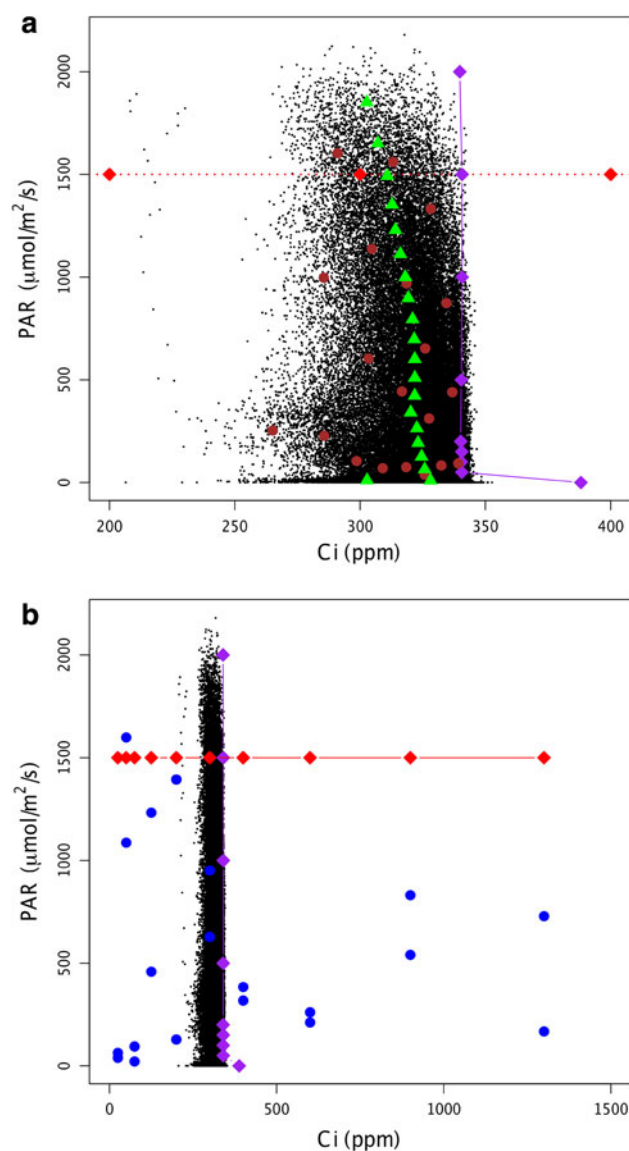
### What is the most efficient approach to sampling?

A common theme of the previous two sections was that, from the perspective of ecosystem models, one of the primary interests is in the parameters of enzyme kinetic models, such as the FvCB C3 model, and how those parameters vary across different scales, across species, and in response to different drivers. The effect of this is that point measurements of  $A_{\text{max}}$  are of considerably lower value to models than response curves, despite the fact that the former is far simpler, quicker to measure, and often adequate to address specific research questions. However, if the goal is to estimate photosynthetic parameters from gas exchange measurements, then an important question becomes how to do this most efficiently. Prior statistical treatments of the FvCB model have focused primarily on

methods for fitting univariate response curves (Dubois et al. 2007; Patrick et al. 2009; Gu et al. 2010) and design has unfortunately been mostly relegated to “rule of thumb” advice (Long and Bernacchi 2003).

One thing common to current measurement approaches is that they have been focused on sampling traditional response curves of either  $\text{CO}_2$  ( $A-C_i$ ) or light ( $A-Q$ ) or both. A reason for this is that there are features of these curves that are of direct interest to physiologists, though perhaps a greater factor is that many of the current statistical approaches to parameterization require these univariate curves. However, if we think about using the parameters of these models to predict GPP, then it is important to remember that we are not using these two curves to make predictions but the full response surface describing net assimilation as a function of light,  $C_i$ , and temperature. Furthermore, in predicting GPP we only use a specific subspace, not all combinations of light,  $C_i$ , and temperature are observed in the environment. Figure 2 (black dots) shows the modeled projection of light and  $C_i$  values experienced by the top of the canopy at the Willow Creek tower over nine years as described above. This was generated using the same parameterization of the FvCB model but coupled to the (Leuning 1995) stomatal conductance model assuming a stomatal slope of 8 and a reference VPD of  $D_0 = 1$  kpa. Onto this surface, we have projected the  $A-C_i$  (red) and  $A-Q$  (purple) response curves that would typically be run to parameterize this model, with the points indicating the default measurement values for the LI-COR 6400 auto-program. A few things can be gleaned from these plots. The first is that only one point on the whole  $A-C_i$  curve intersects with the space over which the model is applied. Second, the light response curve (for which we assumed a typical setting of  $C_a = 400$  ppm) samples a higher fraction prediction space, and in fact the distribution of sampled light levels does a better job of approximating the distribution of observed light levels. It is also obvious that the light response curve only takes a glancing blow past the right-hand edge of the prediction space, which is not surprising given that we assumed a  $\text{CO}_2$  concentration above that experienced by the leaves. More generally, we see that predictions with the FvCB model are largely extrapolations beyond the measurements. We also see that if we want to explore prediction space then light response curves are much more valuable than  $A-C_i$  curves. Since the goal isn't strictly to explore prediction space, but to estimate parameters, there is still reason for making gas exchange measurements outside of this range. However, there is also a strong argument to be made that two orthogonal transects through  $Q-C_i$  space is not a particularly efficient design for either goal.

To begin to explore the options for different “space-filling” designs, I compared four simple alternative sampling designs. In all cases, I fit the FvCB model to



**Fig. 2** **a** Predictive space for photosynthesis modeling. Black dots indicate the observed photosynthetically active radiation (PAR) and predicted substomatal  $\text{CO}_2$  concentration ( $C_i$ ) every half hour for from 1998 to 2006 at the top of the canopy at the Willow Creek Ameriflux Tower, WI. The colored dots indicated different sampling designs over this space:  $A-C_i$  curve (red diamonds)  $A-Q$  curve (purple diamonds), extended  $A-Q$  cluster (green triangles), Normalized cluster (brown dots). **b** Full driver space for photosynthesis modeling. Blue dots indicate an example of a Latin Hypercube design. All other dots are as in (a): predictive space (black dots),  $A-C_i$  curve (red diamonds),  $A-Q$  curve (purple diamonds)

simulated data assuming that the FvCB parameterization described above is the “true” model and that measurement uncertainty has a standard deviation of  $0.1 \mu\text{mol}/\text{m}^2 \text{ s}$ . The first sampling design is the “business as usual” approach of running orthogonal  $A-C_i$  and  $A-Q$  curves, and here we use the sampling illustrated in Fig. 2B with 11 points on the  $A-C_i$  curve (red) and 9 on the  $A-Q$  (purple) for a total sample size of 20. To keep the comparison simple, we preserve this

sample size for all three experiments, though in practice the greater usefulness of alternative designs would be to reduce the handling time per leaf. The next two designs we consider focus on the prediction space. The first applies a k-means cluster analysis to the prediction space in Fig. 2 (green triangles). Because there is much more variability in light than  $\text{CO}_2$ , this approach results in sampling design that is very similar to an  $A-Q$  curve but with a higher sampling intensity, so we will refer to this as the “Extended  $A-Q$  Cluster.” In the second design, “Normalized Cluster,” we again apply cluster analysis, but this time normalize both PAR and  $C_i$  to have a mean of zero and standard deviation of one (Fig. 2a, brown circles). This results in a design that fills the predictive space more evenly. For the final design, we apply a “Latin Hypercube” design to the full  $Q-C_i$  space (Fig. 2b, Blue). A latin hypercube design involves sampling regularly along each axis, but randomly permuting the order, so that the joint distribution fills space more evenly than purely random sampling. For the sequence of twenty light values, we use the center of each 5 % quantile bin so that the sampling is proportional to the observed variability in light, which also means that (like with a traditional  $A-Q$  curve) there is more sampling at low light values than would occur with a regular sequence of light levels. For the  $\text{CO}_2$  values, we use ten points along the “traditional” sequence for an  $A-C_i$  curve, which likewise oversamples low values, but sample each  $\text{CO}_2$  level twice (but at different light levels).

Each of the four design’s pseudo-observations were generated assuming the FvCB model is true but with Normal observation error ( $\text{SD} = 0.1 \mu\text{mol}/\text{m}^2 \text{ s}$ ). All four were fit using the Bayesian approach described in Feng and Dietze (in review). This approach uses standard Markov Chain/Monte Carlo (MCMC) algorithms, as implemented in JAGS 3.2.0 (<http://mcmc-jags.sourceforge.net>), but drops any covariates or random effects since we are only fitting one leaf. Bayesian MCMC methods involve iteratively proposing a new set of parameter values similar to the current parameter set, evaluating the statistical likelihood that the data were generated under the current parameter set, and then moving to the new parameter values with some probability proportional to the likelihood time the a priori probability of that parameter set. Importantly, this approach fits all data at once, rather than segmenting the data, and the output of the analysis is probability distribution for each parameter rather than a single value. Table 1 shows the proportional difference in the standard deviations of different parameters in the FvCB model expressed relative to the “traditional” sampling design of running  $\text{CO}_2$  and light response curves. All of the proposed designs do considerably better at estimating the quantum efficiency (34–44 % reduction in standard deviation) and moderately better at estimating dark respiration

(7–22 % reduction) (Table 1). The alternative designs all did worse at estimating  $J_{\text{max}}$  (8–58 % increase) with the latin hypercube design performing considerably worse than the two cluster designs. This decrease in performance for  $J_{\text{max}}$  is driven by a reduction in sampling, in all of the alternative designs, under conditions where  $J_{\text{max}}$  is the limiting factor. This suggests that alternative designs could also be constructed to focus sampling on the conditions where individual terms are well isolated. That said, in this example, the predictive space was dominated by conditions that were limited by either quantum efficiency or  $V_{\text{cmax}}$  and thus gives a lower predictive uncertainty. If conditions of  $J_{\text{max}}$  limitation had been prevalent, both cluster designs would have increased their sampling of these conditions.

The two cluster based designs also did considerably better at estimating  $V_{\text{cmax}}$  (47–55 % reduction) than the latin hypercube (16 % increase). Of the two cluster designs, the “Extended  $A-Q$  Cluster” performed better than the “Normalized Cluster” design for all four variables. From these results, we can conclude that space-filling designs as a whole definitely appear promising for reducing parameter uncertainties, and that more work would be helpful to flush out these ideas (especially reducing sampling intensity) and empirically validating such designs. Surprisingly, the designs that focused on the prediction space outperformed the design that sampled more widely, which is encouraging because this also suggests lower predictive uncertainties. It is also encouraging that the “Extended  $A-Q$  Cluster” design had the lowest uncertainty because it is also the design that is the most realistic to implement. The other designs involved larger changes in leaf conditions between measurements, and thus potentially larger acclimation times between measurements, which obviously partially defeats the purpose of the alternative designs. This cluster design, in fact, involves smaller changes between adjacent samples than the traditional approach, and thus is likely both faster and lower error than running standard response curves.

Finally, I would like to reiterate that one of the primary reasons to increase the efficiency of sampling at the single-leaf scale is to allow an increase in sampling at the across-leaf scale. At the across-leaf scale, I would like to make a strong argument for regression-based designs over ANOVA-based

**Table 1** Percentage change in FvCB parameter uncertainty for different sampling designs relative to the traditional approach of  $A-C_i$  and  $A-Q$  curves. Negative values indicate that a design results in lower parameter uncertainty for the same sampling intensity

Sampling Design	$\infty$ (%)	$V_{\text{cmax}}$ (%)	$J_{\text{max}}$ (%)	$R$ (%)
Extended $A-Q$ cluster	–44	–55	+8	–14
Normalized cluster	–34	–47	+21	–7
Latin hypercube	–41	16	+58	–20

designs whenever practical. The statistical training that most of us receive places far too much emphasis on replication; however, response *functions* have immensely greater value for modeling and prediction than do discrete treatments. For the same sampling effort, far greater information is gained from measuring 1 leaf at each of  $N$  levels, whatever the treatment is, than is gained from measuring  $N/2$  leaves at each of two treatments. Obviously, a smooth continuum of treatments is not practical for all experiments, for conceptual or logistical reasons, and in this case, replication is the only way to increase power. Nonetheless, we should not hold up high replication as the gold standard for doing good science. Finally, as discussed in the preceding sections, it is important to measure key covariates, such as leaf nitrogen, and to think critically about the scale dependence of observations. However, it is equally important to think critically about how to account for the different sources and scales of uncertainty in experimental designs and our analyses. Modern statistical techniques allow for a robust and explicit partitioning of different types of uncertainty rather than just lumping all variation into a residual (Clark 2005). Thinking about how uncertainty is partitioned should not be just for the statistically obsessive-compulsive, as an understanding of the patterns of what we *don't* know can qualitatively change our inference (Clark et al. 2007) and help direct future research. For example, knowing whether the “residual” error in estimates of  $V_{\text{cmax}}$  are predominantly across space, across time, or actually in the observation error changes the hypotheses about what is driving that variability and how we deal with it. Throwing all the error in a residual does not tell you whether you need to sample more frequently, sample more plots, or design a more precise instrument.

## Conclusions

As stated in the introduction, the goal of this paper was to provide a perspective on the what, when, and how of improving our understanding of photosynthesis at longer and larger scales. Our exploration of *what* are the remaining data needs, what time scales drive uncertainty (*when*), and *how* to improve measurements took us on a whirlwind tour through cyberinfrastructure, modeling, and statistics. To briefly summarize, there is a general need for an open, cross-discipline database on leaf gas exchange. Response curves are more beneficial to models as they allow the parameterization of photosynthetic models. Measuring photosynthetic response surfaces are potentially more efficient and more accurate than univariate response surfaces. Within photosynthetic models, there is a need for a better understanding of how parameters vary across scales, that the variability in parameters may be scale dependent, and to determine to what extent we can generalize these responses rather than relying

on scale-dependent relationships. Understanding responses across scales is not simply a matter of characterizing the change in the mean, but also requires a careful consideration of how uncertainty and variability change across scales. As we move to models that are better able to address the longer time scale dynamics of succession, there is ironically a need to better approximate the instantaneous light responses of leaves to integrate sunfleck responses essential to understory growth and survival. Beyond these challenges that apply globally, there are additional needs focused on specific biomes. At both high and low latitudes, there is a general data limitation and a particular need to better understand the relationships between growing temperature and optimal photosynthetic temperature, as current models of acclimation are largely linear extrapolations from the temperate region. At high latitudes, there is also a need for studies on low-temperature photosynthetic limits, while at low latitudes there is a need for a better understanding of phosphorus limitations on photosynthesis.

While couched in my perspective as a field ecologist and modeler, my hope is that these thoughts have provided fertile ground for synthesis across disciplines as the solutions to many of the problems in global models likely lie as much in the hands of the “omics” researcher as they do among those of us who drag our LI-CORs to the far corners of the world or nurture hundreds of thousands of lines of computer code to maturity. This need for cross-discipline fertilization is essential, but the volume of data and the breadth of scientific understanding are simply too great if the work of assimilating data into models and evaluating models is left solely to a small cadre of modelers. To be successful, it also needs to be paired with a more active engagement of the research community with the modeling processes through the development of tools that make models more accessible and more a part of everyday research (Dietze et al. 2013).

**Acknowledgments** This work was supported by National Science Foundation “Advances in Biological Informatics” Grant #10-62547 to M. Dietze and was part of the North American Carbon Program. Xiaohui Feng and Kevin Wolz contributed to the development to the JAGS code for fitting the FvCB model. Ankur Desai provided data from the Willow Creek Ameriflux tower. Andy Leakey, Evan DeLucia, and Don Ort provided numerous enlightening discussions about photosynthesis during my years in the “Genomic Ecology of Global Change” theme at the University of Illinois’s Institute for Genomic Biology. David LeBauer and two anonymous reviewers provided useful comments and suggestions.

## References

- Albani M, Medvigy DM, Hurtt GC, Moorcroft PR (2006) The contributions of land-use change, CO<sub>2</sub> fertilization, and climate variability to the Eastern US carbon sink. *Glob Change Biol* 12(12):2370–2390. doi:10.1111/j.1365-2486.2006.01254.x



- Amthor J (2000) The McCree–de Wit–Penning de Vries–Thornley Respiration Paradigms: 30 years later. *Ann Bot* 86(1):1–20. doi:[10.1006/anbo.2000.1175](https://doi.org/10.1006/anbo.2000.1175)
- Asner GP, Martin RE (2009) Airborne spectranomics: mapping canopy chemical and taxonomic diversity in tropical forests. *Front Ecol Environ* 7(5):269–276. doi:[10.1890/070152](https://doi.org/10.1890/070152)
- Bernacchi CJ, Rosenthal DM, Pimentel C, Long SP, Farquhar GD (2009) Modeling the temperature dependence of C3 photosynthesis. In: Laisk A, Nedbal L, Govindjee (eds) *Photosynthesis in silico: understanding complexity from molecules to ecosystems*, vol 29. Springer, Netherlands, pp 231–246
- Berry JA, Beerling DJ, Franks PJ (2010) Stomata: key players in the earth system, past and present. *Curr Opin Plant Biol* 13(3):233–240. doi:[10.1016/j.pbi.2010.04.013](https://doi.org/10.1016/j.pbi.2010.04.013)
- Bonan B, Lawrence PJ, Oleson KW, Lewis S, Jung M, Reichstein M, Lawrence DM, Swenson SC (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):1–22. doi:[10.1029/2010JG001593](https://doi.org/10.1029/2010JG001593)
- Bonan GB, Oleson KW, Fisher RA, Lasslop G, Reichstein M (2012) Reconciling leaf physiological traits and canopy flux data use of the TRY and FLUXNET databases in the Community Land Model version 4. *J Geophys Res* 117(25C):1–19. doi:[10.1029/2011JG001913](https://doi.org/10.1029/2011JG001913)
- Booth BBB, Jones CD, Collins M, Totterdell IJ, Cox PM, Sitch S, Huntingford C et al (2012). High sensitivity of future global warming to land carbon cycle processes. *Environ Lett*, 024002. doi:[10.1088/1748-9326/7/2/024002](https://doi.org/10.1088/1748-9326/7/2/024002)
- Clark JS (2005) Why environmental scientists are becoming bayesians. *Ecol Lett* 8(1):2–14. doi:[10.1111/j.1461-0248.2004.00702.x](https://doi.org/10.1111/j.1461-0248.2004.00702.x)
- Clark JS, Dietze MC, Chakraborty S, Agarwal PK, Wolosin MS, Ibanez I, LaDeau S (2007) Resolving the biodiversity paradox. *Ecol Lett* 10(8):647–659. doi:[10.1111/j.1461-0248.2007.01041.x](https://doi.org/10.1111/j.1461-0248.2007.01041.x)
- Davidson CD (2012) The modeled effects of fire on carbon balance and vegetation abundance in Alaskan Tundra. University of Illinois, Urbana-Champaign, p 163
- De Kauwe MG et al (2013) Forest water use and water use efficiency at elevated CO<sub>2</sub>: a model-data intercomparison at two contrasting temperate forest FACE sites. *Glob Change Biol* (in press). doi:[10.1111/gcb.12164](https://doi.org/10.1111/gcb.12164)
- Dietze MC, Clark JS (2008) Changing the gap dynamics paradigm: vegetative regeneration control on forest response to disturbance. *Ecol Monogr* 78(3):331–347. doi:[10.1890/07-0271.1](https://doi.org/10.1890/07-0271.1)
- Dietze M et al (2011) Characterizing the performance of ecosystem models across time scales: A spectral analysis of the North American carbon program site-level synthesis. *J Geophys Res* 116:1–15. doi:[10.1029/2011JG001661](https://doi.org/10.1029/2011JG001661)
- Dietze MC, LeBauer D, Kooper R (2013) On improving the communication between models and data. *Plant Cell Environ* (in press). doi:[10.1111/pce.12043](https://doi.org/10.1111/pce.12043)
- Dios VR et al (2012) Endogenous circadian regulation of carbon dioxide exchange in terrestrial ecosystems. *Glob Change Biol* 18(6):1956–1970. doi:[10.1111/j.1365-2486.2012.02664.x](https://doi.org/10.1111/j.1365-2486.2012.02664.x)
- Dubois J-JB, Fiscus EL, Booker FL, Flowers MD, Reid CD (2007) Optimizing the statistical estimation of the parameters of the Farquhar-von Caemmerer-Berry model of photosynthesis. *New phytol* 176(2):402–414. doi:[10.1111/j.1469-8137.2007.02182.x](https://doi.org/10.1111/j.1469-8137.2007.02182.x)
- Ellis RJ (1979) The most abundant protein in the world. *Trends Biochem Sci* 4(11):241–244. doi:[10.1016/0968-0004\(79\)90212-3](https://doi.org/10.1016/0968-0004(79)90212-3)
- Farquhar G, Caemmerer S, Berry JA (1980) A biochemical model of photosynthetic CO<sub>2</sub> assimilation in leaves of C3 species. *Planta* 149(1):78–90
- Farquhar GD, von Caemmerer S, Berry JA (2001) Models of photosynthesis. *Plant physiol* 125(1):42–45
- Foster DR, Aber JD (2006) *Forests in time: the environmental consequences of 1000 years of change in New England*. Yale University Press, New Haven
- Foster D, Swanson F, Aber J, Burke I, Brokaw N, Tilman D, Knapp A (2003) The importance of land-use legacies to ecology and conservation. *Bioscience* 53(1):77. doi:[10.1641/0006-3568](https://doi.org/10.1641/0006-3568)
- Friend AD (2010) Terrestrial plant production and climate change. *J Exp Bot* 61(5):1293–1309. doi:[10.1093/jxb/erq019](https://doi.org/10.1093/jxb/erq019)
- Goll DS, Brovkin V, Parida BR, Reick CH, Kattge J, Reich PB, Van Bodegom PM et al (2012) Nutrient limitation reduces land carbon uptake in simulations with a model of combined carbon, nitrogen and phosphorus cycling. *Biogeosciences* 9(9):3547–3569. doi:[10.5194/bg-9-3547-2012](https://doi.org/10.5194/bg-9-3547-2012)
- Gross LJ, Kirschbaum MUF, Pearcy RW (1991) A dynamic model of photosynthesis in varying light taking account of stomatal conductance, C3-cycle intermediates, photorespiration and Rubisco activation. *Plant Cell Environ* 14:881–893. doi:[10.1111/j.1365-3040.1991.tb00957.x](https://doi.org/10.1111/j.1365-3040.1991.tb00957.x)
- Gu L, Pallardy SG, Tu K, Law BE, Wullschlegel SD (2010) Reliable estimation of biochemical parameters from C3 leaf photosynthesis-intercellular carbon dioxide response curves. *Plant Cell Environ* 33(11):1852–1874. doi:[10.1111/j.1365-3040.2010.02192.x](https://doi.org/10.1111/j.1365-3040.2010.02192.x)
- Harmer SL (2009) The circadian system in higher plants. *Annu Rev Plant Biol* 60:357–377. doi:[10.1146/annurev.arplant.043008.092054](https://doi.org/10.1146/annurev.arplant.043008.092054)
- Harpole W et al (2011) Nutrient co-limitation of primary producer communities. *Ecol Lett* 14:852–862. doi:[10.1111/j.1461-0248.2011.01651.x](https://doi.org/10.1111/j.1461-0248.2011.01651.x)
- Hennessey TL, Field CB (1991) Circadian rhythms in photosynthesis. *Plant Physiol* 96:831–836
- Hey T, Tansley S, Tolle K (eds) (2009) *The fourth paradigm: data-intensive scientific discovery*. Microsoft Research, Redmond
- Kattge J, Knorr W (2007) Temperature acclimation in a biochemical model of photosynthesis: a reanalysis of data from 36 species. *Plant Cell Environ* 30(9):1176–1190. doi:[10.1111/j.1365-3040.2007.01690.x](https://doi.org/10.1111/j.1365-3040.2007.01690.x)
- Kattge J, Knorr W, Raddatz T, Wirth C (2009) Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global-scale terrestrial biosphere models. *Glob Change Biol* 15(4):976–991. doi:[10.1111/j.1365-2486.2008.01744.x](https://doi.org/10.1111/j.1365-2486.2008.01744.x)
- Kattge J et al (2011) TRY—a global database of plant traits. *Glob Change Biol*. doi:[10.1111/j.1365-2486.2011.02451.x](https://doi.org/10.1111/j.1365-2486.2011.02451.x)
- Leakey ADB, Press MC, Scholes JD, Watling JR (2002) Relative enhancement of photosynthesis and growth at elevated CO<sub>2</sub> is greater under sunflecks than uniform irradiance in a tropical rain forest tree seedling. *Plant Cell Environ* 25(12):1701–1714. doi:[10.1046/j.1365-3040.2002.00944.x](https://doi.org/10.1046/j.1365-3040.2002.00944.x)
- Leakey AD, Bishop KA, Ainsworth EA (2012) A multi-biome gap in understanding of crop and ecosystem responses to elevated CO<sub>2</sub>. *Curr Opin Plant Biol* 15(3):228–236. doi:[10.1016/j.pbi.2012.01.009](https://doi.org/10.1016/j.pbi.2012.01.009)
- LeBauer DS, Wang D, Richter KT, Davidson CC, and Dietze MC (2012), Facilitating feedbacks between field measurements and ecosystem models. *Ecol Monogr* (in press). doi:[10.1890/12-0137.1](https://doi.org/10.1890/12-0137.1)
- Leuning R (1995) A critical appraisal of a combined stomatal-photosynthesis model for C3 plants. *Plant Cell Environ* 18(4):339–355. doi:[10.1111/j.1365-3040.1995.tb00370.x](https://doi.org/10.1111/j.1365-3040.1995.tb00370.x)
- Lin Y-S, Medlyn BE, Ellsworth DS (2012) Temperature responses of leaf net photosynthesis: the role of component processes. *Tree Physiol* 32(2):219–231. doi:[10.1093/treephys/tpr141](https://doi.org/10.1093/treephys/tpr141)
- 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. doi:[10.1093/jxb/erg262](https://doi.org/10.1093/jxb/erg262)

- Medlyn BE, Dreyer E, Ellsworth D, Forstreuter M, Harley PC, Kirschbaum MUF, Roux XLE (2002) Temperature response of parameters of a biochemically based model of photosynthesis. II. A review of experimental data. *Plant Cell Environ* 25: 1167–1179
- Medlyn BE, Duursma RA, Eamus D, Ellsworth DS, Prentice IC, Barton CVM, Crous KY, De Angelis P, Freeman M, Wingate L (2011) Reconciling the optimal and empirical approaches to modelling stomatal conductance. *Glob Change Biol* 17(6): 2134–2144. doi:[10.1111/j.1365-2486.2010.02375.x](https://doi.org/10.1111/j.1365-2486.2010.02375.x)
- Medvigy DM, Wofsy SC, Munger JW, Hollinger DY, Moorcroft PR (2009) Mechanistic scaling of ecosystem function and dynamics in space and time: ecosystem demography model version 2. *J Geophys Res* 114(G1):1–21. doi:[10.1029/2008JG000812](https://doi.org/10.1029/2008JG000812)
- Mohan JE, Clark JS, Schlesinger WH (2007) Long-term CO<sub>2</sub> enrichment of a forest ecosystem: implications for forest regeneration and succession. *Ecol Appl* 17(4):1198–1212
- Moorcroft PR, Hurtt GC, Pacala SW (2001) A method for scaling vegetation dynamics: the ecosystem demography model (ED). *Ecol Monogr* 71(4):557–586. doi:[10.1890/0012-9615\(2001\)071\[0557:AMFSVD\]2.0.CO;2](https://doi.org/10.1890/0012-9615(2001)071[0557:AMFSVD]2.0.CO;2)
- Naumburg E, Ellsworth DS (2002) Short-term light and leaf photosynthetic dynamics affect estimates of daily understory photosynthesis in four tree species. *Tree Physiol* 22:393–401
- Neale D, Aitken S, Dietze MC, Kliebenstein D, Mathews S, Oren R, Wegrzyn J, and Whetten R (2010) Tree biology cyber infrastructure, (online). [http://www.iplantcollaborative.org/sites/default/files/Tree\\_BiologyCI\\_seed\\_proposal\\_FINAL.pdf](http://www.iplantcollaborative.org/sites/default/files/Tree_BiologyCI_seed_proposal_FINAL.pdf). Accessed 10 Oct 2012
- Pacala SW, Canham CD, Saponara J, Silander JA Jr, Kobe RK, Ribbens E (1996) Forest models defined by field measurements: estimation, error analysis and dynamics. *Ecol Monogr* 66(1):1–43
- Patrick LD, Ogle K, Tissue DT (2009) A hierarchical Bayesian approach for estimation of photosynthetic parameters of C(3) plants. *Plant Cell Environ* 32(12):1695–1709. doi:[10.1111/j.1365-3040.2009.02029.x](https://doi.org/10.1111/j.1365-3040.2009.02029.x)
- Reich PB (1987) Quantifying plant response to ozone: a unifying theory. *Tree Physiol* 3(1):63–91
- Ricciuto DM, Thornton PE, Schaefer K, Cook RB, Davis KJ (2009) How uncertainty in gap-filled meteorological input forcing at eddy covariance sites impacts modeled carbon and energy flux. *Eos Trans Am Geophys Union* 90(52):B54A
- Riley WJ, Still CJ, Torn MS, Berry JA (2002) A mechanistic model of H<sub>2</sub>18O and C18OO fluxes between ecosystems and the atmosphere: model description and sensitivity analyses. *Glob Biogeochem Cycles* 16(4):1–14. doi:[10.1029/2002GB001878](https://doi.org/10.1029/2002GB001878)
- Schaefer KM et al (2012) A model-data comparison of gross primary productivity: results from the North American carbon program site synthesis. *J Geophys Res* 117:1–15. doi:[10.1029/2012JG001960](https://doi.org/10.1029/2012JG001960)
- Schwalm CR et al (2010), A model-data intercomparison of CO<sub>2</sub> exchange across North America: results from the North American carbon program site synthesis. *J Geophys Res* 115. doi:[10.1029/2009JG001229](https://doi.org/10.1029/2009JG001229)
- Sitch S, Cox PM, Collins WJ, Huntingford C (2007) Indirect radiative forcing of climate change through ozone effects on the land-carbon sink. *Nature* 448(7155):791–794. doi:[10.1038/nature06059](https://doi.org/10.1038/nature06059)
- Smith NG, Dukes JS (2013) Plant respiration and photosynthesis in global-scale models: incorporating acclimation to temperature and CO<sub>2</sub>. *Glob Change Biol* 19:45–63. doi:[10.1111/j.1365-2486.2012.02797.x](https://doi.org/10.1111/j.1365-2486.2012.02797.x)
- Starr G, Oberbauer SF (2003) Photosynthesis of arctic evergreens under snow: implications for tundra ecosystem carbon balance. *Ecology* 84(6):1415–1420. doi:[10.1890/02-3154](https://doi.org/10.1890/02-3154)
- Townsend AR, Asner GP, Cleveland CC (2008) The biogeochemical heterogeneity of tropical forests. *Trends Ecol Evol* 23(8): 424–431. doi:[10.1016/j.tree.2008.04.009](https://doi.org/10.1016/j.tree.2008.04.009)
- Wang YP, Law RM, Pak B (2010) A global model of carbon, nitrogen and phosphorus cycles for the terrestrial biosphere. *Biogeosciences* 7:2261–2282. doi:[10.5194/bg-7-2261-2010](https://doi.org/10.5194/bg-7-2261-2010)
- Wang D, Maughan MW, Sun J, Feng X, Miguez F, Lee D, Dietze MC (2012) Impact of nitrogen allocation on growth and photosynthesis of *Miscanthus* (*Miscanthus* × *giganteus*). *GCB Bioenergy* 4(6):688–697. doi:[10.1111/j.1757-1707.2012.01167.x](https://doi.org/10.1111/j.1757-1707.2012.01167.x)
- Way DA, Pearcy RW (2012) Sunflecks in trees and forests: from photosynthetic physiology to global change biology. *Tree Physiol* 32(9):1066–1081. doi:[10.1093/treephys/tps064](https://doi.org/10.1093/treephys/tps064)
- Williams W, Gorton H (1998) Circadian rhythms have insignificant effects on plant gas exchange under field conditions. *Physiol Plant* 103:247–256
- Wright IJ et al (2004) The worldwide leaf economics spectrum. *Nature* 428(6985):821–827. doi:[10.1038/nature02403](https://doi.org/10.1038/nature02403)
- Ziehn T, Kattge J, Knorr W, Scholze M (2011) Improving the predictability of global CO<sub>2</sub> assimilation rates under climate change. *Geophys Res Lett* 38:L10404. doi:[10.1029/2011GL047182](https://doi.org/10.1029/2011GL047182)