Max-Planck-Institut für Biogeochemie

MPI-BGC

Jena

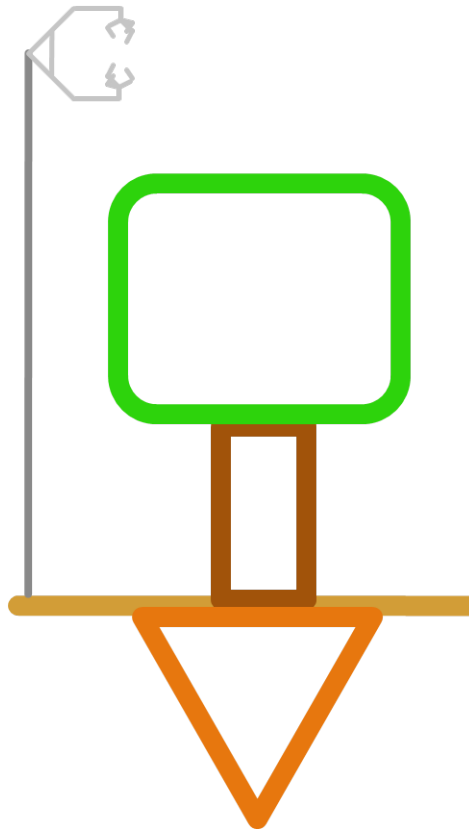# TECHNICAL REPORTS
20

Model data fusion for terrestrial biosphere models
with carbon and water cycle observations

by
Gitta Lasslop

MAX-PLANCK-GESELLSCHAFT

Technical Reports - Max-Planck-Institut für Biogeochemie 20, 2011

Model data fusion for terrestrial biosphere models with carbon and water cycle observations

Dissertation

Zur Erlangung des Doktorgrades der Naturwissenschaften im Department Geowissenschaften der Universität Hamburg

vorgelegt von

Gitta Lasslop

aus

Stuttgart

Hamburg

2010

Als Dissertation angenommen
vom Department Geowissenschaften der Universität Hamburg.

Auf Grund der Gutachten von Prof. Dr. Martin Claußen
und Dr. Markus Reichstein

Hamburg, den 9.11.2010
Prof. Dr. Jürgen Oßenbrügge
Leiter des Departements für Geowissenschaften

Contents

List of Figures

List of Tables

Summary

The modelling of the terrestrial carbon cycle and the future development of carbon sources and sinks predicted by models has considerable uncertainty. Model data fusion (MDF), e.g. combining the models with observations, has become a popular tool to evaluate, improve models, and consequently decrease their uncertainty. This study focuses on the use of eddy covariance (EC) data for model data fusion with terrestrial biosphere models. It first addresses the derivation of photosynthetic and respiratory fluxes from eddy covariance data and the relation between them. A complementary estimate of the two is developed and their reliability is assessed. In the second part, a model data fusion framework, with comprehensive representation of the uncertainties in eddy covariance data, is developed.

Based on the FLUXNET database that assembles EC observations all over the world, a method is developed to partition the net ecosystem exchange of carbon (NEE) into gross primary production (GPP) and ecosystem respiration (Reco). In contrast to the available estimate that is based on potentially biased nighttime EC data, the algorithm developed here uses daytime data in combination with a light response curve extended to account for the temperature dependence of respiration and the water vapour pressure deficit (VPD) limitation of GPP. Including the VPD limitation strongly improved the model's ability to reproduce the asymetric diurnal cycle and reduced biases in Reco for periods with high VPD. Comparing this daytime data based estimate with the available nighttime data based estimate allowed for a derivation of an estimate of potential systematic errors in the flux components. Despite site specific differences between the methods, overall patterns remained robust, adding confidence to statistical studies based on the FLUXNET database. It has been argued often that the relation between Reco and GPP can be spurious when GPP is derived as the residual, since any error in Reco directly propagates into GPP. By using the two quasi independent estimates, i.e. daytime and nighttime based estimates, it was possible to show that the strong correlation between GPP and Reco is not spurious, but holds true.

The careful characterization of errors and uncertainties is a prerequisite for a comprehensive represenation of data streams in MDF studies. A method is proposed to characterize the random error of the eddy covariance flux data, and analyse error distribution, standard deviation, cross- and autocorrelation of CO_2 and H_2O flux errors for four European sites. Moreover, the treatment of those errors and additional systematic errors is examined with respect to their influence on statistical estimates of parameters and the associated uncertainties. The study confirmed that the error standard deviation scales with the flux magnitude. Further analysis showed that the cross correlation between CO_2 and H_2O fluxes was negligible, and the autocorrelation decreased fast. The previously found strongly-peaked error distribution was revealed to be largely due to a superposition of almost Gaussian distributions with standard deviations varying by flux magnitude for the sites used here. A more recent study using the whole FLUXNET database revealed site specific differences most probably due to differences in the data filtering. Using synthetic data, based on model outputs and different types of errors added, we showed that accounting for the varying standard deviation significantly reduced the uncertainty of the parameter estimates. The presence of systematic errors lead to biases in the parameter estimates,

where the uncertainty of the estimate increased only slightly, so that the true parameter was not within the uncertainty range of the estimate. Thus, to avoid biases in parameterizations, the uncertainty due to systematic errors, which cannot be corrected, needs to be addressed.

In the last part, a MDF framework is proposed that combines the uncertainty due to the random error with an estimate for the uncertainty due to systematic errors. For the EC data, we include the uncertainty due to the choice of the u* threshold, used to filter the EC data for low turbulent conditions. For the fAPAR data we use multiple data sources, observed in situ and by satellite. In this study the uncertainty of the fAPAR streams clearly overrode the uncertainty in the eddy covariance data and limited attempts to evaluate and improve the model structure using single sites. Instead more robust characteristics in the spatial domain might have a higher potential to constrain models on the global scale.

Zusammenfassung

Die Modellierung des terrestrischen Kohlenstoffkreislaufs und die Modellvorhersagen über die zukünftige Entwicklung von Kohlenstoffquellen und -senken beinhalten erhebliche Unsicherheiten. Die Model-Daten Fusion (MDF), das heißt die Kombination von Modellen mit Beobachtungen ist zu einer weitverbreiteten Methode geworden um Modelle zu evaluieren, zu verbessern und damit auch deren Unsicherheiten zu vermindern. Diese Arbeit beschäftigt sich mit der Verwendung von Eddy Kovarianz (EC) Daten für die MDF mit terrestrischen biogeochemischen Modellen. Der erste Teil behandelt die Ableitung von photosynthetischen und respiratorischen Flüssen aus Eddy Kovarianz Daten und die Korrelation zwischen den beiden Flusskomponenten. Komplementäre Schätzwerte der beiden Flüsse wurden entwickelt und die Glaubwürdigkeit bewertet. Der zweite Teil entwickelt eine MDF-Methodik mit umfassender Berücksichtigung der Unsicherheiten in Eddy Kovarianz Daten.

Basierend auf der FLUXNET Datenbank, die EC Daten aus der ganzen Welt beinhaltet, wird eine Methode entwickelt, die den Nettokohlenstoffaustausch (NEE) in Bruttoprimärproduktion (GPP) und Ökosystemrespiration (Reco) aufteilt. Im Gegensatz zu den bereits verfügbaren Schätzungen, die auf den potentiell fehlerhaften Nachtdaten beruhen, benutzt der hier entwickelte Algorithmus Tagdaten und eine Lichtabhängigkeitskurve die erweitert wurde, um die Temperatursensitivität der Respiration und die Limitierung der Photosynthese durch hohe Wasserdampfdruckdefizite (VPD) zu berücksichtigen. Das Einfügen der VPD-Limitierung verbessert die Fähigkeit des Modells, den asymmetrischen Tagesgang des NEE zu reproduzieren und vermindert Fehler in der Respiration in Zeitabschnitten mit hohem VPD. Der Vergleich dieser Schätzwerte basierend auf Tagdaten mit der verfügbaren Schätzwerten basierend auf Nachtdaten, ermöglicht eine Prüfung des potentiellen systematischen Fehlers in den Flusskomponenten. Trotz starker Unterschiede zwischen den Methoden für bestimmte Messstandorte, bleiben die übergreifenden Muster in den Daten erhalten und stärken das Vertrauen in statistische Analysen, die auf der FLUXNET Datenbank beruhen.

Die Korrelation zwischen GPP und Reco wurde oft diskutiert und eine künstliche Korrelation wurde erwartet wenn GPP als Residuum berechnet wird, da dann jeder Fehler in Reco direkt in GPP propagiert wird. Durch die Benutzung der zwei quasi unabhängigen Schätzungen war es möglich zu zeigen, dass die starke Korrelation nicht künstlich ist sondern erhalten bleibt.

Eine sorgfältige Charakterisierung der Fehler und Unsicherheiten ist eine Voraussetzung für eine umfassende Repräsentierung der Datenströme in MDF Studien. In dieser Arbeit wird eine Methode zur Charakterisierung des Zufallsfehlers von Eddy Kovarianz Daten vorgeschlagen. Die Verteilung, Standardabweichung, Kreuz- und Autokorrelation der Fehler von CO_2- und H_2O-Flüssen wurde für vier europäische Standorte analysiert. Weiterhin wurde die Behandlung dieser Fehler und zusätzlicher systematischer Fehler untersucht, sowie deren Auswirkung auf Parameterschätzwerte und -unsicherheiten. Die Studie bestätigt die Abhängigkeit zwischen der Fehlerstandardabweichung und der Flusshöhe. Weitere Analysen zeigen, dass die Kreuzkorrelation zwischen Fehlern der CO_2- und H_2O-Flüsse vernachlässigbar ist und dass die Autokorrelation schnell abfällt. Für die Standorte in dieser Studie konnte gezeigt

werden, dass die im Vorfeld gefundene sehr spitze Fehlerverteilung zu einem großen Teil durch die Überlagerung von Verteilungen mit variierender Standardabweichung, die annähernd normalverteilt sind, zustande kam. Eine neuere Analyse, basierend auf der FLUXNET Datenbank, zeigte standortspezifische Unterschiede, die sehr wahrscheinlich durch Unterschiede bei der Filterung der Daten entstehen. Basierend auf synthetischen Daten, bestehend aus modellierten Daten und verschiedenen Typen von addierten Fehlern, wurde gezeigt, dass die Berücksichtigung der variierenden Standardabweichung die Unsicherheit der Parameter signifikant reduziert. Das Vorhandensein von systematischen Fehlern führt zu Fehlern in den Parameter, wobei die Unsicherheit der Schätzung nur leicht erhöht wird, so dass der wahre Parameter nicht im Unsicherheitsbereich des Schätzwertes lag. Um Fehler in der Parameterisierung zu vermeiden, wird daher die Berücksichtigung der Unsicherheit von systematischen Fehlern benötigt.

Der letzte Teil der Arbeit schlägt einen methodischen Rahmen für die MDF vor, der die Unsicherheiten aufgrund des Zufallsfehlers mit einer Schätzung der Unsicherheit aufgrund von systematischen Fehlern kombiniert. Für die EC Daten wurde die Unsicherheit aufgrund der Bestimmung des u*-Filters berücksichtigt. Dieser wird benutzt um die Daten zu filtern, die unter schwach turbulenten Bedingungen gemessen wurden. Für die fAPAR Daten wurden unterschiedliche Datenströme verwendet, in situ gemessene und aus Satellitendaten abgeleitete. In dieser Studie war die Unsicherheit der fAPAR Datenströme dominierend und schränkte Versuche, die Modellstruktur anhand eines Standortes zu evaluieren und zu verbessern, ein. An dieser Stelle könnten robuste Eigenschaften in der räumlichen Domäne ein höheres Potential haben, um Modelle auf dem globalen Maßstab einzuschränken.

List of abbreviations

BETHY	photosynthesis model used in the study
C	carbon
cov	covariance
EC	eddy covariance
fAPAR	fraction of absorbed photosynthetically active radiation
FLUXNET	association of regional eddy covariance observation networks
GPP	gross primary production
HBLR	hyperbolic light response curve
JSBACH	land surface model used in the study
LAI	leaf area index
LM	Levenberg-Marquardt algorithm
MCMC	Monte Carlo Markov chain
MDF	model data fusion
NEE	net ecosystem exchange
PAR	photosynthetically active radiation
rd	rooting depth
Reco	ecosystem respiration
Rg	global radiation
RMSE	root mean squared error
sd	standard deviation
T	temperature
u*	friction velocity
var	variance
Vcmax	maximum carboxylation rate
VPD	water vapour pressure deficit

1. Introduction

The effects of human activities on the carbon cycle and climate change were first studied by the Swedish chemist, Svante Arrhenius, in 1896 (Arrhenius, 1896). He noted that CO_2 accumulates in the atmosphere as a byproduct of fossil fuel combustion, thereby changing the energy balance of the earth and increasing the surface temperature. Since 1750, the atmospheric concentration of CO_2 has risen from around 280 ppm to nearly 390 ppm in 2010 (http://scrippsco2.ucsd.edu). The main input of CO_2 to the atmosphere are the fossil fuel combustion (8.7 ± 0.5 GtC yr^{-1}) and land use change (1.2 GtC yr^{-1}) (Le Quere et al., 2009, estimates for 2008). Although the gross fluxes from ocean (90 GtC yr^{-1}) and land to the atmosphere (120 GtC yr^{-1}) are much larger than the anthropogenic emissions, the nonanthropogenic net fluxes between the earth suface and the atmosphere remove carbon from the atmosphere and only about 45% of the anthropogenic emissions have remained (Denman et al., 2007). While the estimate for the ocean-atmosphere flux is quite well constrained from the atmospheric and oceanic CO_2/O_2 and N_2 signals with a relative uncertainty of 20%, the land-atmosphere flux estimate having an uncertainty of 60% is less understood (Denman et al., 2007). In global climate projections all models included in the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (Solomon et al., 2007) showed a decreased efficiency of the earth system, land and ocean, to take up carbon (Meehl et al., 2007) but the current understanding about the sensitivity of this carbon sequestration is limited (Friedlingstein et al., 2006).

Observations can be used to be better constrain the models (Bonan, 2008) and an increasing amount of studies is making use of atmospheric CO_2, remote sensing, eddy covariance flux and carbon stock data (Knorr and Kattge, 2005; Carvalhais et al., 2008, 2010; Richardson et al., 2010; Wang et al., 2007; Rayner et al., 2005; Knorr and Heimann, 2001b). Observations can be used in model data fusion studies for model evaluation, optimization and improvement of the model structure (Williams et al., 2009). Moreover they help to decrease but also to assess the overall uncertainty of the model output. Carbon cycle observations that can be used for this purpose are increasingly becoming available: databases have been established for CO_2, water and energy fluxes between vegetation and atmosphere by the FLUXNET community (Baldocchi, 2008), for plant functional traits (TRY:www.try-db.org), or for soil respiration (Bond-Lamberty and Thomson, 2010). But not only the data itself are important: in parameter estimation studies it was shown that the characterization of data uncertainties strongly influences the results (Trudinger et al., 2007; Richardson and Hollinger, 2005). Thus Raupach et al. (2005) emphasized that the characterization of uncertainties of the data is as important as the data itself. So far only few studies have comprehensively described and included data uncertainties in model data fusion. This study tries to improve the understanding of the carbon cycle on ecosystem level by the use of eddy covariance data and carbon cycle models of varying complexity. This is achieved on the one hand by using the information that is assembled in the FLUXNET database and assessing their reliability, on the other hand by improving the methodology of model-data fusion for eddy covariance flux observations by improving the uncertainty estimates and including them in the model data fusion framework.

The following sections of the introduction address the main processes of the terrestrial carbon cycle, carbon cycle observations, modelling approaches used in this study and methods for model data fusion.

1.1. The terrestrial carbon cycle

The carbon cycle can be represented in an abstract way by pools, such as the atmosphere, ocean or biomass, and fluxes. Pools store carbon, for instance in the terrestrial part of the carbon cycle the various layers of soil or the different compartments of the biosphere. Fluxes transfer carbon from reservoir to reservoir and are controlled by various factors. Within the global carbon cycle the terrestrial part acts as a sink and removes carbon from the atmospheric reservoir. Mechanisms causing the net uptake of the biosphere are fertilization effects from increased CO_2 concentrations and nitrogen deposition, land use and land management changes, natural climate variability and anthropogenic climatic change (Schimel et al., 2001). The terrestrial biosphere interacts with the climate system through biological, physical and chemical processes that affect the energy balance, the hydrologic cycle and the atmospheric composition (Bonan, 2008). Water and energy exchanges are important complements to the carbon cycle (Chapin et al., 2009): water and carbon cycle are strongly coupled through the limitation of photosynthesis by the availability of water. The two cycles are linked to the energy exchanges by the division of available radiation into latent heat, e.g. transpiration and evaporation, and sensible heat (Chapin et al., 2002). The uncertainty about these interactions between the terrestrial component and the climate system remain high (Friedlingstein et al., 2006), where an important point of divergence between models is the question whether the terrestrial biosphere evolves as a sink or turns into a source for scenarios of continuing climate change.

To increase the understanding of the dynamics of theses sinks or sources it is useful to investigate not only the net ecosystem exchange (NEE), but also the two major underlying processes of the terrestrial carbon cycle: 1) the uptake of CO_2 from the atmosphere by plants via photosynthesis, gross primary production (GPP); and 2) respiration, the CO_2 release by the ecosystem (Reco) to the atmosphere, where organic carbon is used as an energy source for maintaining biotic activity (Fig. 1.1). The ecosystem respiration is an aggregated flux caused by heterotrophic respiration, the respiration of organisms whose primary energy source comes from other organisms and the autotrophic respiration of organisms whose primary energy source is the sun. Autotrophic respiration of plants can be described as being composed of respiration associated with the construction of new tissue (growth respiration), and the maintenance of existing tissue (maintenance respiration) (Cannell and Thornley, 2000) or is attributed to the different parts of a plant, e.g. roots, leaves, stem, branches (see Fig. 1.1. Sometimes a further division is made in above and below ground respiration. Horizontal fluxes like harvest or export of dissolved organic carbon are neglected here.

The two fluxes, GPP and Reco, are linked, as the carbon source for the respiratory process is provided by the photosynthetic uptake of carbon. The plants' biomass is eventually transferred to the ground as litter, and thereby becomes a major substrate for heterotrophic activity in the soil. Additionally plants directly influence soil respiratory processes via root respiration and by providing photo-assimilates feeding the metabolic processes of mycorrhizae, endophytes and microbial populations in the rhizosphere (Hogberg et al., 2001). The assimilates are easily degradable and may also increase microbial breakdown of more complex soil organic matter compounds by providing energy, this effect is known
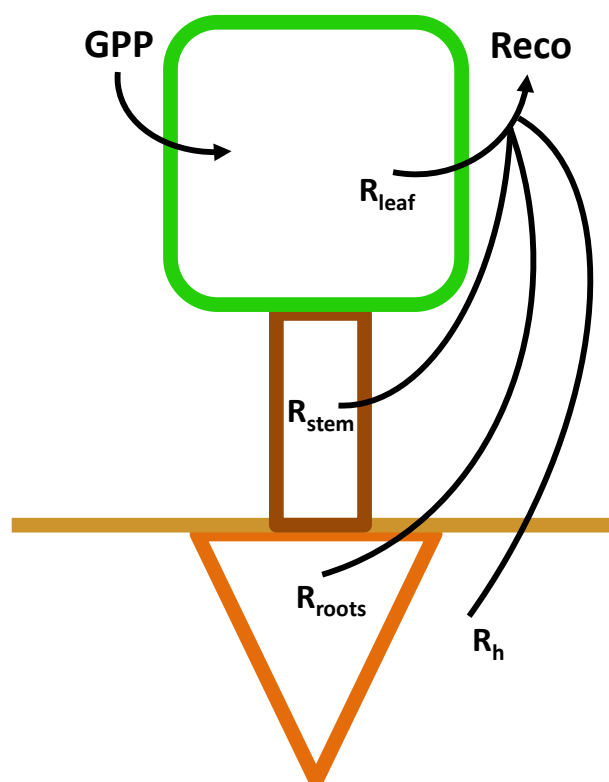
Figure 1.1.: The terrestrial carbon cycle can be simplified to two main fluxes: the gross primary production (GPP), where carbon is taken up and the ecosystem respiration (Reco) where carbon is released from the different carbon stocks in soil by heterotrophic respiration (R_h) and from biomass by roots (R_{roots}), stems and branches (R_{stems}) and leafs (R_{leaf}).

as the 'priming effect' (Bingeman et al., 1953).

Factors controlling the carbon cycle vary on different time scales (Mahecha et al., 2007). The major factors driving photosynthesis are the availability of water, light and nutrients on longer timescales the importance of factors related to changes in the vegetation structure increases (Richardson et al., 2007) that may be induced by temperature (Lucht et al., 2002). Decomposition of the organic material is largely controlled by the availability of material and the activity of the decomposer organisms, which in turn is affected by soil temperature, moisture, nutrient availability. Turnover times of the different biomass components range from minutes for the labile forms of soil organic matter to decades and thousands of years for humus (Chapin et al., 2002). For the net uptake, the seasonal and interannual changes are closely related to climate while on decadal timescale management and disturbance regulate the carbon exchange additionally (Barford et al., 2001). Increases in temperature correlate with both respiration and photosynthesis. A transition of ecosystems normally acting as a carbon sink towards a carbon source was found during the heat wave in 2003 (Ciais et al., 2005) due to increased drought and for the warming in autumn (Piao et al., 2008).

Abiotic sources of CO_2 releases can be photodegradation (Rutledge et al., 2010) and geochemical cycling in Karst systems (Kowalski et al., 2008). Since these sources are restricted to geographically small regions of the world and are hard to quantify, these fluxes are usually neglected.
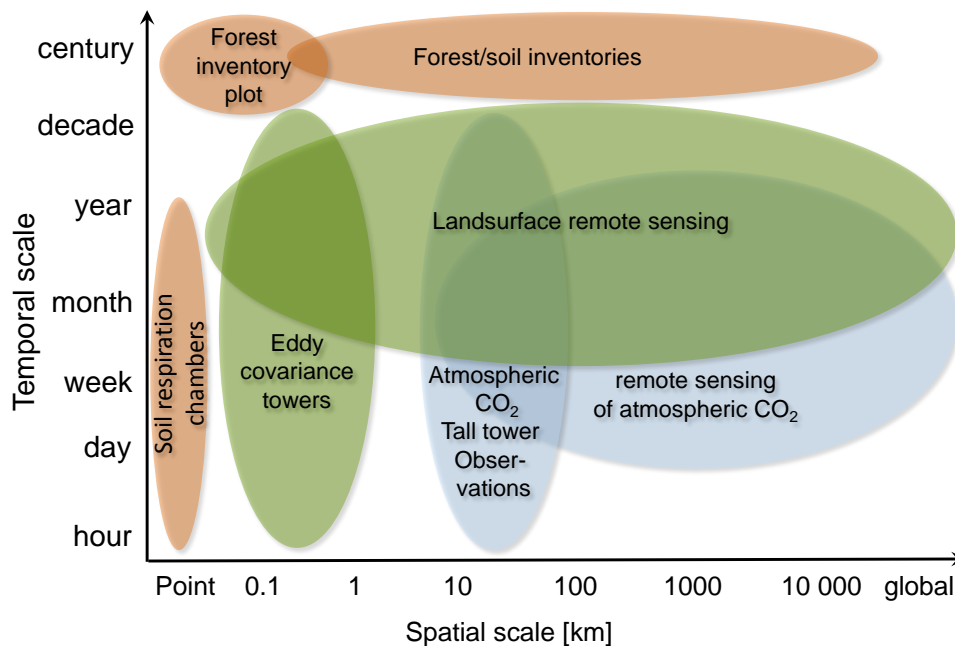
Figure 1.2.: Temporal and spatial scales spanned by available carbon cycle observations.

1.2. Carbon cycle observations

The need to predict future changes in the carbon cycle necessitates knowledge on how the terrestrial biosphere reacts to changing climate. Observations of the system are indispensable to quantify these reactions in forms of functional relationships to be able to project them to the future. Raupach et al. (2005) identify four essential kinds of data for terrestrial carbon observations: remote sensing of land surface properties, atmospheric composition measurements, measurements of the carbon stocks, and direct flux measurements. Integrating all the different observations and identifying consistency or reasons for inconsistency will challenge future research. These observations span a wide range of temporal and spatial scales (Fig. 1.2) and allow because of this to address different aspects of the carbon cycle and to constrain different parts of carbon cycle models. Atmospheric concentration observations integrate information of all carbon fluxes between atmosphere, biosphere and ocean, they are an important constraint for the regional patterns and the overall interannual response of the system (Bousquet et al., 2000). Carbon stock estimates are an important source of information for processes acting on longer time scales, e.g. the turnover of slow carbon pools or lagged effects on tree growth. The strength of remote sensing data are the spatial patterns and the global coverage, while flux data are important to address the direct influence of the meteorology (Friend et al., 2007) due to their high temporal resolution. The following paragraphs describe the eddy covariance flux data and remote sensing data used in this study in detail.

1.2.1. Eddy Covariance flux measurements

The eddy covariance technique revolutionized ecosystem science by providing direct, temporally continuous observations of the carbon, water and energy fluxes on ecosystem scale in a nondestructive way. The development of the theoretical framework can be traced back to the end of the 19th century (Reynolds, 1895). At that time the technical devices were not developed to observe the high frequency fluctuations of wind speed, temperature and humidity or even CO_2. Swinbank (1951) developed devices to measure the vertical transport of heat and water vapor, and conducted the first eddy covariance observations in the lower atmosphere. Fast responding CO_2 sensors were still lacking and the first CO_2 flux measurements were based on the gradient not on the eddy covariance method (Huber, 1950; Inoue, 1957; Lemon, 1960). First eddy covariance CO_2 flux studies emerged in the early 1970s with the aim to understand how weather variations affect crop performance (Desjardins, 1974; Desjardins and Lemon, 1974). The earliest continuous measurements started in the 90s (Wofsy et al., 1993), providing the first measurement of annual carbon uptake of an ecosystem, and continue until today. Regional networks developed in the following years and based on them a network of networks started, FLUXNET, coordinating regional to global analysis of flux observations. In 2006 a dataset across all networks and continents, the La Thuille dataset, was assembled and made available to the scientific community on the website http://fluxdata.org. The database now provides data from half-hourly to multi-annual time scales, across climate and vegetation types. The great amount of information in this database can help to elucidate a variety of questions from short term responses of the ecosystem carbon exchange to climate or to drivers of the interannual variability of the carbon dynamics, specific to or across climate or vegetation types.

In spite of the great potential of direct measurements there are drawbacks stemming from the assumptions on which the method is based. The major assumptions are: homogeneous flat terrain, fully turbulent flux, and stationarity (e.g. mean concentrations don't change significantly in time). Using these assumptions, neglecting density fluctuations and applying the Reynolds decomposition, the Navier-Stokes equation can be simplified. Based on high frequency measurements (usually 10 Hz) the turbulent flux can be derived as the covariance between vertical wind velocity and the concentration of the gas of interest, the horizontal wind velocity for the momentum flux or temperature for the sensible heat (Aubinet et al., 2000). In addition various corrections need to be applied based on these high frequency data: a coordinate transformation of the wind speed to minimize the vertical component (Wilczak et al., 2001), a correction for spectral losses due to path averaging and sensor separation (Moore, 1986; Massman, 2000), a correction for density fluctuations (Webb et al., 1980) and a correction of time lags and high frequency damping if a closed path measurement system is used (Leuning and Moncrieff, 1990). The flux values are usually computed for 30 min intervals, although studies suggest, that this intervall might be too short and missing low frequency components can lead to underestimation of the flux (Lenschow et al., 1994; Mauder et al., 2006). Methods have been developed to filter data where the conditions of stationarity and well developed turbulence are not met to guarantee the quality of the observations (Foken and Wichura, 1996; Papale et al., 2006). For inhomogeneous landscapes footprint models (Göckede et al., 2008; Rebmann et al., 2005) computing the area influencing the measurement help to assure the quality of the measurements, e.g. to assure that the measurement represents the area of interest.

A big effort has been made to standardize the processing of eddy covariance based on the 30 min flux

values with respect to data filtering and storage correction (Papale et al., 2006), gap filling (Falge et al., 2001; Moffat et al., 2007) and flux partitioning of NEE into GPP and Reco (Reichstein et al., 2005; Desai et al., 2008). The methods have been systematically compared for gap filling (Moffat et al., 2007) and flux partitioning (Desai et al., 2008). The standardization of the processing schemes is important to minimize differences between sites due to different processing algorithms, while for some sites a more detailed additional data treatment might be necessary (Rebmann et al., 2010) to account for site specific micrometeorology and topology.

One of the main and most discussed issues is advection that can introduce systematic biases (Aubinet, 2008; Feigenwinter et al., 2004; Kutsch et al., 2008; Marcolla et al., 2005; Montagnani et al., 2009, and many more). This phenomenon is related to low turbulent conditions but also to the heterogeneity and complexity of the measurement site. In complex terrain decoupling of air flow above and within plant canopies that is supported by low turbulence can lead to complex flow patterns (Rebmann et al., 2010). An expected effect of these advective fluxes that are usually not sampled with the eddy covariance system is an underestimation of nighttime fluxes as they usually occur more often during night when the turbulence is low. Extensive measurement campaigns using multiple towers were so far not able to find a general solution to this problem (Aubinet et al., 2010). In practice the data are filtered using a threshold for the friction velocity (u*), a measure of turbulence, data where the u* is below this threshold are filtered (Papale et al., 2006).

Besides meteorological and biological ancillary data that are observed at the site, the FLXUNET dataset includes estimates from flux partitioning methods where NEE is split into the two main components GPP and Reco. Various flux partitioning algorithms are available to estimate these fluxes using eddy covariance data and meteorological drivers (Desai et al., 2008; Reichstein et al., 2005; Stoy et al., 2006). Amongst them are methods that fit a respiration model to nighttime NEE data and extrapolate during daytime as well as methods that fit a light response curve to daytime data and estimate the respiration as the intercept. Methods have been compared for a limited number of sites, but the potential of assessing the impact of biases during nighttime by comparing the estimates based on nighttime data with estimates based on daytime data has not been exploited yet. Moreover the FLUXNET database so far only provides an estimate based on nighttime data.

Within the last years studies started to quantify the data uncertainty due to random errors (Hollinger and Richardson, 2005; Richardson et al., 2006b) or uncertainty of gap filling (Richardson and Hollinger, 2007). These studies provided the statistical tools to estimate the random error of the data, but not all properties important in statistical analysis have been quantified yet. Moreover studies showed that the uncertainty due to systematic errors like the heuristic determination of the threshold to filter low turbulence conditions introduces the largest part of the uncertainty (Papale et al., 2006) and thus need to be quantified and provided to improve the value of the data.

1.2.2. Remote sensing of land surface properties

Remote sensing data capture spatially explicit information about the reflectance of the land surface at large and synoptic scales. These reflectances can be transferred into parameterizations of the land surface that are needed to globally model the carbon cycle (Hilker et al., 2008). Estimates of photosynthetic active radiation were shown early to correlate with atmospheric CO_2 concentration, showing the poten-

tial of remote sensing data to derive the photosynthetic carbon uptake of the biosphere (Tucker et al., 1986). This makes remote sensing a valuable tool in extrapolating ecological knowledge from local to regional and global scale. Remotely observed reflectance data have been used to derive characteristics of the surface like leaf area index (Curran and Steven, 1983), vegetation type (Jung et al., 2006), biomass (Lefsky et al., 2002a) or tree height (Lefsky et al., 2002b), but also soil water content (Gillies et al., 1997), frequency and extent of fire (Robinson, 1991), land (Qin and Karnieli, 1999)and ocean (Schluessel et al., 1990) surface temperature and vegetation stress (Goerner et al., 2009; Krumov et al., 2008), and calculation of various ocean color indices and vegetation indices (Roughgarden et al., 1991). This is possible as plant physiological properties are related to the biochemical composition of the plant foliage, which determines the spectral reflectance properties. The spectral characteristics of a leaf may change as a result of physiological changes, water deficit (Mu et al., 2007) or a change in nutrient status (Jackson, 1986; Peñuelas et al., 1994). An important remotely sensed measure for the modeling of GPP is the fraction of absorbed photosynthetically active radiation (fAPAR), being the driving variable in the light use efficiency approach to model GPP (Hilker et al., 2008; McCallum et al., 2009; Garbulsky et al., 2010). The derivation of fAPAR can be based on empirical or physical models and varies considerably between the products due to different temporal and spatial resolution, retrieval methods and use of LAI and land cover products (McCallum et al., 2010).

The major advantage of remote sensing data is that they are able to cover the whole spatial domain of the globe. The combination of FLUXNET point data with remote sensing data for the upscaling of the direct observations of fluxes to the globe can help to overcome the weak point of global representativness of FLUXNET (Jung et al., 2009; Beer et al., 2010). But for the combination it needs to be assured that the remote sensing pixel agrees with the area of influence, the footprint, of the eddy covariance tower. Moreover the product needs to capture the characteristics that are included in the model, e.g. remote sensing observations of leaf area index potentially include the understorey of a forest, this needs to be considered.

1.3. Modeling the carbon exchange of ecosystems

Carbon cycle models are needed to be able to estimate regional and global carbon budgets from the available observations. The overall scientific objective behind modelling the carbon cycle is to be able to model future changes of atmospheric CO_2 which is needed for climate predictions.

Models used differ in the amount of prior knowledge about processes implemented in their structure and the choice of the model depends on the questions that are to be answered. Process oriented models try to derive the model structure from physical principles. Depending on the purpose, these models include representations of the instantaneous fluxes of energy, water and carbon between land and the atmosphere, seasonal variations in phenology and biomass growth, to long-term dynamics of vegetation structure and composition (Prentice et al. 2007). Such models are required to assess the impact of future environmental changes or anthropogenic influences, and to extrapolate from known conditions to other environments which is necessary for predictions.

Diagnostic models are much less complex they complement process oriented models by being closer to the data and by including only few prior information about processes. The diagnostic models are usually still based on physical concepts and can be distinguished from purely data driven methods like artifi-

cial neural networks (Papale and Valentini, 2003) or model tree ensembles (Jung et al., 2009). These methods derive the relation between target and driving variable from a training dataset involving no assumptions about the link between these variables. In combination with eddy covariance data diagnostic models and data driven methods are used for deriving regional estimates from point measurements (upscaling) (Beer et al., 2010; Jung et al., 2009; Papale and Valentini, 2003), gap filling (Falge et al., 2001; Moffat et al., 2007), flux partitioning of the net exchange into uptake and release of carbon (Desai et al., 2008; Falge et al., 2002a), or exploration of data time series (Reichstein et al., 2002, 2003a; Owen et al., 2007).

The following two sections introduce modelling approaches used in this study.

1.3.1. Flux partitioning

The models applied for the partitioning of NEE into GPP and Reco need to be as close as possible to the data including only few assumptions to be useful as additional constraint for process based models. This is usually achieved by fitting empirical models to small data windows, the complex long term behaviour of the ecosystem is then modelled implicitly by parameter variations. For the flux partitioning of NEE into GPP and Reco two approaches are commonly used (Reichstein et al., 2005; Stoy et al., 2006). Both methods exploit the fact that photosynthesis is zero at night. The first fits a respiration model to nighttime data and using daytime temperature, respiration is extrapolated to the daytime and the difference between modeled Reco and measured NEE yields estimated GPP. The second fits a light-response curve to daytime NEE measurements and the respiration is estimated from the intercept of the ordinate. For respiration many diagnostic formulations of the nonlinear relation between temperature and respiration exist and are in use (Katterer et al., 1998; Kirschbaum, 1995; Lloyd and Taylor, 1994). Here, an Arrhenius-type model after Lloyd and Taylor (1994) is used to describe the temperature dependence of Reco. The light response curve computes NEE as a function of radiation. For this model a rectangular and a more flexible nonrectangular form are in use (Gilmanov et al., 2007; Stoy et al., 2006). Extensions to account for the temperature sensitivity of respiration have been introduced Gilmanov et al. (2003a), but it is usually neglected, that additionally NEE is also a function of VPD, affecting GPP via stomatal regulation.

1.3.2. Process based modelling of photosynthesis

In process-based models, photosynthesis for C3 plants is usually represented by the Farquhar model (Farquhar et al., 1980) or variations of it using different approaches to include the effect of quantum flux density (Harley et al., 1992; McMurtrie et al., 1992), coupled it to a stomatal model (Ball et al., 1987), or dealt with additional constraints, e.g. phosphorus (Harley et al., 1992) or nitrogen (Kattge et al., 2009). For C4 plants photosynthesis is usually implemented after Collatz et al. (1992). To describe the varying response and potentials of different plants to meteorology and resources the biosphere is grouped into plant functional types. Process-based models may be parameterized for individual sites or species, but global models typically represent only a limited amount of functional types to describe vegetation because the availability of data required for parameterization increased only in the recent years. The variations of the parameters need to be investigated within the groups of plant functional types but also in time to identify possible model improvements. A key parameter in the photosyntesis

model is the photosynthetic capacity of a leaf (e.g. the maximum carboxylation rate, Vcmax). A seasonality of Vcmax has been observed on leaf level (Misson et al., 2007; Muraoka et al., 2010) and an influence of drought has been found in an inverse modelling study(Reichstein et al., 2003b). Although this variability is known Vcmax is usually implemented in models as a constant parameter due to a lack of a functional description of the variations. With the increasing availability of data it is possible to quantify these responses as for instance done for the relation between Vcmax and leaf nitrogen content (Kattge et al., 2009).

1.4. Model data fusion

Model data fusion (MDF) is the process of linking models with data (Williams et al., 2009) or in a more specific definition the matching of model predictions and observations by varying model parameters or states (Wang et al., 2009). The purpose of this procedure is to evaluate models, falsify model structures, adjust models to specific systems, estimate parameters that are not directly observable and to reduce the model's uncertainties (Williams et al., 2009). Model data fusion can comprise multiple steps of linking the model with data, e.g. model characterisation, parameter estimation, generalisation, and model validation (Williams et al., 2009). For model improvement these procedures are repeated with a stepwise reformulation of the model (Williams et al., 2009).

Model data fusion methods, were used first in the fields of weather forecasting and hydrology and are then more often referred to as data assimilation (Rutherford, 1972; Bouttier et al., 1993). Due to the necessity to decrease uncertainties in model parameterisations (Bonan, 2008) they are now being used extensively by the C cycle community. This development can be related to the fact that the availability of carbon cycle observations is strongly increasing, for instance the FLUXNET dataset of 2007 includes almost 1000 years of half hourly data measured across sites of different climates and vegetations. New satellites are launched (Yokota et al., 2009), but also extensive measurements like soil carbon stocks increase in density and replications. For a comprehensive analysis of such big and growing datasets available methods need to be adjusted to combine the information from the datasets in a consistent way and to extract reliable information from the data streams.

Trying to link models with data the comparison of model output and observations is always the first step. The central part of the MDF procedure is the minimization of the model data mismatch by varying parameters or state variables. In statistics this part is known as inverse problem, as opposed to the forward problem (Evans and Stark, 2002). In the inverse problem observations of the target variable of the model exist and parameters of the model are optimized to gain the best match between model output and observations. Besides the choice of the model and the type of observations the third question is how to define the distance between them that is to be minimized. The distance between model and observations is a function of the parameter values, named cost function, objective function or misfit. There are various optimization algorithms available for the identification of the parameter set that minimizes the cost function, the choice of the method depends on the problem, mainly the number of parameters, shape of the cost function and feasibility in terms of computer power.

1.4.1. Cost function

The cost function J (also objective function or misfit) defines how the distance between model output and observations is quantified, it is a function of the parameter set p. The most simple example is the ordinary least squares cost function, that consists of the squared differences between model and observation:

$$J(p) = \sum_{i=1}^{N} (x_{d_i} - x_{m_i})^2 \qquad (1.1)$$

where x_{d_i} are the observations, x_{m_i} are the modelled values and N is the number of data points to be compared. To yield the maximum likelihood estimates of the parameters the errors of the data are assumed to be uncorrelated, homoscedastic (uncertainty of each data point is the same) and normally distributed (Tarantola, 1987). This function can be generalized for correlated and heteroscedastic errors by including the error covariance matrix $\mathbf{C}_d$:

$$J(p) = (\vec{x}_d - \vec{x}_m)^T \mathbf{C}_d^{-1} (\vec{x}_d - \vec{x}_m). \qquad (1.2)$$

Here the uncertainty of data due to a random error, e.g. its standard deviation, is included on the diagonals of $\mathbf{C}_d$, the off diagonal elements of the matrix contain the correlations between the errors. A non-Gaussian distribution changes the measure of the distance, for instance a Laplacian distribution of the errors requires the minimization of absolute differences (Hollinger and Richardson, 2005). This textbook approach does not offer a solution to account for uncertainties due to systematic errors, usually they should be removed by calibration or filtering before trying to match observations with model output.

An extension to this "'frequentist approach"' that uses only information about the target variable is based on the Bayesian theorem. Bayesian parameter estimation allows to not only consider the difference between observations and modelled values in the cost function, but also prior information about parameters can be included in the optimization (Omlin and Reichert, 1999).

$$J(p) = (\vec{x}_d - \vec{x}_m)^T \mathbf{C}_d^{-1} (\vec{x}_d - \vec{x}_m) + (\vec{p}_{prior} - \vec{p}_{post})^T \mathbf{C}_p^{-1} (\vec{p}_{prior} - \vec{p}_{post}). \qquad (1.3)$$

The prior information includes the most plausible value for the parameter $\vec{p}_{prior}$ and its uncertainty and correlations between the parameters in form of a covariance matrix $\mathbf{C}_p$. The Bayesian cost function adds the distance between prior and posterior (optimized) parameter ($\vec{p}_{post}$) weighted with $\mathbf{C}_p$. This information can be derived from literature making use of previous estimates of the parameter, it can be a physically plausible range, e.g. many parameters can't be negative or have a natural upper limit, or it can be based on expert knowledge. Including prior knowledge can be necessary for a successful optimization if many parameters are to be optimized. The cost function then often has a complex shape, contains multiple local minima or the same value of the costfunction can be achieved with different combinations of parameter values, often referred to as equifinality (Beven, 2006). Reducing the occurence of local minima can be also approached by including multiple data streams, for instance carbon flux and carbon stock data, in the costfunction (Carvalhais et al., 2010; Luo et al., 2009).

Model errors are commonly neglected, e.g. it is assumed that the model is able to reproduce the

patters in the data except for the random errors with the characteristics specified in the cost function. Neglecting these errors in the model data synthesis the uncertainty is underestimated and parameters can be biased (Carvalhais et al., 2008).

1.4.2. Optimization algorithms

Optimization algorithms are necessary to identify the parameter set of a model that minimizes the cost function, e.g. to find the minimum distance between the model and the observations of the system. While for one or two parameters and one dataset with a good knowledge of the model behaviour the parameters can be possibly tuned by hand, for higher dimensions the process of finding the optimum needs to be automated. Different approaches to finding the minimum of the cost function exist. The available methods can be distinguished into (1) methods that are based on the gradient of the cost function, here the gradient gives the direction of the next parameter set to be explored until a minimum is found, (2) the global search methods explore the whole parameter space to avoid local minima.

Examples for the first type is for instance the Levenberg-Marquard algorithm (Levenberg, 1944; Marquard, 1963) that is based on the Newton method (Tarantola, 1987) but more robust, e.g. the probability of convergence is higher. If a gradient based method is used the construction of the adjoint model that directly computes the sensitivity of the cost function to the parameters, increasing the efficiency of the procedure (Kaminski et al., 2002). The main drawback of gradient descent methods is that they may find local minima instead of the global mimimum. The number of local minima increases with the number of parameters included in the optimization and/or with the complexity of the model Luo et al. (2009). Thus global search methods are advantageous if a large number of parameters needs to be optimized, while the gradient descent methods are more parsimonious concerning computer power.

1.4.3. Challenges and problematic issues

The statistical theory offers standard approaches based on theoretical assumptions about the data and the inherent data errors (Draper and Smith, 1998). These theories and especially their assumptions on data and models need to be verified when applied to real observations. The theory covers the central part of Fig. 1.3 including the optimization of a cost function and the propagation of uncertainties based on a parameter distribution to the model output. While the characterization of model and data errors and uncertainties in reality often requires modifications of the standard approaches. Using synthetic data the limitations of statistical frameworks and existing knowledge have been explored. Trudinger et al. (2007) have shown that while the choice of the optimization algorithm is of minor importance, the characterization of the errors is pivotal to retrieve unbiased estimates of parameters, especially non Gaussian and correlated errors can bias the results.

For eddy covariance data methodologies to estimate the uncertainty have been developed (Hollinger et al., 2004; Hollinger and Richardson, 2005; Richardson et al., 2008). Statistical properties of the random error of eddy covariance data were only partly characterized, the variance of the error has been investigated and found to increase with increasing flux magnitude, but also being influenced by wind speed. The distribution of the random error in eddy covariance data has been characterized as rather Laplacian, although not taking into account the varying standard deviation, that not only scales with
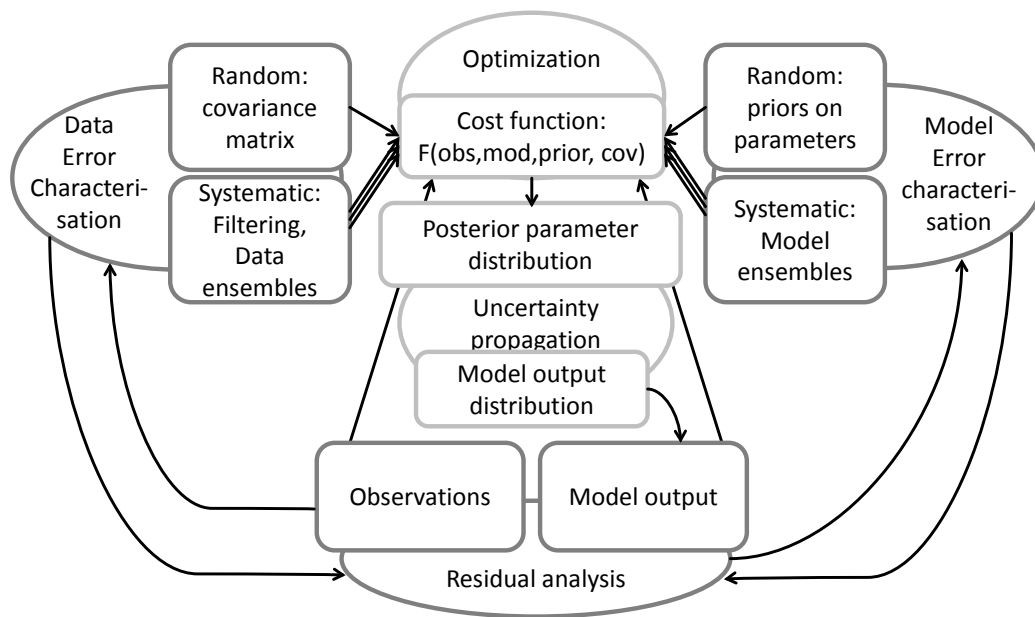
Figure 1.3.: Scheme for model data fusion considering systematic and random errors in data and model. Ellipses represent general procedures, boxes are data or specific calculations within the procedure.

the flux magnitude (Hollinger and Richardson, 2005). A Laplacian distribution requires the minimization of absolute differences and reduces the annual integral of carbon fluxes about 10% (Richardson and Hollinger, 2005), thus the question is of quantitative importance when constraining models with flux observations. Correlations between the errors have not been investigated so far and were always assumed to be zero. The problem of deriving an estimate for the systematic error has been partly addressed for eddy covariance data by using a bootstrapping approach to determine a threshold used for filtering the data (Papale et al., 2006). However this is only one possible source of systematic errors. Moreover an option to include the uncertainty due to systematic errors for eddy covariance data in a model data fusion framework and to propagate the uncertainties to the model output has not been developed so far. These problems need to be addressed and solved before being able to combine multiple data streams in a theoretically sound way to constrain land surface models with all the information that is available in form of observations.

1.5. Goals of this study

The main goal of this study is to improve the understanding of the carbon cycle on ecosystem level by fusing eddy covariance data with carbon cycle models of varying complexity. The study adresses two major topics:

1. the first part improves the understanding of the relation between Reco and GPP: an improved algorithm based on a light response curve is introduced for flux partitioning of NEE into the flux components GPP and Reco and provides a new daytime data based estimate for FLUXNET to complement the available nighttime data based estimate (Reichstein et al., 2005). The light

response curve is extended to account for the temperature sensitivity of respiration and the importance of the effect of water vapour pressure deficit (VPD) on photosynthesis is assessed. The relation between GPP and Reco is explored in greater detail on global scale, e.g. across all FLUXNET sites and potential biases of this relation due to statistical artefacts are discussed.

2. the second part improves the methodology of model-data fusion with eddy covariance data by providing a detailed analysis of the random error and a framework to propagate uncertainties due to both systematic errors and random errors in eddy covariance data to parameter estimates and model output. Here the combined use of daytime data based estimate of GPP and Reco developed in the first part and the already available nighttime data based estimates improves the reliability of the results.

The net flux observed by eddy covariance systems results from two larger fluxes of opposite sign, the uptake of carbon by photosynthesis and the release of carbon by respiration. For further insight about the processes, the derivation of the flux components photosynthesis and respiration is important to be able to interpret the observed variability and attribute it to the underlying process. Within the FLUXNET community usually an estimate based on nighttime NEE data is used (Reichstein et al., 2005). An alternative approach based on light response curves avoids the potentially biased nighttime data, but usually neglects either the influence of water vapour pressure deficit (VPD) or the influence of temperature, or both. In chapter 2 we introduce a flux partitioning algorithm based on daytime eddy covariance data using a light response curve including both, the effect of temperature on respiration and a VPD limitation on GPP. We illustrate the importance of including VPD in the fluxpartitioning algorithm. This estimate is complementary to the standard nighttime estimate (Reichstein et al., 2005), using a different subset of NEE data, a different model and a different data window in the fitting algorithm. Critical issues are discussed and the algorithm is evaluated across FLUXNET sites by comparing it with a nighttime based method at the same time assessing the impact of the potentially biased nighttime data. In the last part of the chapter the global relation between Reco and GPP is revisited using the independent estimates.

In a statistical analysis the use of these two estimates helped to assess the potential occurence of statistical artifacts in the correlation pattern between the two flux components GPP and Reco: The strong correlation between the two flux components has been explored in various studies (Baldocchi, 2008; Fischer et al., 2007; Reichstein et al., 2007; Wang et al., 2008). These studies have relied on the GPP and Reco estimates that were derived from the same NEE data. The correlation of estimates derived in that way has been suggested to be almost completely spurious (Vickers et al., 2009) because GPP is derived as the residual between NEE and Reco and any error in Reco directly propagates into GPP, thus the FLUXNET database would be of limited use to address the relation between the two flux components. In chapter 3 we argue that the spurious correlation is only caused by errors in Reco, as the errors are directly propagated to GPP while the true Reco is only removed from the NEE signal to derive GPP. In appendix B we follow up on chapter 3 exploring how big the effect of the error in Reco can be by using the combination of independent flux components as explained in chapter 2 for a couple of sites on different time scales.

The methodological improvement focuses on the quantification and characterization of uncertainties in eddy covariance data. It is necessary to distinguish between random and systematic errors to be able to include them in model data fusion studies. The characterization of the random error is not only

useful for data assimilation but for most statistical analysis. Statistical methods for testing hypotheses usually rely on assumptions about the random error characteristics, often Gaussian distribution is assumed. This topic was partly addressed for eddy covariance data. The random error has been found to have a non constant variance with a Laplacian distribution (Hollinger and Richardson, 2005; Richardson and Hollinger, 2005). However it remained unclear whether the Laplacian distribution is caused by the varying standard deviation. The correlations between the errors of latent and carbon fluxes and the autocorrelation have not been adressed. In chapter 4 the random error of eddy covariance data is estimated and its standard deviation, distribution and correlations characterized for 4 CarboEurope sites with different climate and vegetation type. Using synthetic data the influence of the random errors and the choice of the cost function on parameter estimation and the parameter uncertainty is illustrated and compared with the effect of systematic errors in observations on parameter retrieval.

In model data fusion studies so far systematic errors were neglected and the uncertainties of parameters and model output therefore underestimated. For annual estimates the uncertainty of the u* threshold has been shown to have a large influence and to cause systematic errors (Papale et al., 2006). In chapter 5 we suggest a model data fusion framework that includes uncertainties due to systematic errors and combines it with the uncertainties arising from random errors. The approach makes use of multiple datasets representing a plausible range of systematic errors in eddy covariance data and fAPAR data. With the combined use of site level, remote sensing and eddy covariance data to estimate parameters of a land surface model the potential to constrain the seasonality of LAI, fAPAR and photosynthetic capacities is explored.

Chapter 6 gives a summary with overall conclusions and implications followed by an outlook of future research following up on the findings of this thesis.

2. Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: critical issues and global evaluation

2.1. Abstract

The measured net ecosystem exchange of CO_2 between the ecosystem and the atmosphere (NEE) reflects the balance between gross CO_2 assimilation (GPP) and ecosystem respiration (Reco). For understanding the mechanistic responses of ecosystem processes to environmental change it is important to separate these two flux components. Two approaches are conventionally employed: 1) respiration measurements made at night are extrapolated to the daytime or 2) light-response curves are fit to daytime NEE measurements and respiration is estimated from the intercept of the ordinate, which avoids the use of potentially problematic night-time data. We demonstrate that this approach is subject to biases if the effect of vapor pressure deficit (VPD) modifying the light response is not included. We introduce an algorithm for NEE partitioning that uses a hyperbolic light response curve fit to daytime NEE, modified to account for the temperature sensitivity of respiration and the VPD limitation of photosynthesis. Including the VPD dependency strongly improved the model's ability to reproduce the asymmetric diurnal cycle during periods with high VPD, and enhances the reliability of Reco estimates given that the reduction of GPP by VPD may be otherwise incorrectly attributed to higher Reco. Results from this improved algorithm are compared against estimates based on the conventional night-time approach. The comparison demonstrates that the uncertainty arising from systematic errors dominates the overall uncertainty of annual sums (median absolute deviation of GPP: 47 g C m^{-2} yr^{-1}), while errors arising from the random error (median absolute deviation: 2 g C m^{-2} yr^{-1}) are negligible. Despite site specific differences between the methods, overall patterns remain robust, adding confidence to statistical studies based on the FLUXNET database. In particular we show that the strong correlation between GPP and Reco is not spurious but holds true when quasi-independent, i.e. day-time and night-time based estimates are compared.

2.2. Introduction

The eddy covariance technique measures the net ecosystem exchange (NEE) of CO_2, the balance between carbon released and taken up by ecosystem respiration (Reco) and gross primary production (GPP). The separation of NEE into its components, which represent underlying processes, helps obtain mechanistic, process-level understanding of the terrestrial carbon cycle. Global, multi-site flux synthesis studies require that NEE be partitioned in a standardized manner, to minimize site-specific biases due to differences in processing (Aubinet et al., 2000; Baldocchi, 2008, 2003; Foken and Wichura, 1996; Göckede et al., 2008; Papale et al., 2006; Rebmann et al., 2005; Reichstein et al., 2005). Various flux partitioning methods are available and have been previously compared using measured or modelled

data from single or multiple sites (Desai et al., 2008; Hagen et al., 2006; Reichstein et al., 2005; Stoy et al., 2006; Yi et al., 2004). Methods that rely on nighttime data for partitioning may be biased due to the frequent nighttime suppression of turbulence and dominance of advective fluxes not measured by conventional EC systems (Aubinet, 2008; Aubinet et al., 2000; Feigenwinter et al., 2004; Goulden et al., 1996). The second common approach, extrapolating respiration from light-response curves conditioned on daytime data, usually does not account for the fact that NEE varies both as a function of temperature (mostly affecting Reco) and vapor pressure deficit (affecting GPP via stomatal regulation), among other factors. Confounding effects introduced by this shortcoming may have contributed to the large observed between-method variability in extracted diurnal cycles of Reco (Desai et al., 2008). The diurnal cycle of NEE observations during dry periods with high VPD often has an asymmetric shape that is partly caused by higher respiration in the afternoon but also due to stomatal limitation of GPP as VPD tends to peak well after maximum diurnal radiation. As a consequence, measured carbon uptake at the same level of insulation may be substantially lower in the afternoon compared to morning hours. This phenomenon has effects on carbon gain and water-use efficiency of the ecosystem as well as partitioning of sensible and latent heat fluxes between the land surface and the atmosphere (Baldocchi, 1997; Reichstein et al., 2003b; Williams et al., 1996). In this study we address the following questions: (1) whether it is necessary to include VPD effects on photosynthesis when partitioning measured NEE using a light-response curve approach, (2) whether estimated annual sums of carbon fluxes based on day-time data show systematic differences compared to those based on night-time data, and (3) whether this affects the strength of the often-noted relation between annual GPP and Reco (Janssens et al., 2001; Reichstein et al., 2007; Wang et al., 2008). Here, we perform these analyses for the first time using a quasi global biosphere-atmosphere carbon dioxide flux data set.

2.3. Methods

2.3.1. Data

We used data from the FLUXNET "La Thuile" database (www.fluxdata.org), where half hourly data had been provided by site managers and further processed in a standardised methodology described in Papale et al. (2006) and Reichstein et al. (2005). While an unprecedented level of standardization has been achieved in this database, one should still note that the derivation of half hourly fluxes from the high frequency raw data still varies from site to site (Aubinet et al., 2000; Foken and Wichura, 1996; Lee et al., 2004; Moore, 1986). We used the database version of December 2007 containing 976 site-years of half hourly eddy covariance data. The data are storage corrected, spike filtered, u*-filtered, and subsequently gap-filled. For the optimization of the model parameters only measured (i.e. non-gapfilled) half hourly data were used. The sites chosen for the first part of the study (Table 2.1) were selected to cover a wide range of climates and vegetation types and to meet the requirement of a high fraction of original non-gapfilled flux observations. The second part of the study, the global comparison of nighttime based and daytime based estimates, included all FLUXNET sites that satisfied the following criteria: (1) data availability for the whole year is higher than 80%. (2) data availability was sufficient to allow the estimation of the light-response curve parameter time series with no gaps larger than 750 hours during the whole year (3) the statistical uncertainty, due to the uncertainty of the estimated parameters (see 2.4) of the annual GPP estimate was below 20 $gCm^{-2}y^{-1}$. The third criterion

Table 2.1.: Eddy covariance sites selected for the first part of this study, EBF: evergreen broadleaf forest, GRA: Grassland, CRO: Crops, DBF: Deciduous broadleaf forest, ENF: Evergreen needle leaf forest, WSA: Woody Savanna. VPD range is the mean diurnal VPD range of the data used in Fig. 2.1.

Site code	Name	Country	Lat	Lon	Veget. type	Year	VPD range
FR-Pue	Puechabon	France	43.74	3.6	EBF	2001	4.1-25.9
US-IB2	IL-Fermi National Accelerator Laboratory-Batavia	USA	41.84	-88.24	GRA	2005	3.4-30
US-Bo1	IL-Bondville	USA	40.01	-88.29	CRO	2000	0.4-20.9
DE-Hai	Hainich	Germany	51.08	10.45	DBF	2003	3.1-17.6
CA-Sj1	Sask.-1994 Harv. Jack Pine	Canada	53.91	-104.66	ENF	2005	1.6-22.7
CA-Oas	Sask.-SSA Old Aspen	Canada	53.63	-106.2	DBF	2003	2.4-21.3
BW-Ma1	Maun-Mopane Woodland	Botswana	-19.92	23.56	WSA	2000	3.7-35.2
BR-Ma2	Manaus-ZF2 K34	Brazil	-2.61	-60.21	EBF	2005	3.2-22.2

added to exclude extrapolation to conditions far from the data used for fitting, but only five site-years were affected additionally by this last criterion. After applying these criteria 417 site-years out of 976 from 145 sites were included in the comparison (site details are given in Appendix B). 511 sites were affected by criterion (2), 273 sites by criterion (1), 265 by both criteria (1 and 2).

2.3.2. Models

In this study we compare three different algorithms to partition NEE into GPP and Reco; we are implicitly assuming that geochemical (i.e. non-biological) processes can be ignored in this partitioning (Kowalski et al., 2008; Hofmeister, 1997). In all cases, models were fit to a short time window (4-15 days) to account for seasonal parameter variability, reflecting changes in the state of the ecosystem that are not represented in the models. The algorithm of the daytime data based estimates is described in detail in Appendix A.

Nighttime data based estimate

This estimate is according to Reichstein et al. (2005), which is currently used to partition data in the FLUXNET database compilation and available as online tool at http://gaia.agraria.unitus.it/database/eddyproc. Briefly, GPP is assumed to be zero during nighttime periods (defined here as global radiation (Rg) < 20 Wm^{-2}) and measured NEE is composed entirely of Reco, to which a model is fit and extrapolated to daytime periods. An Arrhenius-type model after Lloyd and Taylor (1994) is used to describe the temperature dependence of Reco:

$$Reco = rb \cdot exp \left(E_0 \left(\frac{1}{T_{ref} - T_0} - \frac{1}{T_{air} - T_0} \right) \right) \tag{2.1}$$

where rb [μmol C m^{-2} s^{-1}] is the base respiration at the reference temperature (Tref [°C], set to 15°C), E_0 [°C] the temperature sensitivity, and parameter T_0 [°C] is kept constant at -46.02 °C as in Lloyd and Taylor (1994). For E_0 a constant value is used for the whole year while rb was estimated every 5

days using a 15 days window (as in Reichstein et al. (2005)). Using daytime temperature, respiration is extrapolated to the daytime and the difference between modeled Reco and measured NEE yields estimated GPP. We refer to this estimate as "NB" (nighttime data-based).

Daytime data-based estimate including temperature sensitivity of respiration

For the daytime data based estimate NEE was modelled using the common rectangular hyperbolic light response curve:

$$\text{NEE} = -\frac{\alpha \cdot \beta \cdot R_g}{\alpha \cdot R_g + \beta} + \gamma \tag{2.2}$$

where NEE is net ecosystem exchange, α [μmol C J^{-1}] is the canopy light utilization efficiency and represents the initial slope of the light response curve, β [μmol C m^{-2} s^{-1}] is the maximum CO_2 uptake rate of the canopy at light saturation, γ [μmol C m^{-2} s^{-1}] is the ecosystem respiration and Rg is the global radiation [Wm^{-2}]. Although the nonrectangular light response model was shown to improve results, here we preferred the parsimonious rectangular curve. Gilmanov et al. (2003b) found that for the respiration parameter the differences between the two models, rectangular and nonrectangular, are small (<10%). We modified the hyperbolic light response curve to account for the temperature dependency of respiration after Gilmanov et al. (2003a) by replacing the constant respiration γ with a respiration model, in this case the Lloyd & Taylor model Lloyd and Taylor (1994) as given in Eq. (1)

$$\text{NEE} = -\frac{\alpha \cdot \beta \cdot R_g}{\alpha \cdot R_g + \beta} + rb \cdot exp\left(E_0 \left(\frac{1}{T_{ref} - T_0} - \frac{1}{T_{air} - T_0}\right)\right) \tag{2.3}$$

T_{ref} and T_0 were fixed as in the nighttime data based approach. The other parameters (E_0, rb, α, β) of the model were estimated in two different ways: (1) E_0 was estimated using nighttime data (Rg < 4 Wm^{-2}), then E_0 was fixed and rb, α, β were derived from daytime data ("DB noVPD", daytime data based, E_0 estimated with nighttime data). (2) all parameters (E_0, rb, α, β) were estimated using daytime data ("DB noVPD all", daytime data based with all parameters estimated using daytime data). The upper bound of the parameter E_0 as given in Table A1 was not used, as otherwise often the E_0 parameter was rejected during periods with high VPD. For estimates of daily or annual NEE, respiration was extrapolated into the nighttime using T_air measured during the night and the values obtained for E_0 and rb. The threshold for the definition of nighttime data (Rg < 4 Wm^{-2}) is lower here than in the nighttime data based approach, as excluding all data with Rg < 20 Wm^{-2} leads to long gaps for high latitude sites.

Daytime based estimate including temperature sensitivity of respiration and VPD limitation of GPP

The second modification of the hyperbolic light response curve accounts for the VPD limitation of GPP. Here, the fixed parameter β in eq (4) was replaced with an exponential decreasing function (Körner,

1995) for β at high water vapor pressure deficit (VPD):

$$\beta = \begin{cases} \beta_0 \cdot exp\left(-k \cdot (VPD - VPD_0)\right) & \text{if VPD} > \text{VPD}_0 \\ \beta_0 & \text{otherwise} \end{cases} \tag{2.4}$$

Please note that the VPD in the atmosphere is used here, while physiologically more relevant would be the leaf-to-air VPD which is higher or lower than atmospheric VPD when leaf temperatures are higher or lower then air temperature, respectively. For the empirical purpose of this study we deem the use of atmospheric VPD sufficient, given the fact that leaf-to-air VPD (or leaf temperatures) is usually not observed at FLUXNET sites. The k parameter was estimated for each 4 day data window to quantify the response of the maximum carbon uptake to VPD. Since we found that the parameter k was not well constrained after including the VPD_0 in the optimization, the VPD_0 threshold was set to 10 hPa in accordance with earlier findings at the leaf level (Körner, 1995), at this point ignoring potential vegetation specific differences. We will refer to this method as "DB VPD" (daytime data based including VPD). E_0 was estimated using nighttime data as in the "DB noVPD"-method and α , β_0, k and rb were estimated using daytime data (Appendix A).

2.3.3. Parameter estimation

We assume a serially uncorrelated Gaussian distributed random error and a heteroscedastic flux magnitude-varying standard deviation of the random error as found by Lasslop et al. (2008). Hence, parameter estimation made use of this information by applying a weighted least squares cost function (cf.Hollinger et al., 2005). We estimated the error standard deviation of the data (data uncertainty), σ_{meas}, for each data point following Lasslop et al. (2008) and used these estimates to weigh the data in the cost function in Eq. (5). The optimal parameters are found by minimizing the weighted least squares cost function J:

$$J = \sum_{i=1}^{n} \frac{\left(y_{meas,i} - y_{mod,i}\left(p\right)\right)^2}{\sigma_{meas,i}^2} \tag{2.5}$$

where y_{meas} is the observed value and y_{mod} is the parameter dependent modeled value. The parameters were estimated using the Levenberg-Marquardt algorithm of the PV-Wave advantage software package (Visual Numerics, 2005).

2.3.4. Statistical uncertainty of the model output (GPP)

The uncertainty estimate of the model output is based on the classical frequentist approach as described in Omlin and Reichert (1999). The covariance matrix of the model parameters is used to calculate the uncertainty of the model output by linear error propagation:

$$Cov(y_{mod}) = \left(\frac{\partial y_{mod}}{\partial p}\right) Cov(p_{opt}) \left(\frac{\partial y_{mod}}{\partial p}\right)^T \tag{2.6}$$

When interpolating between the model output of two parameter sets (see description of the algorithm, Appendix) the error variance was interpolated as follows:

$$Var(y) = w_1^2 Var(y_1) + w_2^2 Var(y_2),$$ (2.7)

where w_1 and w_2 are the weights representing the temporal distance of y to the middle of the time window of the neighboring parameter sets. When aggregating the variance to annual sums we included the covariance between half hourly values Rüger (1996), y_1,Ě, y_n of the model output:

$$Var\left(\sum y_i\right) = \sum Var(y_i) + \sum Cov(y_i, y_j)$$ (2.8)

Here i and j go from 1 to the number of values being aggregated. The statistically expected differences, err, in annual sums of GPP caused by the random error, assuming a normal distribution of the random error, are computed as:

$$err = Var(GPP_{annual}) * randn,$$ (2.9)

where Var(GPP$_{annual}$) is the variance of the annual sum of GPP and randn is a normally distributed random number with zero mean and unit standard deviation. We draw 100 samples from the distribution for each site.

2.4. Results and discussion

2.4.1. VPD limitation of GPP

Particularly on warm, dry days, the diurnal cycle of NEE is often asymmetric: carbon uptake at comparable insulation is substantially lower in the afternoon compared to morning hours. This behavior could be caused by higher respiration due to higher temperatures or by a limitation of GPP due stomatal closure at high VPDs (Körner, 1995).

The decrease of NEE magnitude with high VPD is evident to varying degrees at each of the eight sites selected for more detailed analysis (see Table 2.1, Fig. 2.1). When the VPD effect is not accounted for in a light-response curve, the consequences are systematic model errors whose magnitude depends on the response of GPP to VPD (Fig. 2.2). Using the DB noVPD approach, the diurnal cycle of the modeled NEE has the symmetric properties of the diurnal cycle of the global radiation, and the model under-predicts the flux magnitude in the morning and over-predicts during the afternoon (see the flux magnitude in Fig. 2.1 and the difference in sign in Fig. 2.3 and Table 2.2). Comparing the biases observed during morning, afternoon and nighttime on annual time scales (Table 2.2) shows that the NEE of the DB noVPD is more negative in the afternoon and more positive at night and during morning hours compared to the observations.

Using the "DB noVPD all" method, the asymmetry of the diurnal cycle can be mimicked by compensating for the absence of a VPD limitation term by increasing the parameter estimate of respiration in the afternoon, caused by an unrealistically high estimate of temperature sensitivity (E_0) up to more than 1000. Such a high value corresponds to a Q10 of 15 between 10 and 20°C. When extrapolated to nighttime periods, the higher E_0 can cause a strong, temperature-related decrease of Reco during the
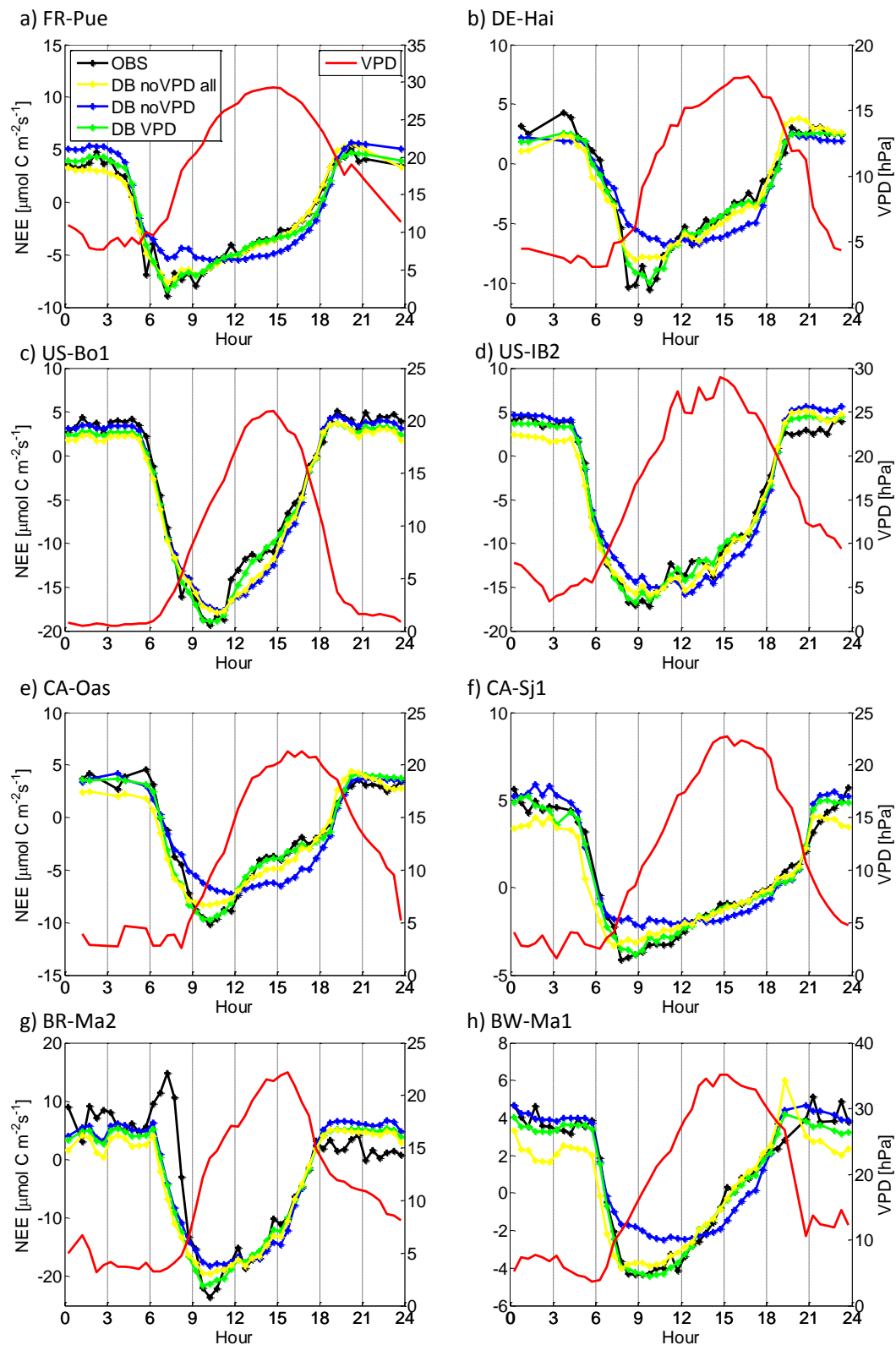
Figure 2.1.: Mean diurnal cycle of NEE observations and the three approaches of the light response curve and VPD for periods with 10 days with high daily maximum VPD ($\geq$ 15 hPa) for sites in different climatic regions and different vegetation types, see Table 2.1 for site details.
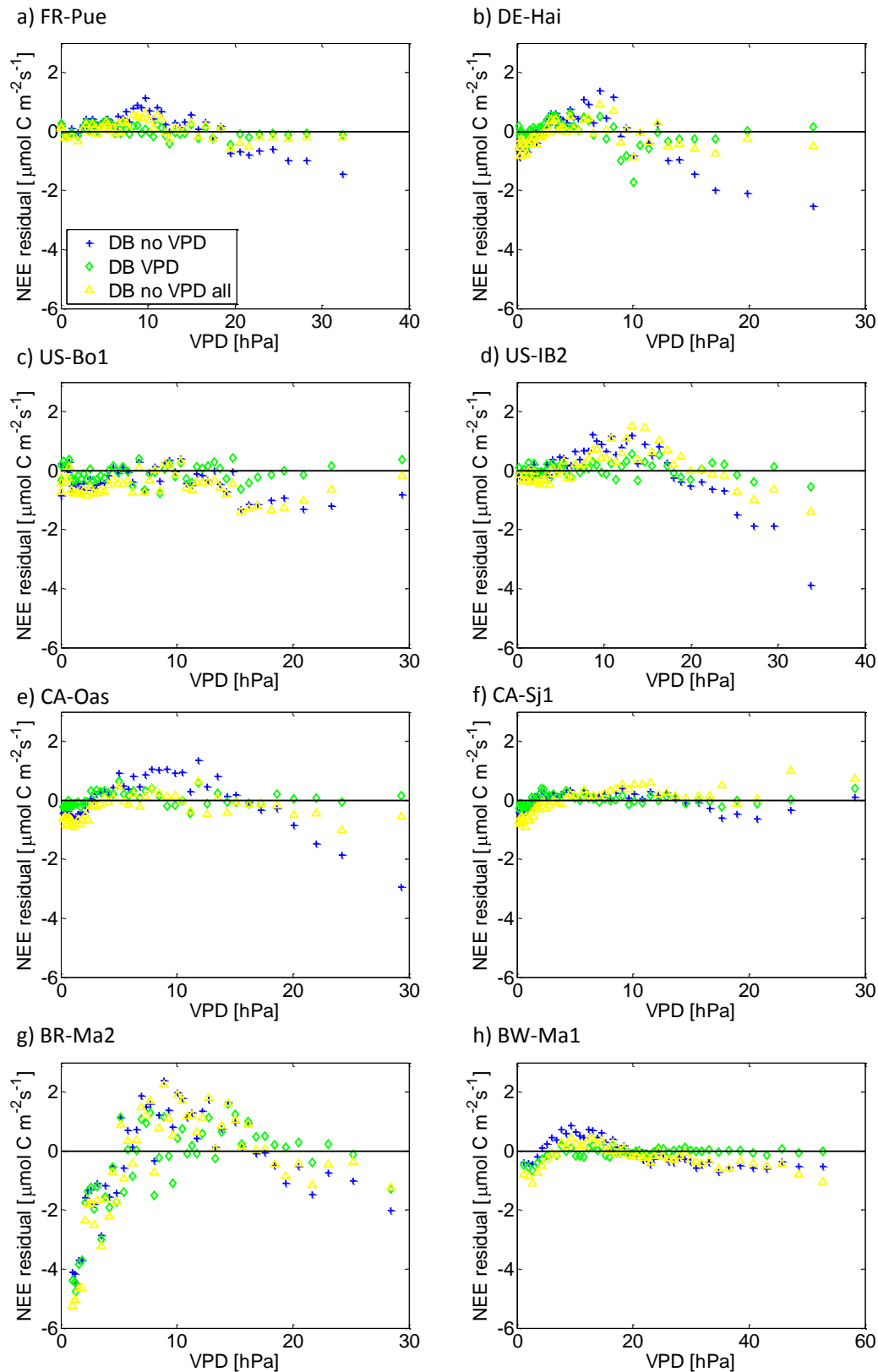
Figure 2.2.: Half hourly NEE residuals assembled into 50 VPD bins, high quality daytime observations of the whole year are used. Positive residuals mean the modeled fluxes are higher (more positive, signifying less ecosystem CO2 uptake) than the observations.
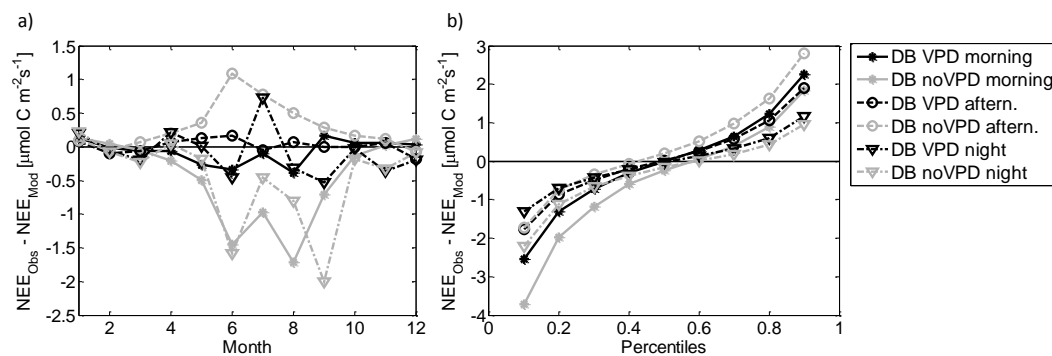
Figure 2.3.: Seasonal cycle of the morning, afternoon and night biases of the DB VPD and DB noVPD estimate (a), percentiles of the morning, afternoon and night biases of the DB VPD and DB noVPD half hourly estimates (b). Data: US-Ib2.
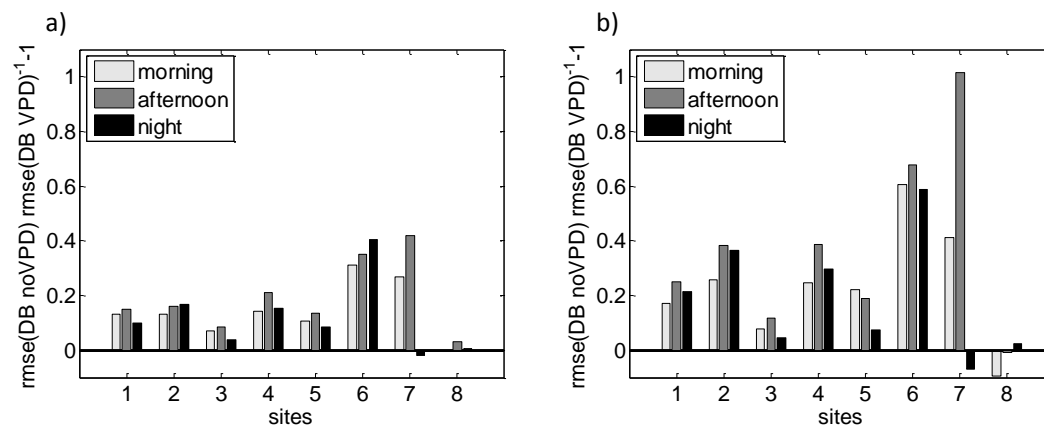
Figure 2.4.: Relative changes in RMSE when including VPD on halfhourly (a) and daily (b) timescale, for the sites 1, FR-Pue; 2, US-IB2; 3,USBo1; 4, DE-Hai; 5, CA-Sj1; 6, CA-Oas; 7, BW-Ma1; 8, BR-Ma2; positive values indicate a lower RMSE for the DB VPD model.
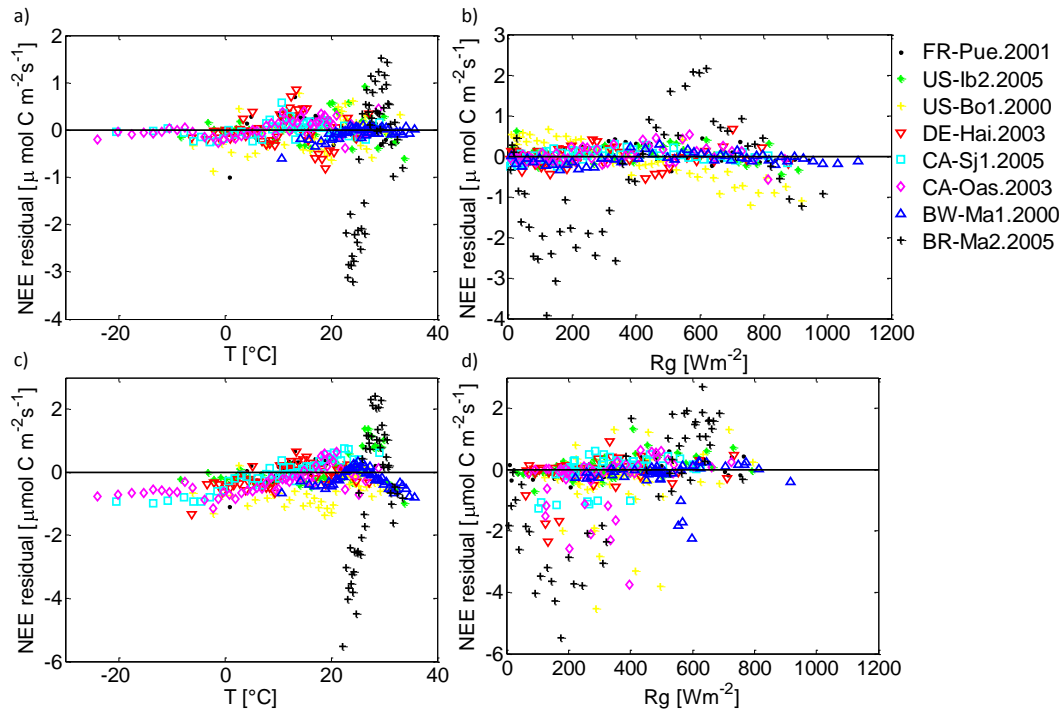
Figure 2.5.: Residuals of the DB VPD derived NEE vs. the model drivers, air temperature (a) and global radiation (b). Residuals of the DB noVPD all vs. air temperature (c) and global radiation (d).

night (BW-Ma1, CA-Oas, US-Ib2, Fig. 2.1), which is inconsistent with the observed data (Fig. 2.1, where nighttime NEE equals Reco). While the pattern in the residuals with respect to VPD is reduced compared to the DB noVPD method (Fig. 2.2), the residuals using the "DB noVPD all" method are biased with respect to Rg and T_{air} (Fig. 2.5).

Different estimates of the temperature sensitivity (E_0) also result in different diurnal amplitudes of Reco; this may explain the earlier reported large disagreement of diurnal Reco courses in the intercomparison of statistical flux-partitioning algorithms (Desai et al., 2008). Our results strongly caution against confounding VPD effects on GPP with temperature effects on Reco; these are fundamentally different mechanisms and must be treated separately. Given the high temperature-dependence of VPD, such confounding effects may be not always easily resolved from the day-time data, we here prefer to derive E_0 (the temperature response of Reco) from night-time data and the magnitude (rb) from day-time data (cf. Appendix A). Contrary to Reichstein et al. (2005) we hence do not rely on the problematic night-time data for estimation of the Reco magnitude. However, our approach did not overcome all issues (see "Limitations of the algorithm" below).

Including a VPD limitation of GPP in the model (DB VPD) generally improves the ability of the model to reproduce the peak before noon and the decrease in the afternoon across the selected sites (Fig. 2.1). The site BR-Ma2 is an exception here, see "Limitations" section below for a discussion.

The annual RMSE is reduced when including VPD in the model, on both half hourly and daily time scales (Fig. 2.4). The model including VPD eliminates the clear systematic bias for the different periods of the day (Table 2.2). The median of the error distribution is closer to zero and the range of the bias is reduced (Fig. 2.3b). Small biases of the model compared to the NEE observations used for fitting can

Table 2.2.: Mean annual bias between modeled and observed NEE [μmolm-2s-1]. Only measured high-quality data were used in the comparison.

Site	morning		afternoon		night	
	HBLR	HBLR VPD	HBLR	HBLR VPD	HBLR	HBLR VPD
FR-Pue	-0.66	-0.31	0.16	0.07	-0.28	0.00
US-IB2	-0.53	-0.09	0.33	0.01	-0.53	-0.12
US-Bo1	-0.03	0.22	0.05	-0.04	-0.39	-0.26
DE-Hai	-0.36	0.03	0.17	-0.05	-0.38	-0.26
CA-Sj1	-0.25	-0.06	0.04	0.04	-0.20	0.05
CA-Oas	-0.54	0.05	0.00	-0.03	-0.27	0.00
BW-Ma1	-0.29	0.16	0.23	0.00	0.11	0.46
BR-Ma2	2.01	2.83	-0.79	-0.90	-0.76	-0.57

be caused by the weighting in the cost function, and the interpolation of fluxes between the different parameter sets.

For DB VPD, the residuals are not correlated with VPD (e.g. maximum R^2=0.02 even for a $3^r d$ degree polynomial for BR-Ma2) and there is no consistent pattern across sites, indicating that systematic biases associated with the revised model tend to be minimal (Fig. 2.2). The bias of the residuals with respect to VPD was reduced by estimating the temperature sensitivity with daytime data: in some sites (FR-Pue and DE-Hai) this bias was removed entirely. At first sight both methods result in similar NEE estimates, however, residual analysis shows that the DB noVPD all method is biased with respect to $T_a ir$ and global radiation (Fig. 2.5). This indicates that the asymmetry in the diurnal cycle is mainly caused by the VPD limitation of GPP. When modeling this behavior by increased respiration, the estimates are biased, with respect to temperature and the temperature sensitivity is too high.

The residuals of the two drivers of the model, temperature and global radiation, do not show consistent patterns across sites for the DB VPD method (Fig. 2.5). As VPD is partly a function of temperature, including the VPD limitation reduces the pattern in the relation between residuals and temperature (not shown). Due to this strong correlation it is not possible to differentiate statistically between VPD-driven and temperature-driven decreases in GPP (Doughty and Goulden, 2008). There is no systematic bias in the residuals for high temperature (Fig. 2.5), suggesting that adding VPD limitation is a logical step for improving estimation of GPP and Reco from daytime data across globally distributed ecosystems.

Limitations of the algorithm

We chose to use a simple, empirical model for this analysis. These models can be applied across a wide range of sites and vegetation types without the need for side-specific data on vegetation structure or C pools. However, complex interactions among physiological processes cannot necessarily be described by a simple equation. Hence, despite the achievement of a good and almost unbiased description of the diurnal NEE course through the inclusion of VPD effects on GPP there remain a number of limitations of the light-response curve approach, namely:

1. It has been reported that canopy assimilation is not only affected by the overall short wave radiation flux density, but also by its "source" i.e. whether dominated by diffuse or direct radiation. With diffuse radiation higher assimilation rates have been observed at the same overall radia-

tion flux density (Baldocchi, 1997; Gu et al., 2003; Hollinger et al., 1994; Jenkins et al., 2007; Knohl and Baldocchi, 2008; Niyogi et al., 2004). This effect is not reflected in our light-response curve. However, two issues remain uncertain: First, the magnitude of the direct effect and the effect of the background correlation of high diffuse radiation with low VPD values (Rodriguez and Sadras, 2007; Wohlfahrt et al., 2008a) , second, and practical limitation that relatively few FLUXNET sites measure diffuse radiation. Currently a specific analysis on this topic is being carried out as part of the FLUXNET synthesis activities (Cescatti et al. in prep.). 2. Circadian rhythms of stomatal conductance are not considered in our approach. They are either endogenous or caused by hydraulic limitations in the afternoon. These patterns in the diurnal cycle can persist for more than a week independent of environmental influences (Hennessey et al., 1991). Although this effect has been widely observed (Gorton et al., 1993; Hennessey et al., 1993; Nardini et al., 2005), the degree to which they affect the carbon exchange under field conditions is less clear. Williams and Gorton (1998) suggested by using a modeling approach that theses circadian rhythms do not significantly affect photosynthesis and stomatal conductance in field conditions.

2. The respiration model is only driven by temperature, but the overall signal of ecosystem respiration originates from different parts of the ecosystem which experience different temperatures. It is not clear which temperature is the appropriate driver for ecosystem respiration; studies suggest that this can vary between sites (Richardson et al., 2006a). We used air temperature as it often explains more variance of the ecosystem respiration (Reichstein et al. (2005) but see Richardson et al. (2006a) and using air temperature more consistent temperature-respiration relationships have been found in some ecosystems (Van Dijk and Dolman, 2004). A large part of soil respiration can be assumed to be derived near the surface across ecosystems, which is better characterized by air temperature than soil temperature at deeper soil layers. Diurnal hysteresis effects are found for respiration when plotted against soil temperature (Bahn et al., 2008; Vargas and Allen, 2008), this hysteresis increases with increasing soil depth (Bahn et al., 2008). Moisture limitation has an significant effect on soil respiration (Irvine and Law, 2002). This limitation is not explicitly included in the model and few FLUXNET sites measure soil moisture, limiting its potential for widespread application at the present. However parameter estimation may account for it by varying rb. Diel patterns in respiration that are not driven by temperature but by soil moisture (Carbone et al., 2008), are not reflected in the model.

3. As the light response curves are fit to daytime NEE, errors in GPP can always be compensated by errors in Reco, resulting in incorrect estimates for both GPP and Reco without compromising NEE model fit. Desai et al. (2008) showed this to occur for synthetic data. This problem occurs in particular if VPD is not included in the model, as the afternoon decrease in NEE is then ascribed to a higher respiration instead of a limited GPP and consequently leads to biased estimates. We reduced this confounding effect by extending the light response approach with a VPD limitation and estimating the temperature sensitivity using nighttime data independent of the NEE response to VPD.

4. The algorithm, as well as other flux partition algorithm strongly depends on the quality of the NEE measurements and an accurate quality assessment (Foken and Wichura, 1996). The positive
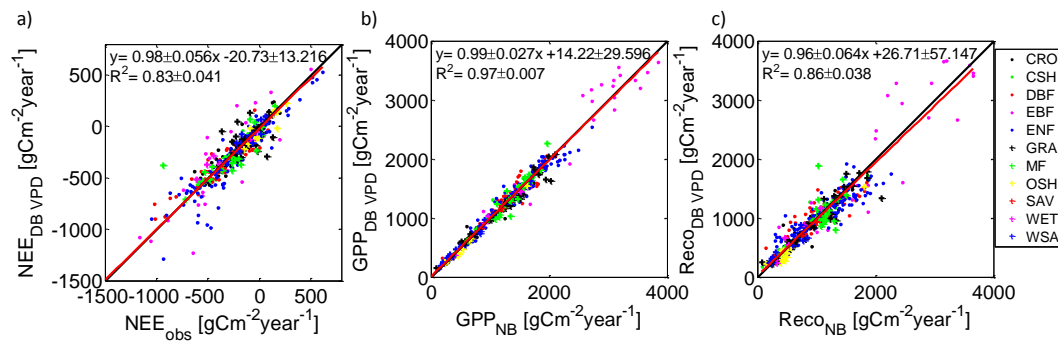
Figure 2.6.: Scatter plots of (a) annual sums of gapfilled observations (x-axis) and DB VPD (y-axis) estimates of NEE, annual sums of nighttime data based (x-axis) and DB VPD (y-axis) estimates of (b) GPP and (c) Reco. CRO, cropland; CSH, closed shrubland; DBF, deciduous broadleaf forest; EBF, evergreen broadleaf forest; ENF, evergreen needleleaf forest; GRA, grassland; MF, mixed forest; OSH, open shrubland; SAV, savanna; WET, wetland; WSA, woody savanna.

peak in measured NEE during the morning at the Brazilian site (Fig. 2.1) and the strong bias in the residuals for low VPD (Fig. 2.2) likely occurs as a result of an incomplete storage correction as documented earlier for this site (Araujo et al., 2008, 2002). Such problems arising from the complexity of site need to be addressed before such simple algorithms can be applied successfully.

2.4.2. Comparison of nighttime and daytime based estimates

We compared annual sums of GPP and Reco of the updated DB VPD and conventional NB partitioning approach for all FLUXNET site-years with sufficient available data (417 site-years, 145 sites, see Appendix B). For NEE we compared the DB VPD estimate with the gap-filled annual sum of observations. The two estimates were strongly correlated (R^2(NEE)=0.83, R^2(GPP)=0.97, R^2(Reco)=0.86), but deviations exceeded 52 g C m^{-2} year^{-1} for NEE, 47 g C m^{-2} year^{-1} for GPP 87 g C m^{-2} year^{-1} for Reco in over 50% of site-years (see Fig. 2.6 and 2.7). These numbers are in a comparable range of the uncertainties reported for the u* threshold, that remain below 100 g C m^{-2} year^{-1} for NEE (Papale et al., 2006). Comparing the gapfilled observations with the DB VPD method does not show systematic differences for the annual NEE estimates throughout the FLUXNET database (Fig. 2.6 and 2.7). For GPP and Reco, the confidence intervals of the regression parameters include a slope of one and an offset of zero, thus there is no systematic bias. For NEE the DB VPD estimate is slightly more negative compared to the observations, indicating greater biosphere C uptake by the model than is apparent in the data. The 95% confidence interval of the NEE offset does not include zero, but the slope is not significantly different from one.

The NB approach produces slightly higher Reco estimates than the DB VPD approach, but the differences are not significant on the annual timescale. Despite being insignificant, differences were thought to be caused by a difference in the diurnal versus seasonal temperature sensitivities of Reco (Gaumont-Guay et al., 2006); the NB approach overestimates daytime Reco because it effectively characterizes the seasonal temperature sensitivity.

The higher NB estimates of Reco are contrary to expectations that are based on the assumption that
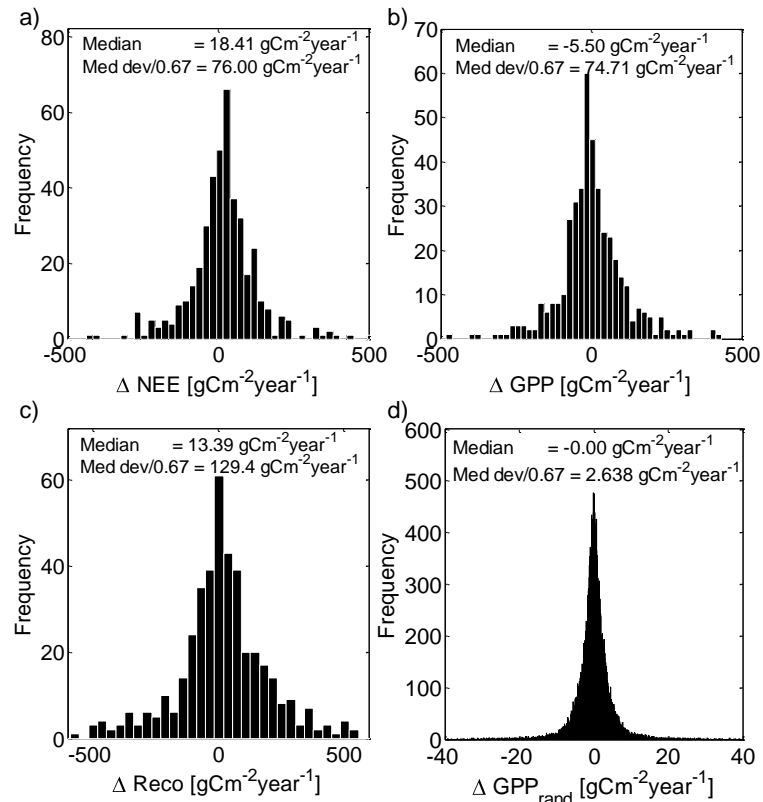
Figure 2.7.: Histograms of the difference between gapfilled annual observations and the daytime data based estimate of NEE (a), annual nighttime data based estimate and daytime data based estimate for (b) GPP, (c) Reco. The Median deviation/0.67 is a robust estimate for the SD, a positive difference denotes a more positive nighttime based estimate. (d) Histogram of the expected difference between GPP estimates based on the statistical uncertainty of GPP caused by the half-hourly random errors.

night-time fluxes would be underestimated due to, e.g. insufficient turbulent exchange or missing low frequency contributions. Possibly, on average, the use of a nighttime u* filter effectively addresses this problem. Nighttime-based approaches could also overestimate daytime Reco because of a reduction of leaf respiration in the light (Atkin et al., 1998; Brooks and Farquhar, 1985), thus the relationship derived from nighttime data could overestimate respiration during daytime and vice versa for the daytime data based estimate, although this day-time reduction of respiration is highly controversial among plant physiologists (Loreto et al., 2001; Pinelli and Loreto, 2003).

The NB and DB VPD estimates of GPP are more strongly correlated than those of Reco, because both approaches estimate GPP from daytime NEE, while NB Reco is estimated independently of the daytime NEE data. While the correlations of Reco and GPP are comparable, the lower correlations for NEE are caused by the smaller range of the data (-1000 to 500 compared to 0-4000 gC m^{-2} year^{-1}), but with the same amount of scatter. This is also reflected in the histograms of the annual differences between nighttime and daytime based estimate in Fig. 2.7.

The median deviation of NEE and GPP are within the same range, while the spread of the differences in Reco is much wider. We chose the median and median deviation to characterize the histogram, as the distribution is not Gaussian but more leptokurtic and the standard deviation does not characterize such distributions appropriately. The median of all three histograms is close to zero, supporting the conclusion that there is no overall systematic difference between day-time and night-time based annual carbon flux estimates.

The deviations between NB and DB VPD represent the uncertainty in the annual estimates caused by inconsistent nighttime data and the choice of the partitioning method. Inconsistencies between day and nighttime data can be caused by low turbulence, advection, insufficient u*- filtering, decoupling of the flow or a difference in the footprint; at night the footprint is smaller than during the day. Comparing the deviations arising from such systematic errors, the deviations arising from statistical uncertainty, is small, in most cases below 20 gCm^{-2} year^{-1}for GPP(Fig. 2.7). The statistical uncertainty is mainly caused by the random error of the data and is relatively small on an annual basis. The uncertainties due to inconsistencies in the data and the partitioning method is one order of magnitude larger (see also Richardson et al., 2006a). The deviations between the methods vary across vegetation types. Table 2.3 characterizes the distribution of differences between daytime data based and nighttime data based estimates of GPP and Reco, and between daytime data based and gapfilled NEE for the different vegetation types. The strongest deviation of the median from zero is found for vegetation types with a small number of sites available, suggesting that increasing the number of sites may remove the apparent bias. The median deviation appears to be higher for tall vegetations (forests). The NEE observations are higher (positive median) for all vegetation types, except wetlands compared to the daytime data based estimates, and the nighttime data based Reco estimate is higher for most vegetation types. For GPP no clear pattern emerged. The strong differences in the median deviation between vegetation types suggest a strongly varying uncertainty between sites. This result supports the necessity of a site and year specific uncertainty estimate, incorporating all sources of uncertainty, to enable scientists to use the data properly to fully exploit the information inherent to the database.

30

Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: critical issues and global evaluation

Table 2.3.: The median, median deviation/0.67 (med dev, i.e. an estimate of the standard deviation) and kurtosis (kurt) of the annual differences between NB and DB VPD estimate (GPP and Reco) and between daytime data based estimate and gapfilled observations (NEE) for different vegetation types.

	NEE			GPP			Reco			
	median	med dev	kurt	median	med dev	kurt	median	med dev	kurt	N
CRO	14.62	79.31	2.66	-9.90	52.00	6.46	11.15	81.22	6.29	37
CSH	33.07	93.19	1.98	-50.61	69.76	3.58	-51.52	153.10	3.21	8
DBF	15.14	76.11	3.68	-15.95	101.66	3.46	-0.62	144.38	3.40	79
EBF	14.36	163.05	3.61	9.21	269.59	2.44	19.25	420.62	2.72	30
ENF	21.43	83.78	6.32	-9.63	77.33	4.07	18.64	147.04	3.74	148
GRA	17.40	59.76	5.72	7.82	55.53	7.54	26.41	82.33	11.78	49
MF	26.32	84.87	10.93	3.53	67.05	5.36	16.83	142.69	8.13	29
OSH	63.46	41.73	3.92	20.66	60.28	1.42	84.12	103.15	2.31	14
SAV	33.05	6.36	1.00	27.28	13.17	1.00	60.33	19.52	1.00	2
WET	-58.49	120.75	1.45	-2.34	20.27	2.76	-60.66	97.39	1.92	6
WSA	4.04	37.96	2.75	-30.73	40.23	2.52	-25.11	85.97	2.61	15

2.4.3. Global relationship between carbon fluxes in the FLUXNET database

For the first time, we can now compare quasi-independent estimates of GPP and Reco across a large data set, since we can use GPP derived from day-time data and Reco derived from night-time data only. Previous studies including Reichstein et al. (2007); Wang et al. (2008); Baldocchi (2008) relied on GPP and Reco estimates which were ultimately derived from the same data. To some extent this may cause spurious correlation between Reco and GPP, since GPP is inferred as Reco minus NEE (Vickers et al. (2009); but see comment by Lasslop et al. (2010a)). Here, we do not compute GPP as a difference, but moreover derive Reco and GPP from quasi-disjoint NEE data subsets. Hence, we minimize spurious correlations and still find a strong and highly significant positive relation between annual GPP and Reco (Figure 8). These results give further evidence to Janssens et al. (2001); Reichstein et al. (2003a, 2007); Baldocchi (2008) - now across and separated into different biomes - that ecosystem assimilation and respiration are strongly coupled on the annual time scale. The overall relationship shown in Fig. 2.8 is dominated by spatial (=between-site) variability - e.g. the spatial coefficient of variation of mean site GPP is 53% while the temporal coefficient of inter-annual GPP variation reaches only 2 to 57%, with a median of 9%. This overall between-site correlation of GPP and Reco can be relatively easily explained by typical ecosystem model concepts that involve carbon pools that are built up by photosynthesis and allocation and subsequently decomposed by autotrophic and heterotrophic respiration (Krinner et al., 2005; Sitch et al., 2003). These concepts predict that after infinite time of constant conditions without disturbance the system will be in steady state and without lateral export of carbon, Reco will consequently approximately equal GPP (i.e. be on the 1:1-line). Several factors may cause deviations from this theoretical state: 1) climate and environmental conditions are never constant, but vary at all time scales causing disequilibrium (e.g. CO_2 and N fertilization effect are thought to increase carbon sinks), 2) in many ecosystems anthropogenic export via wood or crop harvest plays an important role and leads to reduced on-site respiratory fluxes (Ciais et al., 2007; Imhoff et al., 2004), 3) disturbance events (clear cuts, wind throws, fires) temporally reduce productivity while soil carbon is continued to be respired
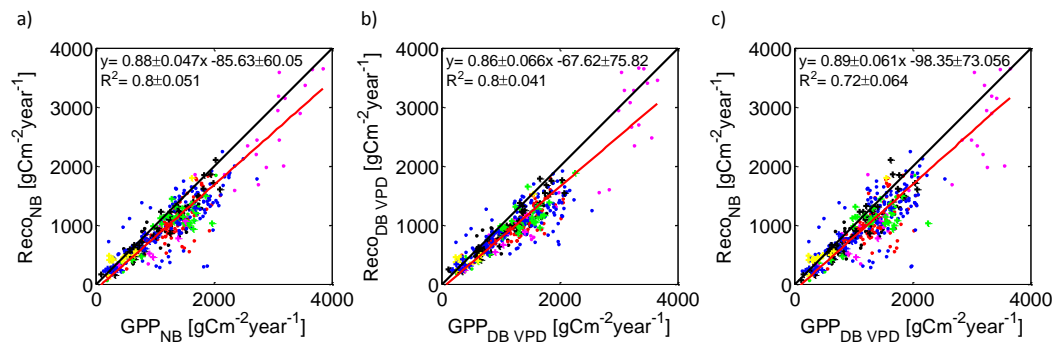
Figure 2.8.: Scatter plot of the annual sums (a) of the nighttime data based estimate of Reco and GPP, (b) of the daytime data based estimate of Reco and GPP, (c) of the daytime data based GPP and the nighttime data based Reco, data: FLUXNET database, legend see Fig. 2.6.

(legacy effect) (Barford et al., 2001; Saleska et al., 2003). Hence the overall tendency of Reco>GPP (slope: 0.86-0.89) should be caused by factor categories 1) and 2), while the site years above the 1:1 line are likely to have been affected by recent disturbance, although also strong interannual variability maybe causes ecosystems to be sources during particular years when GPP is more strongly reduced than Reco (or Reco more strongly enhanced).

We still cannot fully exclude spurious correlation between our GPP and Reco estimates, for instance if the errors in day and night-time data are strongly correlated; this depends also on the temporal resolution used to compute the correlation. However, we can analyze the effect of the spurious correlation expected when using the same data and algorithm (e.g. only the nighttime data based estimate) for GPP and Reco estimation (see above) which should be larger than spurious correlation derived from using different data sets (e.g. Reco nighttime data based, extrapolating the daytime and GPP daytime data based), where the error of Reco does not propagate into the GPP estimate. The relationships of different combinations of GPP and Reco estimates (only nighttime based, daytime based or Reco nighttime and GPP daytime based) are statistically indistinguishable (confidence bounds of the correlation coefficient and the regression parameters overlap). This shows that we can have an increased confidence in the derived global pattern and that the expected effect of the spurious correlation (Vickers et al., 2009; Wang et al., 2008) due to the dependency of Reco and GPP when GPP is computed as the residual and the estimates are based on the same data is rather small on the annual timescale. Hence, from a methodological point of view the robustness of the relationship shows that despite uncertainties and statistical pitfalls inherent to the data global patterns of ecosystem-atmosphere CO_2 exchange can be derived from the eddy covariance method when deployed as a network with standardized processing schemes.

2.5. Concluding discussion

In this study we introduced an algorithm that splits NEE into its main components GPP and Reco using daytime data. Including VPD limitation of CO_2 uptake improved the model's ability to reproduce peak flux before noon and the afternoon decrease in NEE magnitude. Including the VPD limitation removed a systematic pattern in the residuals of the model and improved the models performance. One important finding is that if VPD effects are not explicitly accounted for, they can be easily confounded with temperature effects on ecosystem respiration, resulting in a biased partitioning of the NEE flux

into Reco and GPP including unrealistic diurnal cycles of these quantities. Also our approach is not free of errors, as for instance the u* filtering threshold is uncertain. We provide a Reco/GPP dataset additional to the one generated according to Reichstein et al. (2005) where Reco is based on nighttime respiration. We recommend the combined use of the two datasets to cross check flux estimates and point to sites and periods where carbon flux estimates remain uncertain. The application of further plausible algorithms would be desirable to obtain a better estimate of the possible range of flux estimates derived from eddy covariance flux data. Yet, the comparison should be combined with additional quality and consistency checks based on the comparison with biometric measurements as an additional independent constraint (as for instance Luyssaert et al., 2009; Stoy et al., 2006). The comparison of the two estimates shows a strong correlation and no significant biases for GPP and Reco. Although the overall agreement is good, there can be large deviations for specific sites or years. Comparing these deviations with the deviation that could be caused by the formal statistical uncertainty of GPP arising from the random error of half-hourly values, shows that the uncertainty arising from systematic errors, such as advection, low turbulence, decoupling of the flow, differences in the footprint during the night compared to daytime or the choice of model and extrapolation, clearly dominates the overall uncertainty of the estimates. Hence, these uncertainties should be considered in any statistical analysis, process model evaluation and model data fusion based on the FLUXNET database. Although the annual sums of many sites must be expected to be biased or at least uncertain, the patterns derived from this global dataset, as for instance the correlation between Reco and GPP, are reliable, increasing our confidence in analyses across sites based on the dataset. In spite of this we emphasize that more specific uncertainty estimates for individual sites and years are needed to strengthen the significance of more detailed statistical analysis and to fully exploit the information inherent in the FLUXNET database.

3. Comment on Vickers et al.: Self-correlation between assimilation and respiration resulting from flux partitioning of eddy-covariance CO_2 fluxes

The paper of Vickers et al. (2009) raises the important issue of self-correlation, also named spurious correlation or artificial correlation, in the context of partitioning measured eddy covariance (EC) CO_2 fluxes. Vickers et al. argue that there is a spurious correlation between primary production (GPP) and ecosystem respiration (Reco) when these component fluxes are jointly estimated from the measured net ecosystem exchange (NEE) of CO_2. Spurious correlations can arise, for instance, when an independent variable x is compared with a dependent variable z that is computed from x as x-y (Kenney, 1982). In this example x and y are the "original" variables and z is derived from these. The higher the variance of the shared variable x compared to the variance of y, the higher the spurious correlation between x and z (Kenney, 1982).

Standard flux partitioning methods derive GPP and Reco (here, by definition, both positive quantities, whereas negative NEE indicates net uptake by terrestrial vegetation) by fitting a respiration model to nighttime EC data (where GPP=0, as there is no photosynthesis in the dark), extrapolating the model into daytime, and deriving the estimated GPP as Reco-NEE (Desai et al., 2008; Reichstein et al., 2005). Here, Reco corresponds to x, NEE to y, and GPP to z = x-y. This description can be further split up as described in Vickers et al. (2009) method II: when using daily sums, or data further aggregated, the data can be split into their nighttime and daytime part then GPP=NEEday- Recoday. The correlation between GPP and Reco could be expected to be spurious, as the equation to derive GPP contains Reco or at least the daytime part. Even when splitting the variables into day and nighttime, the estimate of the spurious correlation is not small (Vickers et al., 2009).

A key difference between the flux partitioning and the Kenney (1982) example is that NEE is not an "original" independent random variable but rather consists of two distinct components: Reco and GPP. With the operation Reco-NEE the influence of Reco is removed from NEE thus Reco is not part of GPP. But note that any error in the estimation of Reco would be directly transferred to GPP. Spurious correlation is often illustrated with the example of plotting the weight of the liver against total body weight. Variability of the liver directly causes a variability of the total body weight and thus a correlation (Kainz et al., 2009). In case of the flux partitioning Reco is the analogue of the liver, but NEE is the total body weight and GPP is total body weight minus liver. The correlation between liver and total body weight minus liver would not be expected to be spurious (unless there were measurement errors), as the liver is not part of the second variable and does not contribute to its variability.

Here we are not arguing against the general importance and the necessity to be aware of the problem of self-correlation when estimating GPP and Reco from the same EC data, but self-correlation is important because of the error in Reco rather than because Reco being a shared variable. We show that the spurious correlation estimated using the whole variance of Reco with the approach adopted by Vickers et al. (2009) is high, even when both quantities are in fact correlated and we attribute any spurious correlation
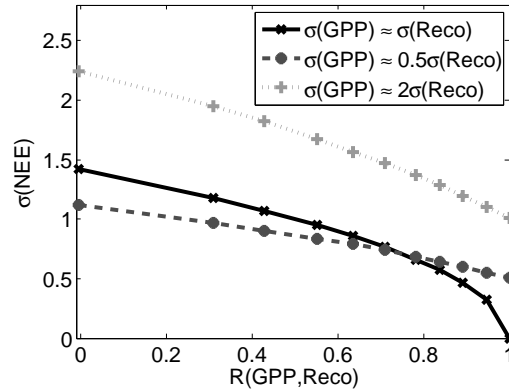
Figure 3.1.: Standard deviation of NEE versus the correlation (R) between GPP and Reco.

to the variance of the error in Reco.

Across FLUXNET sites the variance of Reco on different timescales is not negligible compared to the variance of NEE (Falge et al., 2002b; Stoy et al., 2009). After eq. 2 in (Vickers et al., 2009) the spurious correlation between GPP and Reco is expected to be high. We now examine the implications of the fact that NEE, the net ecosystem exchange, is by definition the difference between GPP and Reco. This identity implies that the variance (var) of NEE decreases with increasing covariance (cov) between GPP and Reco, since

$$var(NEE) = var(GPP) + var(Reco) - 2cov(GPP, Reco). \tag{3.1}$$

This contradicts the concept of spurious correlation following eq. 2 in Vickers et al. (2009), which would predict a high spurious correlation for a low variance of NEE. Note that this equation is based on the positive definition of GPP and Reco. Following the micrometeorological sign convention, where a flux towards the surface is negative, a negative covariance would be added resulting in the same decreased variance of NEE. We illustrated this relationship using artificial data where GPP and Reco are normally distributed random variables with varying variances of GPP and varying correlation between the two, NEE is computed as Reco-GPP (Figure 1). For this constructed dataset our understanding of the correlation between GPP and Reco is, that it is completely real and the "artificial correlation" between GPP and Reco should be zero, although GPP can be computed as Reco-NEE.

Two approaches have been proposed to estimate the self-correlation: (1) based on the variances of the two original variables (x and y) (Kenney, 1982) and (2) a method that decorrelates NEE and Reco (Vickers et al., 2009). GPP is then computed with the randomized Reco, assuming that Reco is part of GPP, the correlation between GPP and the randomized Reco is the self-correlation. They both show the same high values for the "artificial correlation" even if GPP and Reco are in fact uncorrelated (Figure 2a). Using the whole variance of Reco to estimate the spurious correlation can lead to a strong overestimation of spurious correlation and a misleading estimate of the real correlation, if the real correlation is estimated as observed minus spurious correlation as in Vickers et al. (2009) (Figure 2b). So far we assumed perfect observations of NEE and perfect estimates of Reco, which cannot be expected (Desai et al., 2008; Lasslop et al., 2010b; Richardson et al., 2006b). Including the errors the equations can be
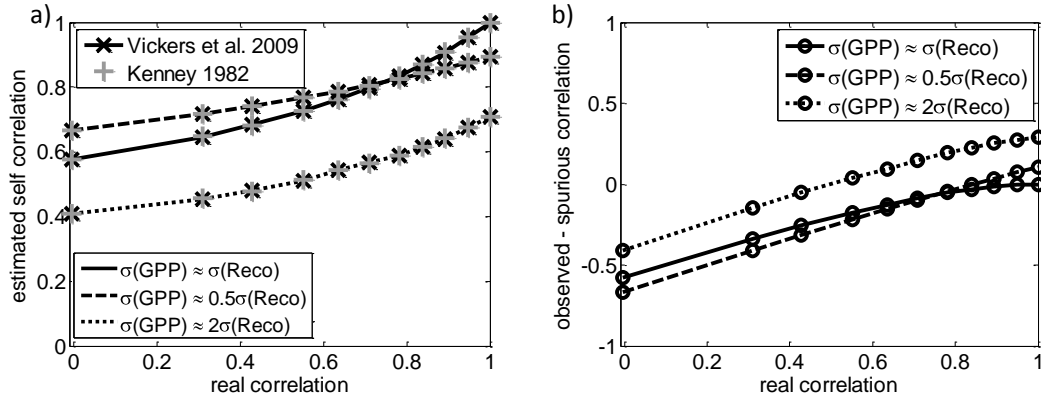
Figure 3.2.: (a) Estimated artificial correlation versus the correlation between GPP and Reco, the different lines indicate the standard deviation of GPP as in (b), markers indicate the approach used to estimate the artificial correlation. (b) Difference of observed and spurious correlation (the estimate for the real correlation of Vickers et al., 2009) versus the real correlation. The observed correlation equals the real correlation of the generated dataset here.

changed into

$$NEE_{obs} = GPP_{true} + Reco_{true} + \epsilon_{NEE}, \tag{3.2}$$

$$Reco_{est} = Reco_{true} + \epsilon_{Reco}, \tag{3.3}$$

$$GPP_{est} = Reco_{est} - NEE_{obs} = GPP_{true} + \epsilon_{Reco} + \epsilon_{NEE}, \tag{3.4}$$

where ϵ denote error terms, the subscript "true" denotes the true value of the system, "obs" and "est" the observed and estimated values, respectively. Thus spurious correlations can be introduced via error terms that are part of both, Reco and GPP, but only if they contribute to the variance, e.g. they are not constant. This explains the high estimates of self-correlation when using the randomized time series of Reco to compute GPP, as the relation between NEE and Reco is destroyed, the Reco error is maximized by the randomization and $Reco_{true}$=0.

To derive relationships between Reco and GPP without having to face the problem of spurious correlations or to explore the extent of a spurious correlation we recommend to combine independent estimates for such analysis. For instance, respiration can be derived from chamber measurements and combined with GPP from eddy covariance measurements. If only eddy covariance measurements are available Reco can be derived using only nighttime data and GPP using only daytime data. The influence of the error in Reco on the correlation between GPP and Reco as well as additional evidence of real correlation between GPP and Reco is further explored in a forthcoming manuscript (Lasslop et al., 2010b).

4. Influences of observation errors in eddy flux data on inverse model parameter estimation

4.1. Abstract

Eddy covariance data are increasingly used to estimate parameters of ecosystem models and for proper maximum likelihood parameter estimates the error structure in the observed data has to be fully characterized. In this study we propose a method to characterize the random error of the eddy covariance flux data, and analyse error distribution, standard deviation, cross- and autocorrelation of CO_2 and H_2O flux errors at four different European eddy covariance flux sites. Moreover, we examine how the treatment of those errors and additional systematic errors influence statistical estimates of parameters and their associated uncertainties with three models of increasing complexity – a hyperbolic light response curve, a light response curve coupled to water fluxes and the SVAT scheme BETHY. In agreement with previous studies we find that the error standard deviation scales with the flux magnitude. The previously found strongly leptokurtic error distribution is revealed to be largely due to a superposition of almost Gaussian distributions with standard deviations varying by flux magnitude. The crosscorrelations of CO_2 and H_2O fluxes were in all cases negligible (R^2 below 0.2), while the autocorrelation is usually below 0.6 at a lag of 0.5 hours and decays rapidly at larger time lags. This implies that in these cases the weighted least squares criterion yields maximum likelihood estimates. To study the influence of the observation errors on model parameter estimates we used synthetic datasets, based on observations of two different sites. We first fitted the respective models to observations and then added the random error estimates described above and the systematic error, respectively, to the model output. This strategy enables us to compare the estimated parameters with true parameters. We illustrate that the correct implementation of the random error standard deviation scaling with flux magnitude significantly reduces the parameter uncertainty and often yields parameter retrievals that are closer to the true value, than by using ordinary least squares. The systematic error leads to systematically biased parameter estimates, but its impact varies by parameter. The parameter uncertainty slightly increases, but the true parameter is not within the uncertainty range of the estimate. This means that the uncertainty is underestimated with current approaches that neglect selective systematic errors in flux data. Hence, we conclude that potential systematic errors in flux data need to be addressed more thoroughly in data assimilation approaches since otherwise uncertainties will be vastly underestimated.

4.2. Introduction

The availability of carbon dioxide and water vapour flux measurements between ecosystems and the atmosphere around the world offers various opportunities to improve our knowledge about processes connected with the global carbon cycle (Friend et al., 2007; Baldocchi et al., 2001). The interplay of models and data gives us insights into the performance of models, our level of understanding the system, but also into the quality of data and the information content therein about the processes repre-

sented in the model. Classically, parameters were often derived from experiments at leaf or plant scale or from expert judgement. If nonlinear relationships are involved the parameters are scale-dependent and cannot be easily transferred to but also not observed on larger scales. An alternative option to obtain parameter estimates is the inversion of models against data. In this case a cost function describing the misfit between model output and observations is minimized by varying the parameters. The inversion of models against Eddy-Covariance (EC) data leads to parameter estimates at ecosystem scale, our scale of interest, thus EC data are increasingly used for model inversions. EC data contain information about the actual ecosystem flux, but a measured quantity is always the sum of the "true" value and errors, these errors need to be addressed in an adequate way. The measurement errors can be distinguished into random errors, fully systematic errors and selective systematic errors (Moncrieff et al., 1996). Fully systematic errors appear constantly and arise for instance from inaccurate calibration or consistently missing high or low frequency components of the cospectrum, while selective systematic errors appear only during special temporal periods, for instance at night under unfavorable micrometeorological conditions. The random error of EC data arises from the measurement instruments, the stochastic nature of turbulence and varying footprint (the area that influences the measurement, it depends primarily on atmospheric stability and surface roughness). Quantification of the random error is a prerequisite for statistical comparisons between models and data and model-data synthesis as it expresses our confidence in the data. The characteristics of the errors play an important role for the parameter estimation, the error distribution, error cross- and autocorrelations or inhomogeneous variance can bias the parameter retrieval if not accounted for (Tarantola, 1987). The study of Trudinger et al. (2007) showed that how data errors and uncertainties are treated in the optimization criterion will have a significant impact on the retrieved parameters. Studies using EC data in inverse modelling often assume constant error variance (Reichstein et al., 2003b; Owen et al., 2007; Wang et al., 2007), use the standard deviation of the model residuals (Sacks et al., 2006; Braswell et al., 2005) or an adhoc fraction of the observations (Knorr and Kattge, 2005). During the last few years approaches for the quantification of random errors of EC data came up, they used paired observations, first spatially separated measurements (Hollinger et al., 2004), but as there are only few appropriately distanced towers available, Hollinger and Richardson (2005) developed a methodology using daily differenced measurements with equivalent environmental conditions that allowed to characterize the univariate distribution for several sites (Hollinger and Richardson, 2005; Richardson et al., 2006b). However, the auto- and crosscorrelation of the errors have so far not been systematically quantified and are assumed to be zero. Moreover, the systematic errors are still under investigation and challenging the scientific community (Wilson et al., 2002; Friend et al., 2007). Hence the aim of this study is

- to fully analyze the random error of EC water and carbon fluxes regarding the properties important for parameter estimation, i.e. beside the univariate distribution, also autocorrelation and multivariate correlations of CO_2 and H_2O fluxes,

- to elucidate the effect of the error model choice on model parameter estimates and their uncertainties,

- and to explore how selective systematic errors influence parameter estimates of models describing carbon and water exchange.

We carry out the parameter estimation experiments with synthetic data based on eddy covariance data from two European sites and with three models of different complexity, a hyperbolic light response curve, a light response curve coupled to water fluxes and the SVAT scheme BETHY, a process-based model that calculates the CO_2, H_2O and energy exchanges of soil, vegetation and atmosphere for the terrestrial land surface (Knorr and Heimann, 2001a).

4.3. Methods

4.3.1. Analysis strategy

The first part of the study deals with the characterization of the random error. We estimate the random error for four different sites, Hainich, Loobos, Puechabon and Hyytiälä, using the gapfilling algorithm of Reichstein et al. (2005). We focus on the statistical properties important for parameter estimation, e.g. variance, distribution, autocorrelation, crosscorrelations. The kurtosis is a measure of peakedness and can be used as an indicator for the type of distribution, the excess kurtosis used here is zero for Gaussian distributions and three for double exponential distributions. To reveal the influence of errors to parameter estimates we designed 20 synthetic data sets with random errors and 20 synthetic data sets with systematic errors for each model that are based on EC data from two sites. We optimized model parameters for three models to match ten periods consisting of 14-day EC data measured at Hainich and Loobos in 2005 from May to September to get a range of reasonable parameter estimates. At a timescale of two weeks the model error can be neglected for a model like the hyperbolic light response curve, as the data error is dominant. The estimated parameters were used to create a reference model output. Then we added a random error and systematic error respectively. The random errors were estimated from the real data in the same way as for the first part of the study, the selective systematic error is a fixed percentage of the averaged observed night time flux subtracted from the modelled night time flux. Afterwards the parameters were reestimated using different ways to account for data uncertainty and error distribution. This strategy offers the advantage that the properties of the error are known and the model error is zero, but the dataset is still realistic. Knowing the true properties of the reference data we could compare estimated parameters with true parameters and model output to a reference model output to reveal the influence of the errors.

4.3.2. Data

We used half hourly EC and meteorological data from the CarboeuropeIP database. In the statistical analysis of the random error we included data from four sites: Hainich in Germany, an unmanaged deciduous broad-leaf beech forest, Loobos in the Netherlands, a planted maritime coniferous forest, Hyytiälä (Finland), an evergreen needleleaf forest and Puechabon (France), an evergreen broadleaf forest. For the parameter retrieval experiments we chose two sites, Hainich and Loobos. The data sets were processed using the standardized methodology described in Papale et al. (2006); Reichstein et al. (2005). CO_2 fluxes are corrected for storage, low turbulence conditions are filtered using the u^* criteria and spikes (outliers) are detected. Subsequently gap filling and fluxpartitioning is applied. For the parameter estimation only filtered and corrected high quality measurements are used.

4.3.3. Observation errors

In this study we assume that the measurement value consists of the actual value and an additive systematic and random error

$$x = F + \delta + \epsilon, \tag{4.1}$$

where δ is a systematic error and ϵ is a random error. The commonly used ordinary least squares (OLS) optimization assumes the random error standard deviation, e.g. the data uncertainty to be constant (homoscedasticity). A constant standard deviation, can usually be provided by the manufacturer of a measurement device or it can be determined with simple tests. For flux data the standard deviation of the random error is not constant in this case tests need to be performed for varying conditions, quantifying the changes of the standard deviation. One option is to perform measurements close to each other, temporally or spatially, provided that the conditions are the same or very similar, then the actual value is equal and the variation is caused by the random error. For the flux data meteorological conditions, the state of the vegetation and if spatially seperated the footprint and topography have to be the comparable. To get an estimate of the random error we used the gapfilling algorithm of Reichstein et al. (2005). This tool computes the expected value of the flux using data measured under the same meteorological conditions in a time window of ±7 days. The small time window is necessary to ensure a similar condition of the ecosystem. The residual of the gap filling algorithm can be used as a random error estimate (Moffat et al., 2007), it is comparable to the paired observations approach used in Hollinger and Richardson (2005), as shown in Richardson et al. (2008). For the parameter estimation an error standard deviation has to be assigned to each observation. For the parameter estimation experiments we compared the different estimates for the standard deviation of the random error:

1. constant weights,

2. the standard deviations of the observations with similar meteorological conditions within a time window of ±7 days is used directly from the gapfilling algorithm (std),this is equal to the standard deviation of the residuals between observations with similar meteorological conditions and expected value,

3. the standard deviations of the residuals of the gapfilling algorithm (res) were obtained grouping the data according to the flux magnitude in 30 groups with an equal number of data points, for each group the standard deviation was computed (see Fig. 4.1). Afterwards the standard deviation was related to the flux magnitude using two linear regression lines to allow for a minimum for net ecosystem exchange of carbon (NEE) and one linear regression line for the latent heat (LE).

For the third method the modeled flux (here the flux derived from the gapfilling algorithm) has to be used to derive the dependency of error standard deviation on flux magnitude, because the relationship between the residuals and measured flux is biased (Draper and Smith, 1998). Furthermore an observation that is accidentally lower is given a higher weight than an overestimated value, which will lead to an underestimation of flux magnitude by the model (Evans, 2003).

4.3.4. Parameter estimation

The procedure of parameter estimation can be described as varying the parameters until the best fit between model and data is found. The fit or misfit between model and data is quantified via the cost-function:

$$J(p) = (\vec{x}_d - \vec{x}_m)^T \, \mathbf{C}_d^{-1} \, (\vec{x}_d - \vec{x}_m) \tag{4.2}$$

$\vec{x}_d$ represents the data vector, $\vec{x}_m$ the model output vector, $\mathbf{C}_d$ the error covariance matrix. The best parameter set is found at the minimum of the cost function. T denotes that the vector is transposed. For uncorrelated errors the function simplifies, as all off diagonal elements of the matrix $\mathbf{C}_d$ are zero, to:

$$J(p) = \sum_{i=1}^{N} \left(\frac{x_{d_i} - x_{m_i}}{\sigma_{d_i}} \right)^2 \tag{4.3}$$

σ_d is the standard deviation of the random errors, N the number of data points. In this study we use synthetic data based on a model output, therefore the model error is zero. To consider the uncertainty of flux measurements is necessary if the errors are hetereoscedastic, e.g. error variance (=squared standard deviation) increases with increasing flux magnitude, or if different data sources are used. From another point of view this means that data with high uncertainty (high error variance) get a lower weight than data with low uncertainty (low error variance). For constant error variance the Eq. (3) simplifies to the OLS method summing up only the squared distances. Given a double exponential distribution as proposed by Hollinger and Richardson (2005) and Richardson et al. (2006b), parameter estimation should be based on the sum of absolute deviations rather than on squares. To find the minimum of the costfunction we used the Levenberg-Marquardt algorithm implemented in the data analysis package "PV-WAVE 8.5 advantage" to (Visual Numerics, 2005) for the simple models. For the complex BETHY model a Bayesian approach was used to determine the a posteriori probability density function (PDF) of parameters including prior information and the Metropolis Markov Chain Monte Carlo (MCMC) technique was used to sample the PDF of parameters, which was then characterised by mean and 95% confidence intervals (Knorr and Kattge, 2005). The Optimisation Intercomparison of Trudinger et al. (2007) compared different algorithms, including the two used here, for the optimisation of a simple coupled model. The optimisation algorithms were found to be comparable with respect to the parameter retrival. We used the MCMC for the complex model with more parameters, since the cost function for the optimization of complex models is more likely to show multiple local minima. For the same reason prior information about the parameters was included for the BETHY model. The LM is suitable for simpler models, as the shape of the cost function does not show many local minima and is then, in spite of the bootstrapping, computationally much more effective.

4.3.5. Evaluation of the parameter estimation performance

The reestimation of the parameters was evaluated through the deviation from the original parameter value, the parameter uncertainty and the root mean squared error between model output computed with the reestimated parameters and the reference model output without noise. The uncertainty of the parameters determined with the Levenberg-Marquardt algorithm, was derived by bootstrapping (n=500),

which is only based on the empirical sample not on assumptions about probability theory of the normal distribution (Wilks, 1995). As a measure of uncertainty for the parameters we used the 95% confidence intervall (=1.96· standard error) of the mean of the parameter distribution. When using the Metropolis algorithm the parameter uncertainties can be directly calculated from the sampling of the MCMC approach. The uncertainty reduction when using the Metropolis algorithm was computed as $1 - \frac{\text{posterior uncertainty}}{\text{prior uncertainty}}$.

The main difference between bootstrapping and MCMC lies in how they derive the uncertainty. Bootstrapping changes the data, e.g. drawing subsamples from the data, and uses the changes in the parameters to derive the uncertainty while MCMC changes the parameters and uses the model-data mismatch to derive the parameter distribution.

4.3.6. Models

Hyperbolic light response curve

The Hyperbolic light response curve (HLRC) computes net ecosystem exchange of CO_2 (NEE) depending on global radiation (R_g, incoming shortwave radiation):

$$\text{NEE} = -\frac{\alpha \cdot \beta \cdot R_g}{\alpha \cdot R_g + \beta} + \gamma \tag{4.4}$$

α is an approximation of the canopy light utilization efficiency, β is GPP (Gross primary production) at light saturation and γ is the ecosystem respiration. Instead of R_g photosynthetic active radiation (PAR) or photosynthetic photon flux density (ppfd) is often used, they are closely related to R_g, but not measured at all EC sites. Using R_g instead of PAR or ppfd changes only the value of α, as PPFD is approximately twice the R_g.

Water use efficiency model

To increase the complexity of the model we coupled NEE with the latent energy (LE) using the HLRC and connecting it to LE via the water use efficiency (WUE), which is the ratio of gross primary production and latent heat. The WUE times water vapour deficit (WUE_VPD) is considered constant (Beer et al., 2007). Using VPD as additional driver NEE and LE can be connected as follows:

$$\text{NEE} = -\frac{\alpha \cdot \beta \cdot R_g}{\alpha \cdot R_g + \beta} + \gamma \tag{4.5}$$

$$\text{LE} = (\gamma - \text{NEE}) \cdot \frac{\text{VPD}}{\text{WUE\_VPD}}. \tag{4.6}$$

This model is inverted against NEE and LE. To make sure, that the LE and NEE misfits contribute to a similar extend to the cost function when using constant weights, as it is the case when using the data derived estimates of the error standard deviation are used, we scaled the residuals with constants c for NEE and LE, respectively. These constants are defined such that the sum of the weighted synthetic error is the same when using the constant and using the std weights, for both NEE and LE:

$$\frac{1}{c} \sum \text{error} = \sum \frac{\text{error}}{\text{std}} \tag{4.7}$$

c denotes the constant weight, the same weight is used for the inversion of the BETHY model. In equation 4.7 'error' is the error estimated from real data, that was added to the synthetic data.

The model underestimates LE, but as we use the model output as reference, by design of the study the model errors are not important. Conclusions about the influence of the data error on model parameterisation are not affected. We used this model to show, that the results derived with the simple model hold for models of various complexities and to mediate between the very simple HLRC and the quite complex SVAT scheme of BETHY.

BETHY

BETHY is a process-based model of the coupled photosynthesis and energy balance system to simulate the exchange of CO_2, water and energy between soil, plant canopy and atmosphere (Knorr and Heimann, 2001a). It computes absorption of PAR in three layers, while the canopy air space is treated as a single, well mixed air mass with a single temperature. Evapotranspiration and sensible heat fluxes are calculated from the Penman-Monteith equation (Monteith, 1965). Carbon uptake is computed with the model by Farquhar et al. (1980) for C3. The stomata and canopy model of Knorr (2000) simulates canopy conductance in response to PAR, VPD and soil water availability. In the version of BETHY applied here, autotrophic respiration is calculated as a temperature modulated fraction of photosynthetic capacity while heterotrophic respiration is based on a basal respiration modulated by soil water availability and air temperature. The inversion set up was the same as in Knorr and Kattge (2005), inverting all 21 parameters simultaneously. The prior uncertainties of the parameters were set to 20% of the prior parameter value. For the synthetic datasets the prior parameter were the parameters used to generate the data.

4.4. Results and discussion

4.4.1. Statistical properties of the error estimates

Heteroscedasticity

The standard deviation of the error has been derived from the residuals of the gapfilling model, e.g. standard deviation of the residuals depending on the flux magnitude (res), and using the standard deviation of the gapfilling algorithm directly (std). Figure 4.1 shows the relationship between flux magnitude and error standard deviation for NEE and LE. The standard deviation is not homogeneous, e.g. the errors are heteroscedastic and increase with increasing flux magnitude. Thus the residuals have to be weighted with the reciprocal of the standard deviation of the random errors as already suggested by previous studies (Richardson et al., 2006b). The magnitude of the error variance is similar for the two methods of deriving the error variance described in the previous section, see Fig. 4.1, for res the observations needed to be grouped to derive the standard deviation. The res standard deviation for NEE ranges from 1 to 5, for LE from 5 to 40. With the std method the ranges are wider because the data were not grouped. For NEE the standard deviation lies between 0.5 and 9.5, for LE between 2.8 and 85. For eddy covariance data it is known, that the error variance increases with increasing flux magnitude, Richardson and Hollinger (2005) showed that the error standard deviation of NEE, not the error itself, not only scales with flux magnitude, but that wind speed also has a fundamental effect on the uncertainty. Thus not
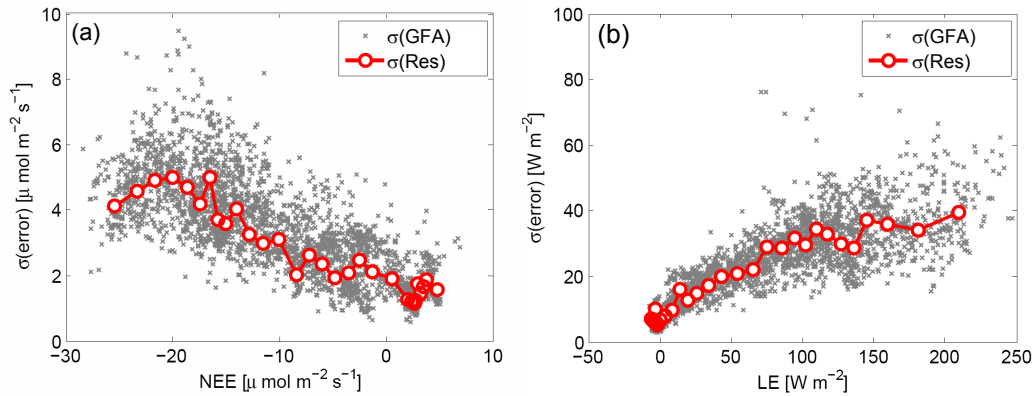
Figure 4.1.: Data uncertainty derived from the gapfilling algorithm directly and standard deviation of the gapfilling algorithm residuals for NEE **(a)** and LE **(b)**.

the whole variability of the standard deviation can be reproduced when only the flux magnitude is used. Another source for the higher scatter of the std results is the uncertainty in the estimation of the standard deviations derived directly (std).

Distribution

Previous studies (Richardson and Hollinger, 2005) showed, that the error distribution of NEE is rather double exponential (Laplace) than normal and this is also found for the data used here (see Fig. 4.2e, f). For LE the distribution is even more peaked than the double exponential distribution. The normal distribution is characterized by the mean and the standard deviation. As error standard deviation increases with increasing flux magnitude the distribution of all error estimates is a superposition of normal distributions with varying standard deviation. If we group the data according to the flux magnitude, we find Gaussian distributions for high flux magnitudes (see Fig. 4.2a, b, exemplary for the Hainich site), adding more data to the distribution plot, we find a rather double exponential distribution (see Fig. 4.2c–f). Another possibility to show the Gaussian distribution is shown in Fig. 4.2g, h , we normalized the errors with the standard deviation derived with the gapfilling algorithm (std) this transforms all error distributions to a standard deviation of unity. For NEE the normalized errors are slightly closer to a normal distribution than for LE. Thus the double exponential distribution is largely due to a superposition of Gaussian distributions and the least squares criteria can be used for the eddy covariance data shown here. Figure 4.3 shows the distribution of errors from the four different sites. The normalization of NEE resulted in a rather Gaussian distribution, for LE the distribution is in between Gaussian and Laplace distribution and is slightly skewed. This indicates, that the distribution of the error varies from site to site or that the error estimation does not perform well for all sites. One indicator for the peakedness of the distribution is the excess kurtosis (=kurtosis-3), it is zero for a normal distribution and 3 for a double exponential distribution, a high kurtosis indicates a strong peak. Figure 4.4 shows the kurtosis for ten two week periods for the four sites. For NEE the normalization of the errors decreases the kurtosis and changes the distribution to a less peaked shape. The kurtosis is in general below the kurtosis for double exponential distributions, but some outliers indicate a much stronger peak (excess kurtosis=5.5). For LE the kurtosis decreases also for HAI, LOO and PUE.
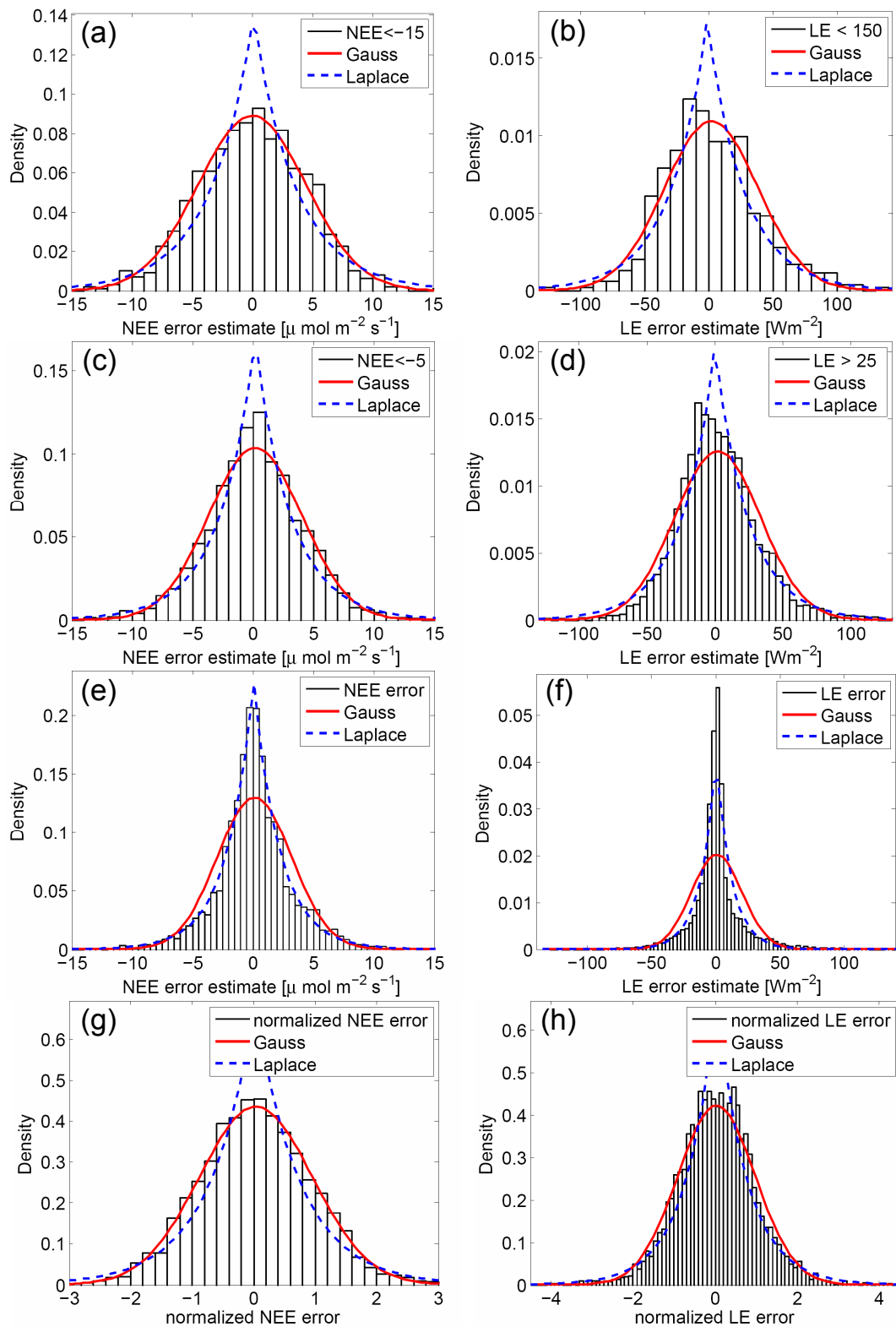
Figure 4.2.: Distributions of NEE (left) and LE (right) error estimated with the gapfilling algorithm. **(a)**, **(b)**: error of high flux magnitudes, **(c)**, **(d)**: error of high and medium flux magnitude, **(e)**, **(f)**: all error estimates, **(g)**, **(h)**: errors estimated with gapfilling algorithm and normalized with std, data: Hainich May–September 2005.
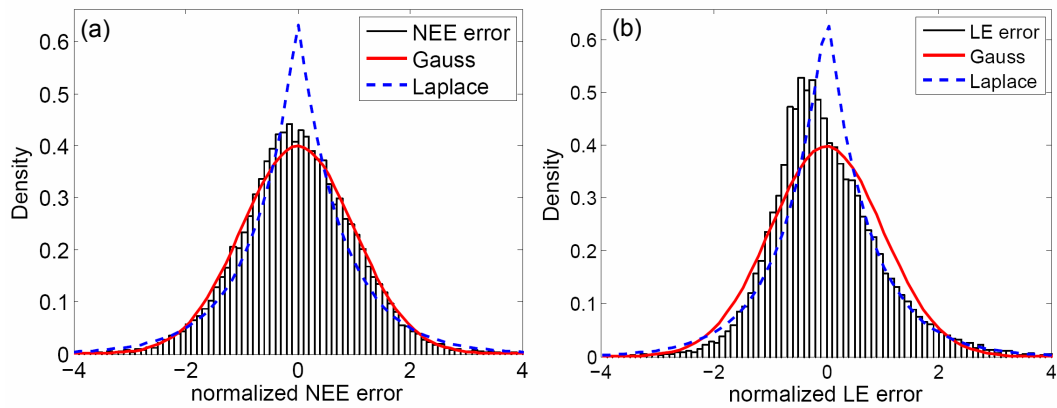
Figure 4.3.: Distribution of NEE **(a)** and LE **(b)** errors normalized with std and z transformed ((error-mean(error))/standarddeviation(error)) using data from HAI, LOO, HYY and PUE.
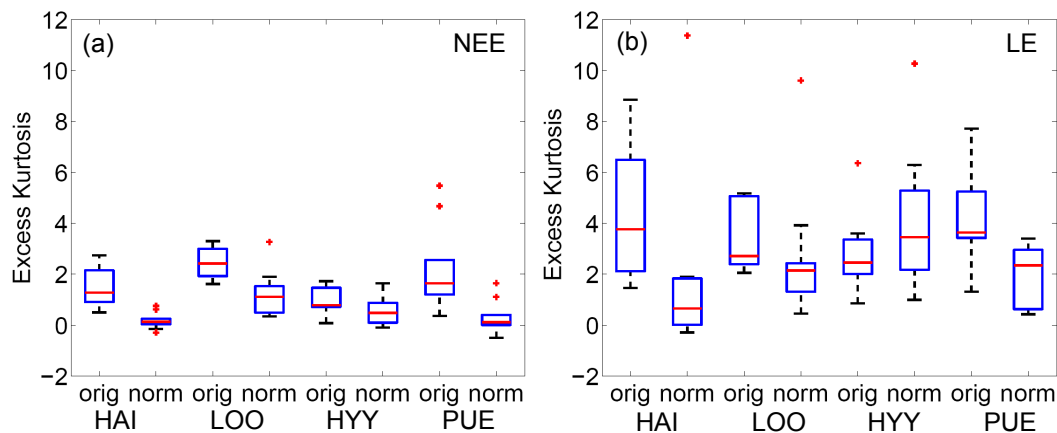
Figure 4.4.: Boxplots with median, upper and lower quartile, minimum and maximum or outliers (points) for the excess kurtosis of the 10 two week periods from May to September 2005 for errors (orig) and normalized errors (norm) of NEE **(a)** and LE **(b)**.

For HYY the kurtosis increases after normalization. The kurtosis shows a high sensibility to outliers, but also to the rule of the detection of outliers, excluding outliers will always decrease the kurtosis. As the outliers are important for the characterisation of the distribution and the data is already prefiltered (spike detection according to Papale et al., 2006) we did not exclude them. For errors of fluxes with high magnitude, normal distribution is still found (excess kurtosis 0.3) and seems to be valid across sites. Overall, we conclude that random error characteristics should be considered on a site-by-site basis. When comparing different studies and different sites regarding their error distribution, a careful documentation of the influence and the treatment of outliers is strongly recommended.

Correlation

Autocorrelation

Figure 4.5 shows the autocorrelation function of the random errors for the four eddy sites. The
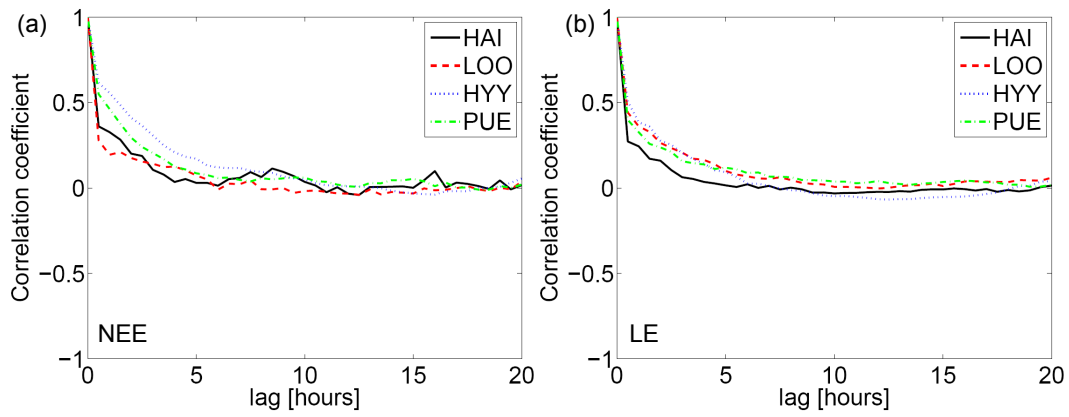
Figure 4.5.: Autocorrelation of the NEE **(a)** and LE **(b)** errors, data: May to September 2005.
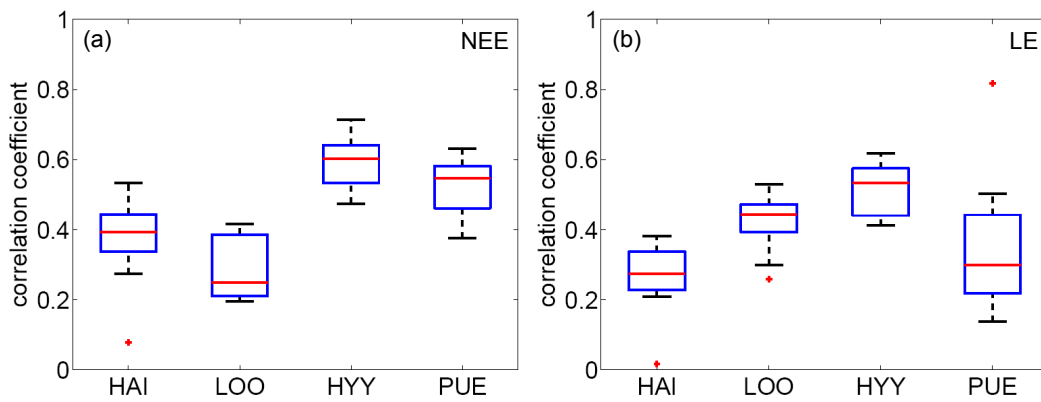
Figure 4.6.: Boxplots of the autocorrelation of lag=1 (0.5 h) for ten two week periods from May to
             September for NEE **(a)** and LE **(b)**.

behaviour of the function is similar for all sites, the autocorrelation decays fast, after 10 hours there is no considerable change in the correlation. Figure 4.6 shows boxplots for the autocorrelation for a lag of 30 min, it is usually below 0.7, with one exception for Puechabon (0.82). Hyytiälä shows the highest autocorrelation for LE and NEE, Loobos the lowest for NEE and Hainich the lowest for LE. Although the gapfilling algorithm provides a reasonable estimate for the random error, the autocorrelation could partly be an artefact of the algorithm, if the deviation from the statistical expection value was not caused by a random error the following and previous value would deviate in a similar way and the actual autocorrelation of the random error would be lower. To make sure that error autocorrelation does not influence the parameter estimation one could prefer to use only every second or third value for the parameter estimation.

Crosscorrelation

R^2 values for the crosscorrelation between NEE and LE errors of the four sites and ten data periods for each site are summarized in Table 4.1. In our study the correlation between NEE and LE

Table 4.1.: Crosscorrelation between NEE and LE errors for ten two week periods between March and September 2005.

	R^2 of NEE and LE errors									
Data period	1.–15.5.	16.–31.5.	1.–15.6.	16.–30.6.	1.–15.7.	16.–31.7.	1.–15.8.	16.–31.8.	1.–15.9.	16.–30.9.
HAI	0.089	0.004	0.176	0.192	0.088	0.136	0.202	0.007	0.077	0.097
LOO	0.004	0.029	0.059	0.086	0.031	0.000	0.004	0.024	0.030	0.010
HYY	0.197	0.033	0.139	0.128	0.021	0.012	0.023	0.049	0.000	0.003
PUE	0.093	0.244	0.038	0.033	0.068	0.003	0.012	0.018	0.019	0.031

errors is close to zero, thus the correlation between NEE and LE errors is of minor importance and does not need to be considered in the error covariance matrix. The highest R^2 was 0.24 for one period for Puechabon, for the same period the outlier of the LE autocorrelation for a lag of 30 min was found (see Fig. 4.6). The measurements of NEE and LE are both based on the vertical wind velocity and errors introduced via the wind velocity measurement, such as errors due to turbulence sampling must show up as a correlation between NEE and LE errors. This indicates that the variation in the measured fluxes under similar meteorological conditions seems (i.e. the flux errors) to be rather caused by changes in concentrations of water and CO_2 than by the measurement of the vertical wind velocity. As auto- and crosscorrelation are low, the generalized least squares method (Eq. 4.2) can be simplified to the weighted least squares method (Eq. 4.3) by setting off-diagonal elements in the error covariance matrix to zero.

4.4.2. Parameter retrieval

Ordinary least squares vs weighted least squares

The parameters were estimated for three models of different complexities, the synthetic data is based on data from two different sites (Loobos and Hainich). We are comparing constant weights with two ways of estimating the standard deviation of the observation errors, which is then used to weight the data for the parameter estimation to account for the non constant error standard deviation. The standard deviation of the errors was estimated as the standard deviation of the observations measured under similar metereorological conditions (std) and as the standard deviation of the gapfilling algorithm residuals related to the modelled flux magnitude (res), see Fig. 4.1. The results of the parameter retrieval experiments (Fig. 4.7) show, as expected, that the random error introduces no systematic error to the parameter estimates and the true parameters are usually within the parameter uncertainty (95% confidence interval) derived from bootstrapping. The mean of the parameter ratios is not significantly different from unity ($\alpha=0.05$). The mean uncertainty of the parameters using a non constant estimate for the error standard deviation as weight is between 10 and 24% lower for the HLRC than using constant weights (see Table 4.2). Due to the stochastic nature of the procedure, these results are true for the mean results but there exist data periods for both sites in which the results are opposite. The std weights decrease the mean uncertainty more than res and therefore describe the error standard deviation better. The root mean squared error (rmse) between reference model output without noise and the model output using
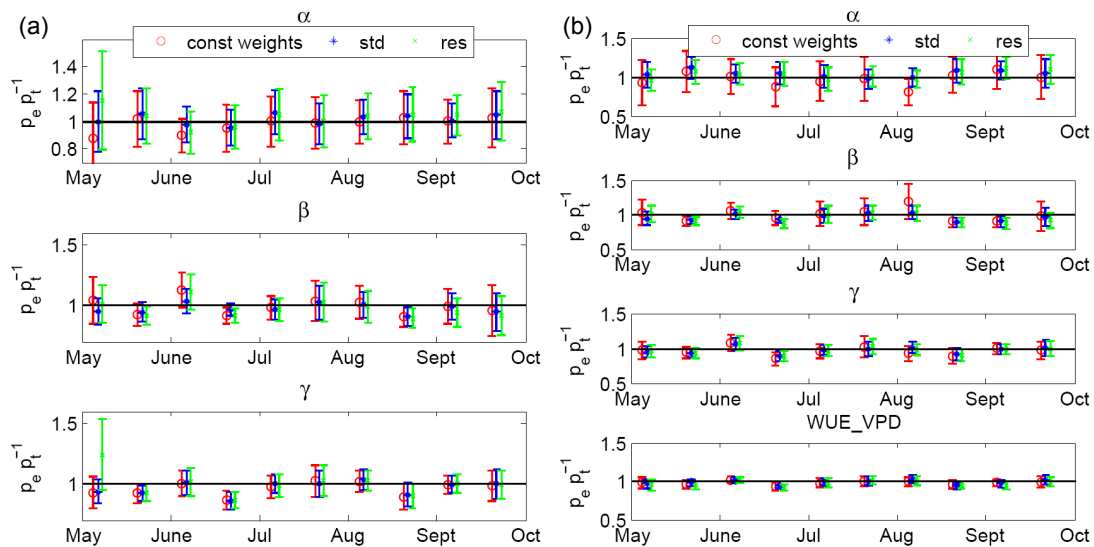
Figure 4.7.: Time series of normalized parameters (estimated/true) based on data from Loobos with a random error for the HLRC (left) and the WUE model.

the reestimated parameters can be decreased using std as weights for the HLRC, for res it increases. This indicates, that the 'std' is a more accurate estimate for the data uncertainty and that a description of the data uncertainty only based on flux magnitude, as 'res', is likely not sufficient.

For the water use efficiency model the results of the model parameterization are similar, estimates of parameter uncertainty decrease between 5% and 60% and the RMSE between reference model output and model output of the reestimated parameters is decreased when using std, while res increases the value (see Table 4.3). For simplicity we focus on the comparison between constant weights and std, since std gave the best results. For the inversion of the BETHY model the distance between retrieved and true parameters can be decreased using std compared to the constant weights (see Table 4.4).

The influence of using varying data uncertainty compared to constant data uncertainty with the MCMC algorithm is less pronounced as the absolute value of the data uncertainty is more important than the relative changes. Nevertheless the reduction of uncertainty for parameters is higher when using 'std' and the rmse between reference and model output is decreased.

Another advantage of weighting the data showed up during the initial fit to real data for the creation of the reference model output, the parameters estimated with std weights resulted in reasonable parameters, whereas using constant weights for some periods negative values for α were estimated for the HLRC and water use efficiency model (not shown). The random error changes the shape of the cost function, it can increase the number of local minima or the minimization can become an ill-posed problem. Using weights representing the data uncertainty seems to improve the behaviour of the cost function and improves the extraction of information inherent to the data. This shows that the standard deviation provided by the gapfilling algorithm is a good measure for the eddy covariance data uncertainty, it improves the parameter retrieval and therefore model performance after optimization, at least for the sites used here. For skewed error distributions we would expect the parameter estimates to be biased.

To explore the power of the Bayesian approach an interesting alternative way to cope with data un-

Table 4.2.: Mean of retrieved normalized parameters and the mean uncertainty for the ten two week periods for the HLRC.

		Loobos 2005
		mean normalized parameter		mean 95% confidence interval	
		Least squares minimization
weights	α	β	γ	α	β	γ		rmse
constant	0.978	0.993	0.961	0.187	0.133	0.102		0.265
std	1.016	0.975	0.967	0.157	0.101	0.091		0.228
res	1.034	0.970	0.998	0.198	0.119	0.115		0.332
		Absolute deviations minimization
constant	0.951	0.991	0.935	0.174	0.145	0.099		0.425
std	0.972	0.976	0.938	0.172	0.133	0.096		0.388
		Hainich 2005
		Least squares minimization
constant	0.983	1.001	0.969	0.173	0.087	0.096		0.231
std	1.048	0.984	0.990	0.149	0.071	0.084		0.218
res	1.046	0.981	0.988	0.176	0.080	0.089		0.239
		Absolute deviations minimization
constant	1.049	0.977	0.981	0.101	0.199	0.107		0.300
std	1.062	0.972	0.986	0.096	0.192	0.103		0.341

certainties would be to include a relationship for the data uncertainty in the Likelihood function. The uncertainty could be represented by a linear dependency and the parameters of the relationship could be estimated in addition to the model parameters. However, since with eddy covariance data one can provide information about the random error in the data independent of the optimization, we expect our method to be more robust, e.g. independent of model errors.

Least squares vs absolute deviations

As the use of absolute deviations in the cost function was suggested previously by Richardson et al. (2006b) we compare least squares and absolute deviations, to illustrate the effect to parameter estimation. Comparing the parameter ratio again shows no significant difference between the methods. For our sites, the parameter uncertainty increases using absolute deviations compared to the ordinary least squares method (see Table 4.2). The rmse increases compared to the OLS using constant weights and for the weighted least squares. Since by normalizing the errors with the standard deviation we get a Gaussian distribution for our selected sites the absolute deviation minimization cannot improve the parameter retrieval. If the errors show a double exponential distribution as a result of the superposition of different Gaussian distributions, then least squares optimization should be applied. If the error distribution is more peaked due to outliers or a different data filtering, robust methods like the minimization of absolute deviations or robust regression techniques, which exclude outliers, may be advantageous. Testing whether the normalized error distribution is Gaussian could support the choice of the cost function.

Table 4.3.: Mean of retrieved normalized parameters and the mean uncertainty for the ten two week periods for the WUE-model using last squares minimization.

| | Loobos 2005 | |
| | \multicolumn{4}{c}{mean normalized parameter} | \multicolumn{4}{c}{mean 95% confidence interval} | |
weights	α	β	γ	wue_vpd	α	β	γ	wue_vpd	rmse
const	0.974	1.003	0.964	0.980	0.249	0.149	0.111	0.058	1.243
std	1.043	0.963	0.973	0.987	0.140	0.088	0.085	0.050	0.930
res	1.032	0.958	0.975	0.973	0.148	0.096	0.090	0.054	1.238
	Hainich 2005	
	\multicolumn{4}{c}{mean normalized parameter}	\multicolumn{4}{c}{mean 95% confidence interval}	
weights	α	β	γ	wue_vpd	α	β	γ	wue_vpd	rmse
const	1.027	0.987	0.986	0.960	0.300	0.135	0.166	0.051	2.926
std	1.049	0.978	0.984	0.978	0.121	0.081	0.092	0.043	1.924
res	1.061	0.964	0.986	0.952	0.143	0.090	0.098	0.045	3.330

Table 4.4.: Sum of the uncertainty reduction, summed absolute deviation of the parameter ratio from 1 and mean rmse between model output and reference output.

site	Loobos		Hainich	
	const	std	const	std
uncertainty reduction	50.93	52.31	47.19	50.08
parameter deviation	14.83	13.30	14.16	12.49
rmse	4.11	3.17	4.34	3.6

Systematic error

Figure 4.8 shows the results of the parameter retrieval based on data with a selective systematic night-time error of 10, 20 and 40%. The α and γ parameters of the HLRC show a systematic bias, estimated parameters underestimate the underlying true parameter. The bias is stronger for higher data error. For β the parameter bias seems to be not systematic, the retrieved parameter is for some periods lower, for some periods higher than the original parameter. β is GPP at light saturation, as NEE at light saturation does not change but only the night time flux, representing the respiration, is lower β should also be lower to sum up to the same NEE values during daytime. The effect on γ seems to be too low to show up in the comparably high values of β. For the water use efficiency model all parameters are biased, all estimates are lower than the true values. Through the interconnection of GPP and LE and the use of water and CO_2 fluxes to constrain the parameters the distance to the true value decreases for all parameters. This illustrates the potential of using multiple constraints for inverse model parameter estimation. The parameter uncertainties increase the higher the error but the real value of the parameter is not within the uncertainty range of the estimated parameter. This means, that the real uncertainty of the parameter is underestimated, projection of the parameter uncertainty to model output will result in uncertainties for the fluxes that are too low. To get the real uncertainty for parameters and fluxes further knowledge about the systematic errors is needed and methods need to be developed to incorporate them into the estimation of uncertainty, if the systematic errors cannot be removed.
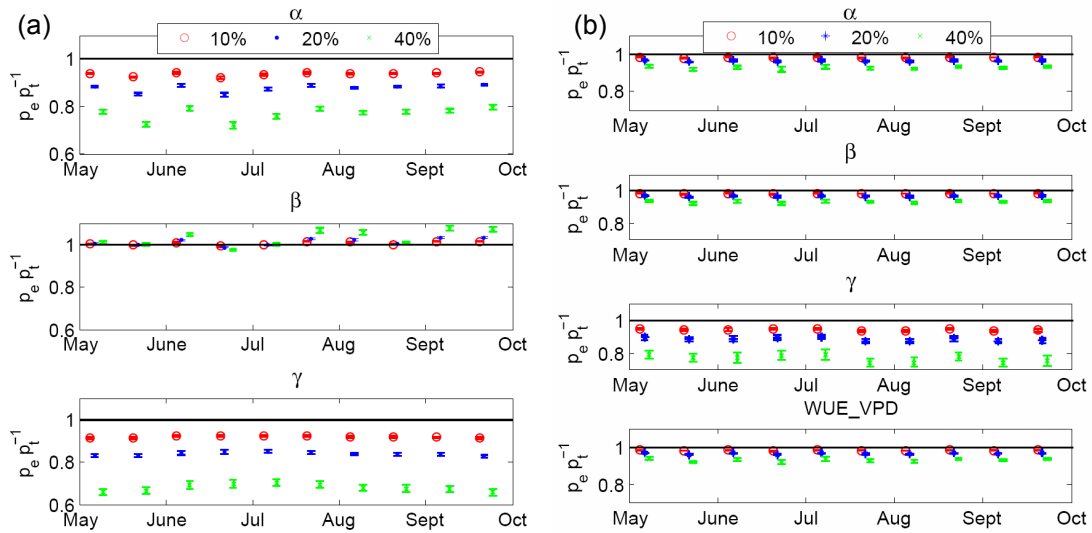
Figure 4.8.: Time series of normalized parameters (estimated/true) based on data from Loobos with a selective systematic nighttime error for the HLRC **(a)** and the WUE model **(b)**.

4.5. Conclusions

Previous work to quantify the random error structure of eddy covariance data (Hollinger et al., 2004; Richardson and Hollinger, 2005; Richardson et al., 2006b) has focused on describing the moments of the distribution of the error, particularly relating the expected magnitude of the error (i.e. its standard deviation) to the flux magnitude, and evaluating whether or not flux errors are Gaussian. Here we have built on these efforts by considering the auto- and crosscorrelation, introducing a new method to quantify the standard deviation of the random errors. We show the effect of the varying standard deviation to the distribution and investigate how random and systematic errors affect parameter estimates.

The analysis of the error distribution shows that the apparently double exponential distribution of the eddy flux data can be almost entirely due to the superposition of Gaussian distributions with inhomogeneous variance. Whether this is the case for a special site can be affirmed by testing the normality of the normalized error distribution. If it cannot be affirmed one should consider using robust methods. The autocorrelation is low, but one might consider to analyse the autocorrelation function and use only every second or third flux for parameter estimation if there is enough data available. As a reason for the low but significant autocorrelation of errors we can not exclude artefacts of the gap filling tool. The crosscorrelation between LE and NEE is low and can be neglected. The assumption for ordinary least squares that is not met is the constant error standard deviation, thus the ordinary least squares method needs to be extended to weighted least squares, using the reciprocal of the standard deviation as weight in the costfunction. We propose a measure for data uncertainty, e.g. the standard deviation of the values used to compute the expected value, that can be used to weight the data in the costfunction. Weighting the data decreases the parameter uncertainty and the parameter retrieval is improved. We showed that this result holds true for a wide range of model complexities. We show that the impact of systematic errors varies by parameter, but the bias is systematic, therefore the interpretation of parameters derived from data with systematic errors might be misleading. The parameter uncertainty slightly increases

when a systematic error is added, but the true parameter is not within the uncertainty range of the estimate. Not considered here but of similar importance is the model error, which was set to zero by using the model output as basis for the synthetic data. For the least squares optimization the model output random error is additive to the data random error and depending on the point of view part of the data random error can also be seen as model errors, e.g. footprint heterogeneity. Model structural problems can also affect parameter estimation in a similar way assystematic data errors, dynamics in the data that are not represented or not sufficiently represented in the model structure can lead to biased parameters (Carvalhais et al., 2008). Using inverse model parameter estimation we can only accept the best fit although if the model reproduces the patterns emerging from these insufficiencies. Hence we conclude that potential systematic errors in flux data or models need to be addressed more thoroughly in data assimilation approaches since otherwise uncertainties will be vastly underestimated.

4.6. Acknowledgements

5. Optimization of a land surface model using fAPAR and eddy covariance data: impact and propagation of uncertainties due to systematic errors

5.1. Abstract

Model data fusion has become a popular tool to optimize and improve carbon cycle models with the goal to reduce the uncertainties of the model output. This method enables us to retrieve estimates for parameters that are not easily observable. For modeling photosynthesis the maximum carboxylation rate (Vcmax), quantifying the photosynthetic capacity of a leaf, is a key parameter. Observations on leaf level indicate a seasonality that has not been implemented in models yet. The observations are rare and not possible on ecosystem scale but the inversion of carbon cycle models against eddy covariance carbon and water fluxes that are available across climate zones and vegetation types allows its estimation. An exact representation of the seasonality of the leaf area index (LAI) and the fraction of absorbed photosynthetically active radiation (fAPAR) is a prerequisite as it has a direct influence on the estimate of Vcmax. Comprehensive error propagation from the data into parameter and model output uncertainties is needed for reliable results in model data fusion studies. Although the uncertainty due to systematic errors in eddy covariance data has been shown to be important it has not been included yet in model data fusion studies. Here we suggest an approach to represent the uncertainty due to systematic errors in eddy covariance data by data ensembles based on the uncertainty of the u* threshold, a filter for unreliable measurements, and to combine it with the uncertainty due to the random errors. To account for systematic differences in fAPAR observations we use in situ and SEAWIFS remote sensing data to constrain the phenology module of the JSBACH model. The optimized fAPAR is then used as input for the photosynthesis module that is constrained with eddy covariance carbon and water flux data of the Hainich flux tower in Germany, optimizing the parameters Vcmax and rooting depth. We find that the Vcmax parameter is well constrained by the eddy covariance data and that the influence of systematic errors due to the uncertainty of the u* threshold when using GPP estimates is small, the rooting depth is not influenced by the uncertainty due to systematic errors in the carbon flux. Limiting in deriving a general estimate of Vcmax is the difference in the fAPAR data streams: when estimating a constant Vcmax the estimate based on the fAPAR constrained with satellite data is more than three times as high as the estimate using the fAPAR constrained with in situ data. To model the seasonality of photosynthesis correctly the seasonality can be either implemented in the fAPAR or Vcmax. While for the in situ fAPAR data an additional seasonality in Vcmax would be needed, the seasonality of the satellite data was too strong, although this could be a site specific effect. In spite of the strong difference in the evolution of fAPAR in summer between the data streams, the information about spring event and maximum LAI were consistent and we propose to exploit the advantage of the spatial availability of remote sensing data to constrain rather these characteristics in the spatial domain than the seasonality.

56

Optimization of a land surface model using fAPAR and eddy covariance data: impact and propagation of uncertainties due to systematic errors

5.2. Introduction

Improving biogeochemical models towards a comprehensive representation of the carbon cycle and the interactions between carbon, water and nutrient cylces is important for a reliable extrapolation of the models in time and in space. Model data fusion has become a popular approach to evaluate, optimize and improve models (Williams et al., 2009) and thereby reducing the uncertainty. The optimization of a model is especially useful to retrieve estimates of parameters that are not easily observable. A key parameter for modeling photosynthesis of ecosystems is the photosynthetic capacity of a leaf (e.g. the maximum carboxylation rate, Vcmax) (Farquhar et al., 1980). Although a seasonality of Vcmax has been observed on leaf level (Muraoka et al., 2010) it is often implemented in models as a constant parameter. Observations of the seasonality of this parameter are rare and only possible on leaf level, a functional quantification of the seasonality does not exist yet. Inverse model parameter estimation provides the opportunity to estimate Vcmax on ecosystem scale by using a carbon cycle model and eddy covariance data (Knorr and Kattge, 2005; Reichstein et al., 2003b) that are available across climate zones and vegetation types (Baldocchi, 2008). An exact representation of the seasonality of the leaf area index (LAI) and the fraction of absorbed photosynthetically active radiation (fAPAR) is a prerequisite as it has a direct influence on the estimate of Vcmax.

To retrieve reliable results in model data fusion studies a comprehensive error propagation of the data into the uncertainty of the parameter and the model output is needed. Parameter estimates are influenced in a different way by random or systematic errors that cannot be removed: Lasslop et al. (2008) showed using synthetic data that while for data with random errors the true parameter is within the uncertainty range of the estimate, for systematic errors in the data the estimate is not within the uncertainty range. For the eddy covariance data random errors have been characterized well for model data fusion (Hollinger and Richardson, 2005; Richardson et al., 2006b; Lasslop et al., 2008). Uncertainties due to systematic errors have been hardly adressed, although for instance the choice of the u*-threshold, that is used to filter the data for unfavourable measurement conditions has been shown to dominate uncertainties in annual estimates of NEE (Papale et al., 2006; Gove and Hollinger, 2006), this uncertainty can cause systematic biases and has not been consistently included in model data fusion studies.

Here we use eddy covariance carbon and water fluxes and the fraction of absorbed photosynthetic radiation in a model-data fusion study to constrain the phenology and photosynthesis module of the JSBACH model (Raddatz et al., 2007; Knorr and Heimann, 2001a) in a two step approach. We explore the possibility to evaluate the model concerning the seasonality of the LAI and Vcmax considering the uncertainties in the data due to random and systematic errors. To better account for uncertainties due to systematic biases in eddy covariance data a bootstrapping approach is used to create data ensembles that represent a range of likely u* thresholds. They are complemented by an additional GPP dataset that is based on daytime eddy covariance measurements to be independent of the rather problematic nighttime observations. We combine this uncertainty due to potential biases with the uncertainty due to random errors to provide a comprehensive estimate of uncertainty for the inversion against eddy covariance data. Moreover we include systematically different fAPAR data streams, observed at the site with sensors above and below the canopy and fAPAR derived from remote sensing, to discuss the

consistency between the fAPAR, flux data streams and the model.

5.3. Methods

5.3.1. Data

Fraction of absorbed photosynthetically active radiation

The fraction of absorbed photosynthetically active radiation is generally defined as the fraction of the incoming solar radiation in the photosynthetically active radiation (PAR) spectral range that is absorbed by the surface. However when it comes to the measurement the realization of this definition is less explicit and systematic differences can be found between the datasets.

Ground observations   Using ground observations fAPAR data are derived with sensors above the canopy measuring the incident PAR (incPAR) and sensors measuring the PAR below the canopy (bc-PAR), using these two measurements the fraction of intercepted PAR can be derived:

$$fiPAR = 1 - \frac{bcPAR}{incPAR} \qquad\qquad (5.1)$$

This observation includes radiation intercepted by branches and stems, but not the understorey. The model is based on the fraction of radiation absorbed by leaves only, other parts of the plants are not considered. To achieve consistency between model and observations the average winter values were scaled to zero to remove the effect of branches and stems, assuming that in summer the canopy is closed and the influence of stems and branches on the observations disappears.

Remote sensing   The fAPAR derived by remote sensing devices conceptually differ from the ground observations: it does not include the fraction of incident PAR reflected by the vegetation, but it includes the fraction of PAR that is reflected by soils and absorbed by the canopy on the way back to space (McCallum et al., 2010). Satellite products differ in definition of fAPAR: some derive a "green" fAPAR including radiation that is absorbed by the green parts of the vegetation only (Cyclopes and SEAWIFS not MODIS), thus incPAR absorbed by soils and stems or branches are not included (McCallum et al., 2010). For an initial comparison of data streams SEAWIFS, MODIS and Cyclopes data were included, MODIS and Cyclopes having a spatial resolution of 1 km, the SEAWIFS data has a spatial resolution of 2 km. In the model data fusion study we used fAPAR data derived from the SEAWIFS satellite for the years 2000-2005 for the following reasons: a comparison study found that SEAWIFS best captures the fAPAR over northern Eurasia (McCallum et al., 2010). More data are available for the SEAWIFS dataset compared to the Cycplopes dataset, compared to the MODIS data the methodology involves less prior assumptions, e.g. MODIS requires LAI as input (McCallum et al., 2010; Gobron and Verstraete, 2008) and the "green fapar" might be more consistent with the model structure. The SEAWIFS data were filtered according to their standard deviation, winter values were filtered, as they mainly reflect the understorey that is not included in the model. Data were used if the fAPAR was greater than 0.3 and the relative standard deviation of the measurement was less than 0.3.

58

Optimization of a land surface model using fAPAR and eddy covariance data: impact and propagation of uncertainties due to systematic errors

Eddy Covariance

We used carbon and water flux data from the FLUXNET site Hainich in Hainich forest Germany (Knohl et al., 2003; Anthoni et al., 2004). The forest is a decidous broadleafed forest with the main species beech (65%), ash (25%) and maple (7%). The understorey consists of Allium ursinum, Mercurialis perennis, Anemone nemorosa and is photosynthetically active between April to October, while the leafed trees start one month later in May (Knohl et al., 2003).

The data were processed in a standardized way after Papale et al. (2006) and Reichstein et al. (2005). The data are storage corrected, spike filtered, u*-filtered, and subsequently gap-filled. For the optimization of the model parameters only measured (i.e. non-gapfilled) half hourly data of water fluxes and gross primary production (GPP) estimates derived from the observed net ecosystem exchange (NEE) were used. As the model does not include interception and evaporation the dataset was further filtered and data were only used if it did not rain on the same day and the day before.

To create the data ensembles of GPP to represent the potential systematic biases due to an underestimation of nighttime fluxes (Goulden et al., 1996; Aubinet et al., 2000; Baldocchi, 2003; Goulden et al., 2006) two complementary flux partitioning algorithms were applied, where one is based on nighttime data (Reichstein et al., 2005), the other on daytime data (Lasslop et al., 2010b). Moreover the uncertainty of the u* threshold, that has been shown to contribute a large part to the uncertainty of annual fluxes (Papale et al., 2006; Gove and Hollinger, 2006), was estimated using a bootstrapping approach (Papale et al., 2006) generating a distribution of likely u* thresholds and subsequently a distribution of NEE and GPP datasets respectively.

5.3.2. Model

We used the photosynthesis (Knorr and Heimann, 2001a) and phenology module of the JSBACH model (Raddatz et al., 2007). The phenology module is needed to provide values of LAI. The spring event is determined based on the alternating model of (Murray et al., 1989), the autumn event happens when the pseudo soil temperature (weighted running mean of the air temperature) first falls below a critical temperature, the autumn event temperature. For summergreen forests the original phenology has three phenological phases: (1) a rest phase in winter, that starts with the autumn event when a certain temperature of the soil is reached and ends with the spring event when an empirical relationship based on growing degree days and chilling days is fulfilled, it is characterized by a high shedding rate, (2) a growth phase that lasts for 60 days and is characterized by the growth rate and the maximum LAI and (3) a vegetative phase that is characterized by a small shedding rate.

The ground based fAPAR observations suggest that the vegetative phase is not needed in the model. Therefore we modified the model removing the vegetative phase and limiting the growth phase by the autumn event.

The photosynthesis module computes PAR as absorption of PAR in three layers, while the canopy air space is treated as a single, well mixed air mass with a single temperature. The stomata and canopy model of Knorr (2000) simulates canopy conductance in response to PAR, VPD and soil water availability. Carbon uptake is computed with the equation by Farquhar et al. (1980) for C3 modified with an inhibition function for high temperature (>55°C) based on (Collatz et al., 1991). Evapotranspiration and sensible heat fluxes are calculated using the Penman-Monteith equation (Monteith, 1965).

5.3.3. Model data fusion framework

Realistic uncertainty estimates on parameters and model output in model data fusion studies can only be achieved including a careful error propagation of the data errors. Random errors have been characterized with respect to their variance, correlation and distribution (Lasslop et al., 2008). The errors due to potentially biased measurements, and the uncertainty in filtering them, result in systematic errors of eddy covariance data, which have not been included and combined with uncertainties due to random errors in uncertainty propagation frameworks to the parameters and model output.

To consider the uncertainty of the fAPAR data we used multiple data streams as input to the photosynthesis module:

1. with the standard parameterization

2. optimized with site data and

3. optimized with SEAWIFS remote sensing data.

In this study we distinguished between uncertainties due to systematic and random errors in eddy covariance data as they affect the parameter estimate in different ways (Lasslop et al., 2008; Vasquez and Whiting, 2006). The random error was estimated and the standard deviation included in the cost function according to Lasslop et al. (2008). The distribution of the random error at the Hainich flux tower has been shown to be approximately Gaussian (Lasslop et al., 2008) and thus requires, together with the non constant standard variation, weighted least squares as distance measure in the cost function.

The uncertainty due to systematic errors of eddy covariance data was represented by multiple datasets generated as described in the eddy covariance data section. By fitting the model to the datasets we retrieved 100 parameter sets, the distribution of these parameters represent the uncertainty of the parameters due to the uncertaity of the u* threshold (systematic error). For each of the parameter sets a parameter uncertainty can be estimated (Omlin and Reichert, 1999), this uncertainty is mainly caused by the random error. To combine these two uncertainties we assumed a normal distribution of the uncertainty due to random errors and generated a distribution for each of the 100 parameter sets according to its uncertainty. Subsequently these distributions were superimposed to retrieve one probability density function quantifying the effect of both error types.

Optimization

We used a two step approach in the optimization: In a first step the phenology module was optimized separately: The purpose here is to explore the differences between the fAPAR data sources. In a second step the photosynthesis module is optimized against GPP and latent heat using the fAPAR data streams as a driver to identify the most consistent fAPAR data source. The phenology module contains "'if"' statements for the beginning of the growth phase and the autumn event. Parameters used in these "'if"' statements can produce steps when included in the cost function for optimization, then gradient based methods like Levenberg-Marquardt fail, but global search methods can be used. Because of this we chose the Metropolis Monte Carlo Markov Chain (MCMC) to optimize the parameters of the phenology module (Knorr and Kattge, 2005). The photosynthesis module can be optimized using the Levenberg-Markquardt algorithm to save computing time. The photosynthesis model was inverted

against latent heat and GPP, optimizing the rooting depth (rd) and the maximum carboxylation rate (Vcmax). The model was fit to the summer period (Julian day 200-280) where strong differences in the fAPAR seasonality were observed, but the influence of the understorey, not included in the model, is rather limited due to the closed canopy of the beech forest. The residuals of the optimized models where compared to explore the consistency of the fAPAR datasets and the seasonality implemented in the model. In a next step Vcmax was estimated for four consecutive data windows of 20 days from Julian day 200 to 280. With this we want to explore if eddy covariance data are informative enough to derive a Vcmax seasonality and assess the influence of the choice of the fAPAR data stream, in spite of the uncertainties inherent to the data.

5.4. Results and discussion

5.4.1. Comparison of fAPAR data streams

The different data streams show systematic differences concerning the absolute level and seasonality of fAPAR. The fAPAR measured at the site is derived from shortwave radiation measurements above and below canopy. These measurements show a strong offset during winter caused by branches and stems that cannot be excluded. To be comparable with the fAPAR computed by the model we scaled the mean winter value to zero assuming that the influence of stems and branches is minimized in summer, where the beech forest has a closed canopy. The second main difference is the seasonality of the fAPAR during summer, while the fAPAR measured at the site and the MODIS fAPAR are constant, the SEAWIFS data shows a strong and the CYCLOPES data a moderate decrease. Possible reasons for the difference between the remote sensing and the site level data are an influence of the surrounding area on the remote sensing data, e.g. a misrepresentation of the remote sensing pixel for the flux tower or a change in reflectance of the canopy due to a change in leaf colour. As the resolution of the SEAWIFS data is 2km compared to a 1km resolution of MODIS and CYCLOPES there could be an influence of the adjacent fields on the SEAWIFS data. The closest change in surface land use to the flux tower is a clearing 800 m in the direction perpendicular to the main wind (Knohl et al., 2003, Fig. 5.2). This wind direction contributes only 5% to the overall wind distribution, the influence on the GPP data is thus likely to be small (Knohl et al., 2003, Fig. 5.2). An early senesence, end of July or early August, of beech leaves and a reduction in photosynthetic activity has been described for the top leaf layer that is fully exposed to sunlight (Schulze, 1970). However the layer directly below is sheltered from the sunlight and has higher photosynthetic capacities and nitrogen contents (Kutsch et al., 2009). Thus while the Cyclopes and SEAWIFS sensor probably sees a too strong decrease due to the senescent top layer, the site data and MODIS miss the effect of the change in leaf colour, reduced nitrogen content and therefore reduced photosynthetic activity.

5.4.2. Optimization of the phenology module of JSBACH

The phenology module of JSBACH was optimized using systematically different data streams, to address the uncertainty in fAPAR data, the optimized model was then used as input for the photosynthesis module. With the use of the different data streams the number of parameters that could be optimized varied. To avoid the large abrupt jump of the phenology output in spring (Fig. 5.1) the seed LAI pa-
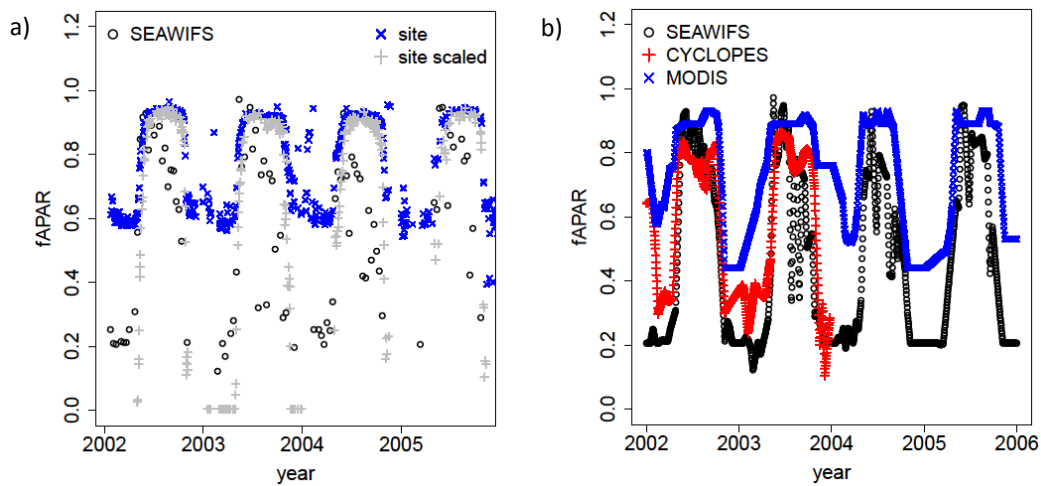
Figure 5.1.: Comparison of different data streams a) site level, scaled site level and SEAWIFS for 2002-2005, b) gapfilled remote sensing data (MODIS, SEAWIFS, Cyclopes) for 2000-2006

Figure 5.2.: Google maps cut out of the Hainich forest flux tower site

rameter (initial LAI value used after the spring event) was reduced from 0.4 to 0.1. The JSBACH prior parameterization is in between the remote and ground based data streams, while the absolute values match better the site observations, the seasonality with the decrease in summer is related to SEAWIFS, but less pronounced. If this decrease in fAPAR is needed, satisfying the model logic and whether it is sufficient to correctly model the carbon uptake needs to be addressed by involving a model of photosynthesis using the different fAPAR data streams as input. For the site level data the model was simplified by removing the shedding phase in summer, the model then consists of only two phases and it has two parameters less, e.g. the growing season length and the shedding rate for the vegetative phase. It was possible to constrain five parameters with the site level data (Tab. 5.1). After the optimization the model reproduces the seasonality of the fAPAR data well (Fig. 5.3). For the maximum LAI the optimized parameter matches the standard parameter (Tab. 5.1, as well as observations at the site (Knohl et al., 2003) increasing the confidence in a reasonable optimization of the model. All 5 parameters are in a similar range as the standard parameters of the JSBACH phenology module, although the model
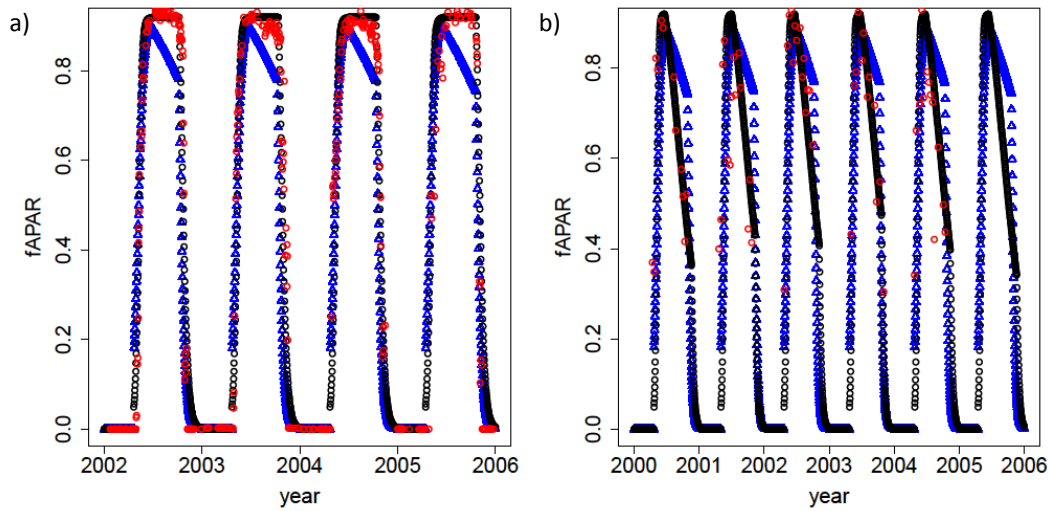
Figure 5.3.: Observed and optimized fAPAR with site data (left) and SEAWIFS (right), red circles show
the observations black shows the optimized model, blue triangles: prior model, standard
JSBACH parameterization

structure was changed to a constant LAI in summer. The strongest deviating optimized parameter is
the growth rate that could be explained with the reduction of the seed LAI. Using the SEAWIFS data
and the original model formulation including the shedding rate during summer only three parameters
could be optimized, including more parameters the MCMC algorithm did not converge. The optimized
parameters deviate stronger from the standard values compared to the optimization with site data (Tab.
5.1). The optimized growth rate increases compared to the standard value, but it is similar to the growth
rate optimized with site level data, indicating that the information on the timing of the spring event
and green up is similar for the two data streams, again the increase can be explained by the reduction
of the seed LAI. The shedding rate is higher and results in a stronger decrease during summer, this in
turn reduces the autumn event temperature as the fAPAR is already strongly reduced during summer.
Although the maximum LAI was not estimated in the optimization the model provides a reasonable fit

Table 5.1.: Optimized parameters of the phenology module and their standard deviation (sd): (r_{growth}:
growth rate, r_{shedv}: shedding rate in the vegetative phase, r_{shedr}: shedding rate in the rest
phase, T_{autumn}:autumn event temperature)

	heat sum range	LAImax	r_{growth}	r_{shedv}	r_{shedr}	T_{autumn}
prior parameters	30	5	0.087	0.004	0.1	10
optimized(site)	28.31	5.00	0.15	–	0.087	9.57
sd(site)	11.7	1.8	0.028	–	0.03	1.41
optimized(SEAWIFS)	prior	prior	0.17	0.011	prior	6.06
sd(SEAWIFS)	–	–	0.01	0.002	–	1.26

(Fig. 5.3), indicating that the information about this parameter is consistent between the data sets.

5.4.3. Uncertainty of Vcmax and rooting depth due to systematic and random errors in GPP and differences in fAPAR data streams

The parameters Vcmax and the rooting depth were estimated using the eddy covariance GPP and latent heat data. The optimization was repeated for the 100 datasets obtained by using a distribution of the u* filter, and the daytime data based estimate of GPP. The parameter Vcmax is well constrained using the different GPP data streams (Fig. 5.4). Subsequently this distribution due to the uncertainty of the u* threshold was combined with the uncertainty due to random errors, yielding a comprehensive uncertainty estimate for the parameter estimates.

The choice of the fAPAR input data has a strong effect on the Vcmax value. The values lie between ca. 20-70, which is within the range of observed leaf level Vcmax values (Kattge et al., 2009). Rayner et al. (2005) use a carbon cycle data assimilation system (CCDAS) based on the BETHY model in combination with fAPAR remote sensing data and atmospheric CO2-concentrations and obtain amongst other parameters a Vcmax value of 34.9 for deciduous temperate forests. The site level fAPAR leads to a lower Vcmax value than SEAWIFS as the fAPAR values are higher in average. The daytime data based estimate generally yields a lower Vcmax value. The uncertainty of the u* threshold causes a parameter uncertainty (variation of the parameter between the different data sets) of ca. half the magnitude as the mean standard error of the parameter for a single dataset for Vcmax. This means that for this site and using the GPP estimates the uncertainty due to the u* threshold is smaller than the uncertainty due to random errors. This uncertainty due to the random error depends on the number of data points used, more data would decrease the standard error of the parameters. Moreover the uncertainty of the u* threshold could be more important when using NEE observations or Reco estimates in the optimization, as usually the nighttime data are more affected by low turbulent conditions. The importance of the u* threshold should vary between sites according to the uncertainty of the u* threshold vs the noise level of the data. The standard nighttime as well as the daytime data based estimate are at the lower edge of the distribution for all three fAPAR data streams. The uncertainty of the parameter increases with the parameter value, this variation is rather not an effect of a better constraint of one of the data streams, but reflects the model structure (fAPAR is multiplied). The influence of the choice of fAPAR data clearly overrides the influence of the different GPP datasets with respect to the parameter values. The combined uncertainty of random and systematic errors increases only slightly compared to the uncertainty of the random error only (Fig. 5.4, Tab. 5.2). Indicating that biases in GPP eddy covariance data are for this site negligible compared to errors occurring in fAPAR data. The differences of the Vcmax parameter is in this case mostly caused by the different seasonality as the maximum values are similar, but the decrease in the SEAWIFS data leads to a higher Vcmax estimate.

The rooting depth parameter is also well constrained (see Fig. 5.5). The distribution of the rooting depth when using the standard JSBACH parameterizations overlaps with the site level observations. The difference between the three different fAPAR data streams is comparably small. The daytime data based estimate yields a higher estimate for rd and the difference between nighttime and daytime based estimate is much higher than for the Vcmax, e.g. the estimate based on the daytime data based GPP is outside the distribution of the 100 differently filtered datasets. The uncertainty of the u* filtering does not increase the overall uncertainty of the rooting depth when combined with the uncertainty due to the

Optimization of a land surface model using fAPAR and eddy covariance data: impact and
propagation of uncertainties due to systematic errors

64
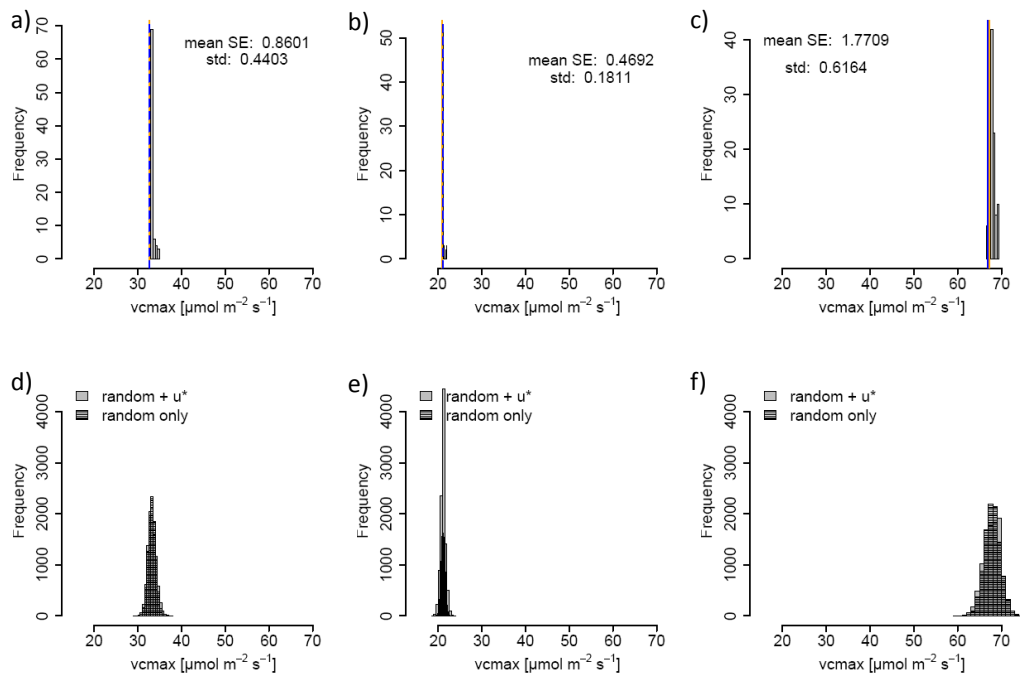
Figure 5.4.: Estimates for Vcmax using the different fAPAR input: JSBACH standard parameterization (first column), site data (second column), SEAWIFS satellite data (third column). The orange vertical line indicates the estimate based on the daytime based flux partitioning, the blue line is based on the standard nighttime based GPP and the histogram is derived from bootstrapping of the u*-threshold (first row). The second row shows the uncertainty of the parameter due to only the random error and the combined uncertainty of random error and u* uncertainty.

random error (Fig. 5.5, Tab. 5.2).

The differences in average GPP for the summer are much smaller than the difference in the Vcmax parameter, e.g. the distributions overlap (Fig. 5.6). This means that the parameter sets need to be tuned according to the fAPAR data stream that was used for the parameterization of the phenology module to model the average GPP right.

Table 5.2.: Standard deviation of the parameters propagating the uncertainties due to random error (r) and due to random and systematic error (r+s).

	Vcmax r	Vcmax r+s	rd r	rd r+s
prior	0.860	0.983	0.0086	0.0085
posterior site	0.469	0.537	0.0088	0.0087
posterior SEAWIFS	1.771	1.937	0.0083	0.0083
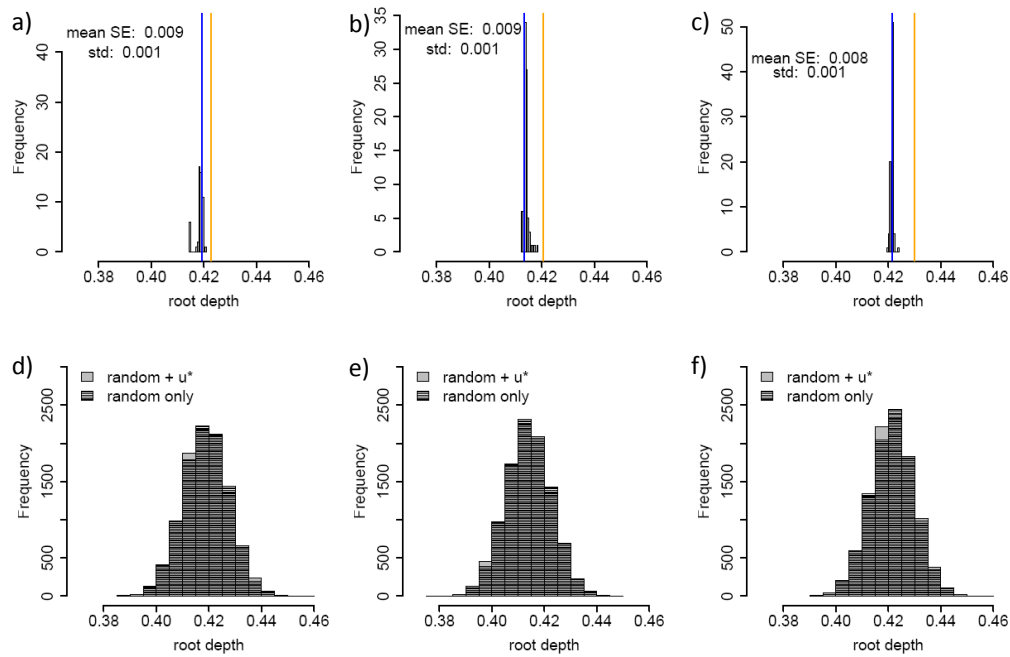
Figure 5.5.: Estimates for the rooting depth using the different fAPAR input a) JSBACH standard parameterization b) site data c) SEAWIFS satellite data. The orange vertical line indicates the estimate based on the daytime based flux partitioning, the blue line is based on the standard nighttime based GPP and the histogram is derived from bootstrapping of the u*-threshold.

5.4.4. Seasonality of GPP model residuals with constant Vcmax

Keeping the Vcmax parameter constant for the whole summer the influence of the different fAPAR seasonalities get obvious in the time series of the residuals. A temporal trend in the residuals indicates model structural errors, e.g. a missing seasonality of Vcmax or inconsistencies between the model structure and the data. According to leaf level observations Vcmax should decrease in the end of summer (Muraoka et al., 2010), thus with a constant Vcmax an overestimation of model in the end of summer could be expected. Using the JSBACH standard parameterization the residuals show no trend in time, while using the site fAPAR GPP is first underestimated then overestimated by the model, the strongest correlation of the residuals is found for the SEAWIFS fAPAR, here the model first overestimates and then underestimates the flux (Fig. 5.7). The JSBACH standard parameterization has no correlation in time of the residuals and seems to cover the overall seasonality of fAPAR and Vcmax best, thus the moderate decrease of fAPAR seems to already include a potential seasonality of Vcmax. The small positive trend when using the site fAPAR could be removed by introducing a seasonality of Vcmax. The strong correlation of the residuals using the SEAWIFS fAPAR indicates an inconsistency of the remote sensing data with the flux tower data, the decay of fAPAR seems to be too strong, an additional decrease in Vcmax would even increase the trend in the residuals. As mentioned before the less pronounced decrease in the MODIS and Cyclopes fAPAR seasonality having a higher resolution (Fig. 5.1) indicate that the inconsistency could be possibly due to the influence of a close by clearing.

Overall the results suggest, that the site level fAPAR (and MODIS) observations represent LAI, while the "green" fAPAR of CYCLOPES and SEAWIFS already include a seasonality of Vcmax, that is for
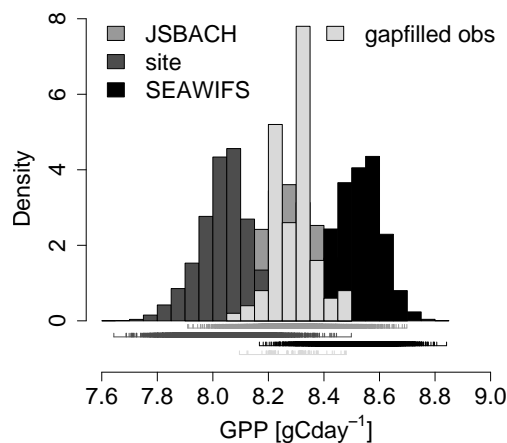
Figure 5.6.: Distribution of average GPP during the summer period using the three fAPAR streams. The orange vertical line indicates the estimate based on the daytime based flux partitioning, the blue line is based on the standard nighttime based GPP. The histogram is derived from optimization against the 100 GPP and LE datasets and a Monte Carlo simulation to include the uncertainty due to the random error.

SEAWIFS possibly too strong due to the influence of an adjacent field. Whether the decrease in fAPAR is caused by a change in leaf colour of the forest or due to the adjacent fields could be adressed by using a radiation transfer model with reflectance data of different resolution. The too strong decrease could also be caused by an early start of senescence of the top leaf layer of the beech forest, while the activity of the shaded leaves directly below remains high.

5.4.5. Seasonal variation of Vcmax

To explore the seasonality of Vcmax the parameters Vcmax and rd were optimized using data periods of 20 days from Julian day 200 to 280. The seasonality of the parameters using the site level and JSBACH standard parameterization are similar with the main difference of the estimates using the site level data being lower, both show a decrease mainly for the last data window. The estimates based on the SEAWIFS fAPAR show a quite strong increase during early summer and then remain rather constant. The seasonality is lowest for the site level fAPAR, as the parameter value is lowest but then multiplied with the highest fAPAR, thus a change in the parameter value causes a larger change in GPP. Comparing these seasonalities with observations of Vcmax on leaf level, the strong increase using the SEAWIFS fAPAR seems unrealistic (Muraoka et al., 2010). The small parameter uncertainties derived from the fit to the GPP data show that a derivation of the Vcmax seasonality could be possible using eddy covariance data, but the confidence in the fAPAR data streams is too low to be able to falsify the model structures of the phenology module.
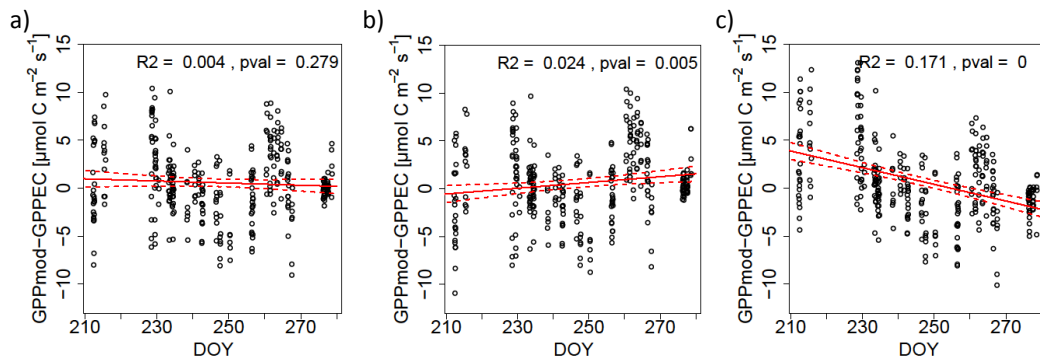
Figure 5.7.: Time series of the residuals with a polynomial fit for a) the standard JSBACH parameterization b) site level fAPAR c) SEAWIFS fAPAR.
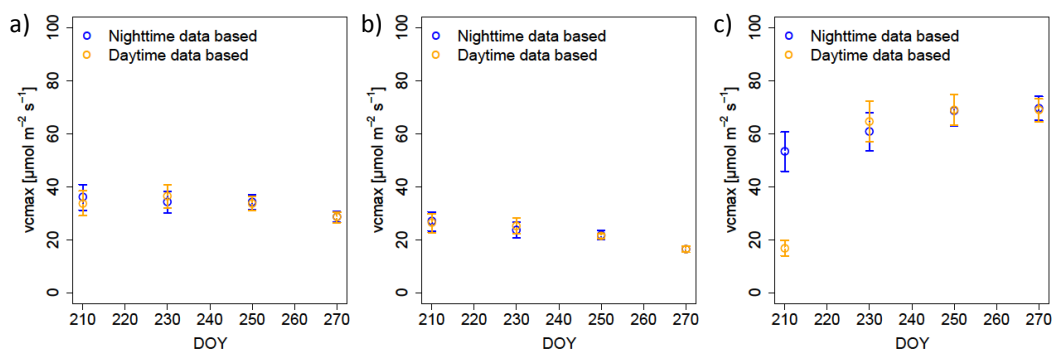
Figure 5.8.: Time series of the parameter Vcmax for a) the standard JSBACH parameterization b) site level fAPAR c) SEAWIFS fAPAR.

5.5. Conclusion

We provide a methodology to include the uncertainty due to the u* threshold on model-data synthesis studies. The potential bias due to the uncertainty in the u* threshold used for filtering can be propagated by creating a likely distribution of u* thresholds using a bootstrapping method. The uncertainty propagated from the u* distribution to the Vcmax parameter is smaller than the standard error of the parameter due to noise in the data, thus the influence of the uncertainty of the u* threshold on the GPP estimates for the Hainich flux tower is only small. The uncertainty of the u* threshold influences the parameters differently, e.g. it increases the uncertainty of Vcmax while the rooting depth is not affected. The choice of the fAPAR data stream influences the optimized Vcmax value much more than the errors and uncertainties in the eddy covariance data. The choice of the fAPAR data stream has a strong influence on the absolute value and the derived seasonality of Vcmax. The strong trend in the residuals when using a constant Vcmax together with the strong seasonality when allowing Vcmax to vary when using the SEAWIFS fAPAR suggests a inconsistency between the remote sensing pixel and the flux tower footprint. Another possible explanation for a too strong decrease of the remote sensing data is the early senescence of the top layer of beech trees while the layer below has still a high activity. The knowledge about the early senescence on the other hand suggests a decrease in fAPAR that is not observed by the in situ data as they actually represent the intercepted PAR. The seasonality of GPP can be either im-

plemented through a variation in fAPAR or in Vcmax. However, to be able to make use of the fAPAR data in model data fusion the model structure needs to be consistent with the fAPAR and the flux data streams. If the fAPAR is constant during summer, as the in situ data suggest, an additional seasonality of Vcmax is needed. Limiting for the derivation of a seasonality is the strong difference between the fAPAR data streams. Nevertheless the remote sensing data provide an important source of information in the spatial domain. The optimization of the phenology module showed, that the information about the spring event, green up and maximum LAI was consistent between the site level observations and the remote sensing data. Deriving such properties from the data set to constrain models in the spatial domain may be more promising than using the time series for one pixel.

6. General discussion and outlook

6.1. Main findings

This thesis has contributed to improve the understanding of the carbon cycle on ecosystem level by the combined use of eddy covariance data with carbon cycle models. The studies are based on the FLUXNET database that assembles eddy covariance data from stations around the world measuring the carbon, water and energy exchange between vegetation and atmosphere. The data were used in combination with carbon cycle models of varying complexity. The main findings of the study address the derivation of and relation between photosynthesis and respiration based on the FLUXNET database and the improvement of model data fusion including a comprehensive representation of uncertainties in eddy covariance data.

Derivation of and relation between photosynthesis and respiration The derivation of the flux components photosynthesis (GPP) and respiration (Reco) from the observed net exchange (NEE) is important to be able to interpret the observed variability of the net carbon flux and attribute it to the underlying process. For this purpose a variety of flux partitioning algorithms are available and have been compared for selected sites (Yi et al., 2004; Reichstein et al., 2005; Stoy et al., 2006; Desai et al., 2008). There are two common approaches to partition the net exchange: the first relies on nighttime data, that are known to be problematic due to low turbulence in the nighttime (Goulden et al., 1996; Feigenwinter et al., 2004; Aubinet, 2008), the second is based on daytime data but usually neglects either the influence of water vapour pressure deficit (VPD) or the influence of temperature, or both.

Chapter 2 describes a flux partitioning algorithm based on mainly daytime eddy covariance data. A light response curve was extended to account for both, the temperature sensitivity of respiration and a VPD limitation of GPP. It shows the importance of including a VPD limitation to not loose the influence of drought effects due to stomatal control in the GPP signal. Moreover the study shows that for sites with influence of VPD on the diurnal cycle, the VPD effect should be included to retrieve unbiased results of respiration. This daytime data based estimate is compared with a nighttime data based estimate (Reichstein et al., 2005) for the whole FLUXNET data satisfying certain quality criteria. The difference between the two is due to methodological differences, e.g. model assumptions, but also because daytime and nighttime eddy covariance data are affected by potential systematic errors in a different way (Goulden et al., 1996; Aubinet et al., 2000; Baldocchi, 2003). Thus the difference between the two can be interpreted as an estimate of the uncertainty due to systematic errors in GPP and Reco estimates. On annual timescale no significant difference between night and daytime based estimate was found across FLUXNET sites. The good agreement between nighttime and daytime NEE based flux partitioning methods increases the confidence in the flux measurements observed with the eddy covariance method and the derived flux components, photosynthesis and respiration. The comparison of the differences between the two estimates compared to uncertainty estimates derived from parameter uncertainties of the flux partitioning model support the necessity of accounting for

systematic errors in model data fusion studies. The strong correlation found between GPP and Reco (Reichstein et al., 2007; Wang et al., 2008; Baldocchi, 2008) was so far suspected to be partly due to a statistical artifact (spurious correlation) as they were ultimately derived from the same NEE data (Vickers et al., 2009). The availability of two quasi independent estimates allowed to revisit this pattern and the strong correlation could be confirmed. Moreover the daytime data based estimate has already contributed to a number of studies providing a basis to account for systematic errors by including both estimates (Beer et al., 2010; Rebmann et al., 2010; Merbold et al., 2010), to minimize spurious correlation between the two flux components (Migliavacca et al., 2010; Mahecha et al., 2010), or the comparison of the two estimates can help to select sites with the most reliable observations to be included in studies (Jung et al., 2009; Beer et al., 2010).

This spurious correlation depends on the ratio of the variances of NEE and the error in Reco. If the whole variance of Reco is used as proposed by Vickers et al. (2009) the estimate of spurious correlation would suggest that the real underlying correlation is close to zero. Using artificial data it is possible to show that this approach fails to identify an existing real correlation between Reco and GPP (see chapter 3). The combined use of the nighttime data based Reco and the daytime data based GPP can minimize the spurious correlation as they are derived from disjoint datasets and errors in Reco cannot directly propagate into the GPP estimate. For selected sites we investigate the effect of the spurious correlation when GPP is derived as a residual (see Appendix B). On different time scales the influence of a spurious correlation was found to be rather small. The correlation between GPP and Reco is not a statistical artifact but a real phenomenon. Nevertheless the results using the eddy covariance estimates of GPP and Reco only show that the two flux components are correlated, but the causality of the correlation needs to be further elucidated by process studies, e.g. it can on the one hand be largely due to a background correlation of drivers or on the other a real link between the processes as the products of assimilation feeds respiratory processes.

Model data fusion for eddy covariance water and carbon fluxes   A careful quantification and characterization of data uncertainties are of major importance for a successful model data fusion (MDF). Chapter 4 provides an analysis of random errors of eddy covariance data, describing the main statistical characteristics important for not only model data fusion but also many other statistical analysis. The study confirms the varying standard deviation found in previous studies (Hollinger and Richardson, 2005; Richardson et al., 2006b). New findings are a low autocorrelation of the random error and a very small crosscorrelation between the random errors of carbon and water fluxes. The distribution of the random error has been investigated as well in previous studies (Hollinger and Richardson, 2005; Richardson et al., 2006b), the overall distribution was found to be double exponential or even more peaked. The finding that the distribution of fluxes pooled by magnitude becomes more Gaussian (Richardson et al., 2008) suggests that the double exponential distribution actually is a superposition of Gaussian distributions with varying standard deviation. The analysis included here suggests a normalization approach to investigate the error distribution without the confounding effect of the varying standard deviation. The Gaussian distribution found for the European sites included in the study cannot be confirmed for all FLUXNET sites (Richardson et al. in prep), probably due to different filtering of the principal investigators of the site as the distribution is often largely influenced by few outliers (Richardson et al., 2008).

But even if the overall distribution is double exponential, the fluxes of high magnitude tend to be approximately Gaussian and the more peaked distributions are found for small fluxes (Richardson et al. in prep.). The description of the random error is important for the choice of statistical methods, in case of parameter optimization it influences the choice of the cost function. Heteroscedasticity requires weights to be included in the cost function. Gaussian distributions require a least squares optimization while a Laplacian distribution requires the minimization of absolute deviations. The strong variability between sites concerning the distributions suggests to analyse the distribution for the specific dataset used and then choose between least squares, absolute deviations or if the distribution is in between trimmed least squares can be an option. Using the tools described in chapter 4 the random error can be characterized for specific sites and support decisions on the statistical method applied.

The effect of errors was illustrated in chapter 4 by the use of sythetic data based on a model output and addition of different types of errors, random and systematic. The effect of random errors on parameter estimates differs from systematic errors: while random errors increase the parameter uncertainty the systematic error biases the estimate. Using standard methods for uncertainty propagation the true parameter is within the uncertainty range for the presence of random errors, for systematic errors the uncertainty range does not include the true value. The strength of the bias differs between parameters and is reduced if the model is inverted against multiple data sources. The results suggest, that systematic errors need to be included to derive reasonable estimates of uncertainty for the model output (see Fig. 6.1). This is further important for the interpretation of inconsistencies between model and data as they might be misinterpreted as model structural errors.

Uncertainties due to systematic errors have been hardly addressed when propagating uncertainties, one example found in literature addresses random and systematic errors for risk assessment. Vasquez and Whiting (2006) obtain the probability distributions for model output variables by performing Monte Carlo simulation coupled with appropriate definitions for the random and systematic errors. For eddy covariance data the choice of the u*-threshold, that is used to filter the data for unfavourable measurement conditions has been shown to dominate uncertainties in annual estimates of NEE (Papale et al., 2006; Gove and Hollinger, 2006), this uncertainty has not been consistently included in model-data fusion studies yet. In chapter 5 the uncertainty of this threshold is included using a bootstrapping approach to create data ensembles that are used in the optimization of the photosynthesis module of the land surface model JSBACH. They are complemented by an additional GPP dataset that was created based on daytime eddy covariance measurements to be independent of the rather problematic nighttime observations. We combine this uncertainty due to potential biases with the uncertainty due to random errors to provide a comprehensive estimate of uncertainty for the parameters and the model output (see Fig. 6.1). In this case the uncertainty of the u* - threshold was of minor importance, the importance is expected to increase if NEE or Reco are used instead of GPP and for sites that are affected by advection. In this study the effect of the errors in eddy covariance data were rather small compared to the differences caused by multiple data streams, observed in situ and by satellite, of the fraction of absorbed photosynthetically active radiation (fAPAR). The standard inverse problem theory needs to be adjusted to the specific study due to errors in both, model and data. The characterization of errors is specific to each data stream. Random and systematic errors should be addressed in a different way, random errors can be included in the cost function in terms of the covariance matrix while uncertainties due to systematic errors are more difficult to represent. The use of multiple data sets (data errors) or

	Data characterization		Implementation in MDF
Chap. 4	**Random error:** Standard deviation Distribution (Correlations	→ → →	error covariance matrix distance measure error covariance matrix)
Chap. 2 & 5	**Systematic error:** Quantify the uncertainty: Plausible range of values	→	data ensembles

Chap. 5	**Combination** of random and systematic uncertainty sources:
	1) Create parameter distribution for each fit to the data
	     ensembles (uncertainty due to random errors)
	2) Superimposition of the distributions of the whole data
	     ensemble		

Figure 6.1.: Treatment of eddy covariance data in model data fusion (MDF) studies

multiple model structures (model errors, see Carvalhais et al. (2008)) allows to include also these errors in the uncertainty propagation. Overall the thesis suggests an approach to quantify uncertainties due to random (chapter 4) and systematic errors (chapter 2) for eddy covariance data and to include them in the MDF procedure (chapter 5) for eddy covariance data (see Fig. 6.1).

6.2. Future directions of model data fusion

To be able to improve models, inconsistencies between models and data need to be attributed to either errors in observations or errors in model structure or the operator between model and data. Pivotal to this is an accurate estimate of potential systematic and random errors for each data stream as well as a good representation of the errors in the model-data fusion procedure. The uncertainties of eddy covariance data have been addressed and quantified as well as incorporated into a model data fusion framework. The errors in remote sensing products need to be further addressed to resolve whether the inconsistencies between data sources are due to errors, that can be removed or if the data streams represent different properties that are not directly comparable. To be able to better understand differences between the data streams on the one hand using the reflectance data directly on the other hand the use of data with high spatial resolution can help. For a full propagation of uncertainties in remote sensing data, the direct use of reflectance data and the coupling of radiation transfer models to land surface model could be an improvement.

Using only eddy covariance data the number of parameters that can be constrained by inverse parameter estimation is limited (Knorr and Kattge, 2005; Luo et al., 2009). Thus including more information either in the form of another data stream to be included in the cost function or by adding prior information on the parameters is necessary. When integrating multiple data streams with a very different quantity of data points the data stream with the fewer data points may have almost no influence on the cost function and therefore on the optimization. One approach to deal with this is to include the number

of observations in the cost function (Carvalhais et al., 2010). On the other hand it seems intuitive that the redundancy or information content of the data should play a role, e.g. the repeated diurnal cycle of the eddy covariance flux data often does not contain new information for the model if the system did not change. Here the analysis of the correlation structure of model residuals could be a starting point to find more general solutions. If the model is optimized the residuals should be uncorrelated, a higher correlation in the residuals for one data stream might be reduced by increasing the weight of this data stream in the cost function, keeping in mind that systematic errors in data, model or operator between data and model also cause correlations. The prior uncertainty of parameters in Bayesian approaches have often been set to arbitrary values, e.g. a percentage of the parameter value (Knorr and Kattge, 2005). A database of plant functional traits has been established and a first database version has been launched (www.try-db.org). This database contains values of parameters used in ecosystem models, this information can be used to provide more objective but also more detailed prior information about parameters in model inversions. Including this information on parameters helps to successfully optimize the model and contributes to the goal of obtaining a model that is consistent with all the available information.

When evaluating the fit of the model to multiple observation streams the strength of each data stream should be emphasized. While remote sensing data are known to be strong in constraining the spatial patterns, eddy covariance data contain excellent information about the temporal dynamics, and observations of the carbon stocks, being the integral of the fluxes, can add information on the long term carbon cycle. In this context geostatistical analysis, time scale separation algorithms or dimensionality reduction methods of the data streams could offer interesting solutions, by preprocessing the data to extract the information we have the highest confidence in, before matching the model to it.

One example of such a preprocessing is the time scale separation of data streams. It has been shown that relations between variables vary between different time scales (Mahecha et al., 2007). One reason for these differences is that processes act on different time scales, e.g. carbon uptake responds to radiation instantaneously while the leaf area index does not change much from one day to the next. Model structural errors can then be addressed by approaches using frequency decomposition techniques if confounding factors act on distinct time scales. This concept of scale dependent parameter estimation (SCAPE) has been applied using the Q10 respiration model that models the respiration based on temperature only. By separating the time scales confounding factors acting on larger time scales, in this case influence of humidity and phenology, can be excluded and the short term temperature sensitivity can be extracted (Mahecha et al., 2010). The problem of correlation between influencing factors can be found quite often in biogeochemical modeling, the SCAPE approach can help to distinguish between these factors and exclude confounding factors or model structural errors to better constrain a specific submodule. The better defined short term variation derived with SCAPE also allows one to reinvestigate the long term variation by excluding the direct response of respiration to temperature on these time scales. Other ways of preprocessing the data should be explored to extract meaningful information from data that might seem biased or too noisy for a direct use in data assimilation. With respect to remote sensing fields this could be approached by using standardization of the data if the absolute values are less trusted than the spatial patterns or the overall seasonality. For the seasonality the data would be normalized for each pixel, for the spatial fields the data would need to be normalized for each time step. Of course also here a careful quality assessment of the data is a prerequisite.

One drawback of the eddy covariance flux data in global model data fusion studies is the only punctual availability of observations. There is now an ensemble of methods available to upscale the information from flux towers to the globe (Beer et al., 2010), providing maps consistent with the spatial scale of models. The use of these global fields in data assimilation must be seen critical as the upscaling already involves assumptions, e.g. they are not observations, but already a model output. On the other hand they are much closer to the data than process based models and the empirical models provide a filter potentially removing the between site noise of the FLUXNET dataset due to site specific characteristics. Provided a critical evaluation of the datasets the assimilation of these global fields can help to evaluate and constrain the models with respect to global patterns. With the increasing availability of data not only the match of one or two target variables can be improved, but also the underlying logical functional relationships of the model can be validated, e.g. is the response of the target variable to the drivers the same in observations and model runs.

7. List of Publications

The following publications were prepared during the work of this thesis.

Publications

- **Lasslop, G.**, Reichstein, M., Kattge, J., Papale, D. (2008) Influences of observation errors in eddy flux data on inverse model parameter estimation. Biogeosciences, 5, 1311-1324.

- **Lasslop, G.**, Reichstein, M., Papale, D., Richardson, A.D., Arneth, A., Barr, A., Stoy, P., Wohlfahrt, G., (2010) Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: critical issues and global evaluation. Global Change Biology, 16, 187–208, doi:10.1111/j.1365-2486.2009.02041.x.

- **Lasslop, G.**, Reichstein, M., Detto, M., Richardson, A. D., Baldocchi, D.D., Comment on Vickers et al.: Self-correlation between assimilation and respiration resulting from flux partitioning of eddy-covariance CO2 fluxes. Agric. Forest Meteorol. (2010), 150, 312–314, doi:10.1016/j.agrformet.2009.11.003

- Rebmann, C., Zeri, M., **Lasslop, G.**, Mund, M., Kolle, O., Schulze, E.-D. and Feigenwinter, C. (2010) Treatment and assessment of the CO2-exchange at a complex forest site in Thuringia, Germany. Agricultural and Forest Meteorology, 150,5,684-691 .

- Migliavacca, M., Reichstein, M., Richardson, A.D., Colombo, R., Sutton, M., **Lasslop, G.**, Wohlfahrt, G., Tomelleri, E., Carvalhais, N., Cescatti, A., Mahecha, M., Montagnani, L., Papale, D., Zaehle, S., Arain, A., Arneth, A., Black, A.T., Carrara,A., Dore, S., Gianelle, D., Helfter, C., Hollinger, D., Kutsch, W., Lafleur, P.M., Nouvellon, Y., Rebmann, C., da Rocha, H.R., Rodeghiero, M., Roupsard, O., Sebastia, M.-T., Seufert, G., Soussana, J.-F. and van der Molen, M.K. (2010, accepted) Semi-empirical modeling of abiotic and biotic factors controlling ecosystem respiration across eddy covariance sites. Global Change Biology.

- Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., Rodenbeck, C., Arain,M. A., Baldocchi, D., Bondeau, G. B. B. A., Cescatti, A., **Lasslop, G.**, Lindroth, A., Lomas, M., Luyssaert, S., Margolis, H., Oleson, K. W., Roupsard, O., Veenendaal, E., Viovy, N., Williams, C., Woodward, F. I., and Papale, D. (2010): Terrestrial Gross Carbon Dioxide Uptake: Global Distribution and Covariation with Climate, Science, 10.1126/science.1184984.

- Mahecha, M. D., Reichstein, M., Carvalhais, N., **Lasslop, G.**, Lange, H., Seneviratne, S. I., Vargas, R., Ammann, C., Arain, M. A., Cescatti, A., Janssens, I. A., Migliavacca, M., Montagnani, L., and Richardson, A. D. (2010): Global Convergence in the Temperature Sensitivity of Respiration at Ecosystem Level, Science, 10.1126/science.1189587.

- **Lasslop, G.**, Reichstein, M., Kattge, J., Tomelleri, E., Zähle, S., Reick, C., (in prep.): Optimization of a land surface model using fAPAR and eddy covariance data: impact and propagation of uncertainties due to systematic errors.

- Reichstein et al. (including **Lasslop, G.**),(in prep.): A new global harmonized eddy covariance data set from FLUXNET: uncertainties, limitations and robust global patterns.

Book chapters

- Reichstein, M., Stoy, P., Desai, A., **Lasslop, G.**, A. Richardson (in prep.) Partitioning net ecosystem carbon exchange into canopy assimilation and ecosystem respiration: basic concepts and future directions, in: Eddy Covariance: A Practical Guide to Measurement and Data Analysis (eds. Aubinet, Vesala and Papale)

- Richardson, A.D., Aubinet, M., Barr, A.G., Hollinger, D.Y., Ibrom, A., **Lasslop, G.**, and Reichstein, M. (in prep.) Uncertainties in surface atmosphere fluxes measured by eddy covariance, in: Eddy Covariance: A Practical Guide to Measurement and Data Analysis (eds. Aubinet, Vesala and Papale)

Conference contributions

- 5th Annual CarboEurope-IP Meeting,8-12 October 2007, Poznan/Poland: **Lasslop, G.**, Reichstein, M.: Influences of observation errors in eddy flux data on inverse model parameter estimation. Poster.

- EGU 2008, Wien: Lasslop, G., Reichstein, M.: Half hourly GPP and Reco estimates derived from a modified light response curve. Oral presentation.

- CarboEurope meeting, 2009, Jena: **Lasslop, G.**, M. Reichstein, D. Papale, A. Richardson, G. Wohlfahrt: Estimation of carbon flux components from eddy covariance data - a modified light response curve approach applied to all FLUXNET data and compared to the standard processing. Oral presentation

- ICDC8, Jena 2009: **Lasslop, G.**, Reichstein, M., Papale, D., Richardson, A.D., Arneth, A., Barr, A., Stoy, P. and Wohlfahrt, G.: Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: critical issues and global evaluation. Poster.

- AGU 2009, San Francisco: **Lasslop, G.**, Reichstein, M., D. Papale, A. Richardson: Derivation and analysis of cross relations of photosynthesis and respiration across FLUXNET sites for model improvement. Oral presentation.

A. Appendix for chapter 3

A.1. Algorithm

The fit parameters of our models showed considerable seasonality to accommodate processes not included in the model. To account for this we use a moving window approach; parameters are estimated every two days with a four day moving window when using the daytime data and a twelve day moving window when using the nighttime data. Only the temperature sensitivity of the Lloyd and Taylor (1994) model is estimated with the nighttime data. The response of NEE measurements to temperature is assumed to be independent of systematic measurement errors during periods of poorly-developed turbulence during nighttime. The parameter determining the magnitude of the respiration, rb, is estimated from daytime data and is thus independent of such biases. The primary support for stable annual estimates of GPP and Reco is a high number of estimates or small parameter gaps throughout the year. The settings for the parameters during the estimation procedure are summarized in Table A.1.

The estimation was sensitive to the chosen initial guess value of β the maximum carbon uptake rate, in the gradient-based parameter estimation routine. To deal with this problem we estimate the parameters three times, changing the initial guess value given in Table A.1 to the half and double value for the second and third estimation. The parameter set with the lowest RMSE was then selected. The parameters were only accepted if they were within a reasonable range (ranges for each parameter are listed in Table A.1). If parameters were outside the range certain parameters were fixed to values defined in the last column of Table A.1 and the others were re-estimated. Fluxes were computed for the two neighboring parameter sets and then linearly interpolated using the reciprocal of the distance to the parameter sets as weight.

Table A.1.: Settings for the parameters during the estimation procedure. If all parameter estimates meet the criteria listed in table, the estimate is accepted. If at least one is outside the pre-defined range, the value is set according to the last column and all other parameters for that time-window are re-estimated or the parameter set is not used (see also last column).

Parameter	Initial guess	Accepted values	If the parameter estimate is not acceptable
E0	100	50-400	Set to value of previous window, if no previous window exists estimates <50 were set to 50, estimates >400 were set to 400
rb	Mean of nighttime NEE	>0	Whole parameter set is not used
α	0.01	>=0, <0.22	Set to value of previous window, if no previous window exists and less than 0, set to zero
β	Abs(0.03quantile - 0.97quantile) of NEE	>=0, <250 If β>100 then $\sigma(\beta) < \beta$	If negative set to zero, else the whole parameter set is not used
k	0	>=0	Set to zero

A.2. Sites

A list of FLUXNET sites used in the global comparison is given in Table 2.1.

Table A.2.: List of FLUXNET sites used in the global comparison

Site	Years	Vegetation type	References
AT-Neu	2003, 2006	GRA	Wohlfahrt et al. (2008b)
AU-Fog	2006	WET	-
AU-Tum	2002, 2003	EBF	van Gorsel et al. (2008)
AU-Wac	2006	EBF	Wood et al. (2008)
BE-Bra	2000, 2002, 2006	MF	Carrara et al. (2004)
BE-Lon	2005, 2006	CRO	Moureaux et al. (2006)
BE-Vie	97-03, 05-06	MF	Aubinet et al. (2001)
BR-Ban	2004	EBF	-
BR-Ji2	2001	GRA	Kruijt et al. (2004)
BR-Ma2	2004,05	EBF	Araujo et al. (2002)
BR-Sa1	2002-04	EBF	Saleska et al. (2003)
BR-Sa2	2001	CRO	Sakai et al. (2004)
BR-Sa3	2001-03	EBF	Goulden et al. (2004)
BR-Sp1	2001	WSA	Santos et al. (2004)
BW-Ma1	2000	WSA	Veenendaal et al. (2004)
CA-Ca1	1998-2005	ENF	Humphreys et al. (2006)
CA-Ca2	2001-2005	ENF	Humphreys et al. (2006)
CA-Ca3	2002-2005	ENF	Humphreys et al. (2006)
CA-Let	1999-2005	GRA	Flanagan et al. (2002)
CA-Man	1995,98,2000,01	ENF	Dunn et al. (2007)
CA-Mer	1999-2005	OSH	Lafleur et al. (2003)
CA-NS1	2003	ENF	Goulden et al. (2006)
CA-NS3	2002-2004	ENF	Goulden et al. (2006)
CA-NS6	2002-04	OSH	Goulden et al. (2006)
CA-NS7	2003-04	OSH	Goulden et al. (2006)
CA-Oas	1997-2005	DBF	Black et al. (2000)
CA-Obs	2000-05	ENF	Bergeron et al. (2007)
CA-Ojp	2000-05	ENF	Howard et al. (2004)
CA-Qcu	2002-06	ENF	Giasson et al. (2006)
CA-Qfo	2004-06	ENF	Bergeron et al. (2007)
CA-SJ1	2003	ENF	Zha et al. (2009)
CA-SJ2	2005	ENF	Zha et al. (2009)
CA-SJ3	2005	ENF	Zha et al. (2009)
CA-TP4	2004-05	ENF	Arain and Restrepo-Coupe (2005)

Table 1 continued

Site	Years	Vegetation type	References
CA-WP1	2004-05	MF	Syed et al. (2006)
CH-Oe1	2003-06	GRA	Ammann et al. (2007)
CH-Oe2	2005	CRO	Dietiker et al. (submitted)
CN-Cha	2003	MF	Guan et al. (2006)
CN-Do1	2005	WET	Yan et al. (2008)
CN-Do2	2005	WET	Yan et al. (2008)
CN-Do3	2005	WET	Yan et al. (2008)
CN-HaM	2002-03	GRA	Kato et al. (2006)
CN-Xfs	2005	GRA	-
DE-Bay	1998-99	ENF	Staudt and Foken (2007)
DE-Geb	2004-06	CRO	Anthoni et al. (2004)
DE-Gri	2006	GRA	Gilmanov et al. (2007)
DE-Hai	2000-05	DBF	Knohl et al. (2003)
DE-Har	2005-06	ENF	Schindler et al. (2006)
DE-Kli	2005-06	CRO	-
DE-Meh	2004-06	MF	Don et al. (2009)
DE-Tha	1997-2006	ENF	Grünwald and Bernhofer (2007)
DE-Wet	2003-06	ENF	Rebmann et al. (2010)
DK-Lva	2005	GRA	Gilmanov et al. (2007)
DK-Ris	2004-05	CRO	Houborg and Soegaard (2004)
DK-Sor	1997-2002,05-06	DBF	Pilegaard et al. (2003)
ES-ES1	99-00,03,05-06	ENF	Reichstein et al. (2005)
ES-ES2	2006	CRO	-
ES-Lma	2004,05	SAV	-
ES-VDA	2004	GRA	Gilmanov et al. (2007)
FI-Hyy	1997-99,2001-04	ENF	Suni et al. (2003b)
FI-Sii	2005	GRA	Aurela et al. (2007)
FI-Sod	2001,05-06	ENF	Suni et al. (2003a)
FR-Fon	2006	DBF	-
FR-Gri	2005	CRO	Hibbard et al. (2005)
FR-Hes	1998-99,01-06	DBF	Granier et al. (2000)
FR-LBr	1998	ENF	Berbigier et al. (2001)
FR-Lq1	2004-05	GRA	Gilmanov et al. (2007)
FR-Lq2	2004	GRA	Gilmanov et al. (2007)
FR-Pue	2001-06	EBF	Rambal et al. (2004)
GF-Guy	2004,06	EBF	Bonal et al. (2008)
HU-Bug	2006	GRA	Nagy et al. (2007)
HU-Mat	2006	GRA	Pinter et al. (2008)
ID-Pag	2002	EBF	Hirano et al. (2007)

continued

Table 1 continued

Site	Years	Vegetation type	References
IE-Dri	2003-04	GRA	-
IL-Yat	2004-06	ENF	Grünzweig et al. (2003)
IT-Amp	2003	GRA	Gilmanov et al. (2007)
IT-Bci	2005,06	CRO	Reichstein et al. (2003a)
IT-Cpz	1997,03,04,06	EBF	Garbulsky et al. (2008)
IT-Lav	2001,04,06	ENF	Marcolla et al. (2003)
IT-Lec	2006	EBF	-
IT-Mbo	2003-06	GRA	Marcolla and Cescatti (2005)
IT-PT1	2003	EBF	Migliavacca et al. (2009)
IT-Ren	2002-04	ENF	Montagnani et al. (2009)
IT-Ro1	2002-06	DBF	Rey et al. (2002)
IT-Ro2	2002,03,06	DBF	Tedeschi et al. (2006)
IT-SRo	2001-03,06	ENF	Chiesi et al. (2005)
JP-Mas	2002-03	CRO	Saito et al. (2005)
JP-Tak	1999,2002-04	DBF	Ito et al. (2006)
JP-Tef	2002	MF	Takagi et al. (2009)
JP-Tom	2001-03	MF	Hirano et al. (2003)
NL-Hor	2005-06	GRA	Jacobs et al. (2007)
NL-Loo	1997,2001-2006	ENF	Dolman et al. (2002)
PL-Wet	2004-05	WET	Chojnicki et al. (2007)
PT-Mi1	2005	EBF	Pereira et al. (2007)
PT-Mi2	2005-06	GRA	Pereira et al. (2007)
RU-Fyo	1999-2006	ENF	Kurbatova et al. (2008)
SE-Deg	2002,04	WET	Sagerfors et al. (2008)
SE-Fla	1997-98,02	ENF	Lindroth et al. (2008)
SE-Nor	1996-97,99	ENF	Lagergren et al. (2008)
UK-Gri	1998, 2001	ENF	Rebmann et al. (2005)
UK-Ham	2004	DBF	-
UK-PL3	2005	DBF	-
US-ARM	2003-04	CRO	Fischer et al. (2007)
US-Aud	2004-05	GRA	-
US-Bar	2004-05	DBF	Jenkins et al. (2007)
US-Bkg	2005-06	GRA	Gilmanov et al. (2005)
US-Blo	2000-03,05-06	ENF	Goldstein et al. (2000)
US-Bn2	2003	DBF	Liu et al. (2005)
US-Bo1	1997-2000,05-06	CRO	Meyers and Hollinger (2004)
US-Bo2	2006	CRO	Meyers and Hollinger (2004)
US-Dk3	2004	MF	Pataki and Oren (2003)
US-Fpe	2004,06	GRA	-

Table 1 continued

Site	Years	Vegetation type	References
US-FR2	2005	WSA	Heinsch et al. (2004)
US-Goo	2006	GRA	-
US-Ha1	1994-96,98-2001	DBF	Urbanski et al. (2007)
US-Ho1	1996-04	ENF	Hollinger et al. (2004)
US-Ho2	1999-01,03-04	MF	Hollinger et al. (2004)
US-IB1	2006	CRO	Allison et al. (2005)
US-IB2	2005-06	GRA	Allison et al. (2005)
US-KS2	2002,04-06	CSH	Powell et al. (2006)
US-LPH	2003-04	DBF	Borken et al. (2006)
US-MMS	1999,2001-05	DBF	Schmid et al. (2000)
US-Moz	2005-06	DBF	Gu et al. (2006)
US-Me4	1999	ENF	Law et al. (2001)
US-NC1	2006	OSH	Sun et al. (2010)
US-NC2	2006	ENF	Noormets et al. (2009)
US-NR1	1999-2000,02-03	ENF	Monson et al. (2002)
US-Ne1	2002-04	CRO	Verma et al. (2005)
US-Ne2	2003-04	CRO	Verma et al. (2005)
US-Ne3	2002-04	CRO	Verma et al. (2005)
US-Pfa	1999	MF	Davis et al. (2003)
US-SO2	2004-06	WSA	Lipson et al. (2005)
US-SO3	2005-06	WSA	Lipson et al. (2005)
US-SO4	2004-06	CSH	-
US-SP1	2005	ENF	Powell et al. (2008)
US-SP2	2001-02,04	ENF	Clark et al. (2004)
US-SP3	2001-04	ENF	Clark et al. (2004)
US-SRM	2004-05	WSA	Potts et al. (2008)
US-Ton	2002-06	WSA	Ma et al. (2007)
US-UMB	2000-03	DBF	Gough et al. (2008)
US-Var	2001-04,06	GRA	Xu and Baldocchi (2004)
US-WBW	1995-99	DBF	Wilson and Baldocchi (2001)
US-WCr	2000,02-03	DBF	Cook et al. (2004)
US-Wrc	2001-02,04	ENF	Falk et al. (2008)
VU-Coc	2002-03	EBF	Roupsard et al. (2006)

B. Crosscorrelations between photosynthesis and respiration derived from eddy covariance data

B.1. Abstract

B.2. Introduction

The correlation between photosynthesis (GPP) and respiration (Reco) has been reported on different spatial and temporal scales (Baldocchi, 2008; Fischer et al., 2007). Recently there has been a debate on the magnitude of self-correlation of GPP and Reco when estimated from eddy covariance net ecosystem exchange (NEE) data using flux partitioning algorithms and suggestions have been made to analyse it (Vickers et al., 2009). Self-correlation, also called spurious or artificial correlation occurs if two variables share a common term, then the correlation between the two variables is partly attributable to this shared term. The degree of spurious correlation depends on the ratio of the variances of the common term compared to the whole variable. Here 1) we estimate the importance of self correlation and validity of a real correlation if the direct propagation of any error in Reco in GPP biases the correlation and 2) we discuss a way to minimize the problem of spurious correlation by using a day-night stratified approach to estimate Reco and GPP from disjoint datasets.

B.3. Methods

B.3.1. Data

In this study we used eddy covariance measurements of NEE from five different FLUXNET (www.fluxdata.org) sites (see table 1), that were processed in a standardized way described in (Papale et al., 2006). The data are storage corrected, spike filtered, u*-filtered and subsequently gap-filled. Based on these data we applied two different flux partitioning algorithms described in the following paragraph. Moreover we use the annual estimates across FLUXNET sites presented in Baldocchi (2008) and the estimates of the flux partitioning of these sites as presented in (Lasslop et al., 2010b).

B.3.2. Flux partitioning

GPP and Reco estimates were derived from the net ecosystem exchange using two approaches: (1) The nighttime data based approach described in (Reichstein et al., 2005), where a respiration model is fit to nighttime data and then extrapolated to the daytime using observations of the temperature. GPP is then derived as GPP= Reco-NEE. (2) A daytime data based approach that uses daytime data to estimate the parameters of a light response curve extended for the temperature sensitivity of Reco and VPD limitation of GPP (Lasslop et al., 2010b). The respiration is then extrapolated into the nighttime using the temperature relationship.

Table B.1.: FLUXNET sites used for the analysis on daily and monthly timescale.

Site	Vegetation type	Latitude	Longitude	Reference
US-WBW	DBF	35.9588	-84.2874	Wilson and Baldocchi (2001)
US-Var	GRA	38.4133	-120.951	Xu and Baldocchi (2004)
US-Ton	WSA	38.4316	-120.966	Ma et al. (2007)
US-Bar	DBF	44.0646	-71.2881	Jenkins et al. (2007)
US-Ho1	ENF	45.2041	-68.7402	Hollinger et al. (2004)

B.3.3. Self-correlation

Self-correlation between two variables occurs if two variables share a common term with non-zero variance. If one variable x is entirely part of the second variable z = x+y or z =x-y, the covariance between x and z can be split in two parts:

$$cov(x, z) = cov(x, x + y) = var(x) + cov(x, y) \qquad \text{(B.1)}$$

The correlation between x and z, e.g. the covariance normalized by the standard deviations of x times z, that is caused by the shared term x (R_{self}(x,z)) can be derived following (Kenney, 1982) from the ratio of the variances of y and x:

$$R_{self}(x, z) = \frac{1}{\sqrt{1 + \frac{var(y)}{var(x)}}} \qquad \text{(B.2)}$$

If x and y are not correlated this self-correlation equals the correlation between x and z. This theoretical concept can be transferred in two different ways to the correlation between GPP and Reco when deriving the two fluxes from eddy covariance NEE measurements:

(1) Vickers et al. (2009) apply it to the equation

$$GPP = Reco - NEE, \qquad \text{(B.3)}$$

where Reco is the shared variable x. We caution against this interpretation, as Reco is only used to derive GPP from NEE but is not part of the flux component (see the comment Lasslop et al. (2010a) for further details). Here instead we are assuming that the true flux components GPP_{true} and $Reco_{true}$ are independent, e.g. they don't share a common term.

(2) When we consider errors that can be present in the flux components using the nighttime data based approach any error in the Reco estimate (ϵ_{Reco}) is part of GPP:

$$NEE_{obs} = GPP_{true} + Reco_{true} + \epsilon_{NEE}, \qquad \text{(B.4)}$$

$$Reco_{est} = Reco_{true} + \epsilon_{Reco} \qquad \text{(B.5)}$$

Combination of equation B.3,B.4,and B.5 yields:

$$GPP_{est} = Reco_{est} - NEE_{obs} = GPP_{true} + \epsilon_{Reco} + \epsilon_{NEE}, \qquad \text{(B.6)}$$

where ϵ denote error terms, the subscript "true" denotes the true value of the system, "obs" and "est" the observed and estimated values, respectively. Thus spurious correlations can be introduced via error terms, here ϵ_{Reco}, but only if they contribute to the variance, e.g. they are not constant. When splitting GPP and Reco in true fluxes and errors, the correlation between GPP$_{est}$ and Reco$_{est}$ is, based on equation (3) and (4) and neglecting ϵ_{NEE}:

$$R(GPP_{est}, Reco_{est}) \leq \frac{cov(GPP_{true}, Reco_{true})}{std(GPP_{est})std(Reco_{est})} + \frac{cov(GPP_{true}, \epsilon_{Reco})}{std(GPP_{est})std(Reco_{est})} + \qquad (B.7)$$
$$\frac{cov(Reco_{true}, \epsilon_{Reco})}{std(GPP_{est})std(Reco_{est})} + \frac{var(\epsilon_{Reco})}{std(GPP_{est})std(Reco_{est})},$$

std is the standard deviation, var is the variance and cov is the covariance. The neglected error in NEE represents the observation noise in NEE and propagates to GPP, if GPP is derived as residual. This noise in GPP reduces the correlation between GPP_{est} and $Reco_{est}$ compared to the correlation between GPP_{true} and $Reco_{true}$. This is the reason for the inequality in equation 4. If the errors are independent and uncorrelated, $cov(Reco_{true}, \epsilon_{Reco}) = 0$ and $cov(GPP_{true}, \epsilon_{Reco}) = 0$ the estimated correlation coefficient is:

$$R(GPP_{est}, Reco_{est}) \leq \frac{cov(GPP_{true}, Reco_{true})}{std(GPP_{est})std(Reco_{est})} + \frac{var(\epsilon_{Reco})}{std(GPP_{est})std(Reco_{est})} \qquad (B.8)$$

B.3.4. Error estimate of $Reco$

To derive an estimate of the error in $Reco$ on the annual timescale we compared a nighttime data based estimate (indicated by the subscript "nb") with a daytime data based estimate (Lasslop et al., 2010b), indicated by the subscript "db"). The two estimates are derived from two disjoint datasets and the differences between them represent uncertainties due to the partitioning method and due to inconsistencies between day- and nighttime data. Following Hollinger et al. (2005) the variance of the error estimate was derived as $var(\epsilon_{Reco}) = 0.5var(Reco_{db} - Reco_{nb})$, here the annual integrals of Reco are used.

B.4. Results

For the Baldocchi (2008) dataset, that contains annual estimates accross FLUXNET sites the observed correlation coefficient was 0.899. The self correlation introduced by the error term would be 0.06 evaluated after eqation B.8 for this correlation across sites based on the difference between the two flux partitioning methods, decreasing the observed correlation to a still strong corrected correlation of 0.84. For single site analysis the reduction in the correlation using these independent estimates is small, see table B.2 for correlations on monthly and daily timescale for five FLUXNET sites.

On monthly timescale the correlation is less reduced than on daily. For US-Bar the correlation even increases on monthly timescale.

Table B.2.: Correlation between GPP and $Reco$ for five FLUXNET sites: observed correlation (R_{obs}) and the correlation between a nighttime data based estimate of $Reco$ and a daytime data based estimate of GPP ($R_{day-night}$).

Daily	Robs	Rday-night
US-WBW	0.86	0.81
US-Var	0.89	0.86
US-Ton	0.73	0.67
US-Bar	0.77	0.72
US-Ho1	0.91	0.89
Monthly	Robs	Rday-night
US-WBW	0.93	0.91
US-Var	0.94	0.93
US-Ton	0.82	0.80
US-Bar	0.86	0.88
US-Ho1	0.97	0.95

B.5. Discussion

The daytime and nighttime based estimates were already shown to agree well and the correlation between GPP and $Reco$ on annual timescale is not reduced significantly when using a combination of independent estimates across siteyears of the FLUXNET database. A possible reason for an increase when using the independent estimates is the reduced noise in the GPP estimate. For the nighttime data based flux partitioning approach the observational noise in NEE is propagated into GPP, while for the daytime data based approach GPP is computed as part of the light response curve. The error variance that we address here does not include errors due to the tower setup or different processing of the high frequency data (Mauder et al., 2008), these can increase the error variance across sites and possibly also for single sites. As the correlation due to the error strongly increases with increasing error variance such relations should always be challenged.

This statistical correlation does not mean that GPP is caused by $Reco$ or vice versa, but a large part of the correlation can be caused by background correlation of driving factors. On the other hand a strong correlation between the two is to be expected, as variation in recent photosynthesis is known to be correlated with root respiration; plants use energy from photosynthesis to drive the metabolism (Ekblad and Hogberg, 2001; Tang et al., 2005). At long time scales, substrate availability (constrained by past productivity) limits the whole-ecosystem respiration (Amthor, 2000; Schulze, 2006).

B.6. Conclusions

In this analysis we explored the effect of the direct propagation of errors in $Reco$ to GPP when using the standard nighttime data based flux partitioning on the correlation between $Reco$ and GPP. To derive relationships between $Reco$ and GPP without having to face the problem of spurious correlations we recommend to combine complementary estimates for such analysis, $Reco$ estimated with nighttime EC data and GPP estimated from daytime data using the light response curve (Lasslop et al., 2010b). Based on our error estimate the impact of the error on the correlation is rather small, but differs from site to site

and depends on the time scales. The combination of different methods to derive the flux components can challenge the robustness of the statistical relationships. In summary we have demonstrated that the correlation between GPP and $Reco$ is not a statistical artifact (either for single sites or across sites). This correlation should be investigated further in terms of ecological process understanding.

Bibliography

Allison, V. J., Miller, R. M., Jastrow, J. D., Matamala, R., and Zak, D. R.: Changes in soil microbial community structure in a tallgrass prairie chronosequence, Soil Science Society of America Journal, 69, 1412–1421, 2005.

Ammann, C., Flechard, C., Leifeld, J., Neftel, A., and Fuhrer, J.: The carbon budget of newly established temperate grassland depends on management intensity, Agriculture Ecosystems and Environment, 121, 5–20, 2007.

Amthor, J. S.: The McCree-de Wit-Penning de Vries-Thornley Respiration Paradigms: 30 Years Later, Ann Bot, 86, 1–20, 2000.

Anthoni, P. M., Knohl, A., Rebmann, C., Freibauer, A., Mund, M., Ziegler, W., Kolle, O., and Schulze, E. D.: Forest and agricultural land-use-dependent CO2 exchange in Thuringia, Germany, Global Change Biology, 10, 2005–2019, 2004.

Arain, M. A. and Restrepo-Coupe, N.: Net ecosystem production in a temperate pine plantation in southeastern Canada, Agricultural and Forest Meteorology, 128, 223–241, 2005.

Araujo, A. C., Nobre, A. D., Kruijt, B., Elbers, J. A., Dallarosa, R., Stefani, P., Randow, C. v., Manzi, A. O., Culf, A. D., Gash, J. H. C., Valentini, R., and Kabat, P.: Comparative measurements of carbon dioxide fluxes from two nearby towers in a central Amazonian rainforest: The Manaus LBA site, Journal of Geophysical Research, 107, 8090, doi:10.1029/2001JD000 676, 2002.

Araujo, A. C., Kruijt, B., Nobre, A. D., Dolman, A. J., Waterloo, M. J., Moors, E. J., and Souza, J. S. d.: Nocturnal accumulation of CO2 underneath a tropical forest canopy along a topographical gradient, Ecological Applications, 18, 1406–1419, 2008.

Arrhenius, S.: On the influence of carbonic acid in the air upon the temperature of the ground, The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science, 5, 237–276, 1896.

Atkin, O. K., Evans, J. R., and Siebke, K.: Relationship between the inhibition of leaf respiration by light and enhancement of leaf dark respiration following light treatment, Australian Journal of Plant Physiology, 25, 437–443, 1998.

Aubinet, M.: Eddy covariance CO2 flux measurements in nocturnal conditions: An analysis of the problem, Ecological Applications, 18, 1368–1378, 2008.

Aubinet, M., Grelle, A., Ibrom, A., Rannik, ., Moncrieff, J., Foken, T., Kowalski, A., Martin, P. H., Berbigier, P., Bernhofer, C., Clement, R., Elbers, J. A., Granier, A., Grunwald, T., Morgenstern, K., Pilegaard, K., Rebmann, C., Snijders, W., Valentini, R., and Vesala, T.: Estimates of the Annual Net Carbon and Water Exchange of Forest: The EUROFLUX Methodology,, Advances in Ecological Research, 30, 114–173, 2000.

Aubinet, M., Chermanne, B., Vandenhaute, M., Longdoz, B., Yernaux, M., and Laitat, E.: Long term carbon dioxide exchange above a mixed forest in the Belgian Ardennes, Agricultural and Forest Meteorology, 108, 293–315, 2001.

Aubinet, M., Feigenwinter, C., Heinesch, B., Bernhofer, C., Canepa, E., Lindroth, A., Montagnani, L., Rebmann, C., Sedlak, P., and Van Gorsel, E.: Direct advection measurements do not help to solve the night-time CO_2 closure problem: Evidence from three different forests, Agricultural and Forest Meteorology, In Press, Corrected Proof, 2010.

Aurela, M., Riutta, T., Laurila, T., Tuovinen, J. P., Vesala, T., Tuittila, E. S., Rinne, J., Haapanala, S., and Laine, J.: CO_2 exchange of a sedge fen in southern Finland - The impact of a drought period, Tellus Series B-Chemical and Physical Meteorology, 59, 826–837, 2007.

Bahn, M., Rodeghiero, M., Anderson-Dunn, M., Dore, S., Gimeno, C., Drosler, M., Williams, M., Ammann, C., Berninger, F., Flechard, C., Jones, S., Balzarolo, M., Kumar, S., Newesely, C., Priwitzer, T., Raschi, A., Siegwolf, R., Susiluoto, S., Tenhunen, J., Wohlfahrt, G., and Cernusca, A.: Soil Respiration in European Grasslands in Relation to Climate and Assimilate Supply, Ecosystems, 11, 1352–1367, 2008.

Baldocchi, D.: Measuring and modelling carbon dioxide and water vapour exchange over a temperate broad-leaved forest during the 1995 summer drought, Plant, Cell and Environment,, 20, 1108–1122, 1997.

Baldocchi, D.: Breathing of the terrestrial biosphere: lessons learned from a global network of carbon dioxide flux measurement systems, Australian Journal of Botany, 56, 1–26, 2008.

Baldocchi, D. D.: Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past, present and future, Global Change Biology, 9, 479–492, 2003.

Baldocchi, D. D., Falge, E., Gu, L., Olson, R., Hollinger, D. Y., Running, S. W., Anthoni, P., Bernhofer, C., Davis, K. J., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B. E., Lee, X., Malhi, Y., Meyers, T. P., Munger, J. W., Oechel, W. C., Paw U, K. T., Pilegaard, K., Schmid, H. P., Valentini, R., Verma, S., Vesala, T., Wilson, K. B., and Wofsy, S. C.: FLUXNET: A New Tool to Study the Temporal and Spatial Variability of Ecosystem-Scale Carbon Dioxide, Water Vapor, and Energy Flux Densities, Bulletin of the American Meteorological Society, 82, 2001.

Ball, J. T., Woodrow, I. E., and Berry, J. A.: A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions, in: Progress In Photosynthesis Research. Proceedings Of the VII International Photosynthesis Congress, edited by Biggins, I., pp. 221–224, 1987.

Barford, C. C., Wofsy, S. C., Goulden, M. L., Munger, J. W., Pyle, E. H., Urbanski, S. P., Hutyra, L., Saleska, S. R., Fitzjarrald, D., and Moore, K.: Factors Controlling Long- and Short-Term Sequestration of Atmospheric CO_2 in a Mid-latitude Forest, Science, 294, 1688–1691, 2001.

Beer, C., Reichstein, M., Ciais, P., Farquhar, G. D., and Papale, D.: Mean annual GPP of Europe derived from its water balance, Geophysical Research Letters, 34, L05 401, doi:10.1029/2006GL029 006, 2007.

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., Rodenbeck, C., Arain, M. A., Baldocchi, D., Bondeau, G. B. B. A., Cescatti, A., Lasslop, G., Lindroth, A., Lomas, M., Luyssaert, S., Margolis, H., Oleson, K. W., Roupsard, O., Veenendaal, E., Viovy, N., Williams, C., Woodward, F. I., and Papale, D.: Terrestrial Gross Carbon Dioxide Uptake: Global Distribution and Covariation with Climate, Science, p. doi:10.1126/science.1184984, 2010.

Berbigier, P., Bonnefond, J., and Mellmann, P.: CO2 and water vapour fluxes for 2 years above Euroflux forest site, Agricultural and Forest Meteorology, 108, 183–197, 2001.

Bergeron, O., Margolis, H., Black, T., Coursolle, C., Dunn, A., Barr, A., and Wofsy, S.: Comparison of carbon fluxes over three boreal black spruce forests in Canada, Global Change Biology, 13, 89–107, 2007.

Beven, K.: A manifesto for the equifinality thesis, Journal of Hydrology, 320, 18–36, 2006.

Bingeman, C. W., Varner, J. E., and Martin, W. P.: The Effect of the Addition of Organic Materials on the Decomposition of an Organic Soil, Soil Science Society of America Proceedings, 17, 34–38, 1953.

Black, T. A., Chen, W. J., Barr, A. G., Arain, M. A., Chen, Z., Nesic, Z., Hogg, E. H., Neumann, H. H., and Yang, P. C.: Increased carbon sequestration by a boreal deciduous forest in years with a warm spring, Geophysical Research Letters, 27, 1271–1274, 2000.

Bonal, D., Bosc, A., Ponton, S., Goret, J., Burban, B., Gross, P., Longdoz, B., Elbers, J., Epron, D., Guehl, J., and Granier, A.: Impact of severe dry season on net ecosystem exchange in the Neotropical rainforest of French Guiana, Global Change Biology, 14, 1917–1933, 2008.

Bonan, G. B.: Forests and climate change: Forcings, feedbacks, and the climate benefits of forests, Science, 320, 1444–1449, 2008.

Bond-Lamberty, B. and Thomson, A.: A global database of soil respiration data, Biogeosciences Discuss., 7, 1321–1344, 2010.

Borken, W., Savage, K., Davidson, E. A., and Trumbore, S. E.: Effects of experimental drought on soil respiration and radiocarbon efflux from a temperate forest soil, Global Change Biology, 12, 177–193, 2006.

Bousquet, P., Peylin, P., Ciais, P., Le Quere, C., Friedlingstein, P., and Tans, P. P.: Regional changes in carbon dioxide fluxes of land and oceans since 1980, Science, 290, 1342–1346, 2000.

Bouttier, F., Mahfouf, J. F., and Noilhan, J.: Sequential Assimilation of Soil-Moisture from Atmospheric Low-Level Parameters .1. Sensitivity and Calibration Studies, Journal of Applied Meteorology, 32, 1335–1351, 1993.

Braswell, B. H., Sacks, W. J., Linder, E., and Schimel, D. S.: Estimating diurnal to annual ecosystem parameters by synthesis of a carbon flux model with eddy covariance net ecosystem exchange observations, Glob Change Biol, 11, 335–355, 2005.

Brooks, A. and Farquhar, G. D.: Effect of Temperature on the Co2/O2 Specificity of Ribulose-1,5-Bisphosphate Carboxylase Oxygenase and the Rate of Respiration in the Light - Estimates from Gas-Exchange Measurements on Spinach, Planta, 165, 397–406, 1985.

Cannell, M. G. R. and Thornley, J. H. M.: Modelling the components of plant respiration: Some guiding principles, Annals of Botany, 85, 45–54, 2000.

Carbone, M. S., Winston, G. C., and Trumbore, S. E.: Soil respiration in perennial grass and shrub ecosystems: Linking environmental controls with plant and microbial sources on seasonal and diel timescales, Journal of Geophysical Research-Biogeosciences, 113, DOI: 10.1029/2007JG000 611, 2008.

Carrara, A., Janssens, I. A., Curiel Yuste, J., and Ceulemans, R.: Seasonal changes in photosynthesis, respiration and NEE of a mixed temperate forest, Agricultural and Forest Meteorology, 126, 15–31, 2004.

Carvalhais, N., Reichstein, M., Seixas, J., Collatz, G. J., Pereira, J. S., Berbigier, P., Carrara, A., Granier, A., Montagnani, L., Papale, D., Rambal, S., Sanz, M. J., and Valentini, R.: Implications of the carbon cycle steady state assumption for biogeochemical modeling performance and inverse parameter retrieval, Global Biogeochemical Cycles, 22, GB2007, doi:10.1029/2007GB003 033, 2008.

Carvalhais, N., Reichstein, M., Ciais, P., Collatz, G. J., Mahecha, M. D., Montagnani, L., Papale, D., Rambal, S., and Seixas, J.: Identification of vegetation and soil carbon pools out of equilibrium in a process model via eddy covariance and biometric constraints, Global Change Biology, early view, doi:10.1111/j.1365–2486.2010.02 173.x, 2010.

Chapin, F. S., Matson, P. A., and Mooney, H. A.: Principles of Terrestrial Ecosystem Ecology, Springer, Berlin, 2002.

Chapin, F. S., McFarland, J., McGuire, A. D., Euskirchen, E. S., Ruess, R. W., and Kielland, K.: The changing global carbon cycle: linking plant-soil carbon dynamics to global consequences, Journal of Ecology, 97, 840–850, 2009.

Chiesi, M., Maselli, F., Bindi, M., Fibbi, L., Cherubini, P., Arlotta, E., Tirone, G., Matteucci, G., and Seufert, G.: Modelling carbon budget of Mediterranean forests using ground and remote sensing measurements, Agricultural and Forest Meteorology, 135, 22–34, 2005.

Chojnicki, B., Urbaniak, M., JÃşzefczyk, D., Augustin, J., and Olejnik, J.: Measurements of gas and heat fluxes at Rzecin wetland, in: Wetlands: Monitoring, Modeling and Management, edited by Okruszko, T., Maltby, E., Szatylowicz, J., Miroslaw-Swiatek, D., and Kotowski, W., pp. 125–131, Taylor & Francis Group, London, 2007.

Ciais, P., Reichstein, M., Viovy, N., Granier, A., Ogee, J., Allard, V., Aubinet, M., Buchmann, N., Bern-hofer, C., Carrara, A., Chevallier, F., De Noblet, N., Friend, A. D., Friedlingstein, P., Grunwald, T., Heinesch, B., Keronen, P., Knohl, A., Krinner, G., Loustau, D., Manca, G., Matteucci, G., Miglietta, F., Ourcival, J. M., Papale, D., Pilegaard, K., Rambal, S., Seufert, G., Soussana, J. F., Sanz, M. J., Schulze, E. D., Vesala, T., and Valentini, R.: Europe-wide reduction in primary productivity caused by the heat and drought in 2003, Nature, 437, 529–533, 2005.

Ciais, P., Bousquet, P., Freibauer, A., and Naegler, T.: Horizontal displacement of carbon associated with agriculture and its impacts on atmospheric CO2, Global Biogeochemical Cycles, 21, DOI: 10.1029/2006GB002 741, 2007.

Clark, K. L., Gholz, H. L., and Castro, M. S.: Carbon dynamics along a chronosequence of slash pine plantationsin north Florida, Ecological Applications, 14, 1154–1171, 2004.

Collatz, G. J., Ball, J. T., Grivet, C., and Berry, J. A.: Physiological and Environmental-Regulation of Stomatal Conductance, Photosynthesis and Transpiration - a Model That Includes a Laminar Boundary-Layer, Agricultural and Forest Meteorology, 54, 107–136, 1991.

Collatz, G. J., Ribas-Carbo, M., and Berry, J. A.: Coupled Photosynthesis-Stomatal Conductance Model for Leaves of C4 Plants, Australian Journal of Plant Physiology, 19, 519–538, 1992.

Cook, B. D., Davis, K. J., Wang, W. G., Desai, A., Berger, B. W., Teclaw, R. M., Martin, J. G., Bolstad, P. V., Bakwin, P. S., Yi, C. X., and Heilman, W.: Carbon exchange and venting anomalies in an upland deciduous forest in northern Wisconsin, USA, Agricultural and Forest Meteorology, 126, 271–295, 2004.

Curran, P. J. and Steven, M. D.: Multispectral Remote Sensing for the Estimation of Green Leaf Area Index, Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences, 309, 257–270, 1983.

Davis, K. J., Bakwin, P. S., Yi, C. X., Berger, B. W., Zhao, C. L., Teclaw, R. M., and Isebrands, J. G.: The annual cycles of CO2 and H2O exchange over a northern mixed forest as observed from a very tall tower, Global Change Biology, 9, 1278–1293, 2003.

Denman, K., Brasseur, G., Chidthaisong, A., Ciais, P., Cox, P., Dickinson, R., Hauglustaine, D., Heinze, C., Holland, E., Jacob, D., Lohmann, U., Ramachandran, S., Dias, P. d. S., Wofsy, S., and Zhang, X.: Couplings Between Changes in the Climate System and Biogeochemistry, in: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, edited by Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K., Tignor, M., and Miller, H., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.

Desai, A. R., Richardson, A. D., Moffat, A. M., Kattge, J., Hollinger, D. Y., Barr, A., Falge, E., Noormets, A., Papale, D., Reichstein, M., and Stauch, V. J.: Cross-site evaluation of eddy covariance GPP and RE decomposition techniques, Agricultural and Forest Meteorology, 148, 821–838, 2008.

Desjardins, R.: Technique to Measure Co2 Exchange under Field Conditions, International Journal of Biometeorology, 18, 76–83, 1974.

Desjardins, R. L. and Lemon, E. R.: Limitations of an eddy-correlation technique for the determination of the carbon dioxide and sensible heat fluxes, Boundary-Layer Meteorology, 5, 475–488, doi:10.1007/BF00123493, 1974.

Dietiker, D., Eugster, W., and Buchmann, N.: Testing CO2 and H2O fluxes under real conditions and yield under changing climatic conditions with the DNDC model, Agriculture Ecosystems Environment, submitted.

Dolman, A. J., Moors, E. J., and Elbers, J. A.: The carbon uptake of a mid latitude pine forest growing on sandy soil, Agricultural and Forest Meteorology, 111, 157–170, 2002.

Don, A., Rebmann, C., Kolle, O., Scherer-Lorenzen, M., and Schulze, E.-D.: Impact of afforestation-associated management changes on the carbon balance of grassland, Global Change Biology, pp. doi:10.1111/j.1365–2486.2009.01 873.x, 2009.

Doughty, C. E. and Goulden, M. L.: Are tropical forests near a high temperature threshold?, J. Geophys. Res., 113, G00B07, doi:10.1029/2007JG000 632, 2008.

Draper, N. and Smith, H.: Applied Regression Analysis, John Wiley and Sons, New York, third edn., 1998.

Dunn, A., Barford, C., Wofsy, S., Goulden, M., and Daube, B.: A long-term record of carbon exchange in a boreal black spruce forest: means, responses to interannual variability and decadal trends, Global Change Biol., 13, 577–590, 2007.

Ekblad, A. and Hogberg, P.: Natural abundance of 13C in CO2 respired from forest soils reveals speed of link between tree photosynthesis and root respiration, Oecologia, 127, 305–308, 2001.

Evans, G. T.: Defining misfit between biogeochemical models and data sets, Journal of Marine Systems, 40-41, 49–54, 2003.

Evans, S. N. and Stark, P. B.: Inverse problems as statistics, Inverse Problems, 18, R55–R97, 2002.

Falge, E., Baldocchi, D., Olson, R., Anthoni, P., Aubinet, M., Bernhofer, C., Burba, G., Ceulemans, R., Clement, R., Dolman, H., Granier, A., Gross, P., Grunwald, T., Hollinger, D., Jensen, N. O., Katul, G., Keronen, P., Kowalski, A., Lai, C. T., Law, B. E., Meyers, T., Moncrieff, H., Moors, E., Munger, J. W., Pilegaard, K., Rannik, U., Rebmann, C., Suyker, A., Tenhunen, J., Tu, K., Verma, S., Vesala, T., Wilson, K., and Wofsy, S.: Gap filling strategies for defensible annual sums of net ecosystem exchange, Agricultural and Forest Meteorology, 107, 43–69, 2001.

Falge, E., Baldocchi, D., Tenhunen, J., Aubinet, M., Bakwin, P., Berbigier, P., Bernhofer, C., Burba, G., Clement, R., Davis, K. J., Elbers, J. A., Goldstein, A. H., Grelle, A., Granier, A., Guomundsson, J., Hollinger, D., Kowalski, A. S., Katul, G., Law, B. E., Malhi, Y., Meyers, T., Monson, R. K., Munger, J. W., Oechel, W., Paw, K. T., Pilegaard, K., Rannik, U., Rebmann, C., Suyker, A., Valentini, R.,

Wilson, K., and Wofsy, S.: Seasonality of ecosystem respiration and gross primary production as derived from FLUXNET measurements, Agricultural and Forest Meteorology, 113, 53–74, 2002a.

Falge, E., Tenhunen, J., Baldocchi, D., Aubinet, M., Bakwin, P., Berbigier, P., Bernhofer, C., Bonnefond, J. M., Burba, G., Clement, R., Davis, K. J., Elbers, J. A., Falk, M., Goldstein, A. H., Grelle, A., Granier, A., Grunwald, T., Gudmundsson, J., Hollinger, D., Janssens, I. A., Keronen, P., Kowalski, A. S., Katul, G., Law, B. E., Malhi, Y., Meyers, T., Monson, R. K., Moors, E., Munger, J. W., Oechel, W., U, K. T. P., Pilegaard, K., Rannik, U., Rebmann, C., Suyker, A., Thorgeirsson, H., Tirone, G., Turnipseed, A., Wilson, K., and Wofsy, S.: Phase and amplitude of ecosystem carbon release and uptake potentials as derived from FLUXNET measurements, Agricultural and Forest Meteorology, 113, 75–95, 2002b.

Falk, M., Wharton, S., Schroeder, M., Ustin, S., and U, K. T. P.: Flux partitioning in an old-growth forest: seasonal and interannual dynamics, Tree Physiology, 28, 509–520, 2008.

Farquhar, G. D., Caemmerer, S., and Berry, J. A.: A biochemical model of photosynthetic CO_2 assimilation in leaves of C3 species, Planta, 149, 78–90, 1980.

Feigenwinter, C., Bernhofer, C., and Vogt, R.: The influence of advection on the short term CO2-budget in and above a forest canopy, Boundary-Layer Meteorology, 113, 201–224, 2004.

Fischer, M. L., Billesbach, D. P., Berry, J. A., Riley, W. J., and Torn, M. S.: Spatiotemporal variations in growing season exchanges of CO_2, H_2O, and sensible heat in agricultural fields of the Southern Great Plains, Earth Interactions, 11, 1–21, 2007.

Flanagan, L. B., Wever, L. A., and Carlson, P. J.: Seasonal and interannual variation in carbon dioxide exchange and carbon balance in a northern temperate grassland, Global Change Biology, 8, 599–615, 2002.

Foken, T. and Wichura, B.: Tools for quality assessment of surface-based flux measurements, Agricultural and Forest Meteorology, 78, 83–105, 1996.

Friedlingstein, P., Cox, P., Betts, R., Bopp, L., von Bloh, W., Brovkin, V., Cadule, P., Doney, S., Eby, M., Fung, I., Bala, G., John, J., Jones, C., Joos, F., Kato, T., Kawamiya, M., Knorr, W., Lindsay, K., Matthews, H. D., Raddatz, T., Rayner, P., Reick, C., Roeckner, E., Schnitzler, K. G., Schnur, R., Strassmann, K., Weaver, A. J., Yoshikawa, C., and Zeng, N.: Climate-Carbon Cycle Feedback Analysis: Results from the C4MIP Model Intercomparison, Journal of Climate, 19, 3337–3353, 2006.

Friend, A. D., Arneth, A., Kiang, N. Y., Lomas, M., Ogee, J., Rodenbeck, C., Running, S. W., Santaren, J. D., Sitch, S., Viovy, N., Woodwards, F. I., and Zaehle, S.: FLUXNET and modelling the global carbon cycle, Global Change Biology, 13, 610–633, 2007.

Garbulsky, M. F., Penuelas, J., Papale, D., and Filella, I.: Remote estimation of carbon dioxide uptake by a Mediterranean forest, Global Change Biology, 14, 2860–2867, 2008.

Garbulsky, M. F., Penuelas, J., Papale, D., Ardo, J., Goulden, M. L., Kiely, G., Richardson, A. D., Rotenberg, E., Veenendaal, E. M., and Filella, I.: Patterns and controls of the variability of radiation

use efficiency and primary productivity across terrestrial ecosystems, Global Ecology and Biogeography, 19, 253–267, 2010.

Gaumont-Guay, D., Black, T., Griffis, T., Barr, A., Morgenstern, K., Jassal, R., and Nesic, Z.: Influence of temperature and drought on seasonal and interannual variations of soil, bole and ecosystem respiration in a boreal aspen stand, Agricultural and Forest Meteorology, 140, 203–219, 2006.

Göckede, M., Foken, T., Aubinet, M., Aurela, M., Banza, J., Bernhofer, C., Bonnefond, J. M., Brunet, Y., Carrara, A., Clement, R., Dellwik, E., Elbers, J., Eugster, W., Fuhrer, J., Granier, A., Grünwald, T., Heinesch, B., Janssens, I. A., Knohl, A., Koeble, R., Laurila, T., Longdoz, B., Manca, G., Marek, M., Markkanen, T., Mateus, J., Matteucci, G., Mauder, M., Migliavacca, M., Minerbi, S., Moncrieff, J., Montagnani, L., Moors, E., Ourcival, J. M., Papale, D., Pereira, J., Pilegaard, K., Pita, G., Rambal, S., Rebmann, C., Rodrigues, A., Rotenberg, E., Sanz, M. J., Sedlak, P., Seufert, G., Siebicke, L., Soussana, J. F., Valentini, R., Vesala, T., Verbeeck, H., and Yakir, D.: Quality control of CarboEurope flux data - Part 1: Coupling footprint analyses with flux data quality assessment to evaluate sites in forest ecosystems, Biogeosciences, 5, 433–450, 2008.

Giasson, M.-A., Coursolle, C., and Margolis, H.: Ecosystem-level carbon fluxes from a boreal cutover in eastern Canada before and after scarification, Agric. Forest Meteorol., 140, 23–40, 2006.

Gillies, R. R., Kustas, W. P., and Humes, K. S.: A verification of the 'triangle' method for obtaining surface soil water content and energy fluxes from remote measurements of the Normalized Difference Vegetation Index (NDVI) and surface e, International Journal of Remote Sensing, 18, 3145–3166, 1997.

Gilmanov, T. G., Johnson, D. A., and Saliendra, N. Z.: Growing season CO2 fluxes in a sagebrush-steppe ecosystem in Idaho: bowen ratio/energy balance measurements and modeling, Basic and Applied Ecology, 4, 167–183, 2003a.

Gilmanov, T. G., Verma, S. B., Sims, P. L., Meyers, T. P., Bradford, J. A., Burba, G. G., and Suyker, A. E.: Gross primary production and light response parameters of four Southern Plains ecosystems estimated using long-term CO2-flux tower measurements, Global Biogeochemical Cycles, 17, 1071, doi:10.1029/2002GB002 023, 2003b.

Gilmanov, T. G., Tieszen, L. L., Wylie, B. K., Flanagan, L. B., Frank, A. B., Haferkamp, M. R., Meyers, T. P., and Morgan, J. A.: Integration of CO2 flux and remotely-sensed data for primary production and ecosystem respiration analyses in the Northern Great Plains: potential for quantitative spatial extrapolation, Global Ecology and Biogeography, 14, 271–292, 2005.

Gilmanov, T. G., Soussana, J. E., Aires, L., Allard, V., Ammann, C., Balzarolo, M., Barcza, Z., Bernhofer, C., Campbell, C. L., Cernusca, A., Cescatti, A., Clifton-Brown, J., Dirks, B. O. M., Dore, S., Eugster, W., Fuhrer, J., Gimeno, C., Gruenwald, T., Haszpra, L., Hensen, A., Ibrom, A., Jacobs, A. F. G., Jones, M. B., Lanigan, G., Laurila, T., Lohila, A., Manca, G., Marcolla, B., Nagy, Z., Pilegaard, K., Pinter, K., Pio, C., Raschi, A., Rogiers, N., Sanz, M. J., Stefani, P., Sutton, M., Tuba, Z., Valentini, R., Williams, M. L., and Wohlfahrt, G.: Partitioning European grassland net ecosystem CO2

exchange into gross primary productivity and ecosystem respiration using light response function analysis, Agriculture Ecosystems & Environment, 121, 93–120, 2007.

Gobron, N. and Verstraete, M.: ECV T10: Fraction of Absorbed Photosynthetically Active Radiation (FAPAR), Essential Climate Variables, 2008.

Goerner, A., Reichstein, M., and Rambal, S.: Tracking seasonal drought effects on ecosystem light use efficiency with satellite-based PRI in a Mediterranean forest, Remote Sensing of Environment, 113, 1101–1111, 2009.

Goldstein, A., Hultman, N., Fracheboud, J., Bauer, M., Panek, J., Xu, M., Qi, Y., Guenther, A., and Baugh, W.: Effects of climate variability on the carbon dioxide, water, and sensible heat fluxes above a ponderosa pine plantation in the Sierra Nevada (CA), Agricultural and Forest Meteorology, 101, 113–129, 2000.

Gorton, H. L., Williams, W. E., and Assmann, S. M.: Circadian-Rhythms in Stomatal Responsiveness to Red and Blue-Light, Plant Physiology, 103, 399–406, 1993.

Gough, C. M., Vogel, C. S., Schmid, H. P., Su, H. B., and Curtis, P. S.: Multi-year convergence of biometric and meteorological estimates of forest carbon storage, Agricultural and Forest Meteorology, 148, 158–170, 2008.

Goulden, M. L., Munger, J. W., Fan, S. M., Daube, B. C., and Wofsy, S. C.: Measurements of carbon sequestration by long-term eddy covariance: Methods and a critical evaluation of accuracy, Global Change Biology, 2, 169–182, 1996.

Goulden, M. L., Miller, S. D., da Rocha, H. R., Menton, M. C., de Freitas, H. C., e Silva Figueira, A. M., and de Sousa, C. A. D.: Diel and seasonal patterns of tropical forest CO2 exchange, Ecological Applications, 14, 42–54, 2004.

Goulden, M. L., Miller, S. D., and da Rocha, H. R.: Nocturnal cold air drainage and pooling in a tropical forest, Journal of Geophysical Research-Atmospheres, 111, doi:10.1029/2005JD006 037, 2006.

Gove, J. H. and Hollinger, D. Y.: Application of a dual unscented Kalman filter for simultaneous state and parameter estimation in problems of surface-atmosphere exchange, Journal of Geophysical Research-Atmospheres, 111, DOI: 10.1029/2005JD006 021, 2006.

Granier, A., Ceschia, E., Damesin, C., Dufrêne, E., Epron, D., Gross, P., Lebaube, S., Dantec, V. L., Goff, N. L., Lemoine, D., Lucot, E., Ottorini, J. M., Pontailler, J. Y., and Saugier, B.: The carbon balance of a young Beech forest, Functional Ecology, 14, 312–325, 2000.

Grünwald, T. and Bernhofer, C.: A decade of carbon, water and energy flux measurement of an old spruce forest at the Anchor Station Tharandt, Tellus B, 59, 387–396, 2007.

Grünzweig, J. M., Lin, T., Rotenberg, E., Schwartz, A., and Yakir, D.: Carbon sequestration in arid-land forest, Global Change Biology, 9, 791–799, 2003.

Gu, L., Baldocchi, D. D., Wofsy, S. C., Munger, J. W., Michalsky, J. J., Urbanski, S. P., and Boden, T. A.: Response of a Deciduous Forest to the Mount Pinatubo Eruption: Enhanced Photosynthesis, Science, 299, 2035–2038, 2003.

Gu, L. H., Meyers, T., Pallardy, S. G., Hanson, P. J., Yang, B., Heuer, M., Hosman, K. P., Riggs, J. S., Sluss, D., and Wullschleger, S. D.: Direct and indirect effects of atmospheric conditions and soil moisture on surface energy partitioning revealed by a prolonged drought at a temperate forest site, Journal of Geophysical Research-Atmospheres, 111, 10.1029/2006JD007 161, 2006.

Guan, D. X., Wu, J. B., Zhao, X. S., Han, S. J., Yu, G. R., Sun, X. M., and Jin, C. J.: CO2 fluxes over an old, temperate mixed forest in northeastern China, Agricultural and Forest Meteorology, 137, 138–149, 2006.

Hagen, S. C., Braswell, B. H., Linder, E., Frolking, S., Richardson, A. D., and Hollinger, D. Y.: Statistical uncertainty of eddy flux-based estimates of gross ecosystem carbon exchange at Howland Forest, Maine, Journal of Geophysical Research (Atmospheres), 111, doi:10.1029/2005JD006 154, 2006.

Harley, P. C., Thomas, R. B., Reynolds, J. F., and Strain, B. R.: Modeling Photosynthesis of Cotton Grown in Elevated Co2, Plant Cell and Environment, 15, 271–282, 1992.

Heinsch, F. A., Heilman, J. L., McInnes, K. J., Cobos, D. R., Zuberer, D. A., and Roelke, D. L.: Carbon dioxide exchange in a high marsh on the Texas Gulf Coast: effects of freshwater availability, Agricultural and Forest Meteorology, 125, 159–172, 2004.

Hennessey, T. L., Freeden, A. L., and Field, C. B.: Environmental-Effects of Circadian-Rhythms in Photosynthesis and Stomatal Opening, Planta, 189, 369–376, 1993.

Hibbard, K. A., Law, B. E., Reichstein, M., and Sulzman, J.: An analysis of soil respiration across northern hemisphere temperate ecosystems, Biogeochemistry, 73, 29–70, 2005.

Hilker, T., Coops, N. C., Wulder, M. A., Black, T. A., and Guy, R. D.: The use of remote sensing in light use efficiency based models of gross primary production: A review of current status and future requirements, Science of The Total Environment, 404, 411–423, 2008.

Hirano, T., Hirata, R., Fujinuma, Y., Saigusa, N., Yamamoto, S., Harazono, Y., Takada, M., Inukai, K., and Inoue, G.: CO2 and water vapor exchange of a larch forest in northern Japan, Tellus Series B-Chemical and Physical Meteorology, 55, 244–257, 2003.

Hirano, T., Segah, H., Harada, T., Limin, S., June, T., Hirata, R., and Osaki, M.: Carbon dioxide balance of a tropical peat swamp forest in Kalimantan, Indonesia, Global Change Biology, 13, 412–425, 2007.

Hofmeister, H.: Lebensraum Wald. Pflanzengesellschaften und ihre Ökologie, Blackwell Wissenschafts-Verlag, 4 edn., 1997.

Hogberg, P., Nordgren, A., Buchmann, N., Taylor, A. F. S., Ekblad, A., Hogberg, M. N., Nyberg, G., Ottosson-Lofvenius, M., and Read, D. J.: Large-scale forest girdling shows that current photosynthesis drives soil respiration, Nature, 411, 789–792, 2001.

Hollinger, D., Aber, J., Dail, B., Davidson, E., Goltz, S., Hughes, H., LeClerc, M., Lee, J., Richardson, A., Rodrigues, C., Scott, N., Achuatavarier, D., and Walsh, J.: Spatial and temporal variability in forest-atmosphere CO2 exchange, Global Change Biology, 10, 1689–1706, 2004.

Hollinger, D. Y. and Richardson, A. D.: Uncertainty in eddy covariance measurements and its application to physiological models, Tree Physiology, 25, 873–885, 2005.

Hollinger, D. Y., Kelliher, F. M., Byers, J. N., Hunt, J. E., Mcseveny, T. M., and Weir, P. L.: Carbon-Dioxide Exchange between an Undisturbed Old-Growth Temperate Forest and the Atmosphere, Ecology, 75, 134–150, 1994.

Houborg, R. M. and Soegaard, H.: Regional simulation of ecosystem CO2 and water vapor exchange for agricultural land using NOAA AVHRR and Terra MODIS satellite data. Application to Zealand, Denmark, Remote Sensing of Environment, 93, 150–167, 2004.

Howard, E. A., Gower, S. T., Foley, J. A., and Kucharik, C. J.: Effects of logging on carbon dynamics of a jack pine forest in Saskatchewan, Canada, Global Change Biology, 10, 1267–1284, 2004.

Huber, B.: Registrierung des CO2-Gefälles und Berechnung des CO2-Stromes über Pflanzengesellschaften mittels Ultrarot-Absorptionsschieber, Deutsche botanische Gesellschaft, 63, 52–63, 1950.

Humphreys, E. R., Black, T. A., Morgenstern, K., Cai, T. B., Drewitt, G. B., Nesic, Z., and Trofymow, J. A.: Carbon dioxide fluxes in coastal Douglas-fir stands at different stages of development after clearcut harvesting, Agricultural and Forest Meteorology, 140, 6–22, 2006.

Imhoff, M. L., Bounoua, L., Ricketts, T., Loucks, C., Harriss, R., and Lawrence, W. T.: Global patterns in human consumption of net primary production, Nature, 429, 870–873, 2004.

Inoue, H.: Studies of the Phenomena of Waving Plants (ŞHONAMIǏ) Caused by Wind Part 4. Turbulent Transfer Phenomena over the Waving Plants, Journal of Agricultural Meteorology, 12, 138–144, 1957.

Irvine, J. and Law, B.: Seasonal soil CO2 effluxes in young and old ponderosa pine forests, Global Change Biology, 8, 1183–1194, 2002.

Ito, A., Muraoka, H., Koizumi, H., Saigusa, N., Murayama, S., and Yamamoto, S.: Seasonal variation in leaf properties and ecosystem carbon budget in a cool-temperate deciduous broad-leaved forest: simulation analysis at Takayama site, Japan, Ecological Research, 21, 137–149, 2006.

Jackson, R. D.: Remote Sensing of Biotic and Abiotic Plant Stress, Annual Review of Phytopathology, 24, 265–287, 1986.

Jacobs, C. M. J., Jacobs, A. F. G., Bosveld, F. C., Hendriks, D. M. D., Hensen, A., Kroon, P. S., Moors, E. J., Nol, L., Schrier-Uijl, A., and Veenendaal, E. M.: Variability of annual CO2 exchange from Dutch grasslands, Biogeosciences, 4, 803–816, 2007.

Janssens, I. A., Lankreijer, H., Matteucci, G., Kowalski, A. S., Buchmann, N., Epron, D., Pilegaard, K., Kutsch, W., Longdoz, B., Grunwald, T., Montagnani, L., Dore, S., Rebmann, C., Moors, E. J., Grelle, A., Rannik, U., Morgenstern, K., Oltchev, S., Clement, R., Gudmundsson, J., Minerbi, S., Berbigier, P., Ibrom, A., Moncrieff, J., Aubinet, M., Bernhofer, C., Jensen, N. O., Vesala, T., Granier, A., Schulze, E. D., Lindroth, A., Dolman, A. J., Jarvis, P. G., Ceulemans, R., and Valentini, R.: Productivity overshadows temperature in determining soil and ecosystem respiration across European forests, Global Change Biology, 7, 269–278, 2001.

Jenkins, J. P., Richardson, A. D., Braswell, B. H., Ollinger, S. V., Hollinger, D. Y., and Smith, M. L.: Refining light-use efficiency calculations for a deciduous forest canopy using simultaneous tower-based carbon flux and radiometric measurements, Agricultural and Forest Meteorology, 143, 64–79, 2007.

Jung, M., Henkel, K., Herold, M., and Churkina, G.: Exploiting synergies of global land cover products for carbon cycle modeling, Remote Sensing of Environment, 101, 534–553, 2006.

Jung, M., Reichstein, M., and Bondeau, A.: Towards global empirical upscaling of FLUXNET eddy covariance observations: validation of a model tree ensemble approach using a biosphere model, Biogeosciences, 6, 2001–2013, 2009.

Kaminski, T., Knorr, W., Rayner, P. J., and Heimann, M.: Assimilating atmospheric data into a terrestrial biosphere model: A case study of the seasonal cycle, Global Biogeochemical Cycles, 16, doi:10.1029/2001GB001 463, 2002.

Kato, T., Tang, Y., Gu, S., Hirota, M., Du, M., Li, Y., and Zhao, X.: Temperature and biomass influences on interannual changes in CO_2 exchange in an alpine meadow on the Qinghai-Tibetan Plateau, Global Change Biology, 12, 1285–1298, 2006.

Katterer, T., Reichstein, M., Andren, O., and Lomander, A.: Temperature dependence of organic matter decomposition: a critical review using literature data analyzed with different models, Biology and Fertility of Soils, 27, 258–262, 1998.

Kattge, J., Knorr, W., Raddatz, T., and Wirth, C.: Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global-scale terrestrial biosphere models, Global Change Biology, 15, 976–991, 2009.

Kenney, B. C.: Beware of Spurious Self-Correlations, Water Resources Research, 18, 1041–1048, 1982.

Kirschbaum, M. U. F.: The Temperature-Dependence of Soil Organic-Matter Decomposition, and the Effect of Global Warming on Soil Organic-C Storage, Soil Biology & Biochemistry, 27, 753–760, 1995.

Knohl, A. and Baldocchi, D. D.: Effects of diffuse radiation on canopy gas exchange processes in a forest ecosystem, J. Geophys. Res., 113, doi:10.1029/2007JG000 663, 2008.

Knohl, A., Schulze, E. D., Kolle, O., and Buchmann, N.: Large carbon uptake by an unmanaged 250-year-old deciduous forest in Central Germany, Agricultural and Forest Meteorology, 118, 151–167, 2003.

Knorr, W.: Annual and interannual CO2 exchanges of the terrestrial biosphere: process-based simulations and uncertainties, Global Ecology and Biogeography, 9, 225 – 252, 2000.

Knorr, W. and Heimann, M.: Uncertainties in global terrestrial biosphere modeling. Part I: a comprehensive sensitivity analysis with a new photosynthesis and energy balance scheme, Global Biogeochemical Cycles, 15, 207–225, 2001a.

Knorr, W. and Heimann, M.: Uncertainties in global terrestrial biosphere modeling, part II: Global constraints for a process-based vegetation model, Global Biogeochemical Cycles, 15, 227–246, 2001b.

Knorr, W. and Kattge, J.: Inversion of terrestrial ecosystem model parameter values against eddy covariance measurements by Monte Carlo sampling, Global Change Biology, 11, 1333–1351, 2005.

Kowalski, A. S., Serrano-Ortiz, P., Janssens, I. A., Sanchez-Moraic, S., Cuezva, S., Domingo, F., Were, A., and Alados-Arboledas, L.: Can flux tower research neglect geochemical CO2 exchange?, Agricultural and Forest Meteorology, 148, 1045–1054, 2008.

Krinner, G., Viovy, N., de Noblet-Ducoudre, N., Ogee, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, Global Biogeochemical Cycles, 19, doi: 10.1029/2003GB002 199, 2005.

Körner, C.: Leaf diffusive conductances in the major vegetation types of the globe, in: Ecophysiology of photosynthesis, edited by Schulze, E.-D. and Caldwell, M., pp. 463–490, Springer Verlag, Berlin Heidelberg New York, 1995.

Kruijt, B., Elbers, J. A., von Randow, C., Araujo, A. C., Oliveira, P. J., Culf, A., Manzi, A. O., Nobre, A. D., Kabat, P., and Moors, E. J.: The robustness of eddy correlation fluxes for Amazon rain forest conditions, Ecological Applications, 14, 101–113, 2004.

Krumov, A., Nikolova, A., Vassilev, V., and Vassilev, N.: Assessment of plant vitality detection through fluorescence and reflectance imagery, Advances in Space Research, 41, 1870–1875, doi: 10.1016/j.asr.2007.11.020, 2008.

Kurbatova, J., Li, C., Varlagin, A., Xiao, X., and Vygodskaya, N.: Modeling carbon dynamics in two adjacent spruce forests with different soil conditions in Russia, Biogeosciences, 5, 969–980, 2008.

Kutsch, W. L., Kolle, O., Rebmann, C., Knohl, A., Ziegler, W., and Schulze, E.-D.: Advection and resulting CO2 exchange uncertainty in a tall forest in central Germany, Ecological Applications, 18, 1391–1405, 2008.

Kutsch, W. L., Wirth, C., Kattge, J., Noellert, S., Herbst, M., and Kappen, L.: Ecophysiological characteristics of large, mature trees and consequences for old-growth functioning, in: Old-growth forests: Function, fate and value. Ecological Studies, edited by Wirth, C., Gleixner, G., and Heimann, M., vol. 207, Springer, New York, Berlin, Heidelberg, 2009.

Lafleur, P. M., Roulet, N. T., Bubier, J. L., Frolking, S., and Moore, T. R.: Interannual variability in the peatland-atmosphere carbon dioxide exchange at an ombrotrophic bog, Global Biogeochemical Cycles, 17, 1036, doi:10.1029/2002GB001 983, 2003.

Lagergren, F., Lindroth, A., Dellwik, E., Ibrom, A., Lankreijer, H., Launiainen, S., Mölder, M., Kolari, P., Pilegaard, K., and Vesala, T.: Biophysical controls on CO_2 fluxes of three Northern forets based on long-term eddy covariance data, Tellus B, 60, 143–152, 2008.

Lasslop, G., Reichstein, M., Kattge, J., and Papale, D.: Influences of observation errors in eddy flux data on inverse model parameter estimation, Biogeosciences, 5, 1311–1324, 2008.

Lasslop, G., Reichstein, M., Detto, M., Richardson, A. D., and Baldocchi, D. D.: 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, 312–314, 2010a.

Lasslop, G., Reichstein, M., Papale, D., Richardson, A. D., Arneth, A., Barr, A., Stoy, P., and Wohlfahrt, G.: Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: critical issues and global evaluation, Global Change Biology, 16, 187–208, 2010b.

Law, B., Thornton, P., Irvine, J., Tuyl, S. V., and Anthoni, P.: Carbon storage and fluxes in ponderosa pine forests at different developmental stages, Global Change Biology, 7, 755–777, 2001.

Le Quere, C., Raupach, M. R., Canadell, J. G., Marland, G., Bopp, L., Ciais, P., Conway, T. J., Doney, S. C., Feely, R. A., Foster, P., Friedlingstein, P., Gurney, K., Houghton, R. A., House, J. I., Huntingford, C., Levy, P. E., Lomas, M. R., Majkut, J., Metzl, N., Ometto, J. P., Peters, G. P., Prentice, I. C., Randerson, J. T., Running, S. W., Sarmiento, J. L., Schuster, U., Sitch, S., Takahashi, T., Viovy, N., van der Werf, G. R., and Woodward, F. I.: Trends in the sources and sinks of carbon dioxide, Nature Geoscience, 2, 831–836, 2009.

Lee, X., Massman, W. J., and Law, B. E.: Handbook of micrometeorology: a guide for surface flux measurement and analysis, Kluwer, Dordrecht, 2004.

Lefsky, M. A., Cohen, W. B., Harding, D. J., Parker, G. G., Acker, S. A., and Gower, S. T.: Lidar remote sensing of above-ground biomass in three biomes, Global Ecology and Biogeography, 11, 393–399, 10.1046/j.1466-822x.2002.00303.x, 2002a.

Lefsky, M. A., Cohen, W. B., Parker, G. G., and Harding, D. J.: Lidar Remote Sensing for Ecosystem Studies, BioScience, 52, 19–30, 2002b.

Lemon, E. R.: Photosynthesis Under Field Conditions. II. An Aerodynamic Method for Determining the Turbulent Carbon Dioxide Exchange Between the Atmosphere and a Corn Field, Agron J, 52, 697–703, 1960.

Lenschow, D. H., Mann, J., and Kristensen, L.: How Long Is Long Enough When Measuring Fluxes and Other Turbulence Statistics?, Journal of Atmospheric and Oceanic Technology, 11, 661–673, 1994.

Leuning, R. and Moncrieff, J.: Eddy-covariance CO2 flux measurements using open- and closed-path CO2 analysers: Corrections for analyser water vapour sensitivity and damping of fluctuations in air sampling tubes, Boundary-Layer Meteorology, 53, 63–76, 1990.

Levenberg, K.: A Method for the Solution of Certain Problems in Least Squares, Quart. Appl. Math., 2, 164–168, 1944.

Lindroth, A., Klemedtsson, L., Grelle, A., Weslien, P., and Langvall, O.: Measurement of net ecosystem exchange, productivity and respiration in three spruce forests in Sweden shows unexpectedly large soil carbon losses, Biogeochemistry, 89, 43Ű60, doi:DOI10.1007/s10533-007-9137-8, 2008.

Lipson, D. A., Wilson, R. F., and Oechel, W. C.: Effects of elevated atmospheric CO2 on soil microbial biomass, activity, and diversity in a chaparral ecosystem, Applied and Environmental Microbiology, 71, 8573–8580, 2005.

Liu, H. P., Randerson, J. T., Lindfors, J., and Chapin, F. S.: Changes in the surface energy budget after fire in boreal ecosystems of interior Alaska: An annual perspective, Journal of Geophysical Research-Atmospheres, 110, D13 101, 2005.

Lloyd, J. and Taylor, J. A.: On the Temperature Dependence of Soil Respiration, Functional Ecology, 8, 315–323, 1994.

Loreto, F., Velikova, V., and Di Marco, G.: Respiation in the light measured by $^{1}2CO_2$ emission in $^{1}3CO_2$ atmosphere in maize leaves, Australian Journal of Plant Physiology, 28, 1103–1108, 2001.

Lucht, W., Prentice, I. C., Myneni, R. B., Sitch, S., Friedlingstein, P., Cramer, W., Bousquet, P., Buermann, W., and Smith, B.: Climatic control of the high-latitude vegetation greening trend and Pinatubo effect, Science, 296, 1687–1689, 2002.

Luo, Y. Q., Weng, E. S., Wu, X. W., Gao, C., Zhou, X. H., and Zhang, L.: Parameter identifiability, constraint, and equifinality in data assimilation with ecosystem models, Ecological Applications, 19, 571–574, 2009.

Luyssaert, S., Reichstein, M., Schulze, E. D., Janssens, I. A., Law, B. E., Papale, D., Dragoni, D., Goulden, M. L., Granier, A., Kutsch, W. L., Linder, S., Matteucci, G., Moors, E., Munger, J. W., Pilegaard, K., Saunders, M., and Falge, E. M.: Toward a consistency cross-check of eddy covariance flux based and biometric estimates of ecosystem carbon balance, Global Biogeochem. Cycles, 23, GB3009, doi:10.1029/2008GB003 377, 2009.

Ma, S. Y., Baldocchi, D. D., Xu, L. K., and Hehn, T.: Inter-annual variability in carbon dioxide exchange of an oak/grass savanna and open grassland in California, Agricultural and Forest Meteorology, 147, 157–171, 2007.

Mahecha, M. D., Reichstein, M., Lange, H., Carvalhais, N., Bernhofer, C., Grunwald, T., Papale, D., and Seufert, G.: Characterizing ecosystem-atmosphere interactions from short to interannual time scales, Biogeosciences, 4, 743–758, 2007.

Mahecha, M. D., Reichstein, M., Carvalhais, N., Lasslop, G., Lange, H., Seneviratne, S. I., Vargas, R., Ammann, C., Arain, M. A., Cescatti, A., Janssens, I. A., Migliavacca, M., Montagnani, L., and Richardson, A. D.: Global Convergence in the Temperature Sensitivity of Respiration at Ecosystem Level, Science, p. 10.1126/science.1189587, 2010.

Marcolla, B. and Cescatti, A.: Experimental analysis of flux footprint for varying stability conditions in an alpine meadow, Agricultural and Forest Meteorology, 135, 291–301, 2005.

Marcolla, B., Pitacco, A., and Cescatti, A.: Canopy Architecture and Turbulence Structure in a Coniferous Forest, Boundary-Layer Meteorology, 108, 39–59, 2003.

Marcolla, B., Cescatti, A., Montagnani, L., Manca, G., Kerschbaumer, G., and Minerbi, S.: Importance of advection in the atmospheric $CO2$ exchanges of an alpine forest, Agricultural and Forest Meteorology, 130, 193–206, 2005.

Marquard, D.: An Algorithm for Least-Squares Estimation of Nonlinear Parameters, SIAM J. Appl. Math., 11, 431–441, 1963.

Massman, W. J.: A simple method for estimating frequency response corrections for eddy covariance systems, Agricultural and Forest Meteorology, 104, 185–198, 2000.

Mauder, M., Liebethal, C., Göckede, M., Leps, J.-P., Beyrich, F., and Foken, T.: Processing and quality control of flux data during LITFASS-2003, Boundary-Layer Meteorology, 121, 67–88, 2006.

Mauder, M., Foken, T., Clement, R., Elbers, J. A., Eugster, W., Grunwald, T., Heusinkveld, B., and Kolle, O.: Quality control of CarboEurope flux data - Part 2: Inter-comparison of eddy-covariance software, Biogeosciences, 5, 451–462, 2008.

McCallum, A., Wagner, W., Schmullius, C., Shvidenko, A., Obersteiner, M., Fritz, S., and Nilsson, S.: Comparison of four global FAPAR datasets over Northern Eurasia for the year 2000, Remote Sensing of Environment, 114, 941–949, 2010.

McCallum, I., Wagner, W., Schmullius, C., Shvidenko, A., Obersteiner, M., Fritz, S., and Nilsson, S.: Satellite-based terrestrial production efficiency modeling, Carbon Balance and Management, 4, 8, 2009.

McMurtrie, R. E., Comins, H. N., Kirschbaum, M. U. F., and Wang, Y. P.: Modifying Existing Forest Growth-Models to Take Account of Effects of Elevated Co2, Australian Journal of Botany, 40, 657–677, 1992.

Meehl, G., Stocker, T., Collins, W., Friedlingstein, P., Gaye, A., Gregory, J., Kitoh, A., Knutti, R., Murphy, J., Noda, A., Raper, S., Watterson, I., Weaver, A., and Zhao, Z.-C.: Global Climate Projections, in: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, edited by Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K., Tignor, M., and Miller, H., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.

Merbold, L., Ziegler, W., Mukelabai, M. M., and Kutsch, W. L.: Spatial and temporal variation of CO_2 efflux along a disturbance gradient in a miombo woodland in Western Zambia, Biogeosciences Discussions, 7, 5757–5800, doi:10.5194/bgd-7-5757-2010, URL http://www.biogeosciences-discuss.net/7/5757/2010/, 2010.

Meyers, T. P. and Hollinger, S. E.: An assessment of storage terms in the surface energy balance of maize and soybean, Agricultural and Forest Meteorology, 125, 105–115, 2004.

Migliavacca, M., Meroni, M., Busetto, L., Colombo, R., Zenone, T., Matteucci, G., Manca, G., and Seufert, G.: Modeling Gross Primary Production of Agro-Forestry Ecosystems by Assimilation of Satellite-Derived Information in a Process-Based Model, Sensors, 9, 922–942, 2009.

Migliavacca, M., Reichstein, M., Richardson, A., Colombo, R., Sutton, M., Lasslop, G., Wohlfahrt, G., Tomelleri, E., Carvalhais, N., Cescatti, A., Mahecha, M., Montagnani, L., Papale, D., Zaehle, S., Arain, A., Arneth, A., Black, A., Carrara, A., Dore, S., Gianelle, D., Helfter, C., Hollinger, D., Kutsch, W., Lafleur, P., Nouvellon, Y., Rebmann, C., da Rocha, H., Rodeghiero, M., Roupsard, O., Sebastia, M.-T., Seufert, G., Soussana, J.-F., and van der Molen, M.: Semi-empirical modeling of abiotic and biotic factors controlling ecosystem respiration across eddy covariance sites, Global Change Biology, p. accepted, 2010.

Misson, L., Baldocchi, D. D., Black, T. A., Blanken, P. D., Brunet, Y., Curiel Yuste, J., Dorsey, J. R., Falk, M., Granier, A., Irvine, M. R., Jarosz, N., Lamaud, E., Launiainen, S., Law, B. E., Longdoz, B., Loustau, D., McKay, M., Paw U, K. T., Vesala, T., Vickers, D., Wilson, K. B., and Goldstein, A. H.: Partitioning forest carbon fluxes with overstory and understory eddy-covariance measurements: A synthesis based on FLUXNET data, Agricultural and Forest Meteorology, 144, 14–31, 2007.

Moffat, A. M., Papale, D., Reichstein, M., Hollinger, D. Y., Richardson, A. D., Barr, A. G., Beckstein, C., Braswell, B. H., Churkina, G., Desai, A. R., Falge, E., Gove, J. H., Heimann, M., Hui, D., Jarvis, A. J., Kattge, J., Noormets, A., and Stauch, V. J.: Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes, Agricultural and Forest Meteorology, 147, 209–232, 2007.

Moncrieff, J. B., Malhi, Y., and Leuning, R.: The propagation of errors in long-term measurements of land-atmosphere fluxes of carbon and water, Global Change Biology, 2, 231–240, 1996.

Monson, R. K., Turnipseed, A. A., Sparks, J. P., Harley, P. C., Scott-Denton, L. E., Sparks, K., and Huxman, T. E.: Carbon sequestration in a high-elevation, subalpine forest, Global Change Biology, 8, 459–478, 2002.

Montagnani, L., Manca, G., Canepa, E., Georgieva, E., Acosta, M., Feigenwinter, C., Janous, D., Kerschbaumer, G., Lindroth, A., Minach, L., Minerbi, S., Molder, M., Pavelka, M., Seufert, G., Zeri, M., and Ziegler, W.: A new mass conservation approach to the study of CO2 advection in an alpine forest, Journal of Geophysical Research-Atmospheres, 114, D07 306, doi:10.1029/2008JD010 650, 2009.

Monteith, J. L.: Evaporation and environment, Symposium of the Society for Experimental Biology, 19, 205–234, 1965.

Moore, C. J.: Frequency-Response Corrections for Eddy-Correlation Systems, Boundary-Layer Meteorology, 37, 17–35, 1986.

Moureaux, C., Debacq, A., Bodson, B., Heinesch, B., and Aubinet, M.: Annual net ecosystem carbon exchange by a sugar beet crop, Agricultural and Forest Meteorology, 139, 25–39, 2006.

Mu, Q., Zhao, M., Heinsch, F. A., Liu, M., Tian, H., and Running, S. W.: Evaluating water stress controls on primary production in biogeochemical and remote sensing based models, J. Geophys. Res., 112, G01 012,doi:10.1029/2006JG000 179, 2007.

Muraoka, H., Saigusa, N., Nasahara, K., Noda, H., Yoshino, J., Saitoh, T., Nagai, S., Murayama, S., and Koizumi, H.: Effects of seasonal and interannual variations in leaf photosynthesis and canopy leaf area index on gross primary production of a cool-temperate deciduous broadleaf forest in Takayama, Japan, Journal of Plant Research, 123, 563–576, 2010.

Murray, M. B., Cannell, M. G. R., and Smith, R. I.: Date of Budburst of 15 Tree Species in Britain Following Climatic Warming, Journal of Applied Ecology, 26, 693–700, ak524 Times Cited:175 Cited References Count:9, 1989.

Nagy, Z., Pinter, K., Czobel, S., Balogh, J., Horvath, L., Foti, S., Barcza, Z., Weidinger, T., Csintalan, Z., Dinh, N. Q., Grosz, B., and Tuba, Z.: The carbon budget of semi-arid grassland in a wet and a dry year in Hungary, Agriculture Ecosystems & Environment, 121, 21–29, 2007.

Nardini, A., Salleo, S., and Andri, S.: Circadian regulation of leaf hydraulic conductance in sunflower (Helianthus annuus L. cv Margot), Plant Cell and Environment, 28, 750–759, 2005.

Niyogi, D., Chang, H. I., Saxena, V. K., Holt, T., Alapaty, K., Booker, F., Chen, F., Davis, K. J., Holben, B., Matsui, T., Meyers, T., Oechel, W. C., Pielke, R. A., Wells, R., Wilson, K., and Xue, Y. K.: Direct observations of the effects of aerosol loading on net ecosystem CO2 exchanges over different landscapes, Geophysical Research Letters, 31, L20 506, doi:10.1029/2004GL020 915, 2004.

Noormets, A., Gavazzi, M. J., Mcnulty, S. G., Domec, J.-C., N., G. S. U., King, J. S., and Chen, J.: Response of carbon fluxes to drought in a coastal plain loblolly pine forest, Global Change Biology, 16, 272 – 287, doi:10.1111/j.1365-2486.2009.01928.x, 2009.

Omlin, M. and Reichert, P.: A comparison of techniques for the estimation of model prediction uncertainty, Ecological Modelling, 115, 45–59, 1999.

Owen, K. E., Tenhunen, J., Reichstein, M., Wang, Q., Falge, E., Geyer, R., Xiao, X., Stoy, P., Ammann, C., Arain, A., Aubinet, M., Aurela, M., Bernhofer, C., Chojnicki, B. H., Granier, A., Gruenwald, T., Hadley, J., Heinesch, B., Hollinger, D., Knohl, A., Kutsch, W., Lohila, A., Meyers, T., Moors, E., Moureaux, C., Pilegaard, K., Saigusa, N., Verma, S., Vesala, T., and Vogel, C.: Linking flux network measurements to continental scale simulations: ecosystem carbon dioxide exchange capacity under non-water-stressed conditions, Global Change Biology, 13, 734–760, 2007.

Papale, D. and Valentini, A.: A new assessment of European forests carbon exchanges by eddy fluxes and artificial neural network spatialization, Global Change Biology, 9, 525–535, 2003.

Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W., Longdoz, B., Rambal, S., Valentini, R., Vesala, T., and Yakir, D.: Towards a standardized processing of Net Ecosystem

Exchange measured with eddy covariance technique: algorithms and uncertainty estimation, Biogeosciences, 3, 571–583, 2006.

Pataki, D. and Oren, R.: Species differences in stomatal control of water loss at the canopy scale in a mature bottomland deciduous forest, Advances in Water Resources, 26, 1267–1278, 2003.

Pereira, J. S., Mateus, J. A., Aires, L. M., Pita, G., Pio, C., David, J. S., Andrade, V., Banza, J., David, T. S., Paco, T. A., and Rodrigues, A.: Net ecosystem carbon exchange in three contrasting Mediterranean ecosystems - the effect of drought, Biogeosciences, 4, 791–802, 2007.

Peñuelas, J., Gamon, J. A., Fredeen, A. L., Merino, J., and Field, C. B.: Reflectance indices associated with physiological changes in nitrogen- and water-limited sunflower leaves, Remote Sensing of Environment, 48, 135–146, dOI: 10.1016/0034-4257(94)90136-8, 1994.

Piao, S. L., Ciais, P., Friedlingstein, P., Peylin, P., Reichstein, M., Luyssaert, S., Margolis, H., Fang, J. Y., Barr, A., Chen, A. P., Grelle, A., Hollinger, D. Y., Laurila, T., Lindroth, A., Richardson, A. D., and Vesala, T.: Net carbon dioxide losses of northern ecosystems in response to autumn warming, Nature, 451, 49–53, 2008.

Pilegaard, K., Mikkelsen, T. N., Beier, C., Jensen, N. O., Ambus, P., and Ro-Poulsen, H.: Field measurements of atmosphere-biosphere interactions in a Danish beech forest, Boreal Environment Research, 8, 315–333, 2003.

Pinelli, P. and Loreto, F.: 12CO2 emission from different metabolic pathways measured in illuminated and darkened C3 and C4 leaves at low, atmospheric and elevated CO2 concentration, Journal of Experimental Botany, 54, 1761–1769, 2003.

Pinter, K., Barcza, Z., Balogh, J., Czobel, S., Csintalan, Z., Tuba, Z., and Nagy, Z.: Interannual variability of grasslands' carbon balance depends on soil type, Community Ecology, 9, 43–48, 2008.

Potts, D. L., Scott, R. L., Cable, J. M., Huxman, T. E., and Williams, D. G.: Sensitivity of Mesquite Shrubland Co2 Exchange to Precipitation in Contrasting Landscape Settings, Ecology, 89, 2900–2910, 2008.

Powell, T. L., Bracho, R., Li, J. H., Dore, S., Hinkle, C. R., and Drake, B. G.: Environmental controls over net ecosystem carbon exchange of scrub oak in central Florida, Agricultural and Forest Meteorology, 141, 19–34, 2006.

Powell, T. L., Gholz, H. L., Clark, K. L., Starr, G., Cropper, W. P., and Martin, T. A.: Carbon exchange of a mature, naturally regenerated pine forest in north Florida, Global Change Biology, 14, 2523–2538, 2008.

Qin, Z. and Karnieli, A.: Progress in the remote sensing of land surface temperature and ground emissivity using NOAA-AVHRR data, International Journal of Remote Sensing, 20, 2367 – 2393, 1999.

Raddatz, T. J., Reick, C. H., Knorr, W., Kattge, J., Roeckner, E., Schnur, R., Schnitzler, K. G., Wetzel, P., and Jungclaus, J.: Will the tropical land biosphere dominate the climate-carbon cycle feedback during the twenty-first century?, Climate Dynamics, 29, 565–574, 2007.

Rambal, S., Joffre, R., Ourcival, J. M., Cavender-Bares, J., and Rocheteau, A.: The growth respiration component in eddy CO2 flux from a Quercus ilex mediterranean forest, Global Change Biology, 10, 1460–1469, 2004.

Raupach, M. R., Rayner, P. J., Barrett, D. J., DeFries, R. S., Heimann, M., Ojima, D. S., Quegan, S., and Schmullius, C. C.: Model-data synthesis in terrestrial carbon observation: methods, data requirements and data uncertainty specifications, Global Change Biology, 11, 378–397, 2005.

Rayner, P. J., Scholze, M., Knorr, W., Kaminski, T., Giering, R., and Widmann, H.: Two decades of terrestrial carbon fluxes from a carbon cycle data assimilation system (CCDAS), Global Biogeochemical Cycles, 19, GB2026, DOI: 10.1029/2004GB002 254, 2005.

Rebmann, C., Gockede, M., Foken, T., Aubinet, M., Aurela, M., Berbigier, P., Bernhofer, C., Buchmann, N., Carrara, A., Cescatti, A., Ceulemans, R., Clement, R., Elbers, J. A., Granier, A., Grunwald, T., Guyon, D., Havrankova, K., Heinesch, B., Knohl, A., Laurila, T., Longdoz, B., Marcolla, B., Markkanen, T., Miglietta, F., Moncrieff, J., Montagnani, L., Moors, E., Nardino, M., Ourcival, J. M., Rambal, S., Rannik, U., Rotenberg, E., Sedlak, P., Unterhuber, G., Vesala, T., and Yakir, D.: Quality analysis applied on eddy covariance measurements at complex forest sites using footprint modelling, Theoretical and Applied Climatology, 80, 121–141, 2005.

Rebmann, C., Zeri, M., Lasslop, G., Mund, M., Kolle, O., Schulze, E.-D., and Feigenwinter, C.: Treatment and assessment of the CO2-exchange at a complex forest site in Thuringia âĂŞ Germany, Agricultural and Forest Meteorology, 150, 684–691, 2010.

Reichstein, M., Tenhunen, J. D., Roupsard, O., Ourcival, J. M., Rambal, S., Miglietta, F., Peressotti, A., Pecchiari, M., Tirone, G., and Valentini, R.: Severe drought effects on ecosystem CO2 and H2O fluxes at three Mediterranean evergreen sites: revision of current hypotheses?, Global Change Biology, 8, 999–1017, 2002.

Reichstein, M., Rey, A., Freibauer, A., Tenhunen, J., Valentini, R., Banza, J., Casals, P., Cheng, Y. F., Grunzweig, J. M., Irvine, J., Joffre, R., Law, B. E., Loustau, D., Miglietta, F., Oechel, W., Ourcival, J. M., Pereira, J. S., Peressotti, A., Ponti, F., Qi, Y., Rambal, S., Rayment, M., Romanya, J., Rossi, F., Tedeschi, V., Tirone, G., Xu, M., and Yakir, D.: Modeling temporal and large-scale spatial variability of soil respiration from soil water availability, temperature and vegetation productivity indices, Global Biogeochemical Cycles, 17, 1104, doi:10.1029/2003GB002 035, 2003a.

Reichstein, M., Tenhunen, J., Roupsard, O., Ourcival, J. M., Rambal, S., Miglietta, F., Peressotti, A., Pecchiari, M., Tirone, G., and Valentini, R.: Inverse modeling of seasonal drought effects on canopy CO2/H2O exchange in three Mediterranean ecosystems, Journal of Geophysical Research-Atmospheres, 108, doi:10.1029/2003JD003 430, 2003b.

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grunwald, T., Havrankova, K., Ilvesniemi, H., Janous, D., Knohl, A., Laurila, T., Lohila, A., Loustau, D., Matteucci, G., Meyers, T., Miglietta, F., Ourcival, J. M., Pumpanen, J., Rambal, S., Rotenberg, E., Sanz, M., Tenhunen, J., Seufert, G., Vaccari, F., Vesala, T., Yakir, D., and Valentini, R.: On the separation of net ecosystem exchange into assimilation

and ecosystem respiration: review and improved algorithm, Global Change Biology, 11, 1424–1439, 2005.

Reichstein, M., Papale, D., Valentini, R., Aubinet, M., Bernhofer, C., Knohl, A., Laurila, T., Lindroth, A., Moors, E., Pilegaard, K., and Seufert, G.: Determinants of terrestrial ecosystem carbon balance inferred from European eddy covariance flux sites, Geophysical Research Letters, 34, L01 402.262,doi:10.1029/2006GL027 880, 2007.

Rey, A., Pegoraro, E., Tedeschi, V., Parri, I. D., Jarvis, P., and Valentini, R.: Annual variation in soil respiration and its components in a coppice oak forest in central italy, Global Change Biology, 9, 851–866, 2002.

Reynolds, O.: On the dynamical theory of incompressible viscous fluids and the determination of the criterion, Philosophical Transactions of the royal society, 186, 123–164, 1895.

Rüger, B.: Induktive Statistik: EinfÃijhrung fÃijr Wirtschafts- u. Sozialwissenschaftler, Oldenbourg Verlag, München ; Wien, 3 edn., 1996.

Richardson, A., Williams, M., Hollinger, D., Moore, D., Dail, D., Davidson, E., Scott, N., Evans, R., Hughes, H., Lee, J., Rodrigues, C., and Savage, K.: Estimating parameters of a forest ecosystem C model with measurements of stocks and fluxes as joint constraints, Oecologia, pp. DOI 10.1007/s00 442–010–1628–y, 2010.

Richardson, A. D. and Hollinger, D. Y.: Statistical modeling of ecosystem respiration using eddy covariance data: Maximum likelihood parameter estimation, and Monte Carlo simulation of model and parameter uncertainty, applied to three simple models, Agricultural and Forest Meteorology, 131, 191–208, 2005.

Richardson, A. D. and Hollinger, D. Y.: A method to estimate the additional uncertainty in gap-filled NEE resulting from long gaps in the CO2 flux record, Agricultural and Forest Meteorology, 147, 199–208, 2007.

Richardson, A. D., Braswell, B. H., Hollinger, D. Y., Burman, P., Davidson, E. A., Evans, R. S., Flanagan, L. B., Munger, J. W., Savage, K., Urbanski, S. P., and Wofsy, S. C.: Comparing simple respiration models for eddy flux and dynamic chamber data, Agricultural and Forest Meteorology, 141, 219–234, 2006a.

Richardson, A. D., Hollinger, D. Y., Burba, G. G., Davis, K. J., Flanagan, L. B., Katul, G. G., Munger, J. W., Ricciuto, D. M., Stoy, P. C., Suyker, A. E., Verma, S. B., and Wofsy, S. C.: A multi-site analysis of random error in tower-based measurements of carbon and energy fluxes, Agricultural and Forest Meteorology, 136, 1–18, 2006b.

Richardson, A. D., Hollinger, D. Y., Aber, J. D., Ollinger, S. V., and Braswell, B. H.: Environmental variation is directly responsible for short- but not long-term variation in forest-atmosphere carbon exchange, Global Change Biology, 13, 788–803, 2007.

Richardson, A. D., Mahecha, M. D., Falge, E., Kattge, J., Moffat, A. M., Papale, D., Reichstein, M., Stauch, V. J., Braswell, B. H., Churkina, G., Kruijt, B., and Hollinger, D. Y.: Statistical properties of random CO2 flux measurement uncertainty inferred from model residuals, Agricultural and Forest Meteorology, 148, 38–50, 2008.

Robinson, J. M.: Fire from space: Global fire evaluation using infrared remote sensing, International Journal of Remote Sensing, 12, 3 – 24, 1991.

Rodriguez, D. and Sadras, V. O.: The limit to wheat water-use efficiency in eastern Australia. I. Gradients in the radiation environment and atmospheric demand, Australian Journal of Agricultural Research, 58, 287–302, 2007.

Roughgarden, J., Running, S. W., and Matson, P. A.: What Does Remote Sensing Do For Ecology?, Ecology, 72, 1918–1922, 1991.

Roupsard, O., Bonnefond, J. M., Irvine, M., Berbigier, P., Nouvellon, Y., Dauzat, J., Taga, S., Hamel, O., Jourdan, C., Saint-Andre, L., Mialet-Serra, I., Labouisse, J. P., Epron, D., Joffre, R., Braconnier, S., Rouziere, A., Navarro, M., and Bouillet, J. P.: Partitioning energy and evapo-transpiration above and below a tropical palm canopy, Agricultural and Forest Meteorology, 139, 252–268, 2006.

Rutherford: Data Assimilation by Statistical Interpolation of Forecast Error Fields, Journal of the Atmospheric Sciences, 29, 809–815, 1972.

Rutledge, S., Campbell, D. I., Baldocchi, D. D., and Schipper, L. A.: Photodegradation leads to increased CO2 losses from terrestrial organic matter, Global Change Biology, pp. 10.1111/j.1365–2486.2009.02 149.x, 2010.

Sacks, W. J., Schimel, D. S., Monson, R. K., and Braswell, B. H.: Model-data synthesis of diurnal and seasonal CO2 fluxes at Niwot Ridge, Colorado, Global Change Biology, 12, 240–259, 2006.

Sagerfors, J., Lindroth, A., Grelle, A., Klemedtsson, L., Weslien, P., and Nilsson, M.: Annual CO2 exchange between a nutrient-poor, minerotrophic, boreal mire and the atmosphere, J. Geophys. Res., 113, G01 001, doi:10.1029/2006JG000 306, 2008.

Saito, M., Miyata, A., Nagai, H., and Yamada, T.: Seasonal variation of carbon dioxide exchange in rice paddy field in Japan, Agricultural and Forest Meteorology, 135, 93–109, 2005.

Sakai, R. K., Fitzjarrald, D. R., Moraes, O. L. L., Staebler, R. M., Acevedo, O. C., Czikowsky, M. J., Da Silva, R., Brait, E., and Miranda, V.: Land-use change effects on local energy, water, and carbon balances in an Amazonian agricultural field, Global Change Biology, 10, 895–907, 2004.

Saleska, S. R., Miller, S. D., Matross, D. M., Goulden, M. L., Wofsy, S. C., da Rocha, H. R., de Camargo, P. B., Crill, P., Daube, B. C., de Freitas, H. C., Hutyra, L., Keller, M., Kirchhoff, V., Menton, M., Munger, J. W., Pyle, E. H., Rice, A. H., and Silva, H.: Carbon in amazon forests: Unexpected seasonal fluxes and disturbance-induced losses, Science, 302, 1554–1557, 2003.

Santos, A. J. B., Quesada, C. A., Da Silva, G. T., Maia, J. F., Miranda, H. S., Miranda, A. C., and Lloyd, J.: High rates of net ecosystem carbon assimilation by Brachiara pasture in the Brazilian Cerrado, Global Change Biology, 10, 877–885, 2004.

Schimel, D. S., House, J. I., Hibbard, K. A., Bousquet, P., Ciais, P., Peylin, P., Braswell, B. H., Apps, M. J., Baker, D., Bondeau, A., Canadell, J., Churkina, G., Cramer, W., Denning, A. S., Field, C. B., Friedlingstein, P., Goodale, C., Heimann, M., Houghton, R. A., Melillo, J. M., Moore, B., Murdiyarso, D., Noble, I., Pacala, S. W., Prentice, I. C., Raupach, M. R., Rayner, P. J., Scholes, R. J., Steffen, W. L., and Wirth, C.: Recent patterns and mechanisms of carbon exchange by terrestrial ecosystems, Nature, 414, 169–172, 2001.

Schindler, D., Turk, M., and Mayer, H.: CO2 fluxes of a Scots pine forest growing in the warm and dry southern upper Rhine plain, SW Germany, European Journal of Forest Research, 125, 201–212, 2006.

Schluessel, P., Emery, W. J., Grassl, H., and Mammen, T.: On the Bulk-Skin Temperature Difference and Its Impact on Satellite Remote Sensing of Sea Surface Temperature, J. Geophys. Res., 95, 13 341–13 356, 1990.

Schmid, H. P., Grimmond, C. S. B., Cropley, F., Offerle, B., and Su, H. B.: Measurements of CO2 and energy fluxes over a mixed hardwood forest in the mid-western United States, Agricultural and Forest Meteorology, 103, 357–374, 2000.

Schulze, E. D.: Der CO2-Gaswechsel der Buche (Fagus silvatica L.) in Abhängigkeit von den Klimafaktoren im Freiland, Flora, 159, 177 – 232, 1970.

Schulze, E.-D.: Biological control of the terrestrial carbon sink, Biogeosciences, 3, 147–166, 2006.

Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J. O., Levis, S., Lucht, W., Sykes, M. T., Thonicke, K., and Venevsky, S.: Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model, Global Change Biology, 9, 161–185, 2003.

Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K., Tignor, M., and Miller, H.: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.

Staudt, K. and Foken, T.: Documentation of reference data for the experimental areas of the Bayreuth Centre for Ecology and Environmental Research (BayCEER) at the Waldstein site, Tech. Rep. 35, University of Bayreuth, 2007.

Stoy, P. C., Katul, G. G., Siqueira, M. B. S., Juang, J.-Y., Novick, K. A., Uebelherr, J. M., and Oren, R.: An evaluation of models for partitioning eddy covariance-measured net ecosystem exchange into photosynthesis and respiration, Agricultural and Forest Meteorology, 141, 2–18, 2006.

Stoy, P. C., Richardson, A. D., Baldocchi, D. D., Katul, G. G., Stanovick, J., Mahecha, M. D., Reich-
    stein, M., Detto, M., Law, B. E., Wohlfahrt, G., Arriga, N., Campos, J., McCaughey, J. H., Montag-
    nani, L., Paw U, K. T., Sevanto, S., and Williams, M.: Biosphere-atmosphere exchange of CO2 in
    relation to climate: a cross-biome analysis across multiple time scales, Biogeosciences Discuss., 6,
    4095–4141, 2009.

Sun, G., Noormets, A., Gavazzi, M. J., McNulty, S. G., Chen, J., Domec, J. C., King, J. S., Amatya,
    D. M., and Skaggs, R. W.: Energy and water balance of two contrasting loblolly pine plantations on
    the lower coastal plain of North Carolina, USA, Forest Ecology and Management, 259, 1299–1310,
    doi: DOI: 10.1016/j.foreco.2009.09.016, 2010.

Suni, T., Berninger, F., Vesala, T., Markkanen, T., Hari, P., Makela, A., Ilvesniemi, H., Hanninen, H.,
    Nikinmaa, E., Huttula, T., Laurila, T., Aurela, M., Grelle, A., A., L., Arneth, A., Shibistova, O., and
    Lloyd, J.: Air temperature triggers the commencement of evergreen boreal forest photosynthesis in
    spring, Global Change Biology, 9, 1410–1426, 2003a.

Suni, T., Rinne, J., Reissell, A., Altimir, N., Keronen, P., Rannik, U., Dal Maso, M., Kulmala, M., and
    Vesala, T.: Long-term measurements of surface fluxes above a Scots pine forest in Hyytiala, southern
    Finland, 1996-2001, Boreal Environment Research, 8, 287–301, 2003b.

Swinbank, W. C.: The Measurement of Vertical Transfer of Heat and Water Vapor by Eddies in the
    Lower Atmosphere, Journal of Meteorology, 8, 135–145, 1951.

Syed, K. H., Flanagan, L. B., Carlson, P. J., Glenn, A. J., and Van Gaalen, K. E.: Environmental control
    of net ecosystem CO2 exchange in a treed, moderately rich fen in northern Alberta, Agricultural and
    Forest Meteorology, 140, 97–114, 2006.

Takagi, K., Fukuzawa, K., Liang, N., Kayama, M., Nomura, M., Hojyo, H., Sugata, S., Shibata, H.,
    Fukazawa, T., Takahashi, Y., Nakaji, T., Oguma, H., Mano, M., Akibayashi, Y., Murayama, T.,
    Koike, T., Sasa, K., and Fujinuma, Y.: Change in CO2 balance under a series of forestry activities
    in a cool-temperate mixed forest with dense undergrowth, Global Change Biology, 15, 1275–1288,
    2009.

Tang, J. W., Baldocchi, D. D., and Xu, L.: Tree photosynthesis modulates soil respiration on a diurnal
    time scale, Global Change Biology, 11, 1298–1304, 2005.

Tarantola, A.: Inverse Problem Theory, Elsevier, Amsterdam, The Netherlands, 1987.

Tedeschi, V., Rey, A., Manca, G., Valentini, R., Jarvis, P. G., and Borghetti, M.: Soil respiration in a
    Mediterranean oak forest at different developmental stages after coppicing, Global Change Biology,
    12, 110–121, 2006.

Trudinger, C. M., Raupach, M. R., Rayner, P. J., Kattge, J., Li, Q., Pak, B., Reichstein, M., Renzullo, L.,
    Richardson, A. D., Roxburgh, S. H., Styles, J., Wang, Y. P., Briggs, P., Barrett, D., and Nikolova, S.:
    OptIC project: An intercomparison of optimization techniques for parameter estimation in terrestrial
    biogeochemical models, Journal of Geophysical Research, Volume 112, Issue G2, CiteID G02027,
    112, doi:10.1029/2006JG000 367, 2007.

Tucker, C. J., Fung, I. Y., Keeling, C. D., and Gammon, R. H.: Relationship between Atmospheric Co2 Variations and a Satellite-Derived Vegetation Index, Nature, 319, 195–199, 1986.

Urbanski, S., Barford, C., Wofsy, S., Kucharik, C., Pyle, E., Budney, J., McKain, K., Fitzjarrald, D., Czikowsky, M., and Munger, J. W.: Factors controlling CO2 exchange on timescales from hourly to decadal at Harvard Forest, J. Geophys. Res., 112, G02 020, doi:10.1029/2006JG000 293, 2007.

Van Dijk, A. I. J. M. and Dolman, A. J.: Estimates of CO2 uptake and release among European forests based on eddy covariance data, Global Change Biology, 10, 1445–1459, 2004.

van Gorsel, E., Leuning, R., Cleugh, H. A., Keith, H., Kirschbaum, M. U. F., and Suni, T.: Application of an alternative method to derive reliable estimates of nighttime respiration from eddy covariance measurements in moderately complex topography, Agricultural and Forest Meteorology, 148, 1174–1180, 2008.

Vargas, R. and Allen, M. F.: Environmental controls and the influence of vegetation type, fine roots and rhizomorphs on diel and seasonal variation in soil respiration, New Phytologist, 179, 460–471, 2008.

Vasquez, V. R. and Whiting, W. B.: Accounting for both random errors and systematic errors in uncertainty propagation analysis of computer models involving experimental measurements with Monte Carlo methods, Risk Analysis, 25, 1669–1681, 2006.

Veenendaal, E. M., Kolle, O., and Lloyd, J.: Seasonal variation in energy fluxes and carbon dioxide exchange for a broad-leaved semi-arid savanna (Mopane woodland) in Southern Africa, Global Change Biology, 10, 318–328, 2004.

Verma, S. B., Dobermann, A., Cassman, K. G., Walters, D. T., Knops, J. M., Arkebauer, T. J., Suyker, A. E., Burba, G. G., Amos, B., Yang, H. S., Ginting, D., Hubbard, K. G., Gitelson, A. A., and Walter-Shea, E. A.: Annual carbon dioxide exchange in irrigated and rainfed maize-based agroecosystems, Agricultural and Forest Meteorology, 131, 77–96, 2005.

Vickers, D., Thomas, C. K., Martin, J. G., and Law, B.: Self-correlation between assimilation and respiration resulting from flux partitioning of eddy-covariance CO2 fluxes, Agricultural and Forest Meteorology, 149, 1552–1555, 2009.

Visual Numerics, I.: PV-Wave 8.5 Reference Guide, http://www.vni.com/books/dod/pdf/wave85Docs/eReferenceGuide85.pdf, last access: 18 December 2008, 2005.

Wang, X., Wang, C., and Yu, G.: Spatio-temporal patterns of forest carbon dioxide exchange based on global eddy covariance measurements, Science in China Series D: Earth Sciences, 51, 1129–1143, 2008.

Wang, Y. P., Baldocchi, D., Leuning, R., Falge, E., and Vesala, T.: Estimating parameters in a land-surface model by applying nonlinear inversion to eddy covariance flux measurements from eight FLUXNET sites, Global Change Biology, 13, 652–670, 2007.

Wang, Y.-P., Trudinger, C. M., and Enting, I. G.: A review of applications of model-data fusion to studies of terrestrial carbon fluxes at different scales, Agricultural and Forest Meteorology, 149, 1829–1842, doi: DOI: 10.1016/j.agrformet.2009.07.009, 2009.

Webb, E. K., Pearman, G. I., and Leuning, R.: Correction of Flux Measurements for Density Effects Due to Heat and Water-Vapor Transfer, Quarterly Journal of the Royal Meteorological Society, 106, 85–100, 1980.

Wilczak, J. M., Oncley, S. P., and Stage, S. A.: Sonic anemometer tilt correction algorithms, Boundary-Layer Meteorology, 99, 127–150, 2001.

Wilks, D. S.: Statistical Methods in the Atmospheric Sciences, Academic Press, 1995.

Williams, M., Rastetter, E. B., Fernandes, D. N., Goulden, M. L., Wofsy, S. C., Shaver, G. R., Melillo, J. M., Munger, J. W., Fan, S. M., and Nadelhoffer, K. J.: Modelling the soil-plant-atmosphere continuum in a Quercus-Acer stand at Harvard forest: The regulation of stomatal conductance by light, nitrogen and soil/plant hydraulic properties, Plant Cell and Environment, 19, 911–927, 1996.

Williams, M., Richardson, A. D., Reichstein, M., Stoy, P. C., Peylin, P., Verbeeck, H., Carvalhais, N., Jung, M., Hollinger, D. Y., Kattge, J., Leuning, R., Luo, Y., Tomelleri, E., Trudinger, C. M., and Wang, Y. P.: Improving land surface models with FLUXNET data, Biogeosciences, 6, 1341–1459, 2009.

Williams, W. E. and Gorton, H. L.: Circadian rhythms have insignificant effects on plant gas exchange under field conditions, Physiologia Plantarum, 103, 247–256, 1998.

Wilson, K., Goldstein, A., Falge, E., Aubinet, M., Baldocchi, D., Berbigier, P., Bernhofer, C., Ceulemans, R., Dolman, H., Field, C., Grelle, A., Ibrom, A., Law, B. E., Kowalski, A., Meyers, T., Moncrieff, J., Monson, R., Oechel, W., Tenhunen, J., Valentini, R., and Verma, S.: Energy balance closure at FLUXNET sites, Agricultural and Forest Meteorology, 113, 223–243, 2002.

Wilson, K. B. and Baldocchi, D. D.: Comparing independent estimates of carbon dioxide exchange over 5 years at a deciduous forest in the southeastern United States, J. Geophys. Res., 106, 34 167–34 178, 2001.

Wofsy, S. C., Goulden, M. L., Munger, J. W., Fan, S. M., Bakwin, P. S., Daube, B. C., Bassow, S. L., and Bazzaz, F. A.: Net Exchange of CO2 in a Mid-Latitude Forest, Science, 260, 1314–1317, 1993.

Wohlfahrt, G., Hammerle, A., Haslwanter, A., Bahn, M., Tappeiner, U., and Cernusca, A.: Disentangling leaf area and environmental effects on the response of the net ecosystem CO2 exchange to diffuse radiation, Geophysical Research Letters, 35, L16 805, doi:10.1029/2008GL035 090, 2008a.

Wohlfahrt, G., Hammerle, A., Haslwanter, A., Bahn, M., Tappeiner, U., and Cernusca, A.: Seasonal and inter-annual variability of the net ecosystem CO2 exchange of a temperate mountain grassland: Effects of weather and management, Journal of Geophysical Research-Atmospheres, 113, D08 110, doi:10.1029/2007JD009 286, 2008b.

Wood, S. A., Beringer, J., Hutley, L. B., Mcguire, A. D., Van Dijk, A., and Kilinc, M.: Impacts of fire on forest age and runoff in mountain ash forests, Functional Plant Biology, 35, 483–492, 2008.

Xu, L. and Baldocchi, D. D.: Seasonal variation in carbon dioxide exchange over a Mediterranean annual grassland in California, Agricultural and Forest Meteorology, 123, 79–96, 2004.

Yan, Y., Zhao, B., Chen, J. Q., Guo, H. Q., Gu, Y. J., Wu, Q. H., and Li, B.: Closing the carbon budget of estuarine wetlands with tower-based measurements and MODIS time series, Global Change Biology, 14, 1690–1702, 2008.

Yi, C. X., Li, R. Z., Bakwin, P. S., Desai, A., Ricciuto, D. M., Burns, S. P., Turnipseed, A. A., Wofsy, S. C., Munger, J. W., Wilson, K., and Monson, R. K.: A nonparametric method for separating photosynthesis and respiration components in CO2 flux measurements, Geophysical Research Letters, 31, L17 107, 2004.

Yokota, T., Yoshida, Y., Eguchi, N., Ota, Y., Tanaka, T., Watanabe, H., and Maksyutov, S.: Global Concentrations of CO2 and CH4 Retrieved from GOSAT: First Preliminary Results, Sola, 5, 160–163, 2009.

Zha, T., Barr, A. G., Black, T. A., McCaughey, J. H., Bhatti, J., Hawthorne, I., Krishnan, P., Kidston, J., Saigusa, N., Shashkov, A., and Nesic, Z.: Carbon sequestration in boreal jack pine stands following harvesting, Global Change Biology, 15, 1475–1487, 2009.

C. Acknowledgements