Does Day and Night Sampling Reduce Spurious Correlation between Canopy Photosynthesis and
Ecosystem Respiration?

Dennis Baldocchi, Cove Sturtevant and Fluxnet Contributors

Department of Environmental Science, Policy and Management
130 Mulford Hall
University of California, Berkeley
Berkeley, CA
94720
Baldocchi@berkeley.edu

April 3, 2015

Agricultural and Forest Meteorology

AGRFORMET-D-14-00701, revised

Key words: eddy covariance, FLUXNET, net ecosystem carbon exchange, biosphere-atmosphere
interactions

Abstract

It is necessary to partition eddy covariance measurements of carbon dioxide exchange into its offsetting gross fluxes, canopy photosynthesis and ecosystem respiration, to understand the biophysical controls on the net fluxes. And, independent estimates of canopy photosynthesis (G) and ecosystem respiration (R) are needed to validate and parameterize carbon cycle models that are coupled with climate and ecosystem dynamics models. Yet, there is concern that carbon flux partitioning methods may suffer from spurious correlation because derived values of canopy photosynthesis and ecosystem respiration both contain common information on net carbon fluxes at annual time scales.

We hypothesized that spurious correlation among canopy photosynthesis and ecosystem respiration can be minimized using day-night conditional sampling of CO₂ exchange; daytime fluxes are dominated by photosynthesis and nighttime fluxes are dominated by respiration. To test this hypothesis, we derived explicit equations that quantify the degree of spurious correlation between photosynthesis and respiration. Theoretically, day and night samples of net carbon exchange share a different common variable, daytime ecosystem respiration, and the degree of spurious-correlation depends upon the variance of this shared variable.

We then applied this theory to ideal measurements of carbon exchange of over a vigorous, irrigated and frequently harvested alfalfa field in the sunny and windy region of the California, the Sacramento-San Joaquin Delta, where soil CO₂ efflux is strong. In this case, we found that the correlation coefficient between canopy photosynthesis and ecosystem respiration was -0.79. This relatively high correlation between canopy photosynthesis and respiration was mostly real as the degree of spurious correlation was only -0.32.

We then expanded this analysis to the FLUXNET database that spans a spectrum of climate and plant functional types. We found, on average, that the correlation between gross photosynthesis and ecosystem respiration, using day-night sampling, was close to minus one (-0.828 +/- 0.130). For perspective, a large fraction of this correlation was real, as the degree of spurious correlation (Eq. 22) was -0.526. We conclude that the potential for spurious correlation between canopy photosynthesis and ecosystem respiration across the FLUXNET database was moderate. Looking across the database, we found that the least negative spurious-correlations coefficients (> -0.3) were associated with seasonal deciduous forests. The most negative spurious correlations coefficients (< -0.7) were associated with evergreen forests found in boreal climates.

Key words: eddy covariance, FLUXNET, net ecosystem exchange, ecosystem respiration, canopy photosynthesis

Introduction

Our ability to simulate ecosystem scale photosynthesis remains poor, despite over 50 years of research on this topic (Dietze et al., 2011; Schaefer et al., 2012; Vargas et al., 2013). Yet, trustworthy measurements of gross CO₂ exchange between vegetation and the atmosphere are needed to improve, and parameterize, the suite of extant ecosystem models that predict gross carbon fluxes (Alton, 2011; Loew et al., 2014; Williams et al., 2009) and couple these fluxes with the climate system (Cox et al., 2000; Friedlingstein et al., 2006).

Eddy covariance measurements of CO₂ exchange between an ecosystem and the atmosphere are direct and represent the net difference between uptake by plant photosynthesis and losses by ecosystem respiration. To understand and diagnose how and why those fluxes vary with time and environmental conditions, practitioners have attempted to partition the net flux into its two, offsetting gross components, canopy photosynthesis and ecosystem respiration (Falge et al., 2002; Lasslop et al., 2010b; Papale et al., 2006; Reichstein et al., 2005).

Several approaches are popular for flux partitioning. One approach estimates canopy photosynthesis by subtracting empirical estimates of daytime respiration from daytime measurements of net carbon dioxide exchange (Reichstein et al., 2005). An empirical function describing the relationship between ecosystem respiration and environmental drivers, like temperature and soil moisture, is derived using night-time eddy covariance flux measurements. There are limitations and uncertainties towards using nocturnal eddy flux measurements of CO₂ exchange to derive this empirical model for daytime respiration. Better models incorporate a seasonal component to a baseline efflux (Reichstein et al., 2005; Xu and Baldocchi, 2004). The other issue revolves around the accuracy and representativeness of nocturnal efflux measurements, when atmospheric turbulence is often low and the atmosphere is stably stratified (Barr et al., 2013; Gu et al., 2005). Under calm, stable conditions, insufficient mixing of the air can occur. This results in either substantial storage of CO₂ in the underlying air space or the pooling and draining of air, which prevent it from crossing the canopy-atmosphere interface. In other circumstances the stable boundary layer causes winds aloft to be decoupled from flow in the surface layer. This condition is untenable and leads to sporadic turbulent transfer (Lee, 2000; Mahrt, 2014; Nappo, 1991). Bursts in CO₂ efflux are non-local in time, as they vent air that was stored in the air space during the preceding calm period. Together, these processes can cause a bias in the eddy covariance flux density measured across the canopy-atmosphere interface (Acevedo et al., 2007; Barr et al., 2013; Finnigan, 2006; Gu et al., 2005).

The second approach derives information on ecosystem respiration from the better-mixed conditions of daytime carbon flux measurements. In one situation, the estimate of ecosystem respiration is produced by extrapolation of the response curve between canopy photosynthesis and light to its zero intercept (Falge et al., 2002; Xu and Baldocchi, 2004). Strength of this method revolves around the fact that it infers night respiration from daytime measurements, when turbulent mixing is greater. A weakness of this approach is the fact it relies on flux measurements during the sunrise and sunset transition periods when the atmosphere is not stationary and atmospheric storage can be great; the

constant flux layer, upon which eddy covariance theory relies on, depends on the assumption of steady-state conditions (Baldocchi, 2003; Foken and Wichura, 1996).

A third approach exploits the fact the respiratory effluxes are relatively capped, when measured a few hours after sundown; then mixing remains relatively vigorous and measurement biases are assumed small (van Gorsel et al., 2009).

Recently there has been constructive criticism about deriving canopy photosynthesis and ecosystem respiration from measurements of net ecosystem carbon exchange. This approach suffers from the closure problem that involves deriving two pieces of information (canopy photosynthesis and ecosystem respiration) from one equation. We acknowledge that efforts are underway to assess canopy photosynthesis with independent proxies like carbonyl sulfide (COS) (Asaf et al., 2013; Berry et al., 2013; Blonquist et al., 2011), with simultaneous fluxes of stable carbon isotopes (^{12}C and ^{13}C) (Bowling et al., 2001; Griffis, 2013), as differences between above and below canopy flux systems (Baldocchi et al., 1987; Baldocchi et al., 1997) and with sun induced fluorescence (Frankenberg et al., 2011; Guanter et al., 2014). But, none of these alternative methods produce measurements of gross canopy photosynthesis, year round, quasi-continuously, and with as few assumptions as direct eddy covariance measurements.

Others have reported on how the partitioning of ecosystem respiration (R) and canopy photosynthesis (G) may suffer from the pitfalls of spurious correlation (Kenney, 1982; Lasslop et al., 2010a; Vickers et al., 2009). Concern rises because estimates of canopy photosynthesis and ecosystem respiration both contain information on net ecosystem carbon exchange (NEE). Additional concern revolves on how well estimates of daytime ecosystem respiration can be represented as a function of temperature, time and soil moisture. A growing number of studies are finding that recent photosynthesis primes soils respiration (Kuzakov, 2010; Tang et al., 2005). While others have shown an inhibition of dark respiration in light by the so-called Kok-effect (Amthor, 1994; Heskell et al., 2013). Consequently, daytime ecosystem respiration, at a given temperature, may be different than nighttime respiration at the same temperature.

The topic of spurious correlation has a long history. The eminent statistician, Karl Pearson (Pearson, 1896), was probably the first researcher to caution about spurious correlations among three inter-related variables. He showed that a correlation between two independent datasets, x and y, would exist if they were divided by another independent variable, z, and the ratios were regressed against with one another. To quantify this measure of spurious correlation, he developed an index:

$$\rho = \frac{(\sigma_z / \bar{z})^2}{\sqrt{(\sigma_x / \bar{x})^2 + (\sigma_z / \bar{z})^2} \sqrt{(\sigma_y / \bar{y})^2 + (\sigma_z / \bar{z})^2}} \quad (1)$$

In micrometeorology, examples of spurious correlation include the relationships between standard deviations in vertical velocity (σ_w / u_*) and Monin-Obukhov length scale ($L = \frac{u_*^3 \theta_v}{k g w' \theta'}$) (Hicks,

1978) and the relationship between nocturnal CO₂ effluxes and friction velocity; in these situations both variables contain information on the standard deviation in vertical velocity.

While we acknowledge the claim of spurious correlation between canopy photosynthesis and ecosystem respiration can be true, we advocate that the problem is more nuanced with respect to the partitioning of net CO₂ exchange. For example, it is important to ask and answer: what is the degree of spurious correlation?

We hypothesize that the degree of spurious correlation depends on the sampling method and time scale upon which the data are summed and integrated (day/night vs. 24 hour vs annual). And, we aim to show that spurious correlation can be minimized if we use conditional sampling of daytime and nighttime flux measurements to construct sums at daily and annual scales. For example, canopy photosynthesis is populated by daytime flux measurements and ecosystem respiration is populated by nighttime measurements. In the case of annual sums, many of the ecosystem respiration measurements are populated with winter-time samples when photosynthesis is nil. Consequently, conditional sampling at these time scales can yield relatively independent measurements of photosynthesis and ecosystem respiration, producing a degree of spurious correlation that may not be as large as previously claimed (Vickers et al., 2009).

In this paper, we derive new theory to examine spurious correlation among day and night measurements of CO₂ exchange, as applied to partitioning net ecosystem carbon exchange into canopy photosynthesis and ecosystem respiration. We use the theory to determine which conditions are most and least prone to spurious correlation. Next, the theory is tested with new measurements over an ideal site; an irrigated alfalfa field, growing in an extremely windy portion of northern California, where the adjustments for lack of wind at night are essentially nil. Finally, we expand the analysis to the larger FLUXNET database and evaluate the degree of spurious correlation that may occur between canopy photosynthesis and ecosystem respiration across a spectrum of climate, plant functional types and landscapes. The objective of the last part is to examine what fraction of the FLUXNET database is subject to spurious correlations, in which biomes and climate zones they may occur, and to what degree.

Theoretical Background

According to Kenney (1982) two variables, A and B, may be dependent upon one another if they share another variable. Let's consider the following case where the variable x is common to both A and B:

$$A = x + y \quad (2)$$

$$B = x \quad (3)$$

From statistics, the correlation between A and B is defined as the ratio between the covariance and the product of the standard deviations of variable A (σ_A) and B (σ_B); overbars denote time averaging and primes denote fluctuations from the mean.

156

$$r_{ab} = \frac{\overline{A'B'}}{\sigma_A \sigma_B} \quad (4)$$

157

158

Substituting x and y into A and B and applying Reynolds' rules of averaging (Reynolds, 1895; Stull, 1988), yields:

159

$$r_{ab} = \frac{\overline{x'x'} + \overline{x'y'}}{\overline{x'x'}^{1/2} [\overline{x'x'} + \overline{y'y'} + 2\overline{x'y'}]^{1/2}} \quad (5a)$$

160

161

Equation 5a can also be expressed in terms of standard deviations, noted by σ_x and σ_y , and the covariance, σ_{xy} :

162

$$r_{ab} = \frac{\sigma_x^2 + \sigma_{xy}}{\sigma_x [\sigma_x^2 + \sigma_y^2 + 2\sigma_{xy}]^{1/2}} \quad (5b)$$

163

Substituting the definition of the correlation coefficient between x and y:

164

$$r_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \quad (6)$$

165

produces a version of r_{ab} that is a function of the correlation between x and y, r_{xy} :

166

$$r_{ab} = \frac{\sigma_x + r_{xy} \sigma_y}{[\sigma_x^2 + \sigma_y^2 + 2r_{xy} \sigma_x \sigma_y]^{1/2}} \quad (7)$$

167

168

According to Kenney (1982) and Vickers et al (2009), spurious-correlation between A and B can be expressed at the limit when the correlation between x and y, r_{xy} , goes to zero.

169

$$r_{sc} = \frac{\sigma_x}{[\sigma_x^2 + \sigma_y^2]^{1/2}} \quad (8)$$

170

171

172

173

We acknowledge that Vickers et al. (2009) analyzed day-night sampling for a ponderosa pine forest. This analysis takes the evaluation of Vickers et al. several steps further, by giving a formal derivation of the conditionally sampled correlation among NEE, G and R and by applying the theory across the larger FLUXNET database.

174

175

176

177

Our goal is to define and formally evaluate the correlation coefficients between canopy photosynthesis and ecosystem respiration by considering the day and night sampling of NEE to define this partitioning; net ecosystem carbon exchange (NEE) equals the sum of gross canopy photosynthesis (G) and ecosystem respiration (R):

$$178 \quad NEE = G + R \quad (9)$$

179 The sign convention adopted here is with respect to the atmosphere. NEE and G are negative when the
 180 carbon dioxide flux is directed to the surface, and from the atmosphere. NEE and R are positive when
 181 the carbon dioxide flux is directed towards the atmosphere. We then use this theory to derive a new
 182 measure for spurious correlation among these three variables.

183 In principle, the correlation coefficient between G and R can be defined as:

$$184 \quad r_{GR} = \frac{\overline{G'R'}}{(\overline{G'G'})^{1/2}(\overline{R'R'})^{1/2}} \quad (10)$$

185 where overbars represent temporal averaging and primes represent fluctuations from the mean. To
 186 evaluate Equation 10, let's break up the problem into day and night sampling and relate these terms to x
 187 (NEE_{day}), y (NEE_{night}) and z ($R_{eco,day}$) for simpler algebraic manipulation.

$$188 \quad G = NEE_{day} - R_{eco,day} = x - z \quad (11)$$

$$189 \quad R = NEE_{night} + R_{eco,day} = y + z \quad (12)$$

190 Note this system has two independent variables (x and y) and a common term z, daytime ecosystem
 191 respiration.

192 Next, we define the covariance between G and R and then expand it in terms of x, y and z; we use the
 193 concept of Reynolds averaging where the fluctuation from the mean is the difference between the
 194 instantaneous value and the ensemble mean:

$$195 \quad \overline{G'R'} = \overline{(G - \bar{G})(R - \bar{R})} = \overline{GR} - \bar{G}\bar{R} \quad (13)$$

196 Applying Reynolds' rules of averaging (Stull, 1988) produces an expression for gross canopy
 197 photosynthesis that equals the sum of the mean and the fluctuation from the mean

$$198 \quad G = \bar{G} + G' = \bar{x} - \bar{z} + x' - z' \quad (14)$$

199 Similarly, ecosystem respiration equals the sum of its mean and the fluctuation from that mean

$$200 \quad R = \bar{R} + R' = \bar{y} + \bar{z} + y' + z' \quad (15)$$

201 The covariance between G and R is derived from the following steps. First we define the mean of the
 202 product:

$$203 \quad \overline{GR} = \overline{(\bar{x} - \bar{z} + x' - z')(\bar{y} + \bar{z} + y' + z')} \quad (16a)$$

204 Means of fluctuation terms drop out so, with time averaging,so only covariances, variances and means
 205 are retained

$$206 \quad \overline{GR} = \overline{xy} + \overline{xz} - \overline{zy} - \overline{zz} + \overline{x'y'} + \overline{x'z'} - \overline{z'y'} - \overline{z'z'} \quad (16b)$$

207 Next we evaluate the product of the means, $\overline{G}\overline{R}$

$$208 \quad \overline{G}\overline{R} = (\overline{x} - \overline{z})(\overline{y} + \overline{z}) = \overline{xy} + \overline{xz} - \overline{yz} - \overline{zz} \quad (17)$$

209 The difference between $\overline{GR}$ and $\overline{G}\overline{R}$ yields the covariance $\overline{G'R'}$:

$$210 \quad \overline{G'R'} = \overline{GR} - \overline{G}\overline{R} = \overline{x'y'} + \overline{x'z'} - \overline{y'z'} - \overline{z'z'} \quad (18)$$

211 Using similar algebraic steps the variance of G is defined as:

$$212 \quad \overline{G'G'} = \overline{GG} - \overline{G}\overline{G} = \overline{x'x'} - 2\overline{x'z'} + \overline{z'z'} \quad (19)$$

213 And the variance of R becomes

$$214 \quad \overline{R'R'} = \overline{RR} - \overline{R}\overline{R} = \overline{y'y'} + 2\overline{y'z'} + \overline{z'z'} \quad (20)$$

215 Finally, from algebraic substitution we can define the correlation coefficient between G and R as:

216

$$217 \quad r_{GR} = \frac{\overline{G'R'}}{\sigma_G \sigma_R} = \frac{\overline{x'y'} + \overline{x'z'} - \overline{y'z'} - \overline{z'z'}}{(\overline{x'x'} - 2\overline{x'z'} + \overline{z'z'})^{1/2} (\overline{y'y'} + 2\overline{y'z'} + \overline{z'z'})^{1/2}} \quad (21)$$

218 Following Pearson (1896), we can derive, from Eq. 21, the spurious-correlation coefficient, r_{sc} . This limit
 219 occurs when the covariances between x and y, x and z and y and z are zero.

$$220 \quad r_{sc} = \frac{-\overline{z'z'}}{(\overline{x'x'} + \overline{z'z'})^{1/2} (\overline{y'y'} + \overline{z'z'})^{1/2}} \quad (22)$$

221

222 Equation 22 is slightly different from Pearson's equation in that we solved for spurious-correlation
 223 between x - z and y + z, while Pearson derived his equation for comparisons between x/z and y/z.

223

224 On inspection of Eq. 22, one can infer that the spurious-correlation coefficient approaches one, in
 225 magnitude, when the variance in the shared variable, z, far exceeds the variance associated with x and y.
 Conversely, spurious correlation is small with the variance in z is small compared to variance in x and y.

Materials and Methods

We used two sets of data in the ensuing analysis. The first dataset is a test case that comes from our measurements of carbon dioxide, water and energy exchange over an alfalfa (*Medicago sativa*, L.) field in the Sacramento-San Joaquin Delta of California, on Twitchell Island (latitude: 38.1159 N; longitude: 121.64666W; elevation: -9 m). The flux system consisted of a three dimensional sonic anemometer (Gill Windmaster Pro, model 1590) mounted at 2.8 m. CO₂ and water vapor fluctuations were measured with an open path infrared gas analyzer (Licor, 7500A), sampled at 20 Hz. The gas analyzer was calibrated on monthly intervals with calibration gases that are linked to the world standards maintained by the NOAA Environmental System Research Laboratory standards (Zhao and Tans, 2006). Canopy phenology and crop management were monitored with a digital camera; we extracted information on greenness index using digital values from red, green and blue (Richardson et al., 2007). Meteorological (four band components of the net radiation budget, air temperature and relative humidity, pressure) and soil (temperature, soil moisture, water table depth) conditions were characterized with standard set of sensors; details are reported in Knox et al. (2014).

This field site and location are nearly ideal for micrometeorological flux measurements and flux interpretation. The alfalfa sits on level terrain, in a windy climate. The fetch in the direction of the prevailing wind (250 degrees) is over 370 m. During much of the year, robust and persistent winds flow through the Carquinez Strait between the cool coast and the warm inland valley; on average wind velocity ranged between 2.5 and 3.5 m s⁻¹ between night and day, respectively. Additional details on turbulence statistics in this region are reported in Detto et al. (2010).

This site is both a strong source and sink for carbon exchange. Respiration rates, at this site, were substantial because the site sits on degraded peat soils with high carbon content (7.4% carbon) (Deverel and Leighton, 2010). The advantage of working with an irrigated alfalfa crop is that irrigation prevents reductions in photosynthetic capacity due to water deficits. And harvesting enabled us to modulate leaf area index and canopy photosynthesis; the alfalfa field was harvested nine times during this study.

The dataset analyzed here consists of 468 days of data, from day 144 in 2013 through day 246 in 2014. Eddy flux measurements are bound to have gaps when the wind direction is directions with limited fetch and instruments are malfunctioning. Applying quality control procedures for stationarity, data spikes, and favorable wind direction (Foken and Wichura, 1996) enabled us to retain 63% of the original measurements. Gap filling was produced using the neural network method, that has gained high approval via previous flux inter-comparison tests (Moffat et al., 2007). To facilitate representativeness, explanatory data were divided into 8-15 natural data clusters via the k-means method. Data used to train, test, and validate the artificial neural network were proportionately sampled from these clusters so that the resultant neural network was not biased toward conditions that have better flux data coverage. To facilitate generalizability of the neural network, several architectures of increasing complexity were tested. The architecture of each neural network was initialized ten times with random starting weights, and the initialization resulting in the lowest mean sampling error is used. The simplest architecture, for which further increases in complexity resulted in less than a 5% reduction in mean standard error, was chosen and the prediction was saved. This procedure was repeated with twenty re-samplings of the

data. The median prediction was used to fill each gap and the spread of the predictions indicates a measure of gap-filling uncertainty.

NEE was gap-filled separately for daytime and nighttime data. The neural network predictions resulting from the nighttime gap-filling were used to model ecosystem respiration for all periods (daytime and night-time). Canopy photosynthesis was then calculated as the difference between the gap-filled NEE time series and modeled respiration. Explanatory variables used to gap-fill daytime NEE included air temperature, photosynthetically active radiation, vapor pressure deficit, soil water content at 10 cm depth, friction velocity, cumulative days between mowing events, and sine and cosine functions to represent seasonal changes. Explanatory variables used to gap-fill nighttime NEE and partition R include the same variables as above with the exception of PAR and VPD. Most often, the optimum neural network architectures for daytime NEE gap-filling included 2 hidden layers with respectively 12 and 6 hidden nodes. Optimum network architectures for nighttime NEE gap-filling and partitioning typically included 1 hidden layer with 6 hidden nodes. Performance of the gap-filling and partitioning routines were excellent. Daytime (night-time) neural network predictions yielded a model r^2 of 0.96 (0.76) with a RMSE of 1.82 (0.87) $\mu\text{mol m}^{-2} \text{s}^{-1}$. The standard deviation of the neural network predictions for each half-hourly estimate of ecosystem respiration averaged 0.69 $\mu\text{mol m}^{-2} \text{s}^{-1}$.

The second dataset comes from the 2007 LaThuile version of the FLUXNET data base (Agarwal et al., 2010; Papale et al., 2012). Data from the FLUXNET database use the eddy covariance method to measure fluxes and deploy three-dimensional sonic anemometers and either open or closed path infrared gas analyzers to measure flux covariances. This harmonized dataset filled data gaps and partitioned net carbon fluxes into respiration and photosynthesis components using similar and well-vetted methods (Papale et al., 2006).

Data submitted to the data archive were first pre-screened and checked for obvious problems, like errors in time stamps, unexplained jumps in the variables due to sensor malfunction, spikes, values out of range, and errors in units; this is accomplished by visualization and producing statistical metrics. After the data pass the pre-screening criteria, data files were assigned names based on a standard convention that describes the site name and year. Data columns within the files were checked for having proper names, labels and units. These data files contained gaps and were assigned version 1, and were otherwise unchanged from what was submitted by the data producers. The next data processing step subjected the data to quality flagging, filtering, and filling. Spikes, extreme outliers and non-physical values in the data were flagged. In a subsequent step, data were filtered and bad data were replaced or filled with highly vetted gap-filling methods; we used the 'marginal distribution sampling, MDS' method of Reichstein et al. (2005).

Information on the climate space and functional traits and other attributes of the network are described elsewhere (Baldocchi, 2008; Reichstein et al., 2014). Out of over 900 site-years of available over 200 sites, we used 216 site-years of data from 92 sites in this analysis; see Appendix 1 for details. Datasets chosen had to comply with stringent data quality flags settings and extensive length of the data record (more than 350 days per year of original data, with gaps less than 5 days). Tests of inter-comparability

across the flux network has been reported elsewhere (Schmidt et al., 2012); in general, independent flux systems agreed within 10% with the reference system.

Results

Alfalfa, the test case

Figure 1 shows the seasonal pattern of daily integrated NEE, canopy photosynthesis (G) and ecosystem respiration (R). This is the body of data used in the ensuing analysis. Over one year period, the net carbon exchange was $-269 \text{ gC m}^{-2} \text{ y}^{-1}$, due the balance between $1682 \text{ gC m}^{-2} \text{ y}^{-1}$ of ecosystem respiration and $1964 \text{ gC m}^{-2} \text{ y}^{-1}$ of canopy photosynthesis. The active growing season ranged between days 70 and 314 of the year. During the growing season, ecosystem respiration rates often exceeded $5 \text{ gC m}^{-2} \text{ d}^{-1}$ and canopy photosynthesis was as large as $-15 \text{ gC m}^{-2} \text{ d}^{-1}$, before cuttings.

How well is the net exchange of carbon partitioned into photosynthesis and respiration on a diel basis? Figure 2 shows the mean diel pattern of net CO_2 exchange, canopy photosynthesis and ecosystem respiration measured at the alfalfa field. It is clear, from these data, that nocturnal respiration rates were robust ($>4 \mu\text{mol m}^{-2} \text{ s}^{-1}$) and steady; they did not suffer from the effects of low winds that plague many nocturnal respiration measurements at other FLUXNET sites (Barr et al., 2013; Gu et al., 2005) because friction velocities at night during the growing season were brisk, on average about 0.28 m s^{-1} . Furthermore, the sunny and warm days of California's weather promoted near ideal patterns of daytime photosynthesis. Windy conditions and a strong source of soil respiration from the moist and aerobic peat soils gave us a strong respiration signal, both day and night.

With the requisite pieces of net carbon flux data, we computed the correlation coefficient between photosynthesis and respiration, using Eq. 21, for day-night sampling. For this dataset, the correlation coefficient, r_{GR} , equaled -0.786 and its coefficient of determination, r^2 , equaled 0.618 . To what degree was this high correlation real, or an artifact of spurious correlation with daytime ecosystem respiration? We found that the spurious correlation coefficient, Eq. 22, equaled -0.320 , a value that is 40% of the total correlation coefficient, r_{GR} . The degree of spurious correlation was on the low side (less than one-half) because the variance in z ($1.86 \text{ gC}^2 \text{ m}^{-4} \text{ d}^{-2}$) was much smaller than the variance in x ($13.0 \text{ gC}^2 \text{ m}^{-4} \text{ d}^{-2}$). Further reduction in spurious correlations were not achieved because the variance in daytime respiration was much greater than that associated with nighttime respiration; the variance in y was $0.41 \text{ gC}^2 \text{ m}^{-4} \text{ d}^{-2}$.

To better understand why G and R were highly correlated with one another, we deconstructed Eq. 21 by looking at the covariances associated with terms in the numerator. First, we inspected the x-y covariance by comparing daytime vs. nighttime net carbon fluxes. Figure 3 shows that these two fluxes were poorly correlated with one another ($r^2 = 0.237$). The repeated cycles of alfalfa canopy development between cuttings modulated daytime carbon uptake markedly, in contrast to the relatively high, but less variable sums of nighttime respiration over the highly organic soils. Consequently, this modulation in leaf area index and photosynthetic potential produced a relatively low x-y covariance in the numerator of Eq. 21, yielding fairly independent measurements of day and night net carbon flux densities. Next, we investigated the x-z covariance in Eq. 21, defined by the plot between daily sums of

net carbon exchanges vs. daily sums of daytime ecosystem respiration (Figure 4). Here, too, the scatter was reasonably large, yielding an r^2 equal to 0.402, but there was a tendency for daytime respiration to increase with more negative daytime net carbon uptake.

Finally, we inspected the y-z covariance, in Eq. 21, by plotting the relation between nighttime net carbon exchange and daytime ecosystem respiration (Figure 5). We find that daytime respiration fluxes were, on average, more than twice the nighttime values. As expected, this association yielded the highest coefficient of determination ($r^2 = 0.63$).

The tendency for G and R to be susceptible to spurious correlation, based on day and night sampling, also depends upon magnitude of the variance of the z term, daytime respiration, and its covariance with day and night NEE. To investigate this point further, we examined the correlation between two sets of random numbers, surrogates for NEE_{day} and NEE_{night} . Figure 6 shows that the correlation between G and R, constructed from random numbers is not always zero, as expected. Instead, it is highly sensitive to the ratio between daytime respiration and nighttime NEE. The G-R correlation coefficient approached minus one (-1) as the ratio between daytime R_{eco} and nighttime NEE reached and exceeded one. For smaller values of this ratio, the link with r_{GR} was linear and proportional, approaching zero as the ratio between $R_{eco, day}$ and NEE_{night} declined. This simulation helps us explain why the observed correlation coefficient between G and R over the alfalfa was near -0.8 even though the correlation coefficients between the x-y, x-z and y-z terms shown in Figures 3, 4 and 5 were relatively low; it was because daytime respiration fluxes over alfalfa were more than twice the nighttime values.

FLUXNET, Synthesis

Next, we applied the conditionally sampled correlation theory to a set of data extracted from the FLUXNET synthesis database. First, we computed the correlation coefficients between G and R using day-night sums of NEE (Eq. 21) and the spurious correlation coefficient (Eq 22). Figure 7 shows the probability distribution of correlation coefficients, fitted with the kernel density estimation, a non-parametric measure of statistical distribution. The central tendency of this probability distribution for the G-R correlation coefficient (Eq. 21) was close to minus one (mean = -0.828; median = -0.879; standard deviation = 0.130; 95% confidence interval = ± 0.0189). About 80% of the cumulative distribution (between the 10th and 90th cumulative percentiles) was concentrated between values of -0.65 and -0.92. In comparison, the central tendency of the spurious correlation coefficient was much lower (mean = -0.526; median = -0.5331; standard deviation = 0.166; 95% confidence interval = ± 0.0223), with 80% of its values between -0.29 and -0.76. We do not claim that the G and R values derived from the FLUXNET database do not suffer from spurious correlation, but we do provide evidence that spurious correlation may not be as large as suspected based on previous analyses, and that some sites are much less susceptible to spurious correlations than others.

Sites with the least negative spurious correlation coefficients (-0.15 to -0.3) tended to be associated with temperate deciduous forests in Italy, France, Germany, Tennessee, Indiana and Michigan. Sites with the most negative spurious correlation coefficients (< -0.70) were associated with boreal conifer forests in Finland, Sweden, Russia and Canada.

Discussion

Day-night sampling of net carbon exchange produced moderate values of spurious correlation between G and R for individual site-years. What happens at longer time scales? Is the relationship between annually integrated fluxes of G and R across the FLUXNET network more or less susceptible to spurious correlation?

To answer this set of questions let's first look at the relation between annually integrated G and R (Figure 8). We see strong correlation between these two variables ($r^2 = 0.767$) yielding suspicion that these two variables may experience a high degree of spurious correlation with NEE. Next, we took the liberty to compute the spurious correlation coefficient, using Eq 22, on the assumption that: $X = G = NEE - R$; $Y = R = NEE - G$; and $Z = NEE = R + G$. We found this spurious correlation coefficient equaled -0.157. This value was low because the variance in NEE ($7.10 \text{e}4 \text{ gC}^2 \text{ m}^{-4} \text{ y}^{-2}$) was low compared to the variance in R ($2.15 \text{e}5 \text{ gC}^2 \text{ m}^{-4} \text{ y}^{-2}$) and the variance in G ($6.40 \text{e}5 \text{ gC}^2 \text{ m}^{-4} \text{ y}^{-2}$), giving us renewed confidence on the relative independence between R and G at annual time scales, when examined across the FLUXNET network. This analysis also indicates that the comparison of annually integrated carbon fluxes between sites is less susceptible to spurious correlation than the comparison of daily-integrated fluxes at a single site and why such an analysis is very robust across regions (Chen et al., 2015).

Does this analysis help us constrain estimates of global primary productivity better? The literature consists of estimates based on satellite remote sensing and light use efficiency models (Maisongrande et al., 1995; Zhao et al., 2005), inference with stable isotopes (Welp et al., 2011), dynamic vegetation models (Cramer et al., 1999), land surface models (Bonan et al., 2011; Denning et al., 1996; Ryu et al., 2011) and statistical upscaling with data from flux networks (Beer et al., 2010; Jung et al., 2011). What is unconstrained is which estimate is better, or worse? A survey of the literature (Figure 9) reveals that the bulk of the estimates of globally-integrated G range between 105 and 170 Pg-C y^{-1} , with a mean of $129 \pm 7.5 \text{ Pg-C y}^{-1}$. It is noteworthy that the estimate derived from eddy flux measurements is 123 Pg-C y^{-1} (Beer et al., 2010). With increased confidence on the independence and the direct nature of eddy flux measurements, shown above, we argue that the estimates of global primary productivity associated with the central tendency of this distribution have more likelihood of representing the true value of global primary productivity than those at the lower and upper extremes of this distribution. We also conclude that selective use of eddy covariance data can continue to serve as independent checks of model validation and as inputs for data assimilation (Keenan et al., 2012).

Conclusion

Statistical theory was developed to examine the susceptibility of using eddy covariance measurements of CO_2 exchange to produce unbiased estimates of canopy photosynthesis and ecosystem respiration, using day-night conditional sampling. For an ideal test case, an alfalfa crop in sunny California, we found the correlation between canopy photosynthesis (G) and ecosystem respiration (R) was -0.79, of which the spurious correlation coefficient equaled -0.32. The spurious correlation coefficient was moderate because the variance in daytime respiration was smaller than the variance in daytime NEE, but greater than the variance in nighttime NEE.

Expansion of this analysis to the global FLUXNET dataset revealed that the average correlation coefficient between photosynthesis and respiration was similar to the alfalfa case study (the mean was-

0.83 +/- 0.02). Across the network, the degree of spurious correlation was moderate (-0.53). The lowest degree of spurious correlation was associated with temperate deciduous forests. The highest degree of spurious correlation was associated with boreal evergreen forests. These results support the hypothesis that long periods of dormancy, associated with temperate deciduous forest helps reduce the degree of spurious correlation. This analysis also establishes confidence levels on the degree to which the eddy covariance method can produce relatively unbiased estimates of ecosystem respiration and canopy photosynthesis.

Acknowledgement

This analysis was supported by funding from the US Department of Energy's Office of Science, which supports our flux towers and activities with the AmeriFlux and FLUXNET projects. We also acknowledge support from the California Department of Water Resources. We thank Joe Verfaillie, Sara Knox, Patty Oikawa and Laurie Koteen for assistance in collection of data at the alfalfa site.

This synthesis study used eddy covariance data acquired by the FLUXNET community and in particular by the following networks: AmeriFlux (US Department of Energy, Biological and Environmental Research, Terrestrial Carbon Program (DE-FG02-04ER63917 and DE-FG02-04ER63911). Afriflux, AsiaFlux, CarboAfrica, CarboEuropeIP, Carbo-Italy, CarboMont, ChinaFlux, Fluxnet-Canada (supported by CFCAS, NSERC, BIOCAP, Environment Canada, and NRCan), GreenGrass, KoFlux, LBA, NECC, OzFlux, TCOS-Siberia, USCC. We acknowledge the financial support to the eddy covariance data harmonization provided by CarboEuropeIP, FAO-GTOS-TCO, ILEAPS, Max Planck Institute for Biogeochemistry, National Science Foundation, University of Tuscia, Universite Laval, Environment Canada and the US Department of Energy and the database development and technical support from the Berkeley Water Center, Lawrence Berkeley National Laboratory, Microsoft eScience, Oak Ridge National Lab, University of California, Berkeley and the University of Virginia.

Finally, we thank the many scientists, technicians and students who contributed to this data system. It is through the open access and sharing of flux data, collected with public funding, that synthesis projects are able to be accomplished and shed new light on processes at different time and space scales.

References

- Acevedo, O.C., Moraes, O.L.L., Fitzjarrald, D.R., Sakai, R.K. and Mahrt, L., 2007. Turbulent carbon exchange in very stable conditions. *Boundary-Layer Meteorology*, 125(1): 49-61.
- Agarwal, D.A. et al., 2010. A data-centered collaboration portal to support global carbon-flux analysis. *Concurrency and Computation: Practice and Experience*, 22(17): 2323-2334.
- Alton, P.B., 2011. How useful are plant functional types in global simulations of the carbon, water, and energy cycles? *Journal of Geophysical Research*, 116(G1).
- Amthor, J.S., 1994. Scaling CO_2 -Photosynthesis Relationships from the Leaf to the Canopy. *Photosynthesis Research*, 39(3): 321-350.
- Asaf, D. et al., 2013. Ecosystem photosynthesis inferred from measurements of carbonyl sulphide flux. *Nat. Geosci.*, 6(3): 186-190.
- Baldocchi, D.D., 2003. Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past, present and future. *Global Change Biol*, 9: 479-492.

- Baldocchi, D.D., 2008. TURNER REVIEW No. 15. 'Breathing' of the terrestrial biosphere: lessons learned from a global network of carbon dioxide flux measurement systems. *Australian Journal of Botany*, 56(1): 1-26.
- Baldocchi, D.D., Verma, S.B. and Anderson, D.E., 1987. Canopy photosynthesis and water use efficiency in a deciduous forest. *Journal Applied Ecology*, 40: 127-146.
- Baldocchi, D.D., Vogel, C.A. and Hall, B., 1997. Seasonal variation of carbon dioxide exchange rates above and below a boreal jack pine forest. *Agricultural and Forest Meteorology*, 83(1-2): 147-170.
- Barr, A.G. et al., 2013. Use of change-point detection for friction–velocity threshold evaluation in eddy-covariance studies. *Agricultural and Forest Meteorology*, 171-172: 31-45.
- Beer, C. et al., 2010. Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate. *Science*, 329(5993): 834-8.
- Berry, J. et al., 2013. A coupled model of the global cycles of carbonyl sulfide and CO₂: A possible new window on the carbon cycle. *Journal of Geophysical Research: Biogeosciences*: n/a-n/a.
- Blonquist, J.M. et al., 2011. The potential of carbonyl sulfide as a proxy for gross primary production at flux tower sites. *Journal of Geophysical Research*, 116(G4).
- Bonan, G.B. et al., 2011. Improving canopy processes in the Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data. *Journal of Geophysical Research*, 116(G2).
- Bowling, D.R., Tans, P.P. and Monson, R.K., 2001. Partitioning net ecosystem carbon exchange with isotopic fluxes of CO₂. *Global Change Biology*, 7(2): 127-145.
- Chen, Z. et al., 2015. Covariation between gross primary production and ecosystem respiration across space and the underlying mechanisms: A global synthesis. *Agricultural and Forest Meteorology*, 203(0): 180-190.
- Cox, P.M., Betts, R.A., Jones, C.D., Spall, S.A. and Totterdell, I.J., 2000. Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature*, 408(6813): 184-187.
- Cramer, W. et al., 1999. Comparing global models of terrestrial net primary productivity (NPP): overview and key results. *Global Change Biol*, 5: 1 - 15.
- Denning, A.S. et al., 1996. Simulations of terrestrial carbon metabolism and atmospheric CO₂ in a general circulation model .1. Surface carbon fluxes. *Tellus Series B-Chemical and Physical Meteorology*, 48(4): 521-542.
- Detto, M., Baldocchi, D. and Katul, G.G., 2010. Scaling Properties of Biologically Active Scalar Concentration Fluctuations in the Atmospheric Surface Layer over a Managed Peatland. *Boundary-Layer Meteorology*, 136(3): 407-430.
- Deverel, S.J. and Leighton, D.A., 2010. Historic, Recent, and Future Subsidence, Sacramento-San Joaquin Delta, California, USA. *San Francisco Estuary and Watershed Science*, 8(2).
- Dietze, M.C. et al., 2011. Characterizing the performance of ecosystem models across time scales: A spectral analysis of the North American Carbon Program site-level synthesis. *Journal of Geophysical Research-Biogeosciences*, 116.
- Falge, E. et al., 2002. Seasonality of ecosystem respiration and gross primary production as derived from FLUXNET measurements. *Agricultural and Forest Meteorology*, 113(1-4): 53-74.
- Finnigan, J., 2006. The storage term in eddy flux calculations. *Agricultural and Forest Meteorology*, 136(3-4): 108-113.
- Foken, T. and Wichura, B., 1996. Tools for quality assessment of surface-based flux measurements. *Agricultural and Forest Meteorology*, 78(1-2): 83-105.
- Frankenberg, C. et al., 2011. New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity. *Geophysical Research Letters*, 38(17): n/a-n/a.

- Friedlingstein, P. et al., 2006. Climate-carbon cycle feedback analysis: Results from the (CMIP)-M-4 model intercomparison. *Journal of Climate*, 19(14): 3337-3353.
- Griffis, T.J., 2013. Tracing the flow of carbon dioxide and water vapor between the biosphere and atmosphere: A review of optical isotope techniques and their application. *Agricultural and Forest Meteorology*, 174–175(0): 85-109.
- Gu, L. et al., 2005. Objective threshold determination for nighttime eddy flux filtering. *Agricultural and Forest Meteorology*, 128(3-4): 179-197.
- Guanter, L. et al., 2014. Global and time-resolved monitoring of crop photosynthesis with chlorophyll fluorescence. *Proceedings of the National Academy of Sciences*, 111(14): E1327-E1333.
- Heskel, M.A., Atkin, O.K., Turnbull, M.H. and Griffin, K.L., 2013. Bringing the Kok effect to light: A review on the integration of daytime respiration and net ecosystem exchange. *Ecosphere*, 4(8): art98.
- Hicks, B.B., 1978. Some limitations of dimensional analysis and power laws. *Boundary Layer Meteorology*, 14: 567-569.
- Jung, M. et al., 2011. Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *Journal of Geophysical Research*, 116.
- Keenan, T.F., Davidson, E., Moffat, A.M., Munger, W. and Richardson, A.D., 2012. Using model-data fusion to interpret past trends, and quantify uncertainties in future projections, of terrestrial ecosystem carbon cycling. *Global Change Biology*, 18(8): 2555-2569.
- Kenney, B.C., 1982. Beware of Spurious Self-Correlations. *Water Resources Research*, 18(4): 1041-1048.
- Knox, S.H. et al., 2014. Agricultural peatland restoration: effects of land-use change on greenhouse gas (CO₂ and CH₄) fluxes in the Sacramento-San Joaquin Delta. *Global Change Biology*: n/a-n/a.
- Kuzyakov, Y., 2010. Priming effects: Interactions between living and dead organic matter. *Soil Biology and Biochemistry*, 42(9): 1363-1371.
- Lasslop, G., Reichstein, M., Detto, M., Richardson, A.D. and Baldocchi, D.D., 2010a. Comment on Vickers et al.: Self-correlation between assimilation and respiration resulting from flux partitioning of eddy-covariance CO₂ fluxes. *Agricultural and Forest Meteorology*, 150(2): 312-314.
- Lasslop, G. et al., 2010b. Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: critical issues and global evaluation. *Global Change Biology*, 16(1): 187-208.
- Lee, X., 2000. Air motion within and above forest vegetation in non-ideal conditions. *Forest Ecology and Management*, 135(1-3): 3-18.
- Loew, A. et al., 2014. Do we (need to) care about canopy radiation schemes in DGVMs? Caveats and potential impacts. *Biogeosciences*, 11(7): 1873-1897.
- Mahrt, L., 2014. Stably Stratified Atmospheric Boundary Layers. In: S.H. Davis and P. Moin (Editors), *Annual Review of Fluid Mechanics*, Vol 46. *Annual Review of Fluid Mechanics*. Annual Reviews, Palo Alto, pp. 23-45.
- Maisongrande, P., Ruimy, A., Dedieu, G. and Saugier, B., 1995. Monitoring Seasonal and Interannual Variations of Gross Primary Productivity, Net Primary Productivity and Net Ecosystem Productivity Using a Diagnostic Model and Remotely-Sensed Data. *Tellus Series B-Chemical and Physical Meteorology*, 47(1-2): 178-190.
- Moffat, A.M. et al., 2007. Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes. *Agricultural and Forest Meteorology*, 147(3-4): 209-232.
- Nappo, C.J., 1991. Sporadic Breakdowns of Stability in the Pbl over Simple and Complex Terrain. *Boundary-Layer Meteorology*, 54(1-2): 69-87.
- Papale, D. et al., 2012. Database Maintenance, Data Sharing Policy, Collaboration. In: M. Aubinet (Editor), *Eddy Covariance: A Practical Guide to Measurement and Data Analysis*. Springer.

- Papale, D. et al., 2006. Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation. *Biogeosciences*, 3(4): 571-583.
- Pearson, K., 1896. Mathematical Contributions to the Theory of Evolution.--On a Form of Spurious Correlation Which May Arise When Indices Are Used in the Measurement of Organs. *Proceedings of the Royal Society of London*, 60(359-367): 489-498.
- Reichstein, M., Bahn, M., Mahecha, M.D., Kattge, J. and Baldocchi, D.D., 2014. Linking plant and ecosystem functional biogeography. *Proceedings of National Academy of Science*.
- Reichstein, M. et al., 2005. On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. *Global Change Biology*, 11(9): 1424-1439.
- Reynolds, O., 1895. On the dynamical theory of incompressible viscous fluids and the determination of the criterion. *Philosophical Transactions of the Royal Society of London, A*, 186: 123-164.
- Richardson, A.D. et al., 2007. Use of digital webcam images to track spring green-up in a deciduous broadleaf forest. *Oecologia*, 152(2): 323-34.
- Ryu, Y. et al., 2011. Integration of MODIS land and atmosphere products with a coupled-process model to estimate gross primary productivity and evapotranspiration from 1 km to global scales. *Global Biogeochemical Cycles*, 25(4): n/a-n/a.
- Schaefer, K. et al., 2012. A model-data comparison of gross primary productivity: Results from the North American Carbon Program site synthesis. *Journal of Geophysical Research*, 117(G3).
- Schmidt, A., Hanson, C., Chan, W.S. and Law, B.E., 2012. Empirical assessment of uncertainties of meteorological parameters and turbulent fluxes in the AmeriFlux network. *Journal of Geophysical Research*, 117(G4).
- Stull, R.B., 1988. *Introduction to Boundary Layer Meteorology*. Kluwer, Dordrecht, The Netherlands, 666 pp.
- Tang, J., Baldocchi, D.D. and Xu, L., 2005. Tree photosynthesis modulates soil respiration on a diurnal time scale. *Global Change Biology*, 11(8): 1298-1304.
- van Gorsel, E. et al., 2009. Estimating nocturnal ecosystem respiration from the vertical turbulent flux and change in storage of CO₂. *Agricultural and Forest Meteorology*, 149(11): 1919-1930.
- Vargas, R. et al., 2013. Drought Influences the Accuracy of Simulated Ecosystem Fluxes: A Model-Data Meta-analysis for Mediterranean Oak Woodlands. *Ecosystems*: 1-16.
- Vickers, D., Thomas, C.K., Martin, J.G. and Law, B., 2009. Self-correlation between assimilation and respiration resulting from flux partitioning of eddy-covariance CO₂ fluxes. *Agricultural and Forest Meteorology*, 149(9): 1552-1555.
- Welp, L.R. et al., 2011. Interannual variability in the oxygen isotopes of atmospheric CO₂ driven by El Niño. *Nature*, 477(7366): 579-82.
- Williams, M. et al., 2009. Improving land surface models with FLUXNET data. *Biogeosciences* J1 - BG, 6(7): 1341-1359.
- Xu, L. and Baldocchi, D.D., 2004. Seasonal variation in carbon dioxide exchange over a Mediterranean annual grassland in California. *Agricultural and Forest Meteorology*, 123(1-2): 79-96.
- Zhao, C.L. and Tans, P.P., 2006. Estimating uncertainty of the WMO mole fraction scale for carbon dioxide in air. *Journal of Geophysical Research*, 111(D8).
- Zhao, M., Heinsch, F.A., Nemani, R.R. and Running, S.W., 2005. Improvements of the MODIS terrestrial gross and net primary production global data set. *Remote Sensing of Environment*, 95(2): 164-176.

List of Figures

Figure 1. Sums of daily integrated net and gross carbon fluxes for net ecosystem exchange (NEE), canopy photosynthesis (G) and ecosystem respiration (R). Day of year refers to after January 1, 2013.

Figure 2. Mean diel pattern of hourly bin-averaged carbon fluxes, net ecosystem exchange (NEE), canopy photosynthesis (G) and ecosystem respiration (R).

Figure 3. Correlation between sums of measurements of net ecosystem exchange measured during the daytime and nighttime periods, representing the x-y terms in Equation 21. These data are from the alfalfa site.

Figure 4. Correlation between sums of measurements of net ecosystem exchange measured during the daytime and estimates of daytime ecosystem respiration, representing the x-z term in Equation 21. These data are from the alfalfa site.

Figure 5. Correlation between nighttime measurements of net carbon exchange and derived estimates of daytime ecosystem respiration, representing the y-z term in Equation 21. These data are from the alfalfa site.

Figure 6. Simulation of the correlation coefficient between G and R using two sets of random number and Eq. 21. We prescribed the daytime respiration flux as a function of the nighttime NEE measurement.

Figure 7. Probability distribution of the correlation between canopy photosynthesis (G) and ecosystem respiration (R) based on the FLUXNET dataset using day-night sampling (eq 21) and the coefficient of spurious correlation (eq. 22). The histograms were fitted with the kernel density estimate to provide a smoothed fit through the data. Correlation coefficients were computed based on variances and covariances from each site-year using the summation of day and night samples.

Figure 8. Comparison of annual sums of canopy photosynthesis (G) and ecosystem respiration (R) from the FLUXNET database. This plot is derived from 216 site years of data. R^2 is 0.76, slope is -0.713 and intercept is $153 \text{ gC m}^{-2} \text{ y}^{-1}$.

Figure 9. Histogram of estimates of global primary productivity from a survey of published papers in the literature.

Fig. 1

Figure 1. Sums of daily integrated net and gross carbon fluxes for net ecosystem exchange (NEE), canopy photosynthesis (G) and ecosystem respiration (R). Day of year refers to after January 1, 2013.

648

649 Gr2

Figure 2. Mean diel pattern of hourly bin-averaged carbon fluxes, net ecosystem exchange (NEE), canopy photosynthesis (G) and ecosystem respiration (R).

Fig. 3

Fig. 4

Fig. 5

Fig. 6

Fig. 7

Fig. 8

Fig. 9

Figure 9. Histogram of estimates of global primary productivity from a survey of published papers in the literature.