Toward a consistency cross-check of eddy covariance flux–based and biometric estimates of ecosystem carbon balance

S. Luysaert,1 M. Reichstein,2 E.-D. Schulze,2 I. A. Janssens,1 B. E. Law,3 D. Papale,4 D. Dragoni,5 M. L. Goulden,6 A. Granier,7 W. L. Kutsch,2 S. Linder,8 G. Matteucci,9 E. Moore,10 J. W. Munger,11 K. Pilegaard,12 M. Saunders,13 and E. M. Falge14

Received 15 September 2008; revised 30 March 2009; accepted 23 April 2009; published 18 July 2009.

[1] Quantification of an ecosystem’s carbon balance and its components is pivotal for understanding both ecosystem functioning and global cycling. Several methods are being applied in parallel to estimate the different components of the CO2 balance. However, different methods are subject to different sources of error. Therefore, it is necessary that site level component estimates are cross-checked against each other before being reported. Here we present a two-step approach for testing the accuracy and consistency of eddy covariance–based gross primary production (GPP) and ecosystem respiration (Re) estimates with biometric measurements of net primary production (NPP), autotrophic (Ra) and heterotrophic (Rh) respiration. The test starts with closing the CO2 balance to account for reasonable errors in each of the component fluxes. Failure to do so within the constraints will classify the flux estimates on the site level as inconsistent. If the CO2 balance can be closed, the test continues by comparing the closed site level Ra/GPP with the Rh/GPP ratio. The consistency of these ratios is then judged against expert knowledge. Flux estimates of sites that pass both steps are considered consistent. An inconsistent ratio is not necessarily incorrect but provides a signal for careful data screening that may require further analysis to identify the possible biological reasons of the unexpected ratios. We reviewed the literature and found 16 sites, out of a total of 529 research forest sites, that met the data requirements for the consistency test. Thirteen of these sites passed both steps of the consistency cross-check. Subsequently, flux ratios (NPP/GPP, Rh/NPP, Rh/Re, and Re/GPP) were calculated for the consistent sites. Similar ratios were observed at sites which lacked information to check consistency, indicating that the flux data that are currently used for validating models and testing ecological hypotheses are largely consistent across a wide range of site productivities. Confidence in the output of flux networks could be further enhanced if the required fluxes are independently estimated at all sites for multiple years and harmonized methods are used.


1. Introduction

[2] Quantification of an ecosystem’s carbon balance and its components is pivotal for understanding both ecosystem functioning and the global carbon cycle. Well defined measurement protocols for carbon cycle component analysis such as inventory and chamber based fluxes have been widely available since the 1960s through the International Biological Program (IBP) [DeAngelis et al., 1981], however technological developments such as the eddy covariance technique provide high temporal estimates of ecosystem-scale carbon dioxide (CO2), water vapor and energy fluxes [Baldocchi et al., 1988].

1Department of Biology, University of Antwerp, Wilrijk, Belgium.
2Max Planck Institute for Biogeochemistry, Jena, Germany.
3College of Forestry, Oregon State University, Corvallis, Oregon, USA.
4Department of Forest Environment Science and Resource, University of Tuscia, Viterbo, Italy.
5Department of Geography, Indiana University, Bloomington, Indiana, USA.
6Department of Ecology and Evolutionary Biology, University of California, Irvine, California, USA.
7Forest Ecology and Ecophysiology Unit, UMR INRA UHP, Champenoux, France.
8Southern Swedish Forest Research Centre, SLU, Alnarp, Sweden.
9CNR-ISAOFM, Rende, Italy.
10ESS-CC, Alterra Wageningen UR, Wageningen, Netherlands.
11Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts, USA.
12Riso National Laboratory for Sustainable Energy, Biosystems Division, Technical University of Denmark, Roskilde, Denmark.
13School of Biology and Environmental Science, University College Dublin, Dublin, Ireland.
14Max Planck Institute for Chemistry, Mainz, Germany.
Delay the number of eddy covariance observation sites has increased, leading to the development of regional networks, such as AmeriFlux and EuroFlux, and the global network FLUXNET, a “network of regional networks” [Baldocchi, 2008]. From the beginning these networks allowed researchers to address a variety of important scientific questions concerning the driving forces of ecosystem productivity and the response of the biosphere to climatic events [Baldocchi et al., 2001; Ciais et al., 2005; Goulden et al., 1998; Law et al., 2002; Luyssaert et al., 2008; Magnani et al., 2007; Reichstein et al., 2007; Valentini et al., 2000].

However, with the eddy covariance technique only observations of net ecosystem exchange are possible, while other quantities such as gross primary production and ecosystem respiration have to be inferred indirectly. Biometric methods are able to observe carbon balance components more or less directly, but not with such high temporal frequency and continuity as the eddy covariance technique. Both the biometric and eddy covariance methods are subject to different error sources and uncertainties irrespective of the methodology applied [Belelli Marchesini et al., 2007; Clark et al., 2001a; Curtis et al., 2002; Gough et al., 2008; Kruijt et al., 2004; Law et al., 2001b; Miller et al., 2004; Moncrieff et al., 1996; Vickers et al., 2009a; Vickers et al., 2009b]. Therefore, it is valuable to produce both biometric and eddy covariance–based estimates at each site for both diagnostic purposes and to understand better the flux components.

The validity of eddy covariance–based NEP estimates can be assessed by checking energy balance closure [Foken, 2008; Twine et al., 2000]. If all energy balance terms are accounted for (net radiation, latent heat, sensible heat and storage fluxes in vegetation, air and soil), and energy balance closure is achieved, this indicates that the CO2 flux measurements are of high quality because (1) the measurements are consistent with the principle of conservation of energy and (2) the daytime latent heat and CO2 fluxes are largely controlled by the same processes, i.e., conductance of the surface. This methodology however generally only works well during daytime hours and therefore only provides confidence that daytime fluxes are accurately accounted for.

Also, eddy covariance–based NEP has been tested against its biometric analog [Curtis et al., 2002; Gough et al., 2008; Law et al., 2001b; Miller et al., 2004]. GPP and Re estimates derived from eddy covariance measurements could be tested against independent estimates of these flux densities for example from chamber measurements [Griffis et al., 2004; Tang et al., 2008; Zha et al., 2007]. This approach is straightforward but strictly methodological. It allows for testing the consistency of eddy covariance–based estimates of NEP, GPP and Re but adds little to our understanding of the CO2 balance of the ecosystem under study unless ecosystem models are being applied.

The objectives of this study are the following: (1) to present a framework for testing the accuracy and consistency of eddy covariance–based GPP and Re estimates with biometric measurements of NPP, Ra, and Rh (see notation section for definitions) and (2) to show the framework’s possibilities, limitations, and potential pitfalls.

2. What Is in a Name?

NEP is defined as the CO2 flux measured by eddy covariance [Chapin et al., 2006] and is often partitioned in its underlying ecological processes, i.e., GPP and Re. The sign convention differs on whether the atmosphere (NEP < 0 denotes a sink) or ecosystem (NEP < 0 denotes a source) is used as the reference. In this study the ecosystem is used as the reference. The different components of ecosystem productivity are related to each other according to three theoretical identities:

\[
Re = Ra + Rh, \quad (1)
\]

\[
NPP = GPP - Ra, \quad (2)
\]

\[
NEP = NPP - Rh = GPP - Re, \quad (3)
\]

where GPP is gross primary production, NPP is net primary production, Ra is autotrophic respiration, NEP is net ecosystem production, Rh is heterotrophic respiration and Re is ecosystem respiration.

The simplicity of the theoretical identities (1)–(3) contrasts with the complexity of measuring these fluxes in the field. Besides NEP, none of the carbon fluxes are physically well defined. In addition, different methods measuring different ecosystem components are being used to estimate the same flux identity thus resulting in different quantities which are, however, addressed by the same terminology. Below we summarize some known methodological issues.

2.1. Eddy Covariance–Based NEP, GPP, and Re

The physical basis for the eddy covariance method has been formalized as a three-dimensional mass conservation equation [Baldocchi et al., 1988]. The CO2 flux between terrestrial ecosystems and the atmosphere equals the sum of (1) rate of CO2 concentration change under the sensor, (2) three-dimensional turbulent transport, and (3) advective transport of CO2.

Under ideal conditions, i.e., a flat terrain covered by a horizontally homogeneous vegetation, the conservation equation can be simplified to components 1 and 2 by assuming the absence of advective and horizontal turbulent transport [Baldocchi et al., 1988]. Typically this simplified experimental setup, is used to estimate the NEP, however, at most forest sites, the ideal conditions do not occur, which may result in biased flux estimates. Horizontal turbulent transport and advection are difficult to measure and currently the most used method to correct the data for advective transport is to filter out measurements acquired at low-turbulence conditions (u* filtering) [Aubinet et al., 2005] although it has been shown that in some cases this correction
is not enough and the magnitude of the bias due to advection remains unknown unless measured [Feigenwinter et al., 2008; Yi et al., 2008].

[12] Half-hourly eddy covariance measurements of NEP can be used to derive an estimate of Re and GPP. At night there is no photosynthesis, so the site-specific relationship between the nighttime NEP and soil or air temperature can be used to estimate the half-hourly respiration during the day given the daytime soil or air temperature. However, because of below canopy CO$_2$ storage and advection, nocturnal NEP measured on calm nights ($u^*$ threshold) is not used to estimate Re. Only data from turbulent episodes are used to fit a relationship between nighttime NEP and air or soil temperature, from which daytime respiration is then estimated. The relationship can be fitted with constant [Falge et al., 2001] or variable parameter values [Reichstein et al., 2005].

[13] Under similar environmental conditions, daytime foliar respiration is less than that at nighttime. This reduction of leaf respiration in the light can be caused by inhibition of nonphotorespiratory CO$_2$ release in the light [Atkin et al., 1998; Brooks and Farquhar, 1985], or by refixation of respired CO$_2$ [Pinelli and Loreto, 2003]. If light-induced inhibition of leaf respiration is the main cause, then extrapolating nighttime NEP measurements to daytime (to estimate daytime respiration in the presence of photosynthesis) results in overestimated daytime and annual respiration rates. However, if soil respiration is the dominant respiration term (typically 60% of Re), the error in using the temperature response function to estimate GPP is correspondingly reduced. Because GPP is typically estimated as the difference between NEP and Re, any error in Re is automatically transferred to GPP. Janssens et al. [2001] and Wohlfart et al. [2005] estimated the upper limit of this overestimation of Re at about 15%. In contrast, if refixation of respired CO$_2$ is the sole cause of the reduced daytime foliar respiration, then Re is not really reduced and hence not overestimated by extrapolating nighttime regressions to daytime periods.

[14] Also, time series are rarely complete because of malfunctioning, maintenance and calibration of the instrumentation and poor instrument performance under certain climatic conditions (i.e., precipitation, dew and low wind speeds). Thus, studies of the carbon budget often require that data gaps are filled so that flux estimates can be cumulated over longer periods, i.e., days, weeks, months or years [Falge et al., 2001; Moffat et al., 2007]. The reliability of the gap filling decreases with increasing gap length [Moffat et al., 2007] especially for gaps occurring during periods of canopy development and senescence. Hence, with gap filling the uncertainty of flux estimates integrated over longer time periods increases [Moffat et al., 2007].

2.2 Biometric NPP

[15] The bulk of NPP is allocated to the production of biomass in different plant components: foliage, wood (including branches and stems), roots (including coarse and fine roots). However, because not all of the biomass produced remains on site, direct measurements of total NPP are very difficult and removed biomass in the year of interest needs to be included. Examples of biomass removal processes include harvest and natural herbivory by insects, mammals but also nematodes that feed on mycorrhizae.

[16] In addition, other components of NPP that are rarely measured include understory plant growth, reproductive organs, carbon lost through the emission of volatile organic compounds (VOC), exudation from roots and carbon transfer to root symbionts. The sum of biomass production, biomass losses/exports and the above mentioned components represents the total ecosystem NPP. However, some of these components are difficult to measure and of minor importance [Clark et al., 2001a]. Typically, these components are ignored when estimating the NPP. This bias is propagated in the Ra estimate when Ra is calculated from equation (2).

[17] As a consequence of ignoring some of the NPP components, biomass-based NPP differs from the flux-based NPP (estimated as the difference between the eddy covariance—derived GPP and chamber-based Ra). The biomass-based NPP is derived from plant growth measurements and correctly excludes the export of carbohydrates to mycorrhizae and soil microorganisms that consume mycorrhizae.

In contrast, the flux-based NPP includes these exports as part of Ra if for example Ra is measured by trenching and girdling techniques [e.g., Högberg et al., 2001].

2.3 Ra and Rh

[18] The flux identities assume that autotrophic and heterotrophic organisms and their respiration are well separated, but in reality, they are interrelated. Consequently, the theoretical separation between auto and heterotrophic respiration does not reflect the physiological reality [Högberg and Read, 2006], yet it is a necessity for logistical and mechanistic reasons.

[19] The most common group of methods used to separate soil respiration into both autotrophic and heterotrophic components include the exclusion of recent photosynthate supply to soil organisms by root trenching or tree girdling [Subke et al., 2006]. However, removing the supply of recent photosynthates also causes changes in the taxonomic diversity of soil organisms [Högberg et al., 2001; Remén et al., 2008; Schulze et al., 2005]. Consequently, soil respiration measurements on trenched or girdled plots represent a different belowground ecosystem than the measurements obtained from the undisturbed plot, particularly over the long-term. This bias is propagated in the Ra estimate when Ra is calculated from equation (1). Further, standard portable soil respiration chambers are too small to include coarse and fine woody debris, hence, Rh estimates often lack decomposition of woody debris.

[20] A general concern when using identities 1 to 3 is that different methods have different spatial and temporal scales, i.e., the eddy covariance—derived estimate of Re represents for a flux that is indicative of a multiple hectare footprint whereas chamber based estimates of Ra typically represents a much smaller area. When these estimates are used to calculate, i.e., Rh (identity 1), the spatial representativeness of Rh will be ill-defined. Similar issues exist for the other component fluxes.

3. Materials and Methods

3.1 Flux Density Estimates

3.1.1 Data Used in This Study

[21] A comprehensive database was designed to store information on CO$_2$ fluxes, ecosystem properties, and site infor-
mation of forest stands [Luyssaert et al., 2007]. Data were sourced from peer-reviewed literature, established databases [e.g., Olson et al., 2001; Papale et al., 2006] and personal communications with research groups involved in regional networks (e.g., AmeriFlux, AsiaFlux, CarboEurope-IP, ChinaFlux, Fluxnet-Canada, NECC, TCOS-Siberia, USCCC), and the Fluxnet project [Baldocchi et al., 2001].

[22] For this study, the database was queried for biometric-based NPP, chamber-based Rs, Rh and Ra and eddy covariance-based NEP, GPP and Re estimates and their uncertainties. Circularity in the consistency check was avoided by excluding flux estimates from methods (i.e., based on identities 1–3) other than those listed below.

[23] Biometric estimates of NPP were included in the analysis when they were based on direct measurements of the main components of NPP [Clark et al., 2001a] and when foliage, stem, branch, coarse and fine root biomass increment were corrected for the annual litterfall of these components. Furthermore, to be included in this study, the following criteria had to be met: (1) the net annual production of leaves or needles was determined by collecting leaf/needle fall throughout the year or by direct assessment of stand leaf biomass in deciduous forests; (2) annual stem and branch increment were determined using biomass harvesting and site-based allometry, species- and region-specific equations relating aboveground woody biomass increment to the change in basal area of individual trees in the plot; (3) coarse root production was determined through species- and region-specific allometric equations relating root mass to basal area; and (4) fine root production was determined by repeated soil coring, isotopic estimates of fine root turnover combined with biomass measurements, upscaled root length production observed in minirhizotrons or the soil respiration and litterfall constraint formulated by Raich and Nadelhoff [1989].

[24] Estimates of Rs and its heterotrophic component Rh were included when based on subtracting upscaled chamber measurements of root respiration [Hanson et al., 2000] from undisturbed Rs measurements or from chamber measurements after trenching or girdling. Also Rh estimates based on isotopes or soil C mineralization [Persson et al., 2000] were included. Directly measured estimates of Ra were included in this study when the estimate was based on upscaled chamber measurements of foliage, stem and root respiration [e.g., see Law et al., 1999; Ryan et al., 1996].

[25] Measurements of annual and multiple-year NEP and the derived GPP and Re estimates were included in the database when based on continuous measurements with a tower-based eddy covariance system. Estimates were accepted when data gaps due to system failure, stable atmospheric conditions or data rejection were filled by means of standardized methods [Falge et al., 2001; Reichstein et al., 2005] to provide complete data sets. We used data sets where gaps covered less than 30% of the data set and adopted the u*-thresholds used in the original publications or harmonized database.

3.1.2. Uncertainty of the Measured CO₂ Fluxes

[26] Recently efforts have been made to quantify the uncertainties of eddy covariance measurements [Black et al., 2007; Dragoni et al., 2007; Hollinger et al., 2004; Hollinger and Richardson, 2005; Richardson et al., 2006; Vickers et al., 2009a] however, publication of the magnitude of uncertainty surrounding CO₂ flux estimates were only rarely reported in the literature and when reported it is often unclear whether the given value denotes instrumental, spatial, temporal and/or other sources of variability. Therefore, we did not use the reported uncertainty and instead estimated the total uncertainty for every flux component contained in the database. The uncertainty was estimated in a uniform way based on expert judgment [Taylor and Kuyatt, 1994]. A detailed description of the uncertainty framework is given in Luyssaert et al. [2007].

[27] The total uncertainty of the flux estimate depends on both the methodology used and its random component which is reciprocal to the length of the time series. Our uncertainty framework resulted in 95% confidence intervals ranging between 105 and 350 g C m⁻² a⁻¹ for NEP. The lower end of this range agrees with reported uncertainties [Griffis et al., 2003; Oren et al., 2006; Richardson and Hollinger, 2005]. The 95% confidence intervals ranged between 110 and 545 g C m⁻² a⁻¹ for NPP. This range compares to uncertainties reported for a single forest [Clark et al., 2001a; Clark et al., 2001b]. The 95% confidence intervals of Rh ranged between 95 and 295 g C m⁻² a⁻¹, of Re between 35 and 670 g C m⁻² a⁻¹, of GPP between 35 and 750 g C m⁻² a⁻¹ and of Ra between 70 and 390 g C m⁻² a⁻¹. We are not aware of observation-based studies that report the uncertainty of Rh, Re, GPP or Ra observations. Therefore, the between-site variability of the flux quantity in the database was used to validate the expert-based assessment. Between-site variability was 250 g C m⁻² a⁻¹ for Rh, 680 g C m⁻² a⁻¹ for Re, 780 g C m⁻² a⁻¹ for GPP, and 570 g C m⁻² a⁻¹ for Ra.

3.1.3. Aggregated Fluxes and Their Uncertainty

[28] According to the analyses presented in this study the data first had to be aggregated by year and then by site. For a given site (i), a single weighted mean flux estimate (Fi) was produced for each available year l. When the flux component was determined with k different methods j in year l, the flux determined by method j for site i was then given as Fijl.

The average flux across methods (Fij) was calculated as the weighted mean:

\[ F_{ij} = \frac{1}{k} \sum_{j=1}^{k} w_{ij} F_{ijl} \]  \hspace{1cm} (4)

where \( w_{ij} = 1/s_{ij}^2 \) and \( s_{ijl} \) is the standard deviation of \( F_{ijl} \). \( s_{ij} \) was estimated from the uncertainty framework. The uncertainty of the weighted mean was estimated by means of error propagation:

\[ s_{ij} = \sqrt{\sum_{j=1}^{k} s_{ijl}^2} \]  \hspace{1cm} (5)

Following, the weighted mean flux component was calculated across years:

\[ F_i = \frac{1}{m} \sum_{l=1}^{m} w_{il} F_{il} \]  \hspace{1cm} (6)
where, \( w_{ii} = 1/s_{ii} \), \( m \) the number of years for which flux estimates are available for site \( i \). The uncertainty of the weighted mean was estimated by means of error propagation:

\[
s_i = \sqrt{\sum_{j=1}^{m} s_{ij}^2}.
\]

### 3.1.4. Availability of the Database


### 3.2. Consistency

[30] Independent measurements of eddy covariance–based estimates of NEP, GPP and Re and biometric based estimates of Ra, Rh and NPP are required for the proposed consistency cross-check of the flux estimates. Site-specific CO₂ balances based on independent measurements are not necessarily closed. However, closure is required before equations (1)–(3) can be used in a meaningful way. Hence, closure of the balances was enforced by introducing six closure terms, one for each flux, to equations (1)–(3). The equations can then be rewritten as follows:

\[
(Re + \delta Re) = (Ra + \delta Ra) + (Rh + \delta Rh),
\]

\[
(NPP + \delta NPP) = (GPP + \delta GPP) - (Ra + \delta Ra),
\]

\[
(NPP + \delta NPP) = (NEP + \delta NEP) + (Rh + \delta Rh).
\]

The CO₂ balance was further constrained using the observed site-specific soil respiration (Rs) data (equations (11)–(13)) and constraining the closure terms to less than 15% of the observed flux for GPP, Ra and Re, 20% for NPP and 33% for NEP (equations (13)–(18)). Arbitrary thresholds were selected to limit the sum of the closure terms. Because relative thresholds were used the threshold of NEP was larger than for the other fluxes because the absolute value of NEP is smaller. Finally, after closing the balance, GPP should still exceed NPP (equation (19)).

\[
Rs > (Rh + \delta Rh),
\]

\[
(Re + \delta Re) > Rs,
\]

\[
(Ra + \delta Ra) > (Re + \delta Re) - Rs,
\]

\[
\delta GPP < [0.15*GPP],
\]

\[
\delta NPP < [0.20*NPP],
\]

\[
\delta NEP < [0.33*NEP],
\]

\[
\delta Ra < [0.15*Ra],
\]

\[
\delta Re < [0.15*Re],
\]

\[
(GPP + \delta GPP) > (NPP + \delta NPP).
\]

The closure terms were optimized by means of quadratic programming (equations (8)–(19)) such that the objective function \( (\delta GPP) + (\delta NPP) + (\delta NEP) + (\delta Ra) + (\delta Re) + (\delta Rh) \) was minimal and the CO₂ balance closed. Failure to close the CO₂ balance within the constraints resulted in classifying the flux estimates on the site level as inconsistent.

[31] If the CO₂ balance can be closed, cross-validation continues by comparing the ratio of (Ra + \( \delta Ra \)) and (GPP + \( \delta GPP \)) with the ratio of (Rh + \( \delta Rh \)) and (GPP + \( \delta GPP \)). A simplified approach was used for sites without independent Ra estimates. For such sites an initial value of Ra* was estimated by the mean of Re – Rh and GPP – NPP and further refined by means of the closure procedure. However, in this case \( \delta Ra \) was removed from the objective function because no independent estimate was available and therefore its correction term should not be accounted for.

[32] Finally, the consistency of the ratios (i.e Ra/GPP, Ra*/GPP and Rh/GPP) is judged against expert knowledge. Thirty experts were asked to give the most likely ranges for the NPP/GPP and Ra/Re ratios for mature forest. We received ten responses and used these ranges to estimate a nonparametric probability surface that is plotted as contour lines in Figures 1 and 2. For a given combination of flux ratios, its value on the probability surface was estimated by the share of experts that considered this combination of flux ratios consistent. Thus, sites that go with a high value on the probability surface are considered consistent by the majority of our expert panel whereas sites that go with a low value are considered inconsistent. It should be stressed that the probability surface is based on expert knowledge rather than theoretical principles. Given that the majority of the respondents have been principal site investigator at one point in their career, their perception of consistent flux ratios is partly shaped by their own experimental results.

### 3.3. Relationships Between Flux Quantities

[33] For sites with consistent flux estimates, we derived estimates for NPP/GPP, Rh/NPP, Rh/Re and Re/GPP from the slopes of four linear regression models fitted to (GPP + \( \delta GPP \)) versus (NPP + \( \delta NPP \)), (NPP + \( \delta NPP \)) versus (Rh + \( \delta Rh \)), (Re + \( \delta Re \)) versus (Rh + \( \delta Rh \)), (GPP + \( \delta GPP \)) versus (Re + \( \delta Re \)), respectively. We used a maximum likelihood estimator, which accounts for the uncertainty in both predictor and predicted variables [Andrews et al., 1996], in contrast to the classical approach where only errors in the predicted variables are considered. Results were accepted when the optimization algorithm [Duan et al., 1993] reached convergence.

[34] Re was derived from eddy covariance measurements but following equations (1)–(3), it may as well be estimated.

5 of 13
as \( \text{GPP} = \text{NPP} + \text{Rh} \). Because the \( \text{CO}_2 \) balance had been closed for the sites under consideration, the mean values of the eddy covariance and component based estimates for \((\Re + \delta\Re)\) are identical. However, both approaches to estimate \( \Re \) have a different uncertainty owing to their different data sources. This difference in uncertainty was used to check the across-sites consistency of the four regression models; \( \Re/\text{GPP} \) (i.e., eddy covariance–based) was compared to the slope of \(((\text{NPP}/\text{GPP}) \ast (\text{Rh}/\text{NPP}) \ast (\text{Rh}/\Re)) \) (i.e., component based).

Following, the same ratios as above were calculated for sites for which we could not test site level consistency because of a lack of data. Again, four regressions were fitted for each ratio in the relevant variable space using maximum likelihood.

3.4. Confidence Intervals

Uncertainties of the flux components (see section 3.1) were propagated throughout the analyses by means of random realizations based on Monte Carlo principles [Rubinstein, 1981]. One thousand realizations of the data set were simulated by adding a normally distributed uncertainty to the observed estimates of all variables. The normally distributed uncertainties were calculated by multiplying the total uncertainty of the flux \((\sigma_f)\) by a normally distributed random number with a mean equal to zero and a variance equal to one. The random realizations were screened for impossible values, which were subsequently removed from the analyses. The following realizations were considered impossible: \( \text{GPP} < 0, \text{NPP} < 0, \text{Rh} < 0, \Re < 0, \Ra < 0, \text{NEP} > \text{GPP}, \text{NPP} > \text{GPP}, \text{Rh} > \Re, \Ra > \Re \) and, on a more arbitrary basis, \( \text{NEP} < -500 \) (source) and \( \text{NEP} > 1100 \) (sink).

Following, each ratio, relationship or processing step (i.e., calculating the closure terms) was repeated a thousand times, once for each random realization. The standard deviation was then estimated as the 16 and 84 percentile of the mean values of the 1000 simulations.

4. Results and Discussion

Owing to the flux identities (equations (1)–(3)), independent \( \text{NPP}, \text{Rh} \) and \( \Ra \) measurements allow to cross-check the consistency of the eddy covariance–based flux estimates of \( \text{GPP} \) and \( \Re \). This approach has the advantage that more terms of the carbon balance are considered, which, besides testing the consistency, increases understanding of ecosystem function.

Figure 1. Consistency test showing the site level \( \Ra/\text{GPP} \) versus \( \text{Rh}/\text{GPP} \) after \( \text{CO}_2 \) balance closure. The contour lines show consistency according to an expert panel. The blue contours indicate high agreement between the experts and thus likely flux consistency on the site level, whereas red contours indicate low agreement between experts and thus ratios that are more likely inconsistent with current knowledge. The isolines show \( \Re/\text{GPP} \) with values <1 indicating a \( \text{CO}_2 \) sink and with values >1 indicating a \( \text{CO}_2 \) source. The color of the marker follows the color bar at the right which shows how well the \( \text{CO}_2 \) balance at the site level closed (\%). Blue indicates that the balance closed well, and red shows sites for which considerable corrections are needed to close the \( \text{CO}_2 \) balance. The sum of the absolute values of the closure terms was used to quantify the flux consistency and ranged between 18 and 24\% of the observed \( \text{GPP} \) for the 4 sites under consideration (Table 1). See color version of this figure in the HTML.
Ideally the Ra/GPP and Rh/GPP ratios are used to test flux consistency at the site level. We identified 4 sites for which all required flux estimates were available (see Table 1), out of 529 sites for which at least one CO2 flux was reported. CO2 balances were closed before calculating the ratios. The closure terms are a numerical way to approach data quality and flux uncertainty on the site level (Table 1). Ideally each individual closure term should be zero; deviations from zero indicate a closure problem. Small deviations indicate a good agreement between the fluxes. Large closure terms (i.e., beyond uncertainties in measured fluxes) or lack of flux balance closure within our constraints could indicate problems with the accuracy of the measurement technique or missing components in the CO2 balance. An underestimation of one flux, i.e., NPP can be accounted for by adding a closure term to NPP but also by decreasing Ra or GPP. Owing to the objective function (see section 3.2), the optimization algorithm will favor many small corrections above a single large correction. Therefore, the sum of the absolute values of the closure terms was used to quantify the flux consistency and ranged between 7 and 29% of the observed GPP for the sites under consideration (Table 1). See color version of this figure in the HTML.

Figure 2. Consistency test showing the site level Ra*/GPP versus Rh/GPP after CO2 balance closure, where Ra* was estimated from GPP, NPP, Re, and Rh. The contour lines show consistency according to an expert panel. The isolines and marker colors are similar as in Figure 1. The sum of the absolute values of the closure terms was used to quantify the flux consistency and ranged between 7 and 29% of the observed GPP for the sites under consideration (Table 1). See color version of this figure in the HTML.

[39] Ideally the Ra/GPP and Rh/GPP ratios are used to test flux consistency at the site level. We identified 4 sites for which all required flux estimates were available (see Table 1), out of 529 sites for which at least one CO2 flux was reported. CO2 balances were closed before calculating the ratios. The closure terms are a numerical way to approach data quality and flux uncertainty on the site level (Table 1). Ideally each individual closure term should be zero; deviations from zero indicate a closure problem. Small deviations indicate a good agreement between the fluxes. Large closure terms (i.e., beyond uncertainties in measured fluxes) or lack of flux balance closure within our constraints could indicate problems with the accuracy of the measurement technique or missing components in the CO2 balance. An underestimation of one flux, i.e., NPP can be accounted for by adding a closure term to NPP but also by decreasing Ra or GPP. Owing to the objective function (see section 3.2), the optimization algorithm will favor many small corrections above a single large correction. Therefore, the sum of the absolute values of the closure terms was used to quantify the flux consistency (Table 1 and Figure 1) instead of reporting individual closure terms.

[40] Although the consistency test in itself (Figure 1) requires only independent estimates of GPP, Ra, and Rh, the procedure for closing the CO2 balance needs additional estimates of NEP, NPP, Re, and Rs (see section 3.2). Currently, data requirements restrict the applicability of the test. However, biometric Rh and NPP are more commonly estimated than biometric Ra. Hence, a simplified consistency test and closure procedure where Ra is substituted by the mean value of Re/Rh and GPP/NPP is proposed and should be more widely applicable. GPP, NPP, Rh, and Re estimates were available for 16 sites of which 13 could be closed (Figure 2 and Table 1).

[41] Sites for which the CO2 balance could be closed and which were plotted within the contours of the probability surface are considered consistent, whereas sites outside the contours are thought to be inconsistent. Because of the use of ratios, consistency does not automatically imply that the absolute fluxes are accurate or precise as errors could compensate each other. Also, an inconsistent ratio is not necessarily incorrect and following careful screening of the data, the fluxes could be reported if the unexpected ratios have a biological reason (i.e., Rh could be large because of decomposition of legacy carbon or NPP could be small when an insect outbreak occurred). It is interesting to note that sites for which the CO2 balance could be closed within our constraints were also considered consistent based on the ratio approach. Hence, successful closure of the CO2 balance is a strong indicator for flux consistency.

[42] An interesting case is the Hainich site (labeled “f” in Figure 2) that passed the consistency test with GPP, Ra and Rh but is most likely affected by advection. Correcting for advection was, however, only possible by using additional data beyond eddy flux measurements [Kutsch et al., 2008]. When we used the eddy current (EC) data that were thought to be advection-corrected, the data points moved closer to the peak of the probability surface (not shown), illustrating that
Table 1. Site Names, Short Site Description, Years That the Eddy Covariance System Has Been Operated, and Data Sources for GPP, Re, NPP, Rh, and Ra Estimates

| Site Description | Years EC | Label | \(\frac{\sum |\Delta Flux|}{GPP}\) (%) | References |
|------------------|----------|-------|----------------------------------|------------|
| Collelongo Th, B, D, Fagus sylvatica L. | 1996 to present | c | NA | M. Bascietto et al. (2003) and CarboEurope-IP database (2006) |
| Doorey Th, N, E, Picea abies L. | 2003 to 2004 | d | 9 | Law et al. [2002] and Lindroth et al. [2008] |
| Flakaliden Bh, B, D, Fagus sylvatica L. | 1996 to present | e | 7 | Black et al. [2007] and Saiz et al. [2006a, 2006b] |
| Harvard Th, B, D, Quercus rubra L. and Acer rubrum L. | 1992 to present | g | 10 | Barford et al. [2001], Gaudinski et al. [2000], Munger et al. [2004], and Urbanski et al. [2007] |
| Hesse Th, B, D, Fagus sylvatica L. | 1996 to present | h | 22 | Davidson et al. [2002], Epron et al. [1999, 2001], and Granier et al. [2008] |
| Loobos Th, N, E, Pinus sylvestris L. | 1996 to present | i | 24 | M. Bascietto et al. (2003), CarboEurope-IP database (2006), and Schelhaas et al. [2004] |
| Metolius young Th, B, D, Picea pungens Laws. | 1999–2002 | n | 15 | Irvine and Law [2002], Law et al. [2001a, 2000], and Williams et al. [2005] |
| Morgan Monroe Th, B, D, Acer saccharum Marsh. and Liriodendron tulipifera L. | 1998 to present | o | 10 | Curtis et al. [2002], Ehman et al. [2002], and Schmidt et al. [2000] |
| Prince Albert SSA SOAS Th, B, D, Populus tremuloides Michx. | 1994 to present | p | 21 | Barr et al. [2007], Gower et al. [2001], and Griffis et al. [2004] |
| Prince Albert SSA SOBS Th, B, D, Picea mariana Mill. | 1994 to present | q | 16 | Gower et al. [2001], Kljun et al. [2006], and O’Connell et al. [2003] |
| Soree Th, B, D, Fagus sylvatica L. | 1996 to present | x | NA | M. Bascietto et al. (2003), CarboEurope-IP database (2006), and Pilsgaard et al. [2003] |
| University of Michigan Th, B, D, Populus grandidentata Michx. and Acer rubrum L. | 1999 to present | y | 10 | Gough et al. [2007] and Schulze et al. [2003] |
| Willow Creek Th, B, D, Acer saccharum Marsh. | 1999–2000 | w | 29 | Curtis et al. [2002] and Falge et al. [2001] |

If the CO2 balance could be closed, the percent ratio between the sum of the absolute values of the closure terms and the observed gross primary production (GPP) is given. NA, sites where the CO2 balance was not closed because of violations with the constraints implied on balance closure.

Bh, boreal humid; Th, temperate humid; Tsa, temperate semiarid; B, broadleaved; N, needle leaved; D, deciduous; E, evergreen.

M. Bascietto et al. (Database of tree stands (Structure, age, biomass, LAI and NPP) of the FORCAST project, 2003) and CarboEurope-IP database (2006, available at http://gaia.agrarian.unitus.it/database/carboeuropeip/).
even consistent flux estimates can be further improved with advection correction.

The proposed simplified test (Figure 2) requires estimates of GPP, NPP, Re, and Rh in addition to NEP and Rs for the CO2 balance closure. An even less demanding approach in terms of data requirements is to calculate Ra* just from GPP – NPP such an approach would only require GPP and Re estimates derived from eddy covariance and an NPP estimate based on biometric measurements. However, with just three fluxes, the CO2 balance cannot be closed and interpretation of the results becomes ambiguous.

A fourth conceivable approach could make use of flux ratios based on soil respiration measurements. Compared to Ra and Rh, soil respiration is widely available and rather straightforward to measure. However, soil respiration does not contain enough information to check the consistency of the flux estimates (because in itself it does not allow the direct quantification of Ra and Rh). Nevertheless, it is an essential flux to constrain the closure of the CO2 balance.

Although we found 13 sites that were consistent within our framework, the presented analyses do not allow generalization of the results. To overcome this issue, we used the GPP, Re, NPP, and Rh data of these 13 sites and estimated their NPP/GPP, Rh/NPP, Rh/Re, and GPP/Re ratios (Table 2 and Figure 3). Most estimates of the flux ratios were within the line of expectation. The NPP/GPP ratio was estimated at 0.51 ± 0.02 which agrees well with the previous independent estimates of 0.47 [Waring et al., 1998], 0.45 [Law et al., 2001a], 0.53 [DeLucia et al., 2007] and 0.43 [Litton et al., 2007].

Our Re/GPP (0.73 ± 0.02) ratio is in between the earlier reported value of 0.65 [Reichstein et al., 2007] and the 0.82 for deciduous broadleaf forests and 0.85 for coniferous forests reported by Law et al. [2002], and 0.80 reported by Janssens et al. [2001]. However, we expect a strong correlation between eddy covariance–based GPP and Re estimates because NEP = GPP – Re, where NEP is an order of magnitude smaller than GPP and Re. Consequently, a very large NEP relative to GPP is required to negate this correlation.

The Rh/NPP ratio of 0.48 ± 0.03 indicates that about half of the NPP is respired by heterotrophs. Heterotrophic respiration from the decomposition of coarse and fine woody debris is often not included in the estimates of Rh, so this ratio could be higher, particularly in forests with a large amount of accumulated debris and climate conditions that may promote decomposition. In managed forests, which dominate this data set, about half of the reported NPP is foliage and root production [Ciais et al., 2008]. Moreover, most of the wood production is frequently removed during thinning and harvests. Hence, our analysis confirms that in managed forests, the short-lived biomass components (roots, foliage) are the main substrate for heterotrophic organisms [Trumbore, 2000, 2006]. In a typical forest, NPP and Ra are of similar magnitude, but Rh is much smaller than NPP because of the wood removals. Hence, the low Rh/Re ratio of 0.34 ± 0.03 is not unexpected and implies that in managed forests on average 66% of the ecosystem respiration is autotrophic. In old-growth forests, where Rh and NPP are expected to be more similar [see Luyssaert et al., 2008], we speculate that the Rh/Re ratio would be closer to 0.5.

In addition, good agreement exists between the eddy covariance–based Re/GPP (0.73 ± 0.02) and the component-based Re/GPP (0.70 ± 0.03). The component-based Re/GPP was obtained by substituting NPP/GPP and Rh/Re in Rh/NPP (see section 3.3). Consequently, the reported flux ratios (Table 2) are consistent with each other across a large range of productivity (GPP from 700 to 2000 g C m⁻² a⁻¹).

We estimated the same flux ratios for sites with an incomplete data set to close the CO2 balance but for which at least two flux estimates were available (Table 2). The ratios derived from the subset of consistent sites were then used to

---

Table 2. Flux Ratios of the Subset of 13 Temperate and Boreal Sites

<table>
<thead>
<tr>
<th></th>
<th>Complete CO2 Balance</th>
<th>Incomplete CO2 Balance</th>
</tr>
</thead>
<tbody>
<tr>
<td>NPP/GPP</td>
<td>0.51 ± 0.02</td>
<td>0.50 ± 0.01</td>
</tr>
<tr>
<td>Rh/NPP</td>
<td>0.48 ± 0.03</td>
<td>0.62 ± 0.02</td>
</tr>
<tr>
<td>Rh/Re</td>
<td>0.34 ± 0.03</td>
<td>0.36 ± 0.01</td>
</tr>
<tr>
<td>Re/GPP</td>
<td>0.74 ± 0.02</td>
<td>0.80 ± 0.00</td>
</tr>
</tbody>
</table>

*With a complete and consistent CO2 balance and all temperate and boreal sites for which at least two flux estimates were available.
evaluate the ratios of the sites with an incomplete CO₂ balance. We found good agreement between all ratios (Table 2) except for RH/NPP. Further, reasonable agreement between the eddy covariance (i.e., 0.80 ± 0.00) and component-based estimate of RE/GPP (i.e., 92 ± 0.05) indicated that the currently available flux estimates are largely consistent across a large range of productivity.

In this study the consistency tests were applied to the mean value of all available flux estimates that met the specified selection criteria (see section 3.1). This approach is suitable to perform a retrospective consistency check of the data that has been used for model validation and hypothesis testing [see Baldocchi et al., 2001; Ciais et al., 2005; Law et al., 2002; Luyssaert et al., 2008; Magnani et al., 2007; Reichstein et al., 2007; Valentini et al., 2000] but is sub-optimal for checking the consistency of individual sites. Principal investigators and their teams are continuously working toward estimates that better represent a specific site than the mean flux estimates used in this study. Site-specific estimates could be improved by for example correcting for advection [Van Gorsel et al., 2007], aiming at more complete NPP estimates [Ehman et al., 2002; Navarro et al., 2008] and including decomposition of coarse woody debris in the Rh estimates [Jomura et al., 2007]. It is our hope that principal investigators will start reporting the outcome of the proposed consistency cross-check when presenting their improved site level flux estimates.

5. Conclusions

Eddy covariance–based flux estimates of NEP, GPP, and RE require independent measurements to be validated. We propose a framework that can be used to cross-check consistency of the eddy covariance data when biometric NPP and RH measurement are available. Out of 529 sites at which at least one CO₂ flux was measured, we identified 4 sites for which all six basic fluxes (GPP, NPP, NEP, Re, Ra, and Rh) were independently measured and 16 sites for which all fluxes except Ra were independently measured.

The consistency cross-check is a two-step test. The test starts by closing the CO₂ balance, failure to do so within the constraints will classify the flux estimates on the site level as inconsistent. However, the individual closing term does not contain information about the accuracy of the individual flux estimates. If the CO₂ balance can be closed, the test continues by confronting the ratios of GPP and Ra (or Ra*), with GPP and Rh (where Ra is measured and Ra* calculated as Re – Rh or GPP – NPP. The ratios are based on the closed flux estimates). Consistency of these ratios is judged against expert knowledge. Because of the use of ratios, consistency does not automatically imply that the absolute fluxes are accurate or precise as errors could compensate each other. Also, an inconsistent ratio is not necessarily incorrect but provides a signal for careful data screening, looking for possible biological reasons of the unexpected ratios. Flux estimates of sites that pass both tests are considered consistent.

Thirteen out of the 16 sites that met the data requirements were found to be consistent. This subset of consistent sites was used to determine flux ratios. Similar ratios, except for Rh/NPP, were obtained from sites with incomplete flux estimates indicating that the currently available flux data are largely consistent across a wide range of site productivities. Confidence in the flux estimates of local, regional and global flux networks could be further enhanced when the required fluxes (at least GPP, RE, NPP, Rh and Ra) would be estimated in a harmonized way and over a longer time frame at more, possibly all, sites.

Notation

- GPP gross primary production (g C m⁻² a⁻¹)
- NPP net primary production (g C m⁻² a⁻¹)
- NEP net ecosystem production (g C m⁻² a⁻¹)
- Re ecosystem respiration (g C m⁻² a⁻¹)
- Ra autotrophic respiration (g C m⁻² a⁻¹)
- Rh heterotrophic respiration (g C m⁻² a⁻¹)
- Rs soil respiration (g C m⁻² a⁻¹)

Acknowledgments. We thank all site investigators, their funding agencies, and the various regional flux networks (Aeriflux, Ameriflux, AsiaFlux, CarboAfrica, CarboEurope-IP, CarboItaly, ChinaFlux, Fluxnet-Canada, KoFlux, LBA, NECC, OzFlux, TCONS-Siberia, and USCCC), and the Fluxnet project, whose support is essential for obtaining the measurements without which the type of integrated analyses conducted in this study would not be possible. Two anonymous reviewers provided substantial improvements to earlier versions of the paper and helped to develop Figures 1 and 2. S.L. was supported by the Centre of Excellence ECO (UA-Methusalem) and the Research Foundation-Flanders (FWO-Vlaanderen) with a postdoctoral fellowship and a research grant. B.E.L. was supported by the Ameriflux Science Team Research grant (U.S. Department of Energy, Terrestrial Carbon program award DE-FG02-04ER65911). D.P. thanks the IMECC EU project for the support. E.-D.S. was supported by the Exploratories of the Deutsche Forschungsgemeinschaft.

References


D. Dragoni, Department of Geography, Indiana University, 701 E. Kirkwood Avenue, Bloomington, IN 47405, USA.


M. L. Goulden, Department of Ecology and Evolutionary Biology, University of California, 3319 Croul Hall, Irvine, CA 92697, USA.

A. Granier, Forest Ecology and Ecophysiology Unit, UMR INRA UHP, F-54280 Champenoux, France.

I. A. Janssens and S. Luyssaert, Department of Biology, University of Antwerp, Universiteitsplein 1, B-2610 Wilrijk, Belgium. (sebastiaan.luyssaert@ua.ac.be)

W. L. Kutsch, M. Reichstein, and E.-D. Schulze, Max Planck Institute for Biogeochemistry, P.O. Box 100164, D-07701 Jena, Germany.

B. E. Law, College of Forestry, Oregon State University, 321 Richardson Hall, Corvallis, OR 97331, USA.

S. Linder, Southern Swedish Forest Research Centre, SLU, P.O. Box 49, SE-230 53 Alnarp, Sweden.

G. Matteucci, CNR-ISAFOM, Via Cavour 4-6, I-87036 Rende, CS, Italy.

E. Moors, ESS-CC, Alterra Wageningen UR, P.O. Box 47, NL-6700 AA Wageningen, Netherlands.

J. W. Munger, Department of Earth and Planetary Sciences, Harvard University, 20 Oxford Street, Cambridge, MA 02138, USA.

D. Papale, Department of Forest Environment Science and Resource, University of Tuscia, DISAFRI, Via Camillo de Lellis, I-01100 Viterbo, Italy.

K. Pilegaard, Risø National Laboratory for Sustainable Energy, Biosystems Division, Technical University of Denmark, P.O. Box 49, DK-4000 Roskilde, Denmark.

M. Saunders, School of Biology and Environmental Science, University College Dublin, Science Centre West, Belfield, Dublin 4, Ireland.