Title:

Influence and predictive capacity of climate anomalies on daily to decadal extremes in canopy photosynthesis

Author:
Ankur R. Desai
Associate Professor
Atmospheric and Oceanic Sciences
University of Wisconsin-Madison
Madison, WI 53706 USA
Tel: +1-608-265-9201
Fax: +1-608-262-0166
Email: desai@aos.wisc.edu

Abstract:
Significant advances have been made over the past decades in capabilities to simulate diurnal and seasonal variation of leaf-level and canopy-scale photosynthesis in temperate and boreal forests. However, long-term prediction of future forest productivity in a changing climate may be more dependent on how climate and biological anomalies influence extremes in interannual to decadal variability of canopy ecosystem carbon exchanges. These exchanges can differ markedly from leaf level responses, especially owing to the prevalence of long lags in nutrient and water cycling. Until recently, multiple long-term (10+ year) high temporal frequency (daily) observations of canopy exchange were not available to reliably assess this claim. An analysis of one of the longest running North American eddy covariance flux towers reveals that single climate variables do not adequately explain carbon exchange anomalies beyond the seasonal timescale. Daily to weekly lagged anomalies of photosynthesis positively autocorrelate with daily photosynthesis. This effect suggests a negative feedback in photosynthetic response to climate extremes, such as anomalies in evapotranspiration and maximum temperature. Moisture stress in the prior season did inhibit photosynthesis, but mechanisms are difficult to assess. A complex interplay of integrated and lagged productivity and moisture-limiting factors indicate a critical role of seasonal thresholds that limit growing season length and peak productivity. These results lead toward a new conceptual framework for improving earth system models with long-term flux tower observations.

Keywords: Eddy covariance; canopy photosynthesis; spectral analysis; carbon cycle

Introduction

Every year in modern times, photosynthetic organisms on land and in the ocean assimilate around 120,000,000,000,000 kg of carbon dioxide from the atmosphere, a process which drives the entire cycle of biosphere metabolism, production and decomposition (Beer *et al*., 2010). Variation of photosynthetic rates across space is strongly a function of adaptation of species to climatic, geological, and biological limiting factors of temperature, light, soil nutrients, moisture, disturbance, and competition. These adaptations are often manifested in differences in plant functional form, such as leaf shape, leaf longevity, tree heights, root depths, etc… Similarly, variation of photosynthesis in time is governed by how species in an ecosystem respond and adapt to diurnal, seasonal, and interannual changes in limiting factors.

Today, society faces a grand challenge as feedbacks between carbon dioxide uptake by photosynthetic organisms and the climate system are a leading source of uncertainty in the magnitude and severity of future climatic change, on the same order as scenarios of future anthropogenic emissions and aerosol or cloud feedbacks (Booth *et al*., 2012). Coupled carbon-climate models show a large range of future climate states depending on assumptions built into models about biospheric uptake, particular in the terrestrial biosphere (Friedlingstein *et al*., 2006). Extreme interannual anomalies in biospheric uptake have been linked to large-scale climate features like El Niño Southern Oscillation (ENSO), and recent increasing trends in the fraction of fossil fuel emissions that remain in the atmosphere point to troubling concerns about the state of the biospheric carbon sink (Le Quéré *et al*., 2009).

Quantifying these variations and improve predictive ecosystem models at the scale of regions to the globe over time periods of days to decades requires careful lab experimentation and long-term field observations (Moorcroft, 2006). Early experiments in the late 1970s and into the 1980s that included careful monitoring of leaf photosynthesis and isotopic discrimination in controlled environments along with theoretical thermodynamic and biochemical arguments led to the first successful representation of leaf photosynthesis through the simplified equations for C3 (and later C4) assimilation as reviewed in Farquhar and Sharkey (1982) and co-occurring development of leaf-atmosphere canopy conductance coupling as reviewed in Collatz *et al* (1991). While major advances have been made on understanding the biochemistry of photosynthesis at the genomic, cellular, and leaf level, many reported in this journal, most of these have not significantly altered these equations and similar formulations that are prevalent in most sophisticated ecosystem models (Schaefer *et al*., 2012).

The reason for the lack of more sophisticated leaf-level photosynthesis models is partly a question of computational resources in that models can not simulate every leaf in an ecosystem, let alone every cell. But a larger source of uncertainty rests in how one goes from the leaf-level model to an ecosystem patch or grid box. Early attempts focused on the issue of scaling of canopy radiative transfer, given that the variation of light through a canopy is the dominant mode of variability of limiting factors within an ecosystem patch to be simulated. Original models include the "big-leaf" representation of the average semi-transparent leaf (e.g., Sellers, 1985), partly owing its success to the ability to characterize vegetation fraction and photosynthesis through satellite remote sensing of visible and infrared canopy reflectance (Kumar and Montieth, 1981). However, field observations

noted that canopy radiative transfer may not necessarily scale so neatly (Baldocchi *et al* ,1985), leading to development of multiple canopy layer models (e.g., De Pury and Farquhar, 1997). At the minimum, sunlit/shaded fractions of the canopy have to be treated separately in models for accurate simulation of photosynthesis (Sprintsin *et al*., 2012). More sophisticated models now allow for multiple cohorts of interactively competing and shading species with varying plant functional types (Medvigy *et al*., 2009).

The primary production model in most ecosystem models now consists of a leaf-level photosynthesis mechanism, embedded within a leaf boundary-layer coupling, leaf energy balance model, canopy scaling algorithm, a soil water and humidity-sensitive transpiration model, and sometimes a nutrient transformation and transport model, primarily for nitrogen. Despite the apparent complete description of canopy photosynthesis, interactions and small changes in parameters of these components causes ecosystem models predict widely divergent estimates of the sensitivity of canopy photosynthesis to climatic and biotic changes, even when they're using the same equations (Schaefer *et al*., 2012). Differences in parameters that controls rates of leaf respiration, canopy architecture, or microclimate variation have large effects on canopy photosynthesis rates and sensitivity.

Uncertainties also arise in our understanding of variation in radiation quality and sun flecks, multiple interacting species, age-dependent changes in photosynthesis and transpiration, moisture and nitrogen availability in soil, transformation of assimilated carbon into storage pools, and canopy-scale stomatal and photosynthetic rate responses to atmospheric CO_2 enrichment (e.g., Fig. 1). Models tend to underestimate variability in canopy photosynthesis in response to climatic anomalies, but overestimate threshold

106 responses to climate and biospheric state shifts. On the one hand, ecosystems, as an
107 assemblage of species, and hence comprised of species that are adapted to have
108 compensating responses to environmental change are more conservative that the single
109 "model" species represented by the typical plant functional type approach in models. On
110 the other hand, overall ecosystem variance is large and since species present on the
111 landscape are partly a function of local climate and soils, they may show additive effects in
112 response to climate anomalies that exceed the a threshold (i.e., an extreme).
113 Consequently, there is a role for long-term canopy scale observations of
114 photosynthesis toward evaluating and improving these kinds of model responses. The rise
115 of canopy-scale observations of net ecosystem exchange of CO_2 (*NEE*) and inference of
116 gross primary productivity (*GPP*) through tower-based eddy covariance methods
117 (Baldocchi, 2008) and other canopy-scale experiments have expanded our ability to make
118 claims of model fidelity and sensitivity. For example, recent articles have focused on
119 dryness (Yi *et al*., 2009), temperature (Niu *et al*., 2012), light response (Schaefer *et al*.,
120 2012), phenology (Richardson *et al*., 2012), short-term climate fluctuations (Medvigy *et al*.,
121 2010), interannual variations (Keenan *et al*., 2012a), disturbance (Amiro *et al*., 2010), and
122 CO_2 enrichment (Norby and Zak, 2011). Still, many of these studies are limited by short
123 time series at most sites. Advanced methods to link anomalies of environmental drivers to
124 fluxes have only been applied in limited cases.
125 In this manuscript, I focus on the value of longer-term observations of high-
126 frequency photosynthetic flux observations at the canopy scale. Are there effects of lagged
127 environmental drivers that are masked at shorter time scales? Can ecosystem adaptation
128 and community reorganization responses from extremes be detected? Hypotheses have

been put forward suggesting that canopy-scale productivity may be linked to prior season carbon storage of non-structural or labile carbohydrates (Carbone *et al*., 2007) and also antecedent moisture availability in mesic forests (Ricciuto *et al*., 2008; Desai *et al*., 2010). These lagged responses would not be seen at the leaf level, where direct controls of moisture availability and leaf carbon content influence photosynthesis strongly by limiting rates of light harvesting and total leaf area.

The purpose of the analysis here is not to put forth brand new models of canopy photosynthesis, but rather to highlight the path forward with long-term observations. As such, my goals are twofold: 1) to demonstrate advanced statistical methods to evaluate modes of variation of long-term environmental data and 2) observe how these methods provide new insight into lags and switches of canopy photosynthesis that make it so hard to model and so different from leaf-level responses. In particular, I use a 15-year record of regional NEE from a very tall tower in the north central US to test which of antecedent soil moisture availability and prior productivity most influence canopy productivity. Moisture lags likely represent the accumulation and storage of available water that allows a canopy to respond to periods of high water demand or physiological stress. Similarly, prior carbon accumulation influences current carbon uptake both in changes in allocation and development of carbohydrate reserves. Can these effects be seen in observations and if so, which is the most relevant to include in models?

I hypothesize that short-term (daily) drivers of photosynthesis are primarily light and temperature, but longer-term (weekly to annual) is limited primarily by moisture and internal storage of prior photosynthate. These hypotheses are tested against the landscape-scale observations of net carbon uptake and associated surface and meteorological forcing.

Methods

Analysis Framework

Many of the analyses described in the prior section attempted to use many short-term (2-5 year) ecosystem flux observations to improve prediction or modeling by substituting space for time (e.g., Yi *et al*., 2010). Many sites each with a few years of data are analyzed to make inferences about long-term evolution of biosphere to climate and environmental drivers. However, there is some evidence that predictions made in this form across sites do not necessarily map well onto long-term predictions at a single site (Desai, 2010; Keenan *et al*., 2012a). The limiting factor is the lack of long-term high-frequency observation of the state of the biosphere.

Short-term measurements are difficult to use for diagnosing anomalies and extremes. Short-term multi-site studies can diagnose mean state and mean variability of carbon fluxes, but may underestimate the true level of variability over years and how extremes in climate and biotic disturbance (including both large short-term pulse and long-term press (steady pressure) disturbances) can drive carbon assimilation differently than short-term responses. For example, many eddy covariance flux tower studies focus on a few years of data to identify particular climate responses (e.g., wet year versus dry year), but are likely to have confounded co-variability among climate factors, to miss multi-year responses, or and have low probability of capturing extreme climate events and the role of pre-conditioning of ecosystem states. Community-scale response from changes in resource availability and competitive advantages to these can occur in forest ecosystems at timescales of years to decades (Gellesch et al., 2013).

While some processes like photosynthetic acclimation have been well captured in lab experiments, they are harder to diagnose with short-term environmental data, even with a decade long record from enriched CO_2 experiments (Norby and Zak, 2011). The statistical sample for environmental observations of low probability, high impact events is too small.

A benefit of an evolving measurement network is that over time there are sites that start having long records where one can look closely at features like memory effects (long lag relationships), decadal trends, and state shifts that would not be easily noted across space. Instead, the challenge is addressing the data deluge. A decade long flux tower record of half-hourly NEE observations and related climate drivers can easily exceed 10^6 observations. Moreover, the data are strongly auto-correlated and may suffer from harmonization issues related to changes in instruments, measurement height, and so forth. As a consequence, many analyses, even at long-term tower sites, limit their analyses to subsets of the data. For example, Keenan *et al* (2012b) found no suitable combination of parameters of a simple model could adequately explain three separate five-year periods in NEE observed over the 18-year record at Harvard forest.

There is ongoing work on improving harmonization of long-term datasets like decadal eddy covariance and the evolving National Ecological Observatory Network (NEON), which will include nearly 60 sites across North America with eventually 30+ years of carbon cycle and biological observations. These observations provide a suitable data testbed if and only if the community first develops reliable and usable statistical metrics and model-data evaluation. Therefore, in this study, I specifically focus on the more than

196 15-year record of eddy covariance carbon and water regional flux observations at a
197 forested site in the north central US (Figs. 2 and 3).

198 ## Site Description and Data

199 I analyzed 15-years of flux tower observations of CO_2 and H_2O flux from one of the
200 longest continuously running eddy covariance flux towers in the U.S., the WLEF Park Falls
201 tower (US-PFa) (Davis *et al*., 2003), where fluxes have been measured since late 1996 with
202 minimal disruptions, except in 2002 (Fig. 2). Meteorological variables were also observed
203 at the site (Fig. 3 and Table 1). WLEF is unique for being the tallest flux tower across the
204 Fluxnet network, allowing us to observe the impact of patchy landscapes and canopy
205 interactions on carbon assimilation. My collaborators and I have observed fluxes at three
206 heights (30 m, 122m, 396 m) and use these to develop a single "preferred" flux product
207 (Davis *et al*., 2003), based on boundary layer turbulence conditions. The tower samples a
208 fetch on the order of 1-5 km depending on atmospheric stability and wind speed.

209 Unlike canopy-scale towers, tall-towers sample fluxes that represent many species
210 and many soil types. However, an advantage of these observations is they are at a similar
211 scale to that which ecosystem models represent canopies and plant functional types (10s to
212 100s of km). Schafer *et al.* (2012) noted that ecosystem model estimates of daily GPP were
213 surprisingly well simulated at this site in a large flux tower-model intercomparison of GPP,
214 either because modelers have used this site significantly for calibration, or, that the fluxes
215 better represent the "model organism" being represented by the single plant functional
216 types used in most models.

217 The site samples carbon and water fluxes from a temperate mixed forest landscape
218 that consists of approximately ¾ forest equal parts young to intermediate age

commercially harvested aspen, mature northern hardwood (sugar maple, ash, basswood), and red pine plantations (Desai *et al*., 2007). The remaining ¼ is primarily a mosaic of wetlands and shrub areas, including black spruce and peat bogs, cedar swamps, sedge wetlands, and shrub fens. Spatial variability occurs in relatively small scales, driven by microtopography and land management, while the overall landscape topography is flat and density of human settlement in the tower footprint is minimal.

Estimating Canopy-Scale Photosynthesis

Eddy covariance towers observe the *net* exchange of trace gases, heat, and momentum from the surface to atmosphere, based on well-established micrometeorlogical theory (Baldocchi, 2008). Turbulence properties of the atmospheric surface layer allows one to take the 30-60 minute mean covariance of high-frequency (>10 Hz) observations of vertical wind and the flux tracer of interest (e.g., carbon dioxide, water, temperature) summed with below-sensor net tracer storage and vertical flux divergence to represent the net surface flux. Sonic anemometry (measuring vertical and horizontal wind components with sound pulses) and infrared gas analyzers sampling air near the anemometer are typically used to measure this net covariance. Contributions from low-frequency transport (advection) are usually neglected, but tend to be small, of the same magnitude as the 10-20% inherent random flux error (Yi *et al*., 2000). Over the years, researchers have instrumented nearly 500 of these sites for carbon and water cycle observations and general quality control approaches have been identified for instrument noise, lag, and spectral corrections, coordinate geometry rotation for wind velocity, low turbulence screening, and other turbulence statistics, which are applied here (Berger *et al*., 2001; Foken *et al*., 2012).

241 The focus of this analysis of the effect of climate anomalies on photosynthesis, not
242 net exchange (which includes respiration and decomposition processes). Therefore, I
243 developed a method to represent this photosynthesis from net ecosystem exchange of CO_2
244 (NEE). Unfortunately, there is no single accepted method for doing so, and all require some
245 level of empirical assumptions or statistical inference that partly takes advantage of the
246 lack of GPP at night. Consequently, methods diverge on estimates of GPP by more than 20%
247 and can include artifacts from fitting NEE to respiration models (Desai *et al*., 2008).
248 Since I want to focus on the value of NEE to models, I developed an alternate metric
249 of the effect of canopy photosynthesis on NEE, termed net photosynthetic drawdown (P_d), a
250 daily metric of canopy photosynthesis that removes assumptions used in many GPP models.
251 P_d was estimated at a daily timescale from the hourly flux data as the difference in
252 nighttime to daytime NEE. Maximum nighttime NEE was identified at night when more
253 than four hours of good observations were available. Maximum is used over mean since it
254 has been shown to be closer to the advection corrected observations at night (Van Gorsel *et*
255 *al*., 2009). This estimate of nighttime NEE is then differenced with the mean daytime gap-
256 filled NEE between 10 and 14 local time if there are more than four hours of good
257 observations during that day (when the sun is up). Here I use gap-filled NEE to avoid
258 biasing the mean NEE, which exhibits a strong diurnal cycle. Gap-filling errors tend to be
259 much smaller than GPP uncertainty (Moffat *et al*., 2007). The P_d time series is shown in Fig.
260 4a. The P_d time series has 5,490 days of data, with 37% of data missing.
261 Further analysis showed that use of *Pd* instead of GPP does not significantly change
262 the results or conclusions of this study and presents a novel way to understand the effect of
263 climate on photosynthesis. The correlation of P_d to *GPP* is high, particularly for maximum

daily GPP (r^2=0.81) and greater at the monthly timescale (r^2=0.96). The fit is linear for GPP, with an intercept of 0 (Fig. 5). Since P_d is a detector of maximum daily photosynthetic uptake and has a greater dynamic range than GPP, it is likely that P_d is better at detecting extreme photosynthesis responses to climate anomalies. While this method is conceptually analogue to atmospheric CO2 "drawdown" (e.g., Desai *et al.,* 2010), it is different as the flux drawdown does not include covariation with boundary layer depth and represents a much smaller footprint.

Statistical Analysis

I tested the hypotheses mentioned above by testing for both direct and lagged relationships between P_d and climate forcing factors (Table 1) at multiple time scales and compared them to the autocorrelation of P_d. A number of studies have identified characteristic timescales of variability in flux data using wavelet, single spectrum, or Fourier time-series analysis (e.g., Baldocchi *et al* ., 2001; Mahecha *et al*., 2007; Sevanta and Williams, 2009; Stoy *et al*., 2009), which have all noted characteristic peaks of variability in NEE especially at the diurnal, synoptic (3-4 day), seasonal, and interannual timescale. Similarly, frequency dependent model-data comparisons (e.g., Dietze *et al*., 2012; Mahecha *et al*., 2010; Keenan *et al*., 2012a) have all found deficiencies of models in representing many of these modes of variability.

I identified these scales in daily P_d and evapotranspiration (*ET*) flux using a similar analysis of empirical model decomposition (EMD), whose results are fed into the Hilbert-Huang spectral transformation (HHT) (Huang and Wu, 2008). EMD is an empirical approach to time series deconvolution that does not require assumptions of cyclical behavior (as needed by Fourier) or stationarity and does not require determination of the

shape of the weighting kernel or wavelet. The discontinuous EMD (Barnhart *et al*., 2012) further extends the application to time series with missing data by applying a mirroring approach to fill the data gaps. EMD decomposes a time series into a series of intrinsic mode functions (IMF) also in the time dimension, which when fed to the HHT algorithm that outputs a time by frequency power spectrum.

Timescales for analysis were determined from the HHT of P_d and ET (Fig. 4). Both signals have a number of similar modes of variability, especially at the synoptic, monthly, and seasonal timescale. ET has greater temporal variations in these modes and greater signal on long-time scales (> 100 days). Longer timescale variability is present in the growing season more than outside of it. HHT identified strong monthly peaks that were not previously identified and suggests that interannual variability explains less of the signal than other methods have previously shown (e.g., Baldocchi *et al*., 2001). Methodologically, I used this analysis select averaging timescales of 1, 3, 8, 15, 30, 90, 180, 360, 720, and 1440 day, as described next.

The HHT analysis also identified the importance of normalizing variability across timescale to best identify climatic and internal controls on P_d. For example, there is a variety of literature that will show high correlation of GPP to other variables, solely because the main modes of variability (e.g., the annual solar cycle) are strong in both, not because one truly explains the other. This method of analysis is disingenuous when it comes to the question I seek to answer here.

One of the benefits of long-term data is the ability to remove much of this co-variability and look at how anomalies or extremes manifest themselves in the data and how they are correlated to anomalies or extremes in another variable. If a daily time series

signal *X(day,year)* is stationary (as it mostly appears to be in this case and discussed more in the discussion), then an anomaly time series $X_{an}(day,year)$ can be simply constructed by removing the ensemble mean:

$$X_{an}(day,year) = X(day,year) \quad \overline{X(day)}\Big|_{year} \tag{1}$$

where $X(day)|_Y$ is the daily time series of variable *X* ensemble averaged across all years. To test across multiple timescales, I applied a forward averaging filter across the time series (e.g., Fig. 6b), avoiding forecasting by removing the end of the data series:

$$X_{an}(day,year)\Big|_{timescale} = \frac{1}{timescale} \sum_{t=day}^{t=day+timescale} X_{an}(t,year) \tag{2}$$

where timescale is the number of days to average. One issue that arises when analyzing this variable across seasonal to interannual timescales is the need for averages to stay aligned with the solar orbital forcing cycle, so that averages in any one year can comparable to other years. To do this, I reduced each year to 360 day length by removing the first few and last few days of data for each year and choosing averaging scales which share divisors with 360. The choice of timing has relatively minimal effect and the choice of winter, where carbon fluxes are near zero is ideal. A second issue involves gaps in the data. For gaps, I sampled the data with replacement, filling gaps linearly across small gaps (days) and taking long-term means for longer gaps (weeks).

To remove the previously discussed solar forcing driven artificial correlation among variables, I normalized the time series. Flux anomalies (e.g., Fig. 6c) display a strong seasonality given the change in variability from winter to summer (Fig. 3). Relative anomalies $X_{rel,an}(day,year)|_{timescale}$ were derived from averaged anomalies by dividing the

time series by the ensemble average standard deviation across all years for a given averaging timescale:

$$X_{rel,an}(day,year)\big|_{timescale} = \frac{X_{an}(day,year)\big|_{timescale}}{\left(X_{an}(day,year)\big|_{timescale}\right)\big|_{year}} \quad (3)$$

The remaining time series appears stationary and random (Fig. 6c), and reflects a statistically defensible view of relative anomalies of the time series as a function of averaging filter. Relative anomalies in this fashion were computed for P_d (Fig. 6) and a variety of observations to test hypotheses including variables related to canopy physiology and structure such as remotely sensed vegetation index (*EVI*) and minimum, maximum, mean, and diurnal range of air temperature, remotely sensed land surface temperature (*LST*), and variables related to canopy moisture availability including *ET*, water use efficiency (*WUE*, *GPP* divided by *ET*), precipitation, and soil moisture, as noted in Table 1 and shown in Fig. 3. Remotely sensed variables were derived from the NASA MODIS TERRA and AQUA reflectance properties and downloaded from the ORNL MODIS land product subset server (http://daac.ornl.gov/MODIS/), while other variables were directly observed by the tower with gaps filled from harmonized daily climate data downloaded from the National Climatic Data Center archive of National Weather Service co-operative observer stations and weather forecast reanalysis from the NOAA North American Regional Reanalysis (NARR).

I compared relative anomalies of all variables to relative anomalies of P_d at all averaging timescales both with direct linear correlation and with lagged correlation where climate factors were lagged against P_d at a range of lags equivalent to the averaging

timescales. A two-tailed t-test for significance was applied to all correlation coefficients and only those coefficients that were significant at the 90% level were saved.

The significance test was modified to account for the autocorrelation present in all environmental time series. Consequently, the degrees of freedom to apply to significance tests should be much smaller than the total number of samples. I reduced the degrees of freedom using a modified effective degrees of freedom (EDOF) approach of Bretherton *et al*. (1999):

$$N_* = \frac{N}{\sum_{t=N/2}^{N}\left[\left(1 - \frac{t}{N}\right)\ {}^{X}_{t}\ {}^{Y}_{t}\right]} \tag{4}$$

where N_* is the reduced degrees of freedom for significance testing of correlation of two time series X and Y with N samples. ${}^{X}_{t}$ represents the autocorrelation of time series X at lag t. Though most of the autocorrelation is in the first few lags, I included all lags to $N/2$ to account for long lead correlations. Further, only those correlations of variables to P_d that exceeded the lagged autocorrelation of P_d are used as a test to compare moisture versus carbon storage control as predictors of current P_d.

The EDOF of P_d (Fig. 7) reveals that while daily P_d has over 3400 observations, the EDOF at daily scale is only slightly above 600, and decreases nearly linearly with logarithmic increases in averaging timescale, such that interannual analysis is limited to EDOF in the few tens to single digits. As shown in the results, this limits the ability of this analysis to diagnose correlations of anomalies at multi-year timescales and highlights how even a 15 year time series may be unreliable for detecting interannual and longer trends and correlations. The results here shows that a few time series were able to meet

significance threshold, but multi-year correlation analysis with flux tower data requires sufficiently long data sets because of high auto-correlation.

A second test was also applied for comparing predictive ability of variables against inherent autocorrelation of P_d, known as the Granger causality analysis (Detto *et al.*, 2012). The method originated from economics, but has recently gained popularity in geophysical time series analysis. The analysis builds a multiple linear regression of lagged values of a time series *X* to predict current values of *X*. This regression is then compared iteratively to including an increasing number of lagged values of time series *Y* that are significantly (two tailed t-test at 90% level) correlated to *X* to predict current values of *X*. When the new regression significantly improves upon (tested with an F-test) the autocorrelation regression, those lags of *Y* are retained and the terminology is that *Y* "Granger causes" *X*, or prior values of *X* and *Y* at certain lags explains a significantly larger fraction of *X* than the prior values of *X* does by itself. A minor modification was made here to include the reduced degrees of freedom for significance testing and the replacement of the F-test with the more empirical Aikake Information Criterion, which incorporates both the likelihood of the regression and penalties for number of parameters. Here, I tested Granger causality for all variables against P_d (Table 2).

Though all of these methods are relatively standard time series analysis, they have only recently been applied to carbon flux data, mostly because long time series of flux data are only now becoming common. The results show how the lagged correlation, spectral, and causality analysis together provide insight on how canopy photosynthesis is different from leaf photosynthesis and how it can be leveraged to improve canopy photosynthesis models.

Results

Modes of variability in observations

Large variability exists in hourly flux data of *NEE* and *ET* (Fig. 2). Outliers exist in most years, but positive anomalies in *ET* and *NEE* decreased after 2005. Diurnal and seasonal variability dominate the signal and trends. The decrease in *NEE* uptake from 2006-2010 is visually evident, but difficult to discern quantitatively against the variability at the hourly scale (Desai *et al*., 2010). Not surprising, hourly data can be challenging when it comes to statistical data assimilation approaches to constraining ecosystem models (Zobitz *et al*., 2011), as these methods tend to force model parameters toward the dominant modes of variability (diurnal and seasonal) and limit excursions away from the mean (over-fitting the mean at the expense of the extremes) (Desai, 2010) especially when data uncertainty is large, as it is the case for hourly flux observations (Raupach *et al*., 2005). Not surprisingly, models have great difficulty simulating other temporal modes of variability in photosynthesis (Dietze *et al*., 2012; Keenan *et al*., 2012a). Moreover, many of the anomalies of photosynthesis that are related to climate anomalies may exist only at longer timescales.

The EMD analysis directly shows the importance of weekly to seasonal variability in P_d and *ET* (Fig. 3), counter to previous wavelet based analyses. For example, Baldocchi *et al*. (2001) found a spectral gap in flux tower *NEE* at the three-four week scale. The transformed data and the more empirical approach of HHT reveal that there are variations present at this scale, perhaps more so than at longer scales, and not consistently at a fixed value. Longer-term variability (90-360 day) is more present in the growing season

418 (summer) and more distributed across frequency for ET compared to P_d, suggesting more
419 coherence in interannual variability of ET than P_d.

420 ## Direct Relationships

421 Given the range of variability in flux tower observations and a general overlap of
422 modes of variability in ET and P_d anomalies at the daily to seasonal timescales, it should be
423 expected that some level of correlation exists among these factors. The strongest direct
424 linear significant correlations between ET and P_d anomalies exist at scales between 30-360
425 days, but persist down to 1 day, where ET explains around 10% of the variability in P_d at
426 short timescales and approaches 40% at seasonal scales (Fig. 8). Longer-term correlations
427 are not significant, but this may partly be the result of insufficient length of time series and
428 the strict degree of freedom constraint. A similar level correlation is seen for *WUE*, but not
429 at annual time scales. This correlation partly stems from self-correlation since *GPP*, which
430 goes into the *WUE* calculation, and P_d are both derived from *NEE*.

431 Temperature is a well-known factor influencing leaf-level photosynthesis. Even
432 though solar forcing leads to a strong correlation between temperature and *NEE* or *GPP* at
433 diurnal and seasonal scales, when the ensemble average is removed and the variables
434 compared as standardized anomalies, the relationship is much weaker, though it remains
435 significant from 1 day out to 30 day, with r around 0.15. None of the temperature factors
436 (average, maximum, minimum, mean, range, or land surface temperature) are particularly
437 better than others at explaining variation in P_d, though stronger correlations exist for
438 maximum daily temperature especially at longer timescales, including a particular strong
439 correlation at the 1440 day scale. It should be noted that *LST* is provided at 8-day intervals,

so 1 and 3 day average correlations are not included. For the purpose of the lag analysis, given similar relationships, only T_{mean} is further analyzed.

For direct correlation, the results are weaker for soil moisture than for temperature, but do provide some evidence to the importance of seasonal moisture budgets on net carbon assimilation, even in mesic forest/wetland systems. This result less consistent with an earlier model calibration study that showed interannual variability in *NEE* at WLEF was best explained by soil moisture (Ricciuto *et al*., 2008), but when lags are taken into account, the results change, as shown below. Though not shown, results with atmospheric humidity variables were similar to those of Q_{soil}. Precipitation anomalies only weakly P_d at 3-day averages. This result may be caused by the intermittent nature of precipitation, the greater error in short term precipitation, and the importance of soil percolation processes prior to plant water uptake. Finally, *EVI* anomalies are found to have no correlation to P_d anomalies at any timescale.

Lagged analysis

Lagged analysis (Fig. 9 and Fig. 10) further supports results shown in Fig. 8, but also reveals subtleties regarding moisture. The non-moisture specific variables support both a strong autocorrelation of P_d as previously shown in Fig. 6 and lack of correlation at many timescales for *EVI* anomalies. Lagged *EVI* (Fig. 9b) does have some weak, but significant negative correlation to P_d for monthly averaging timescales and three month lag, and this correlation exceeds the autocorrelation. This signal represents both the impact of summer vegetation stress (low EVI) on autumn photosynthesis (reduced) (Wu et al., in press) and the impact of phenology (late spring = low EVI) on net carbon uptake (reduced). Spring flush is typically a two-week process and the growing season is around three months.

463 The autocorrelation of P_d (Fig. 9a) is persistent at AR-1 and AR-2 out to 30 days. For
464 any given lag, some amount of autocorrelation exists for all averaging times up to the lag,
465 increasing as the lag approaches the averaging timescale (i.e., AR-1). Some out of phase
466 (180 day) negative correlation exists as it also does at the very long timescale (1440 day),
467 the latter of which is difficult to explain. At one level, ecosystem models incorporate this
468 autocorrelation through the "memory effect" of labile carbon pools, but these are probably
469 not responsible for the observed longer timescale correlations, possibly tied to other non-
470 structural carbohydrate pools (e.g., Carbone *et al*., 2007) or signals of community
471 reorganization, ecosystem dynamics, or climate oscillations. However, the lack of long
472 autocorrelations in P_d do not provide much evidence here for strong long-term internal
473 control through non-structural carbohydrates, shifts in plant allocation, or community
474 reorganization.
475 Lagged temperature is predictive for P_d anomalies at short lags and averaging times
476 out to a few weeks (Fig. 9c and d), with a positive correlation (warm anomalies lead to
477 increased P_d one to two weeks out), but these are much weaker than the autocorrelation.
478 An interesting weak, but stronger than autocorrelation, relationship exists with 90-day
479 prior daily mean temperature and current P_d, hinting at possible long-lag effects, where a
480 short early season warm spell (e.g., false spring), can enhance growth. There is also a weak
481 but significant negative relationship of seasonal average temperature from four years past,
482 once again difficult to explain.
483 Moisture anomalies (*ET*, *WUE*, P_{recip}, and Q_{soil}) also have some predicative ability for
484 P_d anomalies and most all positive – increased *ET* or precipitation enhances carbon
485 assimilation. Correlations of lagged *ET* to P_d (Fig. 9a) exist at short timescales and the first

two lags for each timescale, but they are surprisingly weak compared to the strong direct correlation (Fig. 8). Similar to temperature, a four-year lag negative correlation of *ET* on *Pd* is found for seasonal averaging.

Relationships also exist in weekly to seasonal average precipitation and soil moisture at the 2-3 month lag. Positive anomalies in soil moisture are more predictive for future weekly to seasonal P_d than P_d itself, suggesting a long-term moisture control. For example, early season weekly to seasonal moisture deficits inhibit end of season carbon assimilation. This effect is of slightly greater magnitude and correlation as the effect of direct moisture deficits on P_d. Unlike the direct correlation, these results more strongly support both the work of Ricciuto *et al*. (2008) and the second hypothesis of long-term moisture control on P_d.

Causality analysis

The findings of the Granger causality analysis (Table 2) are consistent with the lag analyses for most except the moisture variables. Daily to weekly temperature and ET both Granger cause P_d, with additional longer term control by seasonal EVI and annual ET. Interestingly, Q_{soil} does not Granger cause P_d at any timescale. Only ET anomalies have predictive ability at long time scales and highlights the difficulty of dissecting the causes of seasonal to interannual variability even with 15 years of flux tower data.

Discussion

Role of lagged forcing in photosynthesis

Internal control (carbon assimilation rates related to prior carbon assimilation rates) is a strong predictor on canopy carbon assimilation at timescales up to a week, while the key direct climatic modifier of this are temperature and available moisture, but

509 primarily on longer timescales and longer lag times. Clearly, this differs from the primarily
510 direct relationships one finds for leaf photosynthesis controls and highlights the
511 complexity of modeling canopy photosynthesis and the value of long-term data.
512 The analysis here is unable to directly identify mechanisms. I assumed that a strong
513 autocorrelation at long time scales implies carbohydrate storage or other mechanisms of
514 buffering that limits ecosystem response to climate extremes. However, the evidence that
515 such occurs here is weak and instead long-lead short-term moisture stress and prior
516 season temperature anomalies appear to have the strongest effects, suggesting that
517 moisture control is stronger than expected. A number of biotic interactions are likely in
518 response to climate anomalies, in addition to changes in internal storage of sugars and
519 starches, there are possible shifts in allocation in response to extremes or aging or
520 community reorganization from shifts in competitive pressure. There was very limited
521 management in the region over the study period and no evidence for a shift in age structure
522 or dominance of certain ecosystems (Gellesch et al., 2013; Scheffer et al., 2001), which
523 allowed this analysis to assume stationary conditions. Additionally, it is unclear from the
524 analysis if any state shifts from multiple or repeating stress, dubbed "ecological stress
525 memory" (Walter et al., 2013), was observed here. Instead, the analysis suggests that
526 modeling and experimental studies should look and evaluate carbon cycle shifts over long
527 time scales (seasons to years) in response to relatively short-term drought manipulation or
528 prior season temperature anomalies. Further, multi-year anomalies that may be related to
529 patterns and oscillations in biology or climate warrant more investigation. For example, the
530 negative autocorrelation of P_d at interannual timescales for annual and longer averages

suggests potential biological oscillations, cycles of herbivory, species successional processes.

Further evidence on the lack of a strong negative feedback or "internal control" is the surprising lack of correlation between relative anomalies of P_d to *EVI*. Anomalies of P_d do not appear to relate to anomalies of *EVI* at any timescale, calling into question how well remote sensing can be used to evaluate how climate anomalies drive productivity anomalies. Many applications have been developed around the ability to apply differences in infrared and visible reflectance of canopies to estimate global photosynthesis from space, ever since early work showed the strong link of absorbed radiation to plant carbon assimilation (e.g., Kumar and Monteith, 1981). For example, remarkably strong monthly to seasonal correlations exist between NASA MODIS derived monthly *GPP* against flux tower estimated *GPP* (Heinsch *et al*., 2006). However, many of these papers find that while satellites can sense large-scale latitudinal variation, significant unexplained variability exists across smaller regions and across longer timescales.

It appears that *EVI* and similar metrics of remotely sensed vegetation greenness or absorbed radiation capture processes like phenology, leaf area, or canopy development, though they likely do not readily capture the anomalies or extremes as formulated in this analysis. Though anomalies do not correlate, *EVI* does explain approximately 75% of the biweekly variation of P_d. EVI has also been argued as a good proxy for carbon uptake phenology, but at this site, dates of start and end of carbon uptake period (period when mean daily smoother P_d is positive) do not correlate strongly to dates of start and end of the "greenness" period as identified in EVI. However, I did find that growing season length as defined by its carbon uptake period has a strong correlation with average growing

554 season EVI (r=-0.88), though with a negative relationship and a small effect size, suggesting
555 that short term and long-term EVI have opposing effects and may partially explain the lack
556 of correlation of anomalies. The strongest relationship with mean growing season *EVI* is
557 with correlation of the end date of this uptake period (r=-0.92), consistent with some
558 recent work that many temperate forest systems have interannual variability in NEE driven
559 by end of season signals (Wu *et al.*, in press). The analysis here does suggest caution is
560 warranted when analyzing anomalies in broadband satellite vegetation indices.

561 Finally, it is apparent that both at the direct timescale, through the high correlation
562 of *ET* to P_d and at the lagged timescale, through the positive association of seasonal soil
563 moisture to P_d and the long lag relationships of temperature to P_d, all imply a variety of
564 moisture retention and moisture use processes influence photosynthesis at a number of
565 timescales. Some of these maybe related to summer droughts influence late season
566 photosynthesis and others may be related the dynamics of the snowpack on soil moisture.
567 The existence of moisture control on plant biogeochemistry in a mesic temperate forest
568 and wetland landscape in and of itself is surprising and opens up a number of new avenues
569 for analysis.

570 **Towards a canopy photosynthesis modeling framework**

571 Compared to lab experiments, both uncontrolled and controlled field observations
572 require greater explicit consideration of time and spatial scale, and the extent to which
573 variability expressed in one dimension truly reflects the signal one seeks to estimate. This
574 paper, like others (e.g., Stoy *et al.*, 2009), demonstrate that frequency dependent analysis is
575 essential for identifying processes over long-time periods or large regions. Otherwise,

576 conclusions can be drawn from short-term or small-scale data that have very limited
577 application to how photosynthesis actually responds to the environment.
578 Statistical analyses for large environmental data sets are still in development.
579 Modern computational speeds, open source libraries for advanced programming languages,
580 and new models of graduate student training have led to continued improvement in these
581 (Zobitz et al., 2011). Of course, whether the methods presented here are useful ultimately
582 depend on the interpretation of results.
583 Moorcroft *et al* (2006) asked if we have reached a predictive ability for the
584 biosphere. Progress has been made, especially with advanced coupled dynamic vegetation
585 and carbon cycling models (e.g., Medvigy *et al*., 2009), but the community may have
586 reached a standstill until we seriously reconsider how we confront models with data.
587 Dietze *et al* (2012) found among more than a dozen ecosystem models, very little
588 confidence present in our ability to simulate both the diurnal cycle and interannual
589 variability, with the latter finding confirmed by Keenan *et al* (2012). New research further
590 finds ever short-term environmental variability can strongly affect long-term carbon
591 cycling (Medvigy *et al*., 2010).
592 The analysis here suggests that models need to be evaluated on the temporal
593 memory of moisture and carbon storage mechanisms. Advances have been made in
594 applying data assimilation or Bayesian inference methods to sift through data and models
595 (Williams *et al*., 2009). Large model-data syntheses as mentioned in the introduction have
596 contributed to our ability to diagnose consistent model errors. Data uncertainty, machine
597 readability, and archival have also gotten greater attention. Uncertainty, in particular, is
598 essential to collect with all these data, given how sensitive model-data comparisons can be

to them (Raupach *et al*., 2005). Finally, recent progress has been made on making modeling and model-data comparisons a routine exercise, or at least, more user-friendly and across a wider range of data sources (LeBauer *et al*., in press).

Multiple lag and time filtering should applied to climate extreme experiments conducted with ecosystem models in the soil moisture and carbon storage domains. Various model structures and parameters (e.g., root exudates, labile carbohydrate storage, community shifts, or soil moisture storage rates) can be investigated not for merely how well they simulate *NEE* or even P_d, but rather how well they simulate the observed relationships among variables across time. Some of these responses span over multiple years. Further, comparisons should be made in anomaly space if we really want to test how climate extremes influence photosynthesis.

Conclusion

I found that neither the carbon storage control or moisture control hypotheses could be falsified with long-term data, once seasonal cycle was removed. The extent to which the hypotheses could be falsified was strongly dependent on the scale of the averaging filter and the lags analyzed. Using Hilbert spectra to identify relevant lags, I found a short-term carbon storage link on the order of weeks and a longer-term seasonal positive soil moisture influence on photosynthesis anomalies. Daily to weekly lagged positive anomalies of photosynthesis positively influence current photosynthesis, inhibiting photosynthetic response to direct climate extremes, primarily anomalies on evapotranspiration and maximum temperature. Moisture stress or surplus in the prior season did inhibit or promote photosynthesis, but mechanisms are difficult to assess.

These results support prior suppositions that spring moisture anomalies and autumn carbon uptake anomalies influence future carbon assimilation rates, not just length of growing season or phenology. Further, the results highlight the difficulty that some commonly used indicators of plant growth such as remotely sensed vegetation indices, can reliably detecting anomalies in net carbon uptake. Finally, multi-year lagged negative relationships of temperature and evapotranspiration anomalies on current photosynthesis are intriguing and suggest new avenues of exploration for the role of long-lead ecosystem responses to extremes.

These findings are not necessarily detectable with shorter-term data or leaf-level analysis, as they involve subtle relationships and canopy and soil level processes. The results are similar to, for example, recent work by Niu *et al*. (2012) who argued that thermal acclimation of NEE (a flux made up of many interacting processes) can occur on interannual timescales in canopies. Also interesting was the lack of relationship between spectral indices and P_d, once converted to anomaly space, similar to some of the results of Heinsch *et al* (2006) that remotely sensed *GPP* is adequate for large spatial scale variation but poor for single pixel interannual variability.

The results here demonstrate the importance of long-term environmental observation of canopy photosynthesis but the caution that has to be taken regarding the high temporal autocorrelation that exists in flux and climate data. Strong covariance of these signals to seasonal orbital forcing requires careful evaluation of spurious correlation. A disconcerting finding was the lack of strong statistical power at detecting many correlations at long timescales, beyond interannual, even with >15 years of data. Methods that seek complementary use of short-term field manipulations, lab observations, and long-

644 term datasets like Fluxnet and the evolving NEON observatory will require continued
645 evolution of model-data comparison tools. Other papers in this special issue point to a
646 number of intriguing new ways to look at photosynthesis in models (Dietze *et al*., this issue;
647 Rogers *et al*., this issue; Sitch *et al*., this issue; Tholen *et al*., this issue). Here, I have shown
648 that a spectral anomaly framework and long-term flux observation network contribute to
649 their evaluation and improvement.

650 ## Acknowledgements

651 This manuscript wouldn't have been possible with the numerous person-hours of support
652 that going into the making of observations at WLEF including J. Thom at UW-Madison, A.
653 Andrews and J. Kofler at NOAA-ESRL, R. Strand and J. Ayers at State of Wisconsin
654 Educational Communications Board, K. Davis and current/former lab members at The
655 Pennsylvania State University, P. Bolstad at University of Minnesota, and B. Cook at NASA
656 GSFC. I also would like to thank A. Leakey for organizing this special issue. Observations
657 and research were supported through NSF Biology Directorate grants #DEB-0845166 and
658 #DBI-1062204.

659 ## References

660 Amiro B, Barr AG, Barr JG, et al (2010) Ecosystem carbon dioxide fluxes after disturbance
661 in forests of North America. *J Geophys Res* 115:G00K02. doi:10.1029/2010JG001390.

662 Baldocchi DD (2008) 'Breathing' of the Terrestrial Biosphere: Lessons Learned from a
663 Global Network of Carbon Dioxide Flux Measurement Systems. *Aust J Bot* 56:1-26

664 Baldocchi DD, Falge E, Wilson K (2001) A spectral analysis of biosphere-atmosphere trace
665 gas flux densities and meteorological variables across hour to year time scales. *Agr*
666 *Forest Meteorol* 107:1-27.

667 Baldocchi DD, Hutchison BA, Matt DR, McMillen RT (1985) Canopy radiative transfer
668 models for spherical and known leaf inclination distribution angles: a test in an oak-
669 hickory forest. *J Applied Ecol* 22:539-555

670 Barnhart BL, Eichinger WE, Preuger JH (2012) Introducing an Ogive method for
671 discontinuous data. *Agr Forest Meteorol* 162:58-62.

672 Beer C *et al* (2010) Terrestrial gross carbon dioxide uptake: Global distribution and
673 covariation with climate. *Science* 329:834-838. doi:10.1126/science.1184984

674 Berger BW, Zhao CL, Davis KJ, Yi C, Bakwin PS (2001) Long-term carbon dioxide fluxes
675 from a very tall tower in a northern forest: Flux measurement methodology. *J Atmos*
676 *Oceanic Technol* 18:529–542

677 Booth BB, *et al* (2012) High sensitivity of future global warming to land carbon cycle
678 processes. *Environ Res Lett* 7:024002. doi:10.1088/1748-9326/7/2/024002

679 Carbone MS, Czimczik CI, McDuffee KE, Trumbore SE (2007) Allocation and residence time
680 of photosynthetic products in a boreal forest using a low-level 14C pulse-chase labeling
681 technique. *Glob Change Biol* 13:466-477

682 Collatz GJ, Ball JT, Grivet C, Berry JA (1991) Physiological and environmental regulation of
683 stomatal conductance, photosynthesis and transpiration: a model that includes a
684 laminar boundary layer. *Agr Forest Meteorol* 54:107–136

685 Davis KJ, Bakwin PS, Yi C, Berger BW, Zhao C, Teclaw RM, Isebrands JG (2003) The annual
686 cycles of CO_2 and H_2O exchange over northern mixed forest as observed from a very tall
687 tower. *Glob Change Biol* 9:1278–93

688 De Pury DGG, Farquhar GD (1997) Simple scaling of photosynthesis from leaves to canopies
689 without the errors of big-leaf models. *Plant Cell Environ* 20:537–557

690 Desai AR (2010) Climatic and phenological controls on coherent regional interannual
691 variability of carbon dioxide flux in a heterogeneous landscape. *J Geophys Res*
692 115:G00J02. doi:10.1029/2010JG001423

693 Desai AR, Helliker BR, Moorcroft PR, Andrews AE, Berry, JA (2010) Interannual variability
694 in regional carbon fluxes from top-down and bottom-up perspectives. *J Geophys Res*
695 115:G02011. doi:10.1029/2009JG001122

696 Desai AR, Moorcroft PR, Bolstad PV, Davis KJ (2007)Regional carbon fluxes from a
697 biometrically-constrained dynamic ecosystem model: Impact of disturbance, CO2
698 fertilization and heterogeneous land cover. *J Geophys Res* 112:G01017.
699 doi:10.1029/2006JG000264

700 Desai AR, Richardson AD, Moffat AM, Kattge J, Hollinger DY, Barr A, Falge E, Noormets A,
701 Papale D, Reichstein M, and Stauch VJ (2008) Cross site evaluation of eddy covariance

GPP and RE decomposition techniques. *Agr Forest Meteorol* 148:821-838. doi:10.1016/j.agrformet.2007.11.012

Detto, M, Molini A, Katul G, Stoy P, Palmroth S, Baldocchi DD (2012) Causality and persistence in ecological systems: A nonparametric spectral granger causality approach. *Am Nat* 179:524-535

Dietze MC, *et al.* (2011) Identifying the time scales that dominate model error: A North American synthesis of the spectral properties of ecosystem models. *J Geophys Res* 116:G04029. doi:10.1029/2011JG001661

Farquhar GD, Sharkey TD (1982) Stomatal conductance and photosynthesis. *Annu Rev Plant Physiol* 1982.33:317-345

Foken, T, Aubinet, M, Leuning, R (2012) The eddy-covariance method. In: Aubinet M *et al*. (eds) *Eddy Covariance: A Practical Guide to Measurement and Data Analysis*. Springer, Dordrecht, pp 1-19

Friedlingstein, P., Cox P, Betts R *et al* (2006) Climate -carbon cycle feedback analysis, results from the C4MIP model intercomparison. *J Climate* 19:3337-3353

Gellesch E, Hein R, Jaeschke A, Beierkuhnlein C, Jentsch A (2013) Biotic interactions in the face of climate change. *Prog. Bot.* 74:321-349. doi:10.1007/978-3-642-30967-0_12.

Heinsch, FA, Zhao M, Running, SW *et al* (2006) Evaluation of remote sensing based terrestrial productivity from MODIS using Ameriflux tower eddy flux network observations. *IEEE T Geosci Remote* 44:1908-1925. doi:10.1109/TGRS.2005.853936

Huang NE, Wu Z (2008) A review on Hilbert-Huang transform: Method and its applications to geophysical studies. *Rev Geophys* 46:RG2006. doi:10.1029/2007RG000228

Keenan TF *et al* (2012a) Evaluation of terrestrial biosphere model performance for land-atmosphere CO_2 exchange on inter-annual time scales: Results from the North American Carbon Program interim site synthesis. *Glob Change Biol*. doi: 10.1111/j.1365-2486.2012.02678.x

Keenan TF, Davidson E, Moffat A, Munger W, Richardson AD (2012b) Using model-data fusion to interpret past trends, and quantify uncertainties in future projections, of forest ecosystem carbon cycling. *Glob Change Biol* 18:2555–2569

Kumar M, Monteith JL (1981) Remote sensing of crop growth. In: Smith H (ed) *Plants and the Daylight Spectrum*. Academic, London, pp 133-144

Le Quéré C, Raupach MR, Canadell JG *et al* (2009) Trends in the sources and sinks of carbon dioxide. *Nat Geosci.* doi:10.1038/ngeo689

735 LeBauer DS, Wang D, Dietze MC (in press) Feedbacks between measurements and models
736 facilitated by scientific workflows: the Predictive Ecosystem Analyzer (PEcAn). *Ecol*
737 *Monogr*.

738 Mahecha MD, Reichstein M, Jung M, Seneviratne SI, Zaehle S, Beer C, Braakhekke MC,
739 Carvalhais N, Lange H, Le Maire G, Moors E (2010) Comparing observations and
740 process-based simulations of biosphere-atmosphere exchanges on multiple time scales.
741 *J Geophys Res* 115:G02003. doi:10.1029/2009JG001016

742 Mahecha MD, Reichstein M, Lange H, Carvalhais N, Bernhofer C, Grünwald T, Papale D,
743 Seufert G (2007) Characterizing ecosystem-atmosphere interactions from short to
744 interannual time scales. *Biogeosciences* 4:743–758

745 Marten S, Carpenter S, Foley JA, Folke C, Walker B (2001) Catastrophic shifts in ecosystems.
746 *Nature* 413:591-596. doi:10.1038/35098000.

747 Medvigy D, Wofsy SC, Munger JW, Hollinger DY, Moorcroft PR (2009) Mechanistic scaling of
748 ecosystem function and dynamics in space and time: the Ecosystem Demography model
749 version 2. *J Geophys Res* 114:G01002. doi:10.1029/2008JG000812

750 Medvigy D, Wofsy SC, Munger JW, Moorcroft PR (2010) Responses of terrestrial ecosystems
751 and carbon budgets to current and future environmental variability. *P Nat Acad Sci*
752 107:8275-8280

753 Moffat AM, Papale D, Reichstein M *et al* (2007) Comprehensive comparison of gap-filling
754 techniques for eddy covariance net carbon fluxes. *Agr Forest Meteorol* 147:209-232.
755 doi:10.1016/j.agrformet.2007.08.011

756 Moorcroft PR (2006) How close are we to a predictive science of the biosphere? *Trends Ecol*
757 *Evol* 21:400-407

758 Niu S *et al* (2012) Thermal optimality of net ecosystem exchange of carbon dioxide and
759 underlying mechanisms. *New Phytol* 194:775-783. doi:10.1111/j.1469-
760 8137.2012.04095.x

761 Norby RJ, Zak DR (2011) Ecological lessons from free-air CO_2 enrichment (FACE)
762 experiments. *Annu Rev Ecol Evol S* 42:181-203

763 Raupach MR, Rayner PJ, Barrett DJ, DeFries RS, Heimann M, Ojima DS, Quegan S, Schmullius
764 CC (2005) Model-data synthesis in terrestrial carbon observation: methods, data
765 requirements and data uncertainty specifications. *Glob Change Biol* 11:378-397

766 Ricciuto DM, Butler MP, Davis KJ, Cook BD, Bakwin P, Andrews A, Teclaw RM (2008) Causes
767 of interannual variability in ecosystem-atmosphere CO2 exchange in a northern
768 Wisconsin forest using a Bayesian model calibration. *Agr Forest Meteorol* 148:309-327.

769 Richardson AR *et al* (2012) Terrestrial biosphere models need better representation of
770 vegetation phenology: results from the North American Carbon Program Site Synthesis.
771 *Glob Change Biol* 18:566-584. doi:10.1111/j.1365-2486.2011.02562.x

772 Schaefer K *et al* (2012) A model-data comparison of gross primary productivity: Results
773 from the North American Carbon Program Site Synthesis. *J Geophys Res* 117:G03010.
774 doi:10.1029/2012JG001960.

775 Sellers, PJ (1985) Canopy reflectance, photosynthesis and transpiration. *Int J Rem Sens*
776 6:1335-1372. doi:10.1080/01431168508948283

777 Sprintsin M, Chen JM, Desai AR, Gough CM (2012) Evaluation of leaf-to-canopy upscaling
778 methodologies against carbon flux data in North America. *J Geophys Res* 117:G01023.
779 doi:10.1029/2010JG001407

780 Stoy PC, Richardson AD, Baldocchi DD, Katul GG, Stanovick J, Mahecha MD, Reichstein M,
781 Detto M, Law BE, Wohlfahrt G, Arriga N, Campos J, McCaughey JH, Montagnani L, Paw U
782 KT, Sevanto S, Williams M (2009) Biosphere-atmosphere exchange of CO2 in relation to
783 climate: a cross-biome analysis across multiple time scales. *Biogeosciences* 6:2297-2312

784 Van Gorsel E, Delpierre N, Leuning R *et al* (2009) Estimating nocturnal ecosystem
785 respiration from the vertical turbulence flux and change in storage of CO_2. *Agr Forest*
786 *Meteorol* 149:1919-1930. doi:10.1016/j.agrformet.2009.06.020

787 Walter J, Beierkuhnlein C, Jentsch A, Kreyling J (2013) Ecological stress memory and cross
788 stress tolerance in plants in the face of climate extremes. *Environ. Exp. Bot*.
789 doi:10.1016/j.envexpbot.2012.02.009.

790 Williams, M, Richardson AD, Reichstein M, Stoy PC, Peylin P, Verbeeck H, Carvalhais N, Jung
791 M, Hollinger DY, Kattge J, Leuning R, Luo Y, Tomelleri E, Trudinger C, Wang Y-P (2009)
792 Improving land surface models with FLUXNET data. *Biogeosciences* 6:1341-1359

793 Wu C, Chen JM, Black T.A., et al (2013) Interannual variability of net ecosystem productivity
794 in forests is explained by carbon flux phenology in autumn. *Global Ecol. Biogeogr.* in
795 press. doi:10.1111/geb.12044.

796 Yi C, Ricciuto DM, Li R *et al* (2010) Climate control of terrestrial carbon exchange across
797 biomes and continents. *Environ Res Lett* 5:034007. doi:10.1088/1748-
798 9326/5/3/034007

799 Yi C, Davis KJ, Bakwin, PS, Berger BW, Marr L (2000) The influence of advection on
800 measurements of the net ecosystem-atmosphere exchange of CO_2 from a very tall
801 tower. *J Geophys Res* 105:9991-9999

802 Zobitz J, Desai AR, Moore DJP, Chadwick MA (2011) A primer for data assimilation with
803 ecological models using Markov Chain Monte Carlo (MCMC). *Oecologia* 167:599-611.
804 doi:10.1007/s00442-011-2107-9.

Figures

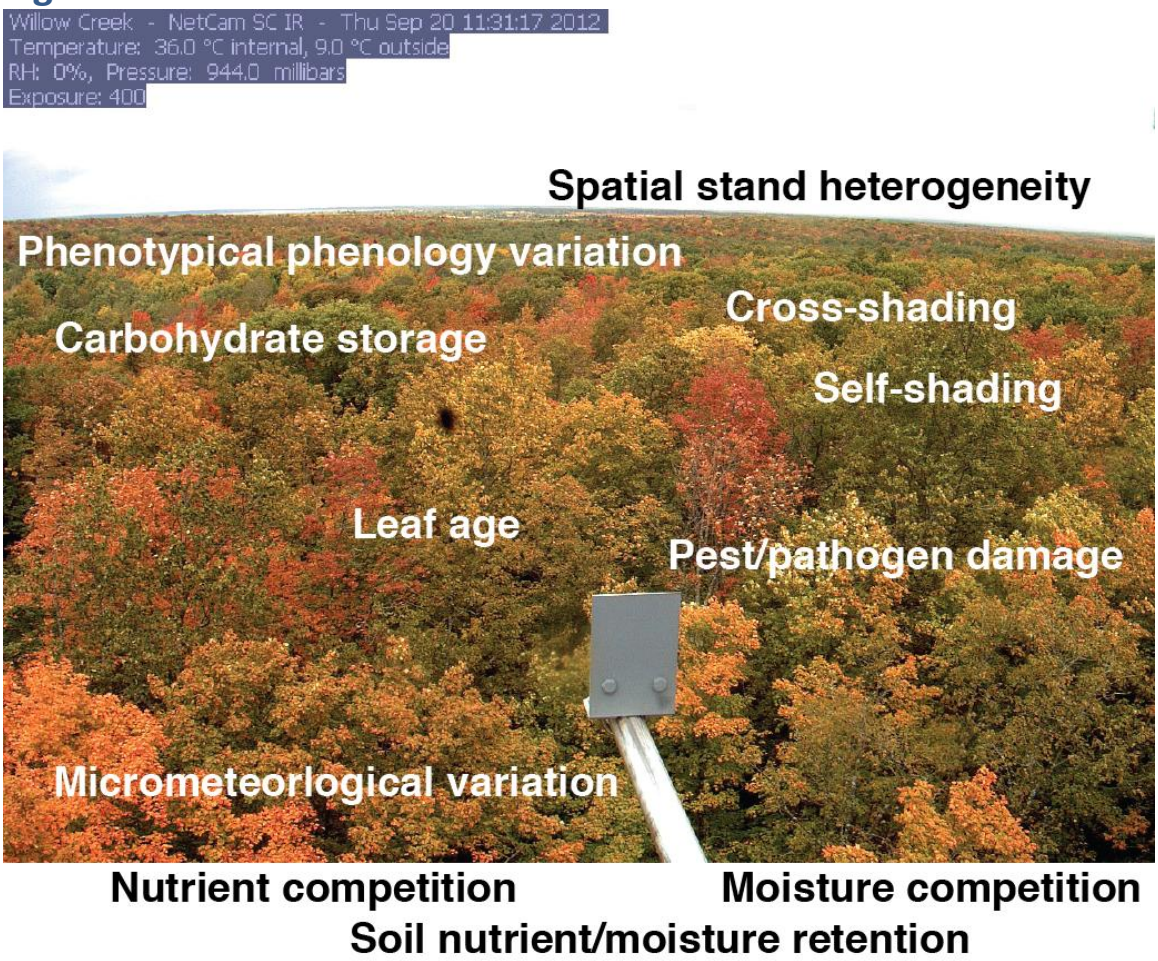

Fig. 1 Examples of processes that affect photosynthesis in the environment that make canopy-scale photosynthesis differ from leaf-level photosynthesis, superimposed on a photo taken from an eddy covariance flux tower overlooking a temperate hardwood forest .

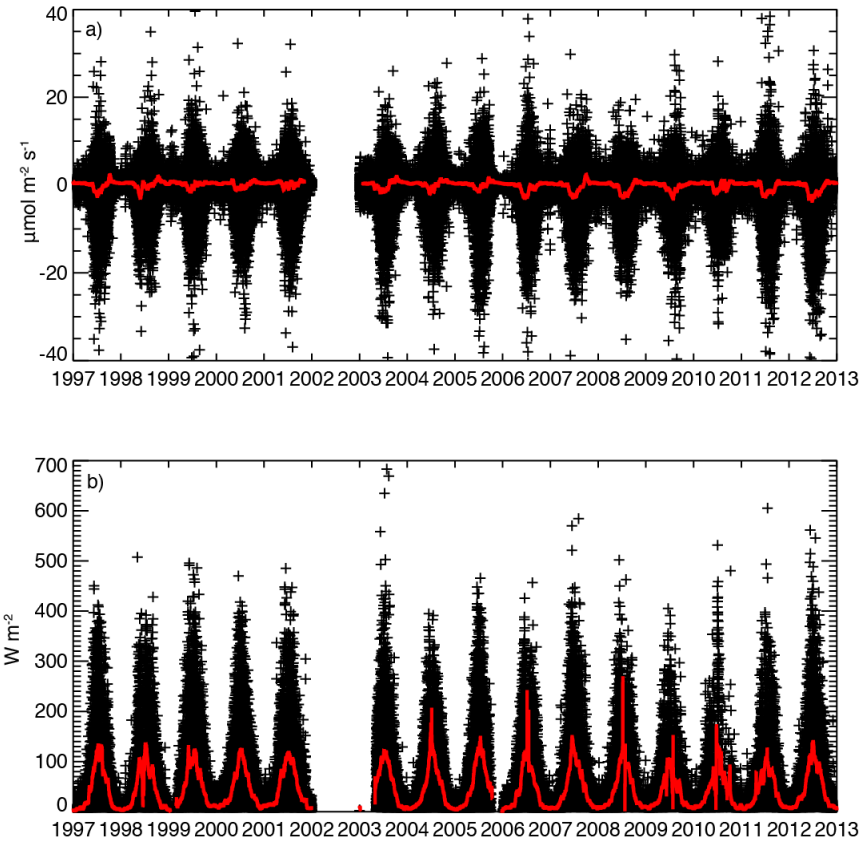

Fig. 2 Nearly fifteen year time series of (a) net ecosystem exchange of CO_2 (*NEE*) and (b) evapotranspiration flux (*ET*) at hourly (black crosses) and biweekly (red line) timescales as observed from the WLEF Park Falls, WI very tall eddy covariance flux tower from 1997-2012. Long-term datasets like these on canopy-scale carbon and water fluxes are starting to be made available in the eddy covariance community. The large variability at multiple timescales presents a significant opportunity and challenge for improving predictions of ecosystem fluxes.

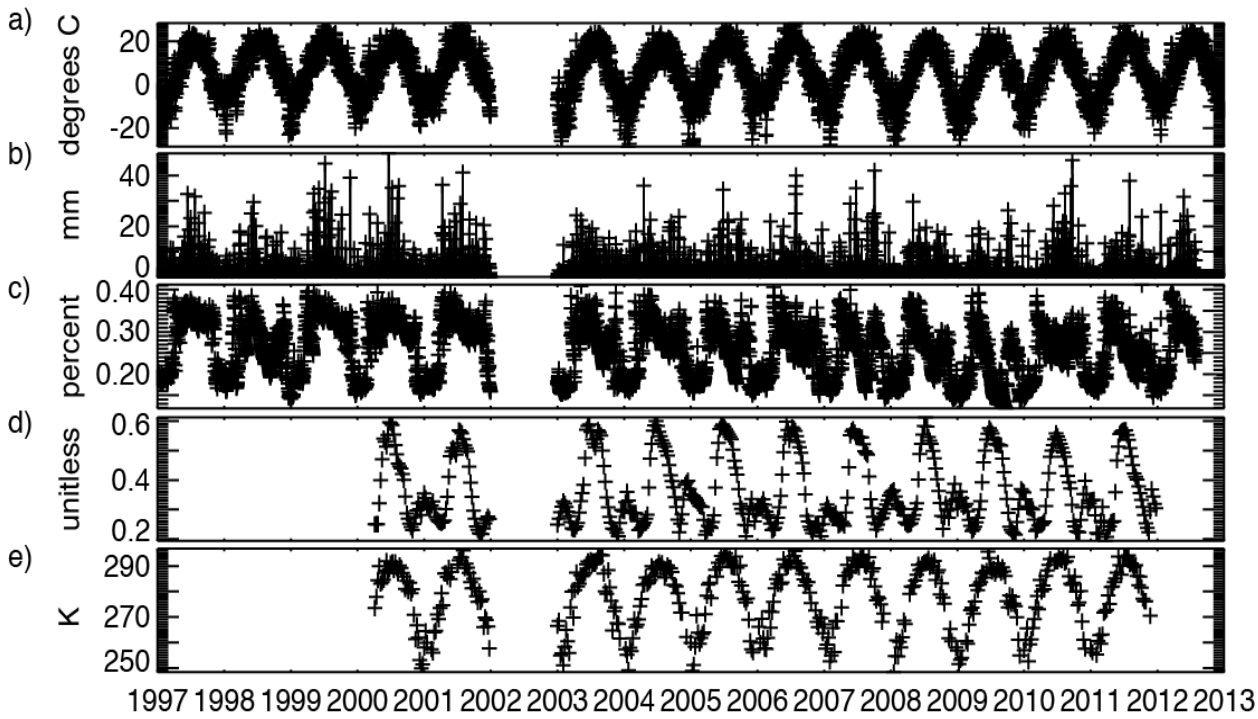

Fig 3. Time series of daily a) air temperature (T_{air}), b) precipitation (P_{recip}), c) soil moisture (Q_{soil}), and biweekly d) MODIS enhanced vegetation index (*EVI*), and e) MODIS land surface temperature (*LST*). To allow comparisons to the quality-controlled flux data, 2002 was also removed from these data.

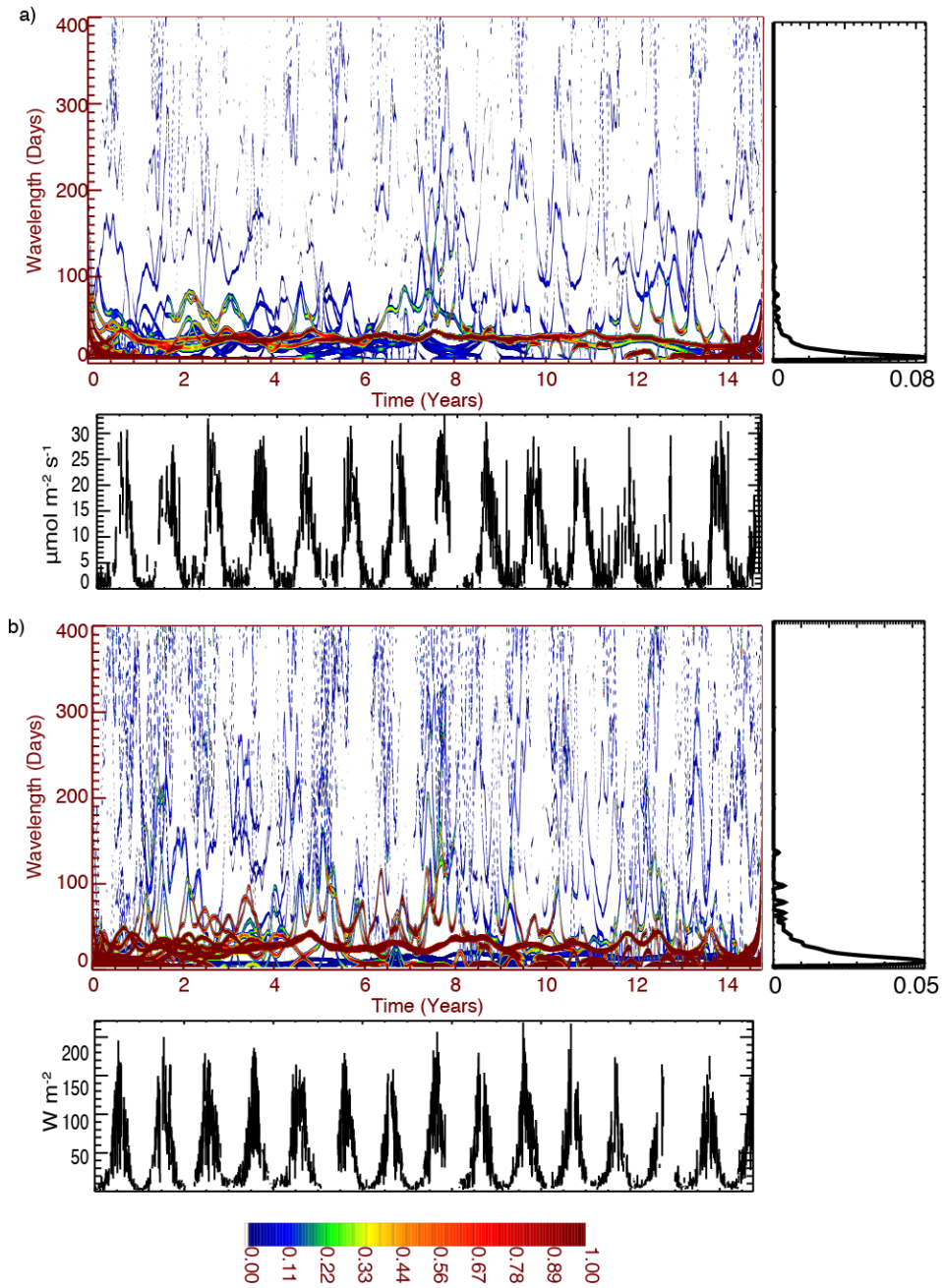

Fig. 4 Hilbert-Huang power spectral transformation of (a) net canopy carbon uptake (P_d) and (b) daily evapotranspiration (ET) derived from empirical mode decomposition of the tall tower fluxes. Raw daily time series are shown below each transformation and marginal power spectrum on the right. To improve clarity of higher frequency variation, the transform is cut-off at 400-day wavelength. The time-frequency decomposition reveals that both carbon and water exchange exhibit a number of characteristic timescales of variability, including substantial peaks at the monthly and seasonal timescale. Longer-term variability is much weaker, especially for carbon, but growing season signals of variability on the 100-200 day wavelength exist and strong coherence among water and carbon are also noted.

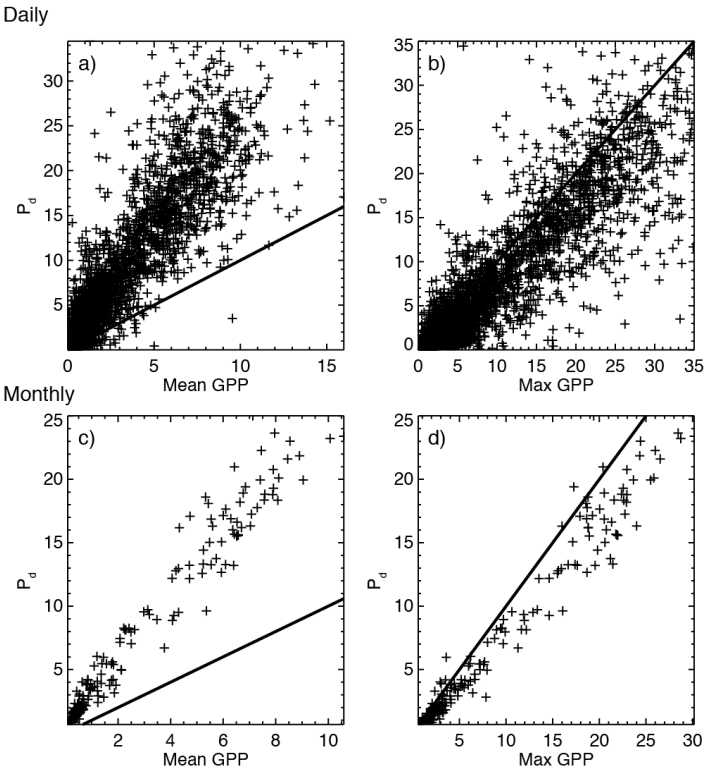

Fig. 5 Scatterplot of flux tower derived GPP (based on fit of nighttime NEE to temperature to estimate respiration) to P_d for a) mean daily GPP, b) maximum daily GPP, c) mean monthly GPP, and d) mean monthly maximum daily GPP. 1:1 fit line shown on all four and shows that P_d is closely related to maximum daily GPP. Correlation at daily scale is r^2=0.81 and monthly r^2=0.96.

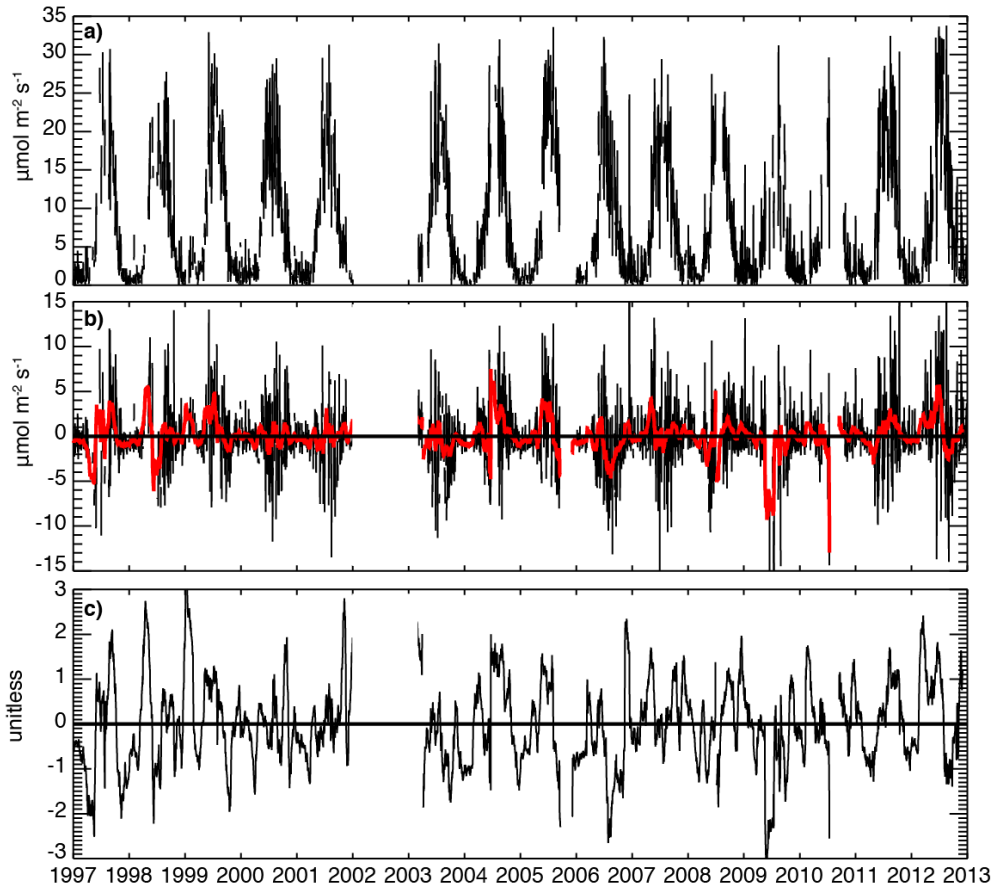

Fig. 6 Example of generating relative anomalies for lag correlation analysis. Raw daily time series of (a) net photosynthetic drawdown (P_d), were (b) de-seasonalized by removal of the ensemble average daily time series and then averaged to the appropriate averaging timescale, in this example, monthly (red line) and finally (c) normalized to relative values. The final signal represents the true anomalies of variation across time and has successfully removed the seasonal variability of solar forcing.

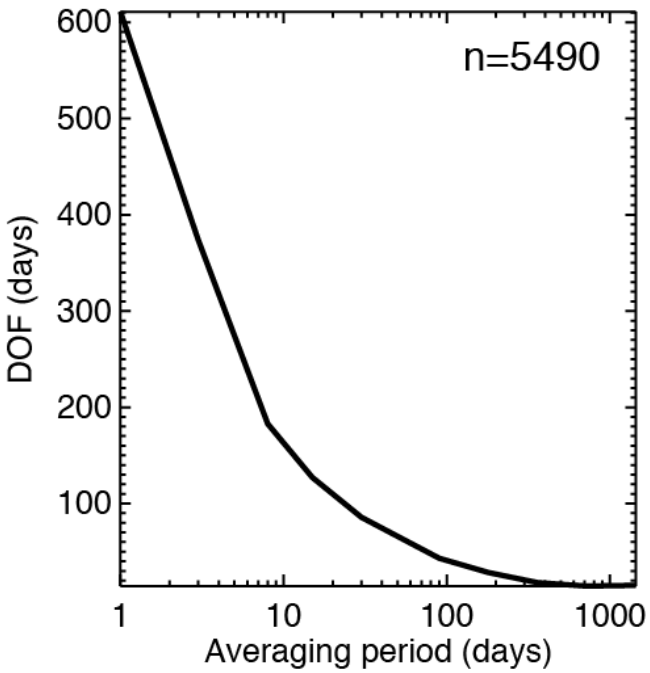

Fig. 7 Degree of freedom analysis for daily net photosynthetic drawdown from the tall tower fluxes. The total record has 5490 observation periods (with 37% of observations missing), but the full autocorrelation analysis reveals an exponentially declining true degree of freedom from slightly over 600 for no-averaging, to near zero above 360 day smoothing of the time series. This reduced N is used for all correlation significance tests.

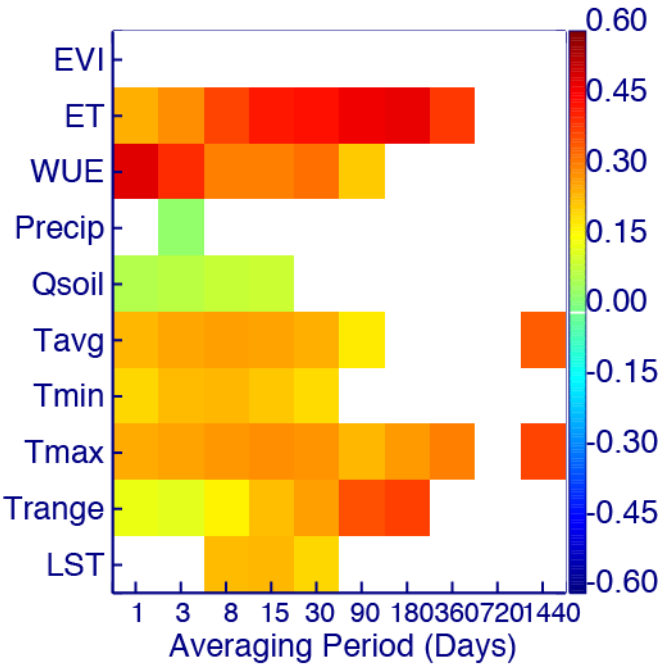

Fig. 8 Direct correlation coefficients of standardized anomalies of climate variables to net photosynthetic drawdown (P_d) anomalies at daily to multi-year filter scales. Temperature variables explain a small fraction of weekly to monthly anomalies in P_d, while moisture variability explains a greater fraction of long-term seasonal variability in anomalies of P_d. Interestingly, remotely sensed vegetation index anomalies do not significantly explain P_d anomalies at any timescale.

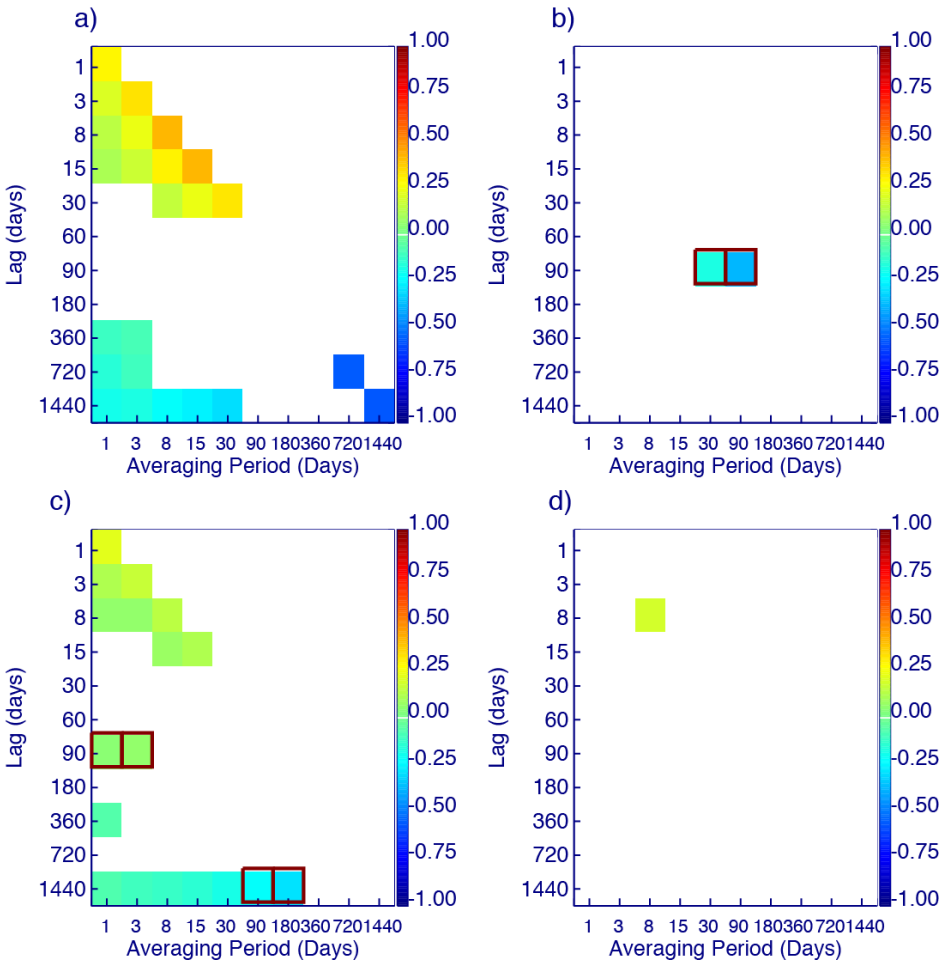

Fig. 9 Analysis of lagged standardized anomalies of (a) P_d, (b) *EVI*, (c) T_{mean}, and (d) *LST* to P_d at a range of averaging times and lag times. Only significant correlations are displayed, after correction of autocorrelated degrees of freedom. Red squares indicate significant correlations that exceed P_d autocorrelation at that averaging and lag scale. Strong AR-1 correlation persist in P_d at lags up to one month and averaging periods to one month, while small negative correlations show predictive ability for temperature at lags exceeding one year and *EVI* at the two-month lag.

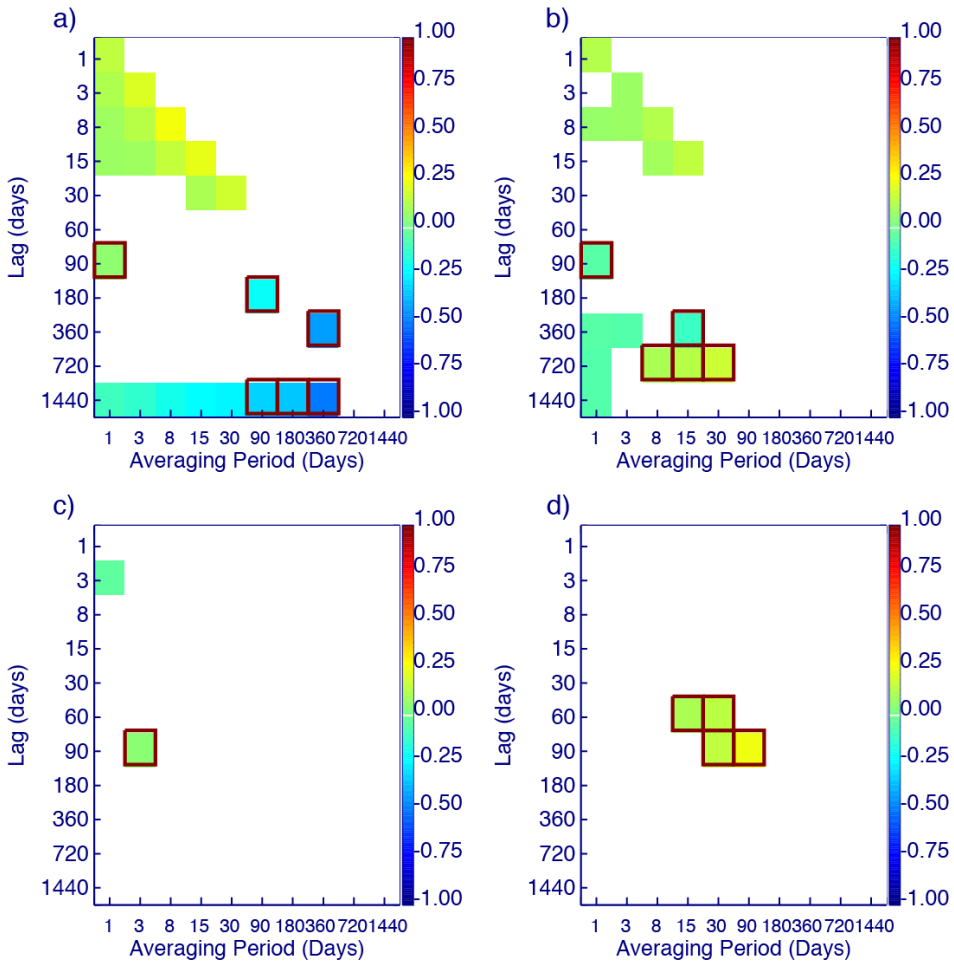

Fig. 10 Same as Fig. 9 but for moisture variables of (a) ET, (b) WUE, (c) P_{recip}, and (d) Q_{soil}. While ET and WUE have significant direct correlations as shown in Fig. 5, lagged correlations are small. Stronger positive correlations exist for precipitation and soil moisture at the weekly to seasonal timescale for seasonal scale lags.

Tables

Table 1 Abbreviations used in this paper

Abbreviation	Description	Source
NEE	Net ecosystem exchange of CO_2	Flux tower
GPP	Gross primary production	Flux tower
P_d	Photosynthetic drawdown	Flux tower
EVI	Enhanced Vegetation Index, 8-day average	MODIS TERRA/AQUA
ET	Evapotranspiration	Flux tower
WUE	Water Use Efficiency (P_d/ET)	Flux tower
P_{recip}	Daily precpitation	NCDC + NARR Reanalysis
Q_{soil}	10 cm soil moisture	NARR Reanalysis
T_{mean}	Daily temperature	Flux tower + NCDC
T_{min}	Minimum daily temperature	Flux tower + NCDC
T_{max}	Maximum daily temperature	Flux tower + NCDC
T_{range}	Daily temperature range (max - min)	Flux tower + NCDC
LST	Land Surface Temperature, 8-day day/night average	MODIS TERRA/AQUA

Table 2 Granger causality analysis for P_d as a function of averaging period reveals best predictive models generally include temperature and transpiration observations for daily to weekly averages but precipitation, soil moisture, and water use efficiency for monthly time scales, primarily within the first or two lags. No variables could exceed P_d autoregression in explaining longer averaging scales (> 30 days).

Variable/Averaging period (Days)	1	3	8	15	30	90	360
EVI						90	
T_{mean}	1	3	8				
LST	1	3					
ET	1-3	3	8				360
WUE	1-3	3		15-360			
P_{recip}							
Q_{soil}							