Biogeosciences Discuss., 12, 9393–9441, 2015
www.biogeosciences-discuss.net/12/9393/2015/
doi:10.5194/bgd-12-9393-2015

Biogeosciences Discussions
Open Access

An objective prior error quantification for regional atmospheric inverse applications

P. Kountouris[1], C. Gerbig[1], K.-U. Totsche[2], A.-J. Dolman[3], A.-G.-C.-A. Meesters[3], G. Broquet[4], F. Maignan[4], B. Gioli[5], L. Montagnani[6,7], and C. Helfter[8]

[1]Max Planck Institute for Biogeochemistry, Jena, Germany
[2]Institute of Geosciences, Department of Hydrology, Friedrich-Schiller-Universitaet, Jena, Germany
[3]VU University Amsterdam, Amsterdam, the Netherlands
[4]Laboratoire des Sciences du Climat et de l'Environnement, CEA-CNRS-UVSQ, UMR8212, IPSL, Gif-sur-Yvette, France
[5]Institute of Biometeorology, IBIMET CNR, Firenze, Italy
[6]Faculty of Science and Technology, Free University of Bolzano, Piazza Università 5, 39100 Bolzano, Italy
[7]Forest Services, Autonomous Province of Bolzano, Via Brennero 6, 39100 Bolzano, Italy
[8]Centre for Ecology and Hydrology, Edinburgh, UK

Received: 05 May 2015 – Accepted: 02 June 2015 – Published: 23 June 2015

Correspondence to: P. Kountouris (pkount@bgc-jena.mpg.de)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Assigning proper prior uncertainties for inverse modeling of CO_2 is of high importance, both to regularize the otherwise ill-constrained inverse problem, and to quantitatively characterize the magnitude and structure of the error between prior and "true" flux. We use surface fluxes derived from three biosphere models VPRM, ORCHIDEE, and 5PM, and compare them against daily averaged fluxes from 53 Eddy Covariance sites across Europe for the year 2007, and against repeated aircraft flux measurements encompassing spatial transects. In addition we create synthetic observations to substitute observed by modeled fluxes to explore the potential to infer prior uncertainties from model-model residuals. To ensure the realism of the synthetic data analysis, a random measurement noise was added to the tower fluxes which were used as reference. The temporal autocorrelation time for tower model-data residuals was found to be around 35 days for both VPRM and ORCHIDEE, but significantly different for the 5PM model with 76 days. This difference is caused by a few sites with large model-data bias. The spatial correlation of the model-data residuals for all models was found to be very short, up to few tens of km. Long spatial correlation lengths up to several hundreds of km were determined when synthetic data were used. Results from repeated aircraft transects in south-western France, are consistent with those obtained from the tower sites in terms of spatial autocorrelation (35 km on average) while temporal autocorrelation is markedly lower (13 days). Our findings suggest that the different prior models have a common temporal error structure. Separating the analysis of the statistics for the model data residuals by seasons did not result in any significant differences of the spatial correlation lengths.

BGD
12, 9393–9441, 2015

1 Introduction

Atmospheric inversions are widely used to infer surface CO_2 fluxes from observed CO_2 dry mole fractions with a Bayesian approach (Ciais et al., 2000; Gurney et al., 2002; Lauvaux et al., 2008). In this approach a limited number of observations of atmospheric CO_2 mixing ratios are used to solve for generally a much larger number of unknowns, making this an ill-posed problem. By using prior knowledge of the surface–atmosphere exchange fluxes and by using an associated prior uncertainty, the information retrieved in the inversion from the observations is spread out in space and time corresponding to the spatiotemporal structure of the prior uncertainty. In this way, the solution of the otherwise ill-posed problem is regularized. This prior knowledge typically comes from process-oriented or diagnostic biosphere models that simulate the spatiotemporal patterns of terrestrial fluxes, as well as from inventories providing information regarding anthropogenic fluxes such as energy consumption, transportation, industry, and forest fires.

The Bayesian formulation of the inverse problem is a balance between the a priori and the observational constraints. It is crucial to introduce a suitable prior flux field and assign to it proper uncertainties. When prior information is combined with inappropriate prior uncertainties, this can lead to poorly retrieved fluxes (Wu et al., 2011). Here, we are interested in biosphere–atmosphere exchange fluxes and their uncertainties, and make the usual assumption that the uncertainties in anthropogenic emission fluxes are not strongly affecting the atmospheric observations at the rural sites that are used in the regional inversions of biosphere–atmosphere fluxes.

Typically inversions assume that prior uncertainties have a normal and unbiased distribution, and thus can be represented in the form of a covariance matrix. The covariance matrix is a method to weigh our confidence of the prior estimates. The prior error covariance determines to what extent the posterior flux estimates will be constrained by the prior fluxes. Ideally the prior uncertainty should reflect the mismatch between the prior guess and the actual (true) biosphere–atmosphere exchange fluxes.

BGD
12, 9393–9441, 2015

In this sense it needs to also have the corresponding error structure with its spatial and temporal correlations.

A number of different assumptions of the error structure have been considered by atmospheric CO_2 inversion studies. Coarser scale inversions often neglect spatial and temporal correlations as the resolution is low enough for the inverse problem to be regularized (Bousquet et al., 1999; Rödenbeck et al., 2003a) or assume large spatial correlation lengths (several hundreds of km) over land (Houweling et al., 2004; Rödenbeck et al., 2003b). For regional scale inversions, with higher spatial grid resolutions which are often less than 100 km, the spatial correlations are decreased (Chevallier et al., 2012) and the error structure need to be carefully defined. A variety of different assumptions exist. This is because only recently an objective approach to define prior uncertainties was followed (Chevallier et al., 2006 and 2012). In some regional studies, correlations are derived from large scale inversions in order to regularize the problem, although the change of resolution could lead to different correlation scales (Schuh et al., 2010). Alternatively, they are defined with a correlation length representing typical synoptic meteorological systems (Carouge et al., 2010). In other cases, ad-hoc solutions are adopted, where the correlation lengths are assumed to be smaller than in the case of global inversions (Peylin et al., 2005), or derived from climatological and ecological considerations (Peters et al., 2007) where correlation lengths only within the same ecosystem types have a value of 2000 km. In addition some studies use a number of different correlation structures in order to analyze which seems to be the most appropriate one based on some evaluation of the resulting inverted fluxes (Lauvaux et al., 2012). Michalak et al. (2004) applied a geostatistical approach based on the Bayesian method, in which the prior probability density function is based on an assumed form of the spatial and temporal correlation and no prior flux estimates are required. It optimizes the prior error covariance parameters, the variance and the spatial correlation length by maximizing the probability density function of the observations with respect to these parameters.

BGD
12, 9393–9441, 2015

A recent study by Broquet et al. (2013) obtained good agreements between the statistics at the European and 1 month scale of both the prior and posterior uncertainties from their inversions of the biosphere fluxes and that of the average misfits of the prior and posterior estimates of the fluxes to the local flux measurements. These good agreements relied in large part on their definition of the prior uncertainties based on the statistics derived in an objective way from model-data mismatch by Chevallier et al. (2006, 2012). In these studies, modeled daily fluxes from a site scale configuration of the ORCHIDEE model are compared with flux observations made within the global FLUXNET site network, based on the eddy covariance method (Baldocchi et al., 2001), and a statistical upscaling technique is used to derive estimates of the uncertainties in ORCHIDEE simulations at lower resolutions. While typical inversion systems have a resolution ranging from tens of kilometers up to several degrees (hundreds of km), the spatial representativity of the flux observations is typically around a kilometer. Considering also the scarcity of the observing sites in the flux network, the spatial information they bring is limited without methods for up-scaling such as the one applied by Chevallier et al. (2012). Nevertheless these measurements provide a unique opportunity to infer estimates of the prior uncertainties by examining model-data misfits for spatial and temporal autocorrelation structures.

In this study, we augment the approach of Chevallier et al. (2006 and 2012), to a multi-model – data comparison, investigating among others a potential generalization of the error statistics, suitable to be applied by inversions using different biosphere models as priors. This expectation is derived from the observation that the biosphere models, despite their potential differences typically have much information in common, such as driving meteorological fields, land use maps, or remotely sensed vegetation properties, and sometimes even process descriptions. We evaluate model – model mismatches to (i) investigate intra-model autocorrelation patterns and (ii) to explore whether they are consistent with the spatial and temporal correlation lengths of the model – data mismatch comparisons. Model comparisons have been used in the past to infer the structure of the prior uncertainties. For example, Rödenbeck et al. (2003b)

BGD
12, 9393–9441, 2015

used prior correlation lengths based on statistical analyses of the variations within an ensemble of biospheric models. This approach is to a certain degree questionable, as it is unclear how far the ensemble of models actually can be used as representative of differences between modeled and true fluxes. However, if a relationship between model – data and model – model statistics can be established for a region with dense network of flux observations, it could be used to derive prior error structure also for regions with a less dense observational network.

Moreover, to improve the knowledge of spatial flux error patterns, we make use of a unique set of aircraft fluxes measured on 2 km spatial windows along intensively sampled transects of several tens of km, ideally resolving spatial and temporal variability of ecosystem fluxes across the landscape without the limitation of the flux network with spatial gaps in between measurement locations. Lauvaux et al. (2009) compared results of a regional inversion against measurements of fluxes from aircraft and towers, while this is the first attempt to use aircraft flux measurements to assess spatial and temporal error correlation structures.

This study focuses on the European domain for 2007 (tower data) and 2005 (aircraft data), and uses output from high-resolution biosphere models that have been used for regional inversions. Eddy covariance tower fluxes were derived from the FLUXNET ecosystem network (Baldocchi et al., 2001), while aircraft fluxes were acquired within the CarboEurope Regional Experiment (CERES) in southern France. The methods and basic information regarding the models are summarized in Sect. 2. The results from model-data and model-model comparisons are detailed in Sect. 3. Discussion and conclusions are following in Sect. 4.

2 Data and methods

Appropriate error statistics for the prior error covariance matrix are derived from comparing the output of three biosphere models which are used as priors for regional scale inversions with flux data from the ecosystem network and aircraft. We investigate spa-

BGD
12, 9393–9441, 2015

tial and temporal autocorrelation structures of the model-data residuals. The temporal autocorrelation is a measure of similarity between residuals at different times but at the same location as a function of the time difference. The spatial autocorrelation refers to the correlation, at a given time, of the model-data residuals at different locations as a function of spatial distance. With this analysis we can formulate and fit an error model such as an exponentially decaying model, which can be directly used in the mesoscale inversion system to describe the prior error covariance.

2.1 Observations

A number of tower sites within the European domain, roughly expanding from −12° E to 35° E and 35° N to 61° N (see also Fig. 1), provide us with direct measurements of CO_2 biospheric fluxes using the eddy covariance technique. This technique computes fluxes from the covariance between vertical wind velocity and CO_2 dry mole fraction (Aubinet et al., 1999). We use Level 3, quality checked, half hourly observations of net ecosystem exchange fluxes (NEE), downloaded from the European Flux Database (www.europe-fluxdata.eu), and listed by site in Table 1. Each site is categorized into different vegetation types (Table 1). A land cover classification is used to label the sites as crop (17 sites), deciduous forest (4), evergreen forest (17), grassland (8), mixed forest (3), savannah (1 site), and shrub land (1). For the current study we focus on observations from these 53 European sites during the year 2007 (Fig. 1).

Additionally, aircraft fluxes are used, obtained with an eddy covariance system installed onboard a SkyArrow ERA aircraft (Gioli et al., 2006). Flights were made in southern France during CERES (CarboEurope Regional Experiment) from 17 May to 22 June 2005. Eddy covariance fluxes were computed on 2 km length spatial windows along transects of 69 km above forest and 78 km above agricultural land, flown 52 and 54 times across entire daily course, respectively. Exact routes are reported in Dolman et al. (2006).

BGD
12, 9393–9441, 2015

2.2 Biosphere models

We simulate CO_2 terrestrial fluxes for 2007 with three different biosphere models described in the following. The "Vegetation Photosynthesis and Respiration Model" (VPRM) (Mahadevan et al., 2008), used to produce prior flux fields for inverse studies (Pillai et al., 2012), is a diagnostic model that uses EVI – enhanced vegetation index and LSWI – land surface water index from MODIS, a vegetation map (Synmap, Jung et al., 2006) and meteorological data (temperature at 2 m and downward shortwave radiative flux extracted from ECMWF short term forecast fields at 0.25° resolution) to derive gross biogenic fluxes. VPRM parameters controlling respiration and photosynthesis for different vegetation types (a total of four parameters per vegetation type) were optimized using eddy covariance data for the year 2005 collected during the CarboEuropeIP project (Pillai et al., 2012). For this study, VPRM fluxes are provided at hourly temporal resolution and at two spatial resolutions of 1 and 10 km (referred to as VPRM1 and VPRM10). The difference between the 1 and 10 km resolution version is the aggregation of MODIS indices to either 1 or 10 km, otherwise the same meteorology and VPRM parameters are used. At 10 km resolution VPRM uses a tiled approach, with fractional coverage for the different vegetation types, and vegetation type specific values for MODIS indices. For the comparison with the aircraft data VPRM produced fluxes for 2005 at 10 km spatial resolution.

The "Organizing Carbon and Hydrology In Dynamic Ecosystems", ORCHIDEE, model (Krinner et al., 2005) is a process based site scale to global land surface model that simulates the water and carbon cycle using meteorological forcing (temperature, precipitation, humidity, wind, radiation, pressure). The water balance is solved at a half-hourly time step while the main carbon processes (computation of a prognostic LAI, allocation, respiration, turnover) are called on a daily basis. It uses a tiled approach, with fractional coverage for 13 Plant Functional Types (PFT). It has been extensively used as prior information in regional and global scale inversions (Piao et al., 2009; Broquet et al., 2013). For the present simulation, we use a global configuration of the

BGD
12, 9393–9441, 2015

version 1.9.6 of ORCHIDEE, where no parameter has been optimized against eddy covariance data. The model is forced with 0.5° WFDEI meteorological fields (Weedon et al., 2014). The PFT map is derived from an Olson land cover map (Olson, 1994) based on AVHRR remote sensing data (Eidenshink and Faundeen, 1994). The fluxes are diagnosed at 3 hourly temporal resolution and at 0.5° horizontal resolution.

The "5 parameter model" (5PM) (Groenendijk et al., 2011), also used in atmospheric inversions (Tolk et al., 2011; Meesters et al., 2012), is a physiological model describing transpiration, photosynthesis, and respiration. It uses MODIS LAI (leaf area index) at 1 km resolution, meteorological data (temperature, moisture, and downward shortwave radiative flux, presently from ECMWF at 0.25° resolution), and differentiates PFTs for different vegetation types and climate regions. 5PM fluxes are provided at 0.25° spatial and hourly temporal resolution. The optimization has been done with EC-data from Fluxnet as described in Groenendijk et al. (2011). Regarding the optimization of the heterotrophic respiration, EC-data from 2007 were used here.

Modeled fluxes for all above mentioned sites have been provided by the different models by extracting the fluxes from the grid cells which encompass the EC station location using vegetation type specific simulated fluxes, i.e. using the vegetation type within the respective grid cell for which the eddy covariance site is assumed representative. For most of the sites the same vegetation type was used for model extraction as long as this vegetation type is represented within the grid-cell. As VPRM uses a tile approach, for two cases ("IT-Amp", "IT-MBo") the represented vegetation type (crop) differ from the actual one (grass). For these cases, the fluxes corresponding to crop were extracted. Fluxes were aggregated to daily fluxes in the following way: first, fluxes from VPRM and 5PM as well as the observed fluxes were temporally aggregated to match with the ORCHIDEE 3 hourly resolution; in a second step we created gaps in the modeled fluxes where no observations were available; the last step aggregated to daily resolution on the premise that (a) the gaps covered less than 50 % of the day, and (b) the number of gaps (number of individual 3 hourly missing values) during day and during night were similar (not different by more than a factor two) to avoid biasing.

BGD
12, 9393–9441, 2015

Spatial and temporal correlation structures and the standard deviation of flux residuals (model–observations) were examined for daily fluxes over the year 2007. For the aircraft analysis, only the VPRM model was used since it is the only having a sufficiently high spatial resolution (10 km) comparable with aircraft flux footprint and capable of resolving spatial variability in relatively short flight distances. Aircraft NEE data have been grouped into 10 km segments along the track, to match the VPRM grid, obtaining 6 grid points in forest transects and 8 in agricultural land transects. Each grid point was sampled 52 times in forests, and 54 in agricultural land. VPRM fluxes at each aircraft grid cell were extracted, and then linearly interpolated to the time of each flight.

3 Analysis of model–observation differences

Observed and modeled fluxes are represented as the sum of the measured or simulated values and an error term, respectively. When we compare modeled to observed data this error term is a combination of model (the prior uncertainty we are interested in) and observation error. Separating the observation error from the model error in the statistical analysis of the model–observation mismatch is not possible; therefore we first neglect the impact of the observation error term on the correlation lengths. Nevertheless later in the analysis of model-model differences we assess the impact of the observation error on estimated correlation lengths.

The tower temporal autocorrelation is computed between the time series of model–observations differences $x_{l,i}$ at site l and the same series lagged by a time unit k (Eq. 1), where $\overline{x}$ is the overall mean and N the number of observations:

$$r_l(k) = \frac{\sum_{i=1}^{N-k} (x_{l,i} - \overline{x_l}) \cdot (x_{l,i+k} - \overline{x_l})}{\sum_{i=1}^{N} (x_{l,i} - \overline{x_l})^2} \quad (1)$$

BGD
12, 9393–9441, 2015

In order to reduce boundary effects in the computation of the autocorrelation at lag times around one year, the one-year flux time series data (model and observations) for each site was replicated four times. This follows the approach of Chevallier et al., (2012), where sites with at least three consecutive years of measurements have been used.

In the current analysis we introduce the all-site temporal autocorrelation by simultaneously computing the autocorrelation for all the observation sites, with M the number of the sites:

$$r(k) = \frac{\sum_{l=1}^{M}\sum_{i=1}^{N-k}(x_{l,i} - \overline{x_l}) \cdot (x_{l,i+k} - \overline{x_l})}{\sum_{l=1}^{M}\sum_{i=1}^{N}(x_{l,i} - \overline{x_l})^2} \quad (2)$$

Temporal correlation scales τ were derived by fitting an exponentially decaying model:

$$r = \alpha \cdot e^{-\frac{t}{\tau}} \quad (3)$$

Here t is the time lag. For the exponential fit, lags up to 180 days were used (thus the increase in correlations for lag times larger than 10 months is excluded). At zero lag time the correlogram has a value of one (fully correlated), however for even small lag times this drops to values smaller than one, also known as the nugget effect. For this we include the nugget effect variable α.

The aircraft temporal autocorrelation was similarly computed according to Eq. (1) using VPRM, and the same exponentially decaying model (Eq. 3) was used to fit the individual flight flux data, distributed along the entire daily course. The temporal interval was limited at 36 days by the experiment duration.

For the spatial analysis the correlation between model–observation residuals at two different locations (i.e sites or aircraft grid points) separated by a specific distance was computed in a way similar to the temporal correlation, and involved all possible pairs

BGD
12, 9393–9441, 2015

of sites and aircraft grid points. Additional data treatment for the spatial analysis was applied to reduce the impact of tower data gaps, as it is possible that the time series for two sites might have missing data at different times. Thus in order to have more robust results, we also examined spatial structures by setting a minimum threshold of 150 days of overlapping observations within each site pair. Furthermore spatial correlation was investigated for seasonal dependence, where seasons are defined as summer (JJA), fall (SON), winter (DJF for the same year), and spring (MAM). In those cases a different threshold of 20 days of overlapping observations was applied.

To estimate the spatial correlation scales, the pairwise correlations were grouped into bins of 100 km distance for towers and 10 km for aircraft data, respectively (dist). Following the median for each bin was calculated, and a model similar to Eq. (3) was fitted, but omitting the nugget effect variable:

$$r = e^{-\frac{\text{dist}}{d}} \tag{4}$$

The nugget effect could not be constrained simultaneously with the spatial correlation scale d, given the relatively coarse distance groups, the fast drop in the median correlation from one at zero distance to small values for the first distance bin combined with the somewhat variations at larger distances. Note that this difference between the spatial and the temporal correlation becomes obvious in the results Sect. 3.

As aircraft fluxes cannot obviously be measured at the same time at different locations, given the relatively short flight duration (about one hour) we treated aircraft flux transect as instantaneous "snapshots" of the flux spatial pattern across a landscape, neglecting temporal variability that may have occurred during flight.

4 Analysis of model-model differences

We evaluate both model-data flux residuals and model-model differences in a sense of pairwise model comparisons, in order to assess if model-model differences can be used as proxy for the prior uncertainty, assuming that the involved prior errors for each model

BGD
12, 9393–9441, 2015

are identical in a sense that they share the same statistics and not correlated. Similar to the model–observation analysis, the statistical analysis gives a combined effect of both model errors. We assess the impact in the error structure between model–observation and model-model comparisons caused by the observation error by adding a random measurement error to each model-model comparison. This error has the same characteristics as the observation error which is typically associated with eddy covariance observations; the error characteristics were derived from the paired observation approach (Richardson et al., 2008). Specifically, we implement the flux observation error as a random process (white noise) with a double-exponential probability density function. This can be achieved by selecting a random variable u drawn from the uniform distribution in the interval $(-1/2, 1/2)$, and then applying Eq. (5) to get a Laplace distribution (also referred to as the double-exponential)

$$x = \mu - \frac{\sigma}{\sqrt{2}} \cdot \mathrm{sgn}(u) \cdot \ln(1 - 2 \cdot |u|) \tag{5}$$

Here $\mu = 0$ and σ is the standard deviation of the double-exponential. We compute the σ according to Richardson et al. (2006) as

$$\sigma = \alpha_1 + \alpha_2 \cdot |F| \tag{6}$$

where F is the flux and α_1, α_2 are scalars specific to the different vegetation classes. Lasslop et al. (2008) found that the autocorrelation of the half hourly random errors is below 0.7 for a lag of 30 min, and falls off rapidly for longer lag times. Thus we assume the standard deviation for hourly random errors to be comparable with the half hourly errors. Hourly random errors specific for each reference model are generated for each site individually. With ORCHIDEE as reference with fluxes at 3 hourly resolution, a new ensemble of 3 hourly random noise was generated with σ for the 3 hourly errors modified (divided by the square root of three to be coherent with the hourly σ). As both modeled and observed fluxes share the same gaps, the random errors were aggregated to daily resolution, with gaps such to match those of observed fluxes. Finally the daily random errors were added to the modeled fluxes.

BGD
12, 9393–9441, 2015

5 Results

5.1 Model-data comparison for tower and aircraft fluxes

Observed daily averaged NEE fluxes, for all ground sites and the full time-series, yield a standard deviation of 3.01 $\mu mol\,m^{-2}\,s^{-1}$, while the modeled fluxes were found to be less spatially varying and with a standard deviation of 2.84, 2.80, 2.53, 2.64 $\mu mol\,m^{-2}\,s^{-1}$ for VPRM10, VPRM1, ORCHIDEE and 5PM respectively.

The residual distribution of the models defined as the difference between simulated and observed daily flux averages for the full year 2007 was found to have a standard deviation of 2.47, 2.49, 2.7 and 2.25 $\mu mol\,m^{-2}\,s^{-1}$ for VPRM10, VPRM1, ORCHIDEE and 5PM respectively. Those values are only slightly smaller than the standard deviations of the observed or modeled fluxes themselves, which is in line with the generally low fraction of explained variance with r-square values of 0.31, 0.27, 0.12, and 0.25 for VPRM10, VPRM1, ORCHIDEE and 5PM respectively. When using site-specific correlations, the average fraction of explained variance increases to 0.38, 0.36, 0.35, and 0.42, for VPRM10, VPRM1, ORCHIDEE and 5PM, respectively. This indicates better performance for the models to simulate temporal changes at the site level, and the differences to the overall r-square values indicate limitation of the models to reproduce observed spatial (site to site) differences. Figure 2 shows the correlation between modeled and observed daily fluxes as a function of the vegetation type characterizing each site. All models exhibit a significant scatter of the correlation ranging from 0.9 for some sites to 0 or even negative correlation for some crop sites, with the highest correlation coefficients for deciduous and mixed forest.

The distribution is biased by −0.07, 0.26, 0.92 and 0.25 $\mu mol\,m^{-2}\,s^{-1}$ for VPRM10, VPRM1, ORCHIDEE and 5PM, respectively. Figure 3 shows the bias distribution for different vegetation types. Bias and standard deviation seem to depend on the vegetation type for all models, without a clear general pattern.

The temporal autocorrelation was calculated for model-data residuals for each of the flux sites (“site data” in Fig. 4), but also for the full dataset (“all-site” in Fig. 4).

BGD
12, 9393–9441, 2015

The "all site" temporal autocorrelation structure of the residuals appears to have the same pattern for all models. It decays smoothly for time lags up to 3 months and then remains constant near to 0 or to some small negative values. The temporal autocorrelation increases again for time lags > 10 months, which is caused by the seasonal cycle. The all-sites correlation for the VPRM model at 10 km resolution remains positive for lags <104 days and for lags > 253 days. Weak negative correlations were found in between with minimum value −0.03. In contrast we found only positive correlation for VPRM at 1 km resolution for the whole year with a minimum value of 0.002. Similarly, ORCHIDEE follows the same patterns with positive correlations for lags < 76 days and for lags > 291. Minimum correlation was found to be −0.09. For 5PM model we also found only positive correlations. The minimum value was found to be 0.08. These temporal autocorrelation results agree with the findings of Chevallier et al. (2012).

The exponentially decaying model in Eq. (3) was used to fit the data. At zero separation time (t = 0) the correlogram value is 1. However the correlogram exhibits a nugget effect (i.e. a value of 0.39 for VPRM10) as a consequence of an uncorrelated part of the error. For the current analyses we fit the exponential model with an initial correlation different from 1. The fit has a root mean square error of 0.041. The e-folding time (defined as the lag required for the correlation to decrease by a factor of e (63 % of its initial value) ranged between 26–70 days for the different models (see Table 2). Specifically, for VPRM10 and VPRM1 the e-folding time is 32 and 33 days respectively (30–34 days within 95 % confidence interval for both). Confidence intervals for the e-folding time were calculated by computing the confidence intervals of the parameter in the fitted model. For ORCHIDEE best fit was 26 days (23–28 days within 95 % confidence interval). In contrast, 5PM yields a significantly longer correlation time between 65–75 days (95 % confidence interval) with the best fit being 70 days.

For a number of sites a large model-data bias was found. In order to assess how the result depends on individual sites where model-data residuals are more strongly biased the analysis was repeated under exclusion of sites with an annual mean of model-data flux residuals larger than 2.5 $\mu mol\, m^{-2}$ s. This threshold value is roughly

BGD
12, 9393–9441, 2015

half of the most deviant bias. In total 9 sites ("CH-Lae", "ES-ES2", "FR-Pue", "IT-Amp", "IT-Cpz", "IT-Lav", "IT-Lec", "IT-Ro2", "PT-Esp") across all model-data residuals were excluded. From these sites "CH-Lae" appears to have serious problems related to the steep terrain, where the basic assumptions made for eddy covariance flux measurements are not well applicable (Göckede et al., 2008). The rest of the sites are located in the Mediterranean region, and suffer from summer drought according to the Köppen–Geiger climate classification map (Kottek et al., 2006); in those cases a large model – data bias is expected as existing models tend to have difficulties to estimate carbon fluxes for drought prone periods (Keenan et al., 2009). The model-data bias at those sites does not necessarily exceed the abovementioned threshold of 2.5 µmol m^{-2} s for each individual model, but a larger bias than the average was detected. After exclusion of those sites the temporal correlation times were found to be between 33–35 days within 95 % confidence interval for 5PM with the best fit value being 34 days. The rest of the models had temporal e-folding times of 27, 29 and 24 days (1st row of Table 2), while the all-site correlation remains positive for lags < 76, < 79, < 66 days for VPRM10, VPRM1 and ORCHIDEE respectively. Some weak negative correlations exist, with a minimum value of −0.06, −0.02, −0.09, −0.005 for VPRM10, VPRM1, ORCHIDEE and 5PM respectively.

The temporal correlation of differences between VPRM and aircraft flux measurements could be computed for time intervals up to 36 days (Fig. 5) corresponding to the duration of the campaign. The correlation shows an exponential decrease, and levels off after about 25 days with an e-folding correlation time of 13 days. Whilst the general behavior is consistent with results obtained for VPRM-observation residuals for flux sites, the correlation time is two times smaller.

Regarding spatial error correlations, results for all models show a dependence on the distance between pairs of sites. The median correlation drops within very short distances (Fig. 6). Fitting the simple exponentially decaying model (Eq. 4) to the correlation as a function of distance we find an e-folding correlation length d of 40, 37, 32 and 31 km with a root mean square error (RMSE) of 0.14, 0.09, 0.05 and 0.07 for

BGD
12, 9393–9441, 2015

VPRM10, VPRM1, ORCHIDEE and 5PM, respectively. Spatial correlation scales are also computed for a number of different data selections (cases) in addition to the standard case shown in Fig. 6 (case S): using only pairs with at least 150 overlapping days of non-missing data (case S*), using only pairs with identical PFT (case I), using only pairs with different PFT (case D), and using only pairs with at least 150 overlapping days for the D and I cases (cases D*, I*). The results for these cases are summarized in Fig. 7. Also 95 % confidence intervals were computed, and the spread spatial correlation was found to be markedly more critical than for the time correlations. Note that for some cases the 2.5 % (the lower bound of the confidence interval) hit the lower bound for correlation lengths (0 km).

Interestingly, if we restrict the analysis to pairs with at least 150 overlapping days between site pairs, larger correlation scales are found (case S* in Fig. 7). Considering only pairs with different PFT (case D), consistently, all correlation lengths are found to be smaller compared to the standard case (S). This is expected to a certain degree, as model errors should be more strongly correlated between sites with similar PFTs than between sites with different PFTs. By considering only pairs within the same vegetation type (case I) we observe a significant increase of the correlation length relative to case S for VPRM at 10 and 1 km resolution to values of 432 and 305 km, respectively. The ORCHIDEE and 5PM models show some (although not significant) increase in correlation length. Restricting again the analysis to pairs with at least 150 overlapping days for the D and I cases (D*, I*) we observe an increase of the correlation lengths that is however significant only for VPRM at 10 and 1 km.

Seasonal dependence of the correlation lengths for at least 20 overlapping days per season and for all site-pairs is also shown in Fig. 7. VPRM showed somewhat longer correlation lengths during spring and summer, ORCHIDEE had the largest lengths occurring during summer and fall, and 5PM correlation lengths show slightly enhanced values during spring and summer. However, none of these seasonal differences are significant with respect to the 95 % confidence interval.

BGD
12, 9393–9441, 2015

The spatial error correlation between VPRM10 model and aircraft fluxes measured during May–June along continuous transects at forest and agriculture land use (Fig. 8) shows an exponential decay up to the maximum distance that was encompassed during flights (i.e. 70 km). Of note is that only two measurements were available at 60 km distance and none for larger distances making difficult to identify where the asymptote lying. Nevertheless fitting the decay model (Eq. 4) leads to d = 35 km (26–46 km within the 95 % confidence interval), which is in good agreement with the spatial correlation scale derived for VPRM10 using flux sites during both spring and summer (Fig. 7).

6 Model-model comparison

We investigate the model-model error structure of NEE estimates by substituting the observed fluxes which were used as reference, with simulated fluxes from all the biosphere models. Note that for consistency with the model-data analysis, the simulated fluxes contained the same gaps as the observed flux time series. The e-folding correlation time is found to be slightly larger compared to the model-data correlation times, for most of the cases. An exception are the 5PM-VPRM10 and 5PM-VPRM1 pairs which they produced remarkably larger correlation times (Table 2). Specifically, VPRM10-ORCHIDEE and VPRM10-5PM residuals show correlation times of 30 days (range between 27–33 days within 95 % confidence interval) and 131 (range between 128–137 days within 95 % confidence interval), respectively. Repeating the analysis excluding sites with residual bias larger than 2.5 $\mu mol\, m^{-2}\, s$, correlation times of 31 and 100 days for VPRM10-ORCHIDEE and VPRM10-5PM are found, respectively. If we use ORCHIDEE as reference the e-folding correlation times are 30, 28 and 38 days with respect to VPRM10, VPRM1 and 5PM comparisons respectively.

Although the e-folding correlation times show but minor differences compared to the model-data residuals, this is not the case for the spatial correlation lengths (Fig. 9). The standard case (S) was applied for the annual analysis, with no minimum number of days with overlapping non-missing data for each site within the pairs. Taking VPRM10

BGD
12, 9393–9441, 2015

as reference, much larger e-folding correlation lengths of 312 km with a range of 206–421 km within 95 % confidence interval yielded for VPRM10-ORCHIDEE comparisons, and 306 km for VPRM10-5PM were found. With ORCHIDEE as reference the correlation length for the ORCHIDEE-5PM comparison is 271 km with a range of 183–360 km within 95 % confidence interval. Seasonal correlation lengths, using a minimum of 20 days overlap in the site-pairs per season (Fig. 9), are also significantly larger compared with those from the model-data analysis.

When we add the random measurement error to the modeled fluxes used as reference (crosses in Fig. 9), we observe only slight changes in the annual correlation lengths, without a clear pattern. The correlation lengths show a random increase or decrease but limited up to 6 %. Interestingly, the seasonal correlation lengths for most of the cases show a more clear decrease. For example, the correlation length of the VPRM10-5PM residuals during winter, decreases by 22 % or even more for spring season. Despite this decrease, the seasonal correlation lengths remain significantly larger in comparison to those from the model-data analysis. Overall, all models when used as reference show the same behavior with large e-folding correlation lengths that mostly decrease slightly when the random measurement error is included. Although the random measurement error was added as "missing part" to the modeled fluxes to better mimic actual flux observations, it did not lead to correlation lengths similar to those from the model-data residual analysis. To investigate if a larger random measurement error could cause spatial correlation scales in model-model differences, we repeated the analysis with artificially increased random measurement error (multiplying with a factor between 1 and 15). Only for very large random measurement errors did the model-model correlation lengths start coinciding with those of the model-data residuals (Fig. 10).

BGD
12, 9393–9441, 2015

7 Discussion and conclusions

We analyzed the error structure of a-priori NEE uncertainties derived from a multi-model – data comparison by comparing fluxes simulated by three different vegetation models to daily averages of observed fluxes from 53 sites across Europe, categorized into 7 land cover classes. The different models showed comparable performance with respect to reproducing the observed fluxes; we found mostly insignificant differences in the mean of the residuals (bias) and in the variance. Site-specific correlations between simulated and observed fluxes are significantly higher than overall correlations for all models, which suggest that the models struggle with reproducing observed spatial flux differences between sites. Furthermore, the site-specific correlations reveal a large spread even within the same vegetation class, especially for crops (Fig. 2). This is likely due to the fact that none of the models uses a specific crop model that differentiates between the different crop types and their phenology. The models using remotely sensed vegetation indices (VPRM and 5PM) better capture the phenology; ORCHIDEE is the only model that differentiates between C_3 and C_4 plants, but shows the largest spread in correlation for the crop. Differences in correlations between the different vegetation types were identified for all the biosphere models, however it must be noted that the number of sites per vegetation type is less than 10 except for crop and evergreen forests.

Model-data flux residual correlations were investigated to give insights regarding prior error temporal scales which can be adopted by atmospheric inversion systems. Autocorrelation times were found to be in line with findings of Chevallier et al. (2012). The model-data residuals were found to have an e-folding time of 32 and 26 days for VPRM and ORCHIDEE respectively, and 70 days for 5PM. This significant difference appears to have a strong dependence on the set of sites used in the analysis. Excluding nine sites with large residual bias, the autocorrelation time from the 5PM-data residuals drastically decreased and became coherent with the times of the other biosphere models. The all-models and all-sites autocorrelation time was found to be 39

BGD
12, 9393–9441, 2015

days, which reduces to 30 days (28–31 days within 95 % confidence interval), when excluding the sites with large residual bias, coherent with the single model times. From the model-model residual correlation analysis, the correlation time appear to be consistent with the above-mentioned results, and lies between 28 and 46 days for most of the ensemble members. However model-model pairs consisting of the VPRM and 5PM models produced larger times up to 136 days; omitting sites with large residual biases this is reduced to 105 days (102–108 days within 95 % confidence interval). This finding could be attributed to the fact that despite the conceptual difference between those models, they do have some common properties. Both models were optimized against eddy covariance data although for different years (2005 and 2007 respectively), while no eddy covariance data were used for the optimization of ORCHIDEE. In addition, VPRM and 5PM both use data acquired from MODIS, although they estimate photosynthetic fluxes by using different indices of reflectance data. Summarizing the temporal correlation structure, it appears reasonable to (a) use same error correlation in atmospheric inversions regardless which biospheric model is used as prior, (b) use an autocorrelation length of around 30 days.

Only weak spatial correlations for model-data residuals were found, limited to short lengths up to 40 km without any significant difference between the biospheric models (31–40 km). Although the estimated spatial scales are shorter than the spatial resolution that we are solving for (100 km bins), the autocorrelation analysis of aircraft measurements made during CERES supports the short scale correlations. These measurements have the advantage of providing continuous spatial flux transects along specific tracks that were sampled routinely (in this case over period of 36 days at various times of the day), thus resolving flux spatial variability also at small scales, where pairs of eddy covariance sites may not be sufficiently close. On the other hand, aircraft surveys are necessarily sporadic in time. Of note is that the impact of the eddy covariance observation error on the estimated prior error and its structure had to be initially neglected as it is not possible to subtract the unknown error from the observations. However we do not expect this to have a significant impact on the error structure, as the addition of

BGD
12, 9393–9441, 2015

an observation error to the analysis of model-model differences had only minor influence on the error structure.

Model-data residual correlation lengths show a clear difference, between the cases where pairs only with different (D) or identical (I) PFT were considered, with the latter resulting in longer correlation lengths, but only identified for the VPRM model at both resolutions. The "D" case has slightly shorter lengths for all models than the standard case (S). One could argue that as VPRM uses PFT specific parameters that were optimized against 2005 observations, the resulting PFT specific bias could lead to longer spatial correlations. However ORCHIDEE and 5PM also show comparable biases (Fig. 3), but long correlation scales were not found. Moreover we repeated the spatial analysis after subtracting the PFT specific bias from the fluxes, and the resulting correlation lengths showed no significant change. The impact of data gaps was also investigated by setting a threshold value of overlapping observations between site pairs. Setting this to 150 days results in an increase for the "S" case up to 60 km, but only for the VPRM model. For the "D" and "I" cases when setting the same threshold value (D^* and I^*) we only found an insignificant increase, indicating that data gaps are hardly affecting the "D" and "I" cases. These findings suggest that high-resolution diagnostic models might be able to highlight the increase of the spatial correlation length between identical PFTs vs. different PFTs. Note that the Chevallier et al. (2012) study concluded that assigning vegetation type specific spatial correlations is not justified, based on comparisons of eddy covariance observations with ORCHIDEE simulated fluxes. The current study could not further investigate this dependence, as the number of pairs within a distance bin is not large enough for statistical analyses, when using only sites within the same PFT. With respect to the seasonal analysis, spatial correlations are at the same range among all models and seasons. Although in some cases (VPRM10 and VPRM1 spring) the scales are larger, they suffer from large uncertainties. Hence, implementing distinct and seasonally dependent spatial correlation lengths in inversion systems cannot be justified.

BGD
12, 9393–9441, 2015

The analysis of model-model differences did not reproduce the same spatial scales as those from the model-data differences, but instead spatial correlation lengths were found to be dramatically larger. Adding a random measurement error to the modeled fluxes used as reference slightly reduced the spatial correlation lengths to values ranging from 86 to 320 km. Even when largely inflating the measurement error, the resulting spatial correlation lengths (Fig. 10) still do not approach those derived from model-data residuals. Only when the measurement error is scaled up by a factor of 8 or larger (which is quite unrealistic as this corresponds to a mean error of 1.46 $\mu mol\, m^{-2}\, s^{-1}$ or larger, which is comparable to the model-data mismatch where a standard deviation of around 2.5 $\mu mol\, m^{-2}\, s^{-1}$ was found), the correlation lengths are consistent with those based on model-data differences. Whilst the EC observations are sensitive to a footprint area of about 1 km^2, the model resolution is too coarse to capture variations at such a small scale. This local uncorrelated error has not been taken into account by the analysis of model-data residuals as the error model could not be fitted with a nugget term included, favoring therefore smaller correlation scales. The analysis of differences between two coarser models (excluding VPRM at 1 km for the reason mentioned in the next paragraph) does not involve such a small scale component, thus resulting in larger correlation scales. This would suggest that for inversion studies targeting scales much larger than the eddy covariance footprint scale, the statistical properties of the prior error should be derived from the model-model comparisons.

A special case in the context of the model-model study is the comparison between VPRM1 and VPRM10, which is the only case that produced short spatial correlation scales. These two models only differ in the spatial resolution of MODIS indices EVI and LSWI (1 vs. 10 km). Thus differences between those two models are only related to variability of these indices at scales below 10 km, which is not expected to show any spatial coherence. Indeed the results show only very short correlation scales (Fig. 9) with an exception during fall, however there the uncertainty is also large.

The large correlation lengths yielded from this model-model residual analysis suggest that the models are more similar to each other than to the observed terrestrial

BGD
12, 9393–9441, 2015

fluxes, at least on spatial scales up to a few hundred kilometers regardless of their conceptual differences. This might be expected at some extent due to elements that the models share. Respiration and photosynthetic fluxes are strongly driven by temperature and downward radiation, respectively, and those meteorological fields have significant commonalities between the different models. VPRM and 5PM both use temperature and radiation from ECMWF analysis and short-term forecasts. Also the WFDEI temperature and radiation fields used in ORCHIDEE are basically from the ERA-Interim reanalysis, which also involves the integrated forecasting system (IFS) used at ECMWF (Dee et al., 2011). Regarding the vegetation classification all models are site specific and therefore are using the same PFT for each corresponding grid-cell. Photosynthetic fluxes are derived with the use of MODIS indices in VPRM (EVI and LSWI) and in 5PM (LAI and albedo).

Using full flux fields from the model ensemble (rather than fluxes at specific locations with observation sites only) to assess spatial correlations in model-model differences is not expected to give significantly different results, as the sites are representative for quite a range of geographic locations and vegetation types within the domain investigated here.

The current study intended to provide insight on the error structure that can be used for atmospheric inversions. Typically, inversion systems have a pixel size ranging from 10 to 100 km for regional and continental inversions, and as large as several degrees (hundreds of km) for global inversions. If a higher resolution system assumes such small-scale correlations (as those found in the current analysis), in the covariance matrix, of note is that this leads to very small prior uncertainties when aggregating over large areas and over longer time periods. For example, with a 30 km spatial and a 40 day temporal correlation scale, annually and domain-wide (Fig. 1) aggregated uncertainties are around 0.06 Gt C. This is about a factor ten smaller than uncertainties typically used e.g. in the Jena inversion system (Rödenbeck et al., 2005). In addition, the aggregated uncertainties using the VPRM10-ORCHIDEE error structure (32 days and 320 km temporal and spatial correlation scales) are found to be 0.46 Gt C yr^{-1} which is

BGD
12, 9393–9441, 2015

also much smaller than the difference between VPRM10 (NEE = $-1.45\,\mathrm{Gt\,C\,yr^{-1}}$) and ORCHIDEE (NEE= $-0.2\,\mathrm{Gt\,C\,yr^{-1}}$), when aggregated over the domain shown in Fig. 1. Although this analysis does capture the dominating spatiotemporal correlation scale in the error structure, it fails in terms of the error budget, suggesting that also other parts of the error structure are important as well. Therefore additional degrees of freedom (e.g. for a large-scale bias) need to be introduced in the inversion systems to fully describe the error structure.

Exponentially decaying correlation models are a dominant technique among atmospheric inverse studies to represent temporal and spatial flux autocorrelations (Rödenbeck et al., 2009; Broquet et al., 2011, 2013). However, regarding the temporal error structure we need to note the weakness of this model to capture the slightly negative values at 2–10 months lags and, more importantly, the increase in correlations for lag times larger than about 10 months. Error correlations were parameterized differently by Chevalier et al. (2012) where the prior error was investigated without implementing it to atmospheric inversions. Polynomial and hyperbolic equations were used to fit temporal and spatial correlations respectively. Nevertheless, we use here e-folding lengths not only for their simplicity in describing the temporal correlation structure with a single number, but also because this error model ensures a positive definite covariance matrix (as required for a covariance). This is crucial for atmospheric inversions as otherwise negative, spatially and temporally integrated uncertainties may be introduced. In addition it can keep the computational costs low; this is because the hyperbolic equation has significant contributions from larger distances: for the case of the VPRM1 model, at 200 km distance the correlation according to Eq. (7) is 0.16, compared to 0.004 for the exponential model. As a consequence, more none-zero elements are introduced to the covariance matrix, which increases computational costs in the inversion systems. Using the parameterization from Eq. (7) for the spatial correlation, d values of 73, 39, 12 and 20 km were found with a RMSE of 0.11, 0.07, 0.05, 0.07 for VPRM10, VPRM1, ORCHIDEE and 5PM respectively. A similar RMSE was found when using the expo-

BGD
12, 9393–9441, 2015

nential (0.14, 0.09, 0.05 and 0.07), indicating similar performance of both approaches with respect to fitting the spatial correlation.

Whilst temporal scales found from this study have already been used in inversion studies, this is not the case to our best knowledge for the short spatial scales. The impact of the prior error structure derived from this analysis, on posterior flux estimates and uncertainties will be assessed in a subsequent paper. For that purpose, findings from this study are currently implemented in three different regional inversion systems aiming to focus on network design for the ICOS atmospheric network.

Acknowledgements. The research leading to these results has received funding from the European Community's Seventh Framework Program ([FP7/2007–2013]) under grant agreement no. 313169 (ICOS-INWIRE) and under grant agreement no. 283080 (GEOCARBON). The EC data used in this study were funded by the European Community's sixth and seventh framework program and by national funding agencies. In the following sites are listed, sorted by project/funding agency: "BE-Bra", "BE-Lon", "BE-Vie", "CH-Lae", "CH-Oe1", "CH-Oe2", "CZ-BK1", "DE-Geb", "DE-Gri", "DE-Hai", "DE-Kli", "DE-Tha", "DK-Lva", "ES-ES2", "ES-LMa", "FI-Hyy", "FR-Fon", "FR-Hes", "FR-LBr", "FR-Lq1", "FR-Lq2", "FR-Pue", "IT-SRo", "NL-Dij", "NL-Loo", "PT-Esp", "PT-Mi2", "SE-Kno", "SE-Nor", "SE-Sk1", "SK-Tat", "UK-AMo", "UK-EBu" funded by CarboEuropeIP (grand agreement no. GOCE-CT-2003-505572); "UK-AMo", "UK-EBu" also co-funded by EU FP7 ECLAIRE project and NERC-CEH; "IT-BCi", "IT-Cas", "IT-Lav", "IT-LMa" funded by CarboItaly (IT-FISR), "IT-Amp", "IT-Col", "IT-Cpz", "IT-MBo", "IT-Ren", "IT-Ro2" co-funded by CarboEuropeIP and by CarboItaly, "CH-Cha", "CH-Dav", "CH-Fru" co-funded by CarboExtreme (grant agreement no. 226701) and GHG-Europe (grant agreement no. 244122), "ES-Agu" funded by GHG-Europe, "FR-Mau" co-funded by CNRM/GAME (METEO-FRANCE, CNRS), CNES and ONERA. We also acknowledge funding agencies for the sites "FR-Aur", "FR-Avi", "HU-Mat".

The article processing charges for this open-access publication were covered by the Max Planck Society.

BGD
12, 9393–9441, 2015

References

Albergel, C., Calvet, J.-C., Gibelin, A.-L., Lafont, S., Roujean, J.-L., Berne, C., Traullé, O., and Fritz, N.: Observed and modelled ecosystem respiration and gross primary production of a grassland in southwestern France, Biogeosciences, 7, 1657–1668, doi:10.5194/bg-7-1657-2010, 2010.

Allard, V., Ourcival, J.-M., Rambal, S., Joffre, R., and Rocheteau, A.: Seasonal and annual variation of carbon exchange in an evergreen Mediterranean forest in southern France, Glob. Change Biol., 14, 714–725, 2008.

Ammann, C., Spirig, C., Leifeld, J., and Neftel, A.: Assessment of the nitrogen and carbon budget of two managed grassland fields, Agr. Ecosyst. Environ., 133, 150–162, 2009.

Aubinet, M., Grelle, A., Ibrom, A., Rannik Ü., Moncrieff, J., Foken, T., Kowalski, A.-S., Martin, P.-H., Berigier, P., Bernhofer, C., Clement, R., Elbers, J., Granier, A., Grünwald T., Morgenstern, K., Pilegaard, K., Rebmann, C., Snijders, W., Valentini, R., and Vesala, T.: Estimates of the Annual Net Carbon and Water Exchange of Forests: the EUROFLUX Methodology, Adv. Ecol. Res., 30, 113–175, 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, Agr. Forest Meteorol., 108, 293–315, 2001.

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Mahli, Y., Meyers, T., Munger, W., Oechel, W., Paw, K. T., Pilegaard, K., Schmid, H. P., Valentini, R., Verma, S., Vesala, T., Wilson, K., and Wofsy, S.: FLUXNET: a new tool to study the temporal and spatial variability of ecosystem–scale carbon dioxide, water vapor, and energy flux densities, B. Am. Meteorol. Soc., 82, 2415–2434, 2001.

Barcza, Z., Weidinger, T., Csintalan., Zs., Dinh, N.-Q., Grosz, B., and Tuba, Z.: The carbon budget of a semiarid grassland in a wet and a dry year in Hungary, Agr. Ecosyst. Environ., 121, 21–29, 2007.

Bates, D. M. and Watts, D. G.: Nonlinear Regression: Iterative Estimation and Linear Approximations, in Nonlinear Regression Analysis and Its Applications, John Wiley & Sons, Inc., Hoboken, NJ, USA, doi:10.1002/9780470316757.ch2, 1988.

BGD
12, 9393–9441, 2015

Bousquet, P., Ciais, P., Peylin, P., Ramonet, M., and Monfray, P.: Inverse modeling of annual atmospheric CO_2 sources and sinks: 1. Method and control inversion, J. Geophys. Res.-Atmos., 104, 26161–26178, 1999.

Broquet, G., Chevallier, F., Rayner, P., Aulagnier, Cé., Pison, I., Ramonet, M., Schmidt, M., Vermeulen, A.-T., and Ciais, P.: A European summertime CO_2 biogenic flux inversion at mesoscale from continuous in situ mixing ratio measurements, J. Geophys. Res.-Atmos., 116, D23303, doi:10.1029/2011JD016202, 2011.

Broquet, G., Chevallier, F., Bréon, F.-M., Kadygrov, N., Alemanno, M., Apadula, F., Hammer, S., Haszpra, L., Meinhardt, F., Morguí, J. A., Necki, J., Piacentino, S., Ramonet, M., Schmidt, M., Thompson, R. L., Vermeulen, A. T., Yver, C., and Ciais, P.: Regional inversion of CO_2 ecosystem fluxes from atmospheric measurements: reliability of the uncertainty estimates, Atmos. Chem. Phys., 13, 9039–9056, doi:10.5194/acp-13-9039-2013, 2013.

Carouge, C., Bousquet, P., Peylin, P., Rayner, P. J., and Ciais, P.: What can we learn from European continuous atmospheric CO_2 measurements to quantify regional fluxes – Part 1: Potential of the 2001 network, Atmos. Chem. Phys., 10, 3107–3117, doi:10.5194/acp-10-3107-2010, 2010.

Casals, P., Lopez-Sangil, L., Carrara., A., Gimeno, C., and Nogues, S.: Autotrophic and heterotrophic contributions to short-term soil CO_2 efflux following simulated summer precipitation pulses in a Mediterranean dehesa, Global Biogeochem. Cy., 25, GB3012, doi:10.1029/2010GB003973, 2011.

Chevallier, F., Viovy, N., Reichstein, M., and Ciais, P.: On the assignment of prior errors in Bayesian inversions of CO_2 surface fluxes, 33, L13802, doi:10.1029/2006GL026496, 2006.

Chevallier, F., Wang, T., Ciais, P., Maignan, F., Bocquet, M., Altaf, A.-M., Cescatti, A., Chen, J., Dolman, A.-J., Law, B.-E., Margolis, H. A., Montagnani, L., and Moors, E.: What eddy-covariance measurements tell us about prior land flux errors in CO_2 flux inversion schemes, Global Biogeochem. Cy., 26, GB1021, doi:10.1029/2010GB003974, 2012.

Chiesi, M., Fibbi, L., Genesio, L., Gioli, B., Magno, R., Maselli, F., Moriondo, M., and Vaccari, F.-P.: Integration of ground and satellite data to model Mediterranean forest processes. Int. J. Appl. Earth Obs., 13, 504–515, doi:10.1016/j.jag.2010.10.006, 2011.

Ciais, P., Peylin, P., and Bousquet, P.: Regional biospheric carbon fluxes as inferred from atmospheric CO_2 measurements, Ecol. Appl., 10, 1574–1589, 2000.

Dee, D.-P., Uppala, S.-M., Simmons, A.-J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.-A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.-C.-M., van de Berg, L.,

BGD
12, 9393–9441, 2015

Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.-J., Haimberger, L., Healy, S.-B., Hersbach, H., Hólm, E.-V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., Mcnally, A.-P., Monge-Sanz, B.-M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteor. Soc., 137, 553–597, doi:10.1002/qj.828, 2011.

Delpierre, N., Soudani, K., François, C., Köstner, B., Pontailler, J.-Y., Nikinmaa, E., Misson, L., Aubinet, M., Bernhofer, C., Granier, A., Grünwald, T., Heinesch, B., Longdoz, B., Ourcival, J.-M., Rambal, S., Vesala, T., and Dufrêne, E.: Exceptional carbon uptake in European forests during the warm spring of 2007: a data-model analysis, Glob. Change Biol., 15, 1455–1474, doi:10.1111/j.1365-2486.2008.01835.x, 2009.

Dietiker, D., Buchmann, N., and Eugster, W.: Testing the ability of the DNDC model to predict CO_2 and water vapour fluxes of a Swiss cropland site, Agr. Ecosyst. Environ., 139, 396–401, 2010.

Dolman, A.-J., Noilhan, J., Durand, P., Sarrat, C., Brut, A., Piquet, B., Butet, A., Jarosz, N., Brunet, Y., Loustau, D., Lamaud, E., Tolk, L., Ronda, R., Miglietta, F., Gioli, B., Magliulo, V., Esposito, M., Gerbig, C., Korner, S., Glademard, R., Ramonet, M., Ciais, P., Neininger, B., Hutjes, R- W.-A., Elbers, J.-A., Macatangay, R., Schrems, O., Perez-Landa, G., Sanz, M.-J., Scholz, Y., Facon, G., Ceschia, E., and Beziat, P.: The CarboEurope Regional Experiment Strategy, B. Am. Meteorol. Soc., 87, 1367–1379, 2006.

Eidenshink, J. C. and Faundeen, J. L.: The 1 km AVHRR global land data set: first stages in implementation, Int. J. Remote Sens., 15, 3443–3462, 1994.

Elbers, J.-A., Jacobs, C.-M.-J., Kruijt, B., Jans, W.-W.-P., and Moors, E.-J.: Assessing the uncertainty of estimated annual totals of net ecosystem productivity: a practical approach applied to a mid latitude temperate pine forest, Agr. Forest Meteorol., 151, 1823–1830, doi:10.1016/j.agrformet.2011.07.020, 2011.

Etzold, S., Buchmann, N., and Eugster, W.: Contribution of advection to the carbon budget measured by eddy covariance at a steep mountain slope forest in Switzerland, Biogeosciences, 7, 2461–2475, doi:10.5194/bg-7-2461-2010, 2010.

Farquhar, G., G., von Caemmerer, S., and Berry, J.-A.: A biochemical model of photosyntetic CO_2 assimilation in leaves of C_3 species, Planta, 149, 78–90, 1980.

BGD
12, 9393–9441, 2015

Gabriel, P., Gielen, B., Zona, D., Rodrigues, A., Rambal, S., Janssens, I., and Ceulemans, R.: Carbon and water vapor fluxes over four forests in two contrasting climatic zones, Agr. Forest Meteorol., 180, 211–224, 2013.

Garbulsky, M.-F., Penuelas, J., Papale, D., and Filella, I.: Remote estimation of carbon dioxide uptake by a Mediterranean forest, Glob. Change Biol., 14, 2860–2867, doi:10.1111/j.1365-2486.2008.01684.x, 2008.

Garrigues, S., Olioso, A., Calvet, J.-C., Martin, E., Lafont, S., Moulin, S., Chanzy, A., Marloie, O., Desfonds, V., Bertrand, N., and Renard, D.: Evaluation of land surface model simulations of evapotranspiration over a 12 year crop succession: impact of the soil hydraulic properties, Hydrol. Earth Syst. Sci. Discuss., 11, 11687–11733, doi:10.5194/hessd-11-11687-2014, 2014.

Gielen, B., B De Vos, Campioli, M., Neirynck, J., Papale, D., Verstraeten, A., Ceulemans, R., and Janssens, I.: Biometric and eddy covariance-based assessment of decadal carbon sequestration of a temperate Scots pine forest. Agr. Forest Meteorol., 174, 135–143, 2013.

Gioli, B., Miglietta, F., Vaccari, F.-P., Zaldei, A., and De Martino, B.: The Sky Arrow ERA, an innovative airborne platform to monitor mass, momentum and energy exchange of ecosystems, Ann. Geophys., 49, 109–116, 2006, http://www.ann-geophys.net/49/109/2006/.

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, doi:10.5194/bg-5-433-2008, 2008.

Groenendijk, M., Dolman, A.-J., van der Molen, M.-K., Leunung, R., Arneth, A., Delpierre, N., Gash, J.-H.-C., Lindroth, A., Richardson, A.-D., Verbeeck, H., and Wohlfahrt, G.: Assessing parameter variability in a photosynthesis model within and between plant functional types using global Fluxnet eddy covariance data, Agr. Forest Meteorol., 151, 22–38, doi:10.1016/j, 2011.

BGD
12, 9393–9441, 2015

Guidolotti, G., Rey, A., D'Andrea, E., Matteucci, G., and De Angelis, P.: Effect of environmental variables and stand structure on ecosystem respiration components in a Mediterranean beech forest, Tree Physiol., 33, 960–972, doi:10.1093/treephys/tpt065, 2013.

Gurney, K.-R., Law, R.-M., Denning, A.-S., Rayner, P.-J., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., Fan, S., Fung, I. Y., Gloor, M., Heimann, M., Higuchi, K., John, J., Maki, T., Maksyotov, S., Masarie, K., Peylin, P., Prather, M., Pak, B. C., Randerson, J., Sarmiento, J., Taguchi, S., Takahashi, T., and Yuen, C. W.: Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models, Nature, 415, 626–630, 2002.

Helfter, C., Campbell, C., Dinsmore, K.-J., Drewer, J., Coyle, M., Anderson, M., Skiba, U., Nemitz, E., Billett, M.-F., and Sutton, M.-A.: Drivers of long-term variability in CO_2 net ecosystem exchange in a temperate peatland, Biogeosciences, accepted, 2015.

Houweling, S., Breon, F.-M., Aben, I., Rödenbeck, C., Gloor, M., Heimann, M., and Ciais, P.: Inverse modeling of CO_2 sources and sinks using satellite data: a synthetic inter-comparison of measurement techniques and their performance as a function of space and time, Atmos. Chem. Phys., 4, 523–538, doi:10.5194/acp-4-523-2004, 2004.

Jans, W.-W.-P., Jacobs, C.-M.-J., Kruijt, B., Elbers, J.-A., Barendse, S., and Moors, E.-J.: Carbon exchange of a maize (*Zea mays* L.) crop: influence of phenology, Agr. Ecosyst. Environ., 139, 316–324, 2010.

Jarosz, N., Brunet, Y., Lamaud, E., Irvine, M., Bonnefond, J.-M., and Loustau, D.: Carbon dioxide and energy flux partitioning between the understorey and the overstorey of a maritime pine forest during a year with reduced soil availability, Agr. Forest Meteorol., 148, 1508–1523, 2008.

Jongen, M., Pereira, J., Saires., L.-M.-I., and Pio, C.-A.: The effects of drought and timing of precipitation on the inter-annual variation in ecosystem-atmosphere exchange in a Mediterranean grassland, Agr. Forest Meteorol., 151, 595–606, doi:10.1016/j.agrformet.2011.01.008, 2011.

Jung, M., Henkel, K., Herold, M., and Churkina, G.: Exploiting synergies of global land cover products for carbon cycle modeling, Remote Sens. Environ., 101, 534–553, doi:10.1016/j.rse.2006.01.020, 2006.

Keenan, T., García, R., Friend, A. D., Zaehle, S., Gracia, C., and Sabate, S.: Improved understanding of drought controls on seasonal variation in Mediterranean forest canopy CO_2 and water fluxes through combined in situ measurements and ecosystem modelling, Biogeosciences, 6, 1423–1444, doi:10.5194/bg-6-1423-2009, 2009.

BGD
12, 9393–9441, 2015

Klumpp, K., Tallec, T., Guix, N., and Soussana, J.-F.: Long-term impacts of agricultural practices and climatic variability on carbon storage in a permanent pasture, Glob. Change Biol., 17, 3534–3545, 2011.

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

Kottek, M., Grieser, J., Beck, C., Rudolf, B., and Rubel, F.: World Map of the Köppen–Geiger climate classification updated, Meteorol. Z., 15, 259–263, 2006.

Krinner, G., Viovz, N., Ogee, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, C.-I.: A dynamic global vegetation model for studies of the coupled atmosphere–biosphere system, Global Biogeochem. Cy., 19, GB1015, doi:10.1029/2003GB002199, 2005.

Kutsch, W.-L., Aubinet, M., Buchmann, N., Smith, P., Osborne, B., Eugster, W., Wattenbach, M., Schrumpf, M., Schulze, E. D., Tomelleri, E., Ceschia, E., Bernhofer, C., Beziat, P., Carrara, A., Tommasi, P., Grünwald, T., Jones, M., Magliulo, V., Marloie, O., Moureaux, C., Olioso, A., Sany, M. J., Saunders, M., Sogaard, H., and Ziegler, W.: The net biome production of full crop rotations in Europe, Agr. Ecosyst. Environ., 139, 336–345, 2010.

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, doi:10.5194/bg-5-1311-2008, 2008.

Lauvaux, T., Uliasz, M., Sarrat, C., Chevallier, F., Bousquet, P., Lac, C., Davis, K. J., Ciais, P., Denning, A. S., and Rayner, P. J.: Mesoscale inversion: first results from the CERES campaign with synthetic data, Atmos. Chem. Phys., 8, 3459–3471, doi:10.5194/acp-8-3459-2008, 2008.

Lauvaux, T., Gioli, B., Sarrat, C., Rayner, P.-J., Ciais, P., Chevallier, F., Noilhan, J., Miglietta, F., Brunet, Y., Ceschia, E., Dolman, H., Elbers, J.-A., Gerbig, C., Hutjes, R., Jarosz, N., Legain, D., and Uliasz, M.: Bridging the gap between atmospheric concentrations and local ecosystem measurements, Geophys. Res. Lett., 36, L19809, doi:10.1029/2009GL039574, 2009.

Lauvaux, T., Schuh, A.-E., Bocquet, M., Wu, L., Richardson, S., Miles, N., and Davis, K.-J.: Network design for mesoscale inversions of CO_2 sources and sinks, Tellus B, 64, 17980, doi:10.3402/tellusb.v64i0.17980, 2012.

BGD
12, 9393–9441, 2015

Longdoz, B., Gross, P., and Granier, A.: Multiple quality tests for analysing CO_2 fluxes in a beech temperate forest, Biogeosciences, 5, 719–729, doi:10.5194/bg-5-719-2008, 2008.
Mahadevan, P., Wofsy, S.-C., Matross, D.-M., Xiao, X., Dunn, A.-L., Lin, J.-C., Gerbig, C., Munger, J.-W., Chow, V.-Y., and Gottlieb, E.-W.: A satellite-based biosphere parameterization for net ecosystem CO_2 exchange: vegetation Photosynthesis and Respiration Model (VPRM), Global Biogeochem. Cy., 22, GB2005, doi:10.1029/2006GB002735, 2008.
Marcolla, B., Pitacco, A., and Cescatti, A.: Canopy architecture and turbulence structure in a coniferous forest, Bound.-Layer Meteorol., 108, 39–59, 2003.
Marcolla, B., Cescatti, A., Montagnani, L., Manca, G., Kerschbaumer, G., and Minerbi, S.: Importance of advection in the atmospheric CO_2 exchanges of an alpine forest, Agr. Forest Meteorol., 130, 193–206, 2005.
Marcolla, B., Cescatti, A., Manca, G., Zorer, R., Cavagna, M., Fiora, A., Gianelle, D., Rodeghiero, M., Sottocornola, M., and Zampedri, R.: Climatic controls and ecosystem responses drive the inter-annual variability of the net ecosystem exchange of an alpine meadow, Agr. Forest Meteorol., 151, 1233–1243, 2011.
Matteucci, M., Gruening, C., Goded., B., I., and Cescatti, A.: Soil and ecosystem carbon fluxes in a Mediterranean forest during and after drought, Agrochimica, LVIII, 91–115, 2014.
Meesters, A.-G.-C.-A., Tolk, L.-F., Peters, W., Hutjes, R.-W.-A., Vellinga, O.-S., Elbers, J.-A., Vermeulen, A.-T., van der Laan, S., Neubert, R.-E.-M., Meijer, H.-A.-J., and Dolman, A. J.: Inverse carbon dioxide flux estimates for the Netherlands, J. Geophys. Res.-Atmos., 117, 1984–2012, doi:10.1029/2012jd017797, 2012.
Meijide, A., Manca, G., Goded, I., Magliulo, V., di Tommasi, P., Seufert, G., and Cescatti, A.: Seasonal trends and environmental controls of methane emissions in a rice paddy field in Northern Italy, Biogeosciences, 8, 3809–3821, doi:10.5194/bg-8-3809-2011, 2011.
Michalak, A.-M., Bruhwiler, L., and Tans, P.-P.: A geostatistical approach to surface flux estimation of atmospheric trace gases, J. Geophys. Res., 109, D14109, doi:10.1029/2003JD004422, 2004.
Montagnani, L., Manca, G., Canepa, E., Georgieva, E., Acosta, M., Feigenwinter, C., Janous, D., Kerschbaumer, G., Lindroth, A., Minach, L., Minerbi, S., Mölder, M., Pavelka, M., Seufert, G., Zeri, M., and Ziegler, W.: A new mass conservation approach to the study of CO_2 advection in an alpine forest, J. Geophys. Res.-Atmos., 114, D07306, doi:10.1029/2008JD010650, 2009.

BGD
12, 9393–9441, 2015

Moors, E.-J., Jacobs, C., Jans, W., Supit, I., Kutsch, W.-l, Bernhofer, C., Bezat, P., Buchmann, N., Carrara, A., Ceschia, E., Elbers, J., Eugster, W., Kruijt, B., Loubet, B., Magliulo, E., Moureaux, C., Olioso, A., Saunders, M., and Soegaard, H.: Variability in carbon exchange of European croplands, Agr. Ecosyst. Environ., 139, 325–335, 2010

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

Nagy, Z., Pintér, K., Czóbel, Sz., Balogh, J., Horváth, L., Fóti, Sz., Barcza, Z., Weidinger, T., Csintalan, Zs., Dinh, N., Q., Grosz, B., and Tuba, Z.: The carbon budget of a semiarid grassland in a wet and a dry year in Hungary, Agr. Ecosyst. Environ., 121, 21–29, 2007.

Olson, J.-S.: Global Ecosystem Framework-Definitions, USGS EROS Data Cent. Intern. Rep., Sioux Falls, SD, 37, 1994.

Peters, W., Jacobson, A.-R., Sweeney, C., Andrews, A.-E., Conway, T.-J., Masarie, K.-B., Miller, J., Bruhwiler, L.-M.-P., Petron, G., Hirsch, A.-I., Worthy, D.-E.-J., van der Werf, G.-R., Wennberg, J.-T.-R.-P.-O., Krol, M.-C., and Tans, P.-P.: An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, P. Natl. Acad. Sci. USA, 104, 18925–18930, doi:10.1073pnas.0708986104, 2007.

Peylin, P., Rayner, P. J., Bousquet, P., Carouge, C., Hourdin, F., Heinrich, P., Ciais, P., and AEROCARB contributors: Daily CO_2 flux estimates over Europe from continuous atmospheric measurements: 1, inverse methodology, Atmos. Chem. Phys., 5, 3173–3186, doi:10.5194/acp-5-3173-2005, 2005.

Piao, S., Fang, J., Ciais, P., Peylin, P., Huang, Y., Sitch, S., and Wang, T.: The carbon balance of terrestrial ecosystems in China, Nature, 458, 1009–1013, doi:10.1038/nature07944, 2009.

Pillai, D., Gerbig, C., Kretschmer, R., Beck, V., Karstens, U., Neininger, B., and Heimann, M.: Comparing Lagrangian and Eulerian models for CO_2 transport – a step towards Bayesian inverse modeling using WRF/STILT-VPRM, Atmos. Chem. Phys., 12, 8979–8991, doi:10.5194/acp-12-8979-2012, 2012.

Pita, G., Gielen, B., Zona, D., Rodrigues, A., Rambal, S., Janssens, I., and Ceulemans, R.: Carbon and water vapor fluxes over four forests in two contrasting climatic zones, Agr. Forest Meteorol., 180, 211–224, 2013.

Prescher, A., K., Grünwald, T., and Bernhofer, C.: Land use regulates carbon budgets in eastern Germany: from NEE to NBP, Agr. Forest Meteorol., 150, 1016–1025, 2010.

BGD
12, 9393–9441, 2015

Rey, A., Belelli-Marchesini, L., Were, A., Serrano-Ortiz, P., Etiope, G., Papale, D., Domingo, F., and Pegoraro, E.: Wind as a main driver of the net ecosystem carbon balance of a semiarid Mediterranean steppe in the South East of Spain, Glob. Change Biol., 18, 539–554, 2012.

Richardson, A.-D., Hollinger, D.-Y., Burba, G.-G., Davis, K.-J., Flanagan, L.-B., Katul, G.-G., William, M.-J., 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, Agr. Forest Meteorol., 136, 1–18, 2006.

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 CO_2 flux measurement uncertainty inferred from model residuals, Agr. Forest Meteorol., 148, 38–50, doi:10.1016/j.agrformet.2007.09.001, 2008.

Rödenbeck, C.: Estimating CO_2 sources and sinks from atmospheric mixing ratio measurements using a global inversion of atmospheric transport, Technical Report 6, Max Planck Institute for Biogeochemistry, Jena, 2005.

Rödenbeck, C., Houweling, S., Gloor, M., and Heimann, M.: Time-dependent atmospheric CO_2 inversions based on interannually varying tracer transport, Tellus B, 55, 488–497, 2003a.

Rödenbeck, C., Houweling, S., Gloor, M., and Heimann, M.: CO_2 flux history 1982–2001 inferred from atmospheric data using a global inversion of atmospheric transport, Atmos. Chem. Phys., 3, 1919–1964, doi:10.5194/acp-3-1919-2003, 2003b.

Rödenbeck, C., Gerbig, C., Trusilova, K., and Heimann, M.: A two-step scheme for high-resolution regional atmospheric trace gas inversions based on independent models, Atmos. Chem. Phys., 9, 5331–5342, doi:10.5194/acp-9-5331-2009, 2009.

Schuh, A. E., Denning, A. S., Corbin, K. D., Baker, I. T., Uliasz, M., Parazoo, N., Andrews, A. E., and Worthy, D. E. J.: A regional high-resolution carbon flux inversion of North America for 2004, Biogeosciences, 7, 1625–1644, doi:10.5194/bg-7-1625-2010, 2010.

Skiba, U., Jones, S. K., Drewer, J., Helfter, C., Anderson, M., Dinsmore, K., McKenzie, R., Nemitz, E., and Sutton, M. A.: Comparison of soil greenhouse gas fluxes from extensive and intensive grazing in a temperate maritime climate, Biogeosciences, 10, 1231–1241, doi:10.5194/bg-10-1231-2013, 2013.

Soussana, J.-F., Allard, V., Pilegaard, K., Ambus, P., Amman, C., Campbell, C., Ceschia, E., Clifton-Brown, J., Czóbel, Sz., Domingues, R., Flechard, C., Fuhrer, J., Henseh, A., Horvath, L., Jones, M., Kasper, G., Martin, C., Nagy, Z., Neftel, A., Raschi, A., Baronti, S., Rees, R. M., Skiba, U., Stefani, P., Manca, G., Sutton, M., Tuba, Z., and Valentini, R.: Full accounting of

BGD
12, 9393–9441, 2015

the greenhouse gas (CO_2, N_2O, CH_4) budget of nine European grassland sites, Agr. Ecosyst. Environ., 121, 121–134, 2007.
Suni, T., Rinne, J., Reissel, A., Altimir, N., Keronen, P., Rannik, Ü., Dal Maso, M., Kulmala, M., and Vesala, T.: Long-term measurements of surface fluxes above a Scots pine forest in Hyytiälä, southern Finland, 1996–2001, Boreal Environ. Res. 4, 287–301, 2003.
Tallec, T., Béziat, P., Jarosz, N., Rivalland, V., and Ceschia, E.: Crops water use efficiencies: comparison of stand, ecosystem and agronomical approaches, Agr. Forest. Meteorol., 168, 69–81, 2013.
Taufarova, K., Havrankova, K., Dvorská, A., Pavelka, M., Urbaniak, M., and Janous, D.: Forest ecosystem as a source of CO_2 during growing season: relation to weather conditions, Int. Agrophys., 28, 239–249, doi:10.2478/intag-2014-0013, 2014.
Tolk, L. F., Dolman, A. J., Meesters, A. G. C. A., and Peters, W.: A comparison of different inverse carbon flux estimation approaches for application on a regional domain, Atmos. Chem. Phys., 11, 10349–10365, doi:10.5194/acp-11-10349-2011, 2011.
Weedon, G.-P., Balsamo, G., Bellouin, N., Gomes, S., Best, M.-J., and Viterbo, P.: The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERAInterim reanalysis data, Water Resour. Res., 50, 7505–7514, 2014.
Wei, S, Yi, C., Hendrey, G., Eaton, T., Rustic, G., Wang, S., Liu, H., Krakauer, N.-Y., Wang, W., Desai, A.-R., Montagnani, L., Paw, K. T., Falk, M., Black, A., Bernhofer, C., Grünwald, T., Laurila, T., Cescatti, A., Moors, E., Bracho, R., and Valentini, R.: Data-based perfect-deficit approach to understanding climate extremes and forest carbon assimilation capacity, Environ. Res. Lett., 9, 065002, doi:10.1088/1748-9326/9/6/065002, 2014.
Wu, L., Bocquet, M., Lauvaux, T., Chevallier, F., Rayner, P., and Davis, K.: Optimal representation of source-sink fluxes for mesoscale carbon dioxide inversion with synthetic data, J. Geophys. Res., 116, D21304, doi:10.1029/2011JD016198, 2011.
Zeeman, M., J., Hiller, R., Gilgen, A.-K., Michna, P., Plüss, P., Buchmann, N., and Eugster, W.: Management and climate impacts on net CO_2 fluxes and carbon budgets of three grasslands along an elevational gradient in Switzerland, Agr. Forest Meteorol., 150, 519–530, doi:10.1016/j.agrformet.2010.01.011, 2010.
Zweifel, R., Eugster, W., Etzold, S., Dobbertin, M., Buchmann, N., and Häsler, R.: Link between continuous stem radius changes and net ecosystem productivity of a subalpine Norway spruce forest in the Swiss Alps, New Phytol., 187, 819–830, doi:10.1111/j.1469-8137.2010.03301.x, 2010.

BGD
12, 9393–9441, 2015

Table 1. Eddy covariance sites measuring CO_2 fluxes that were used in the analysis. The land cover classification which is used, is coded as follows; CRO, DCF, EVG, MF, GRA, OSH, SAV for crops, deciduous forest, evergreen forest, mixed forest, grass, shrub and savanna respectively.

Site code	Site name	Land cover classification	Latitude	Longitude	Citation
BE-Bra	Brasschaat	MF	51.31	4.52	Gielen et al. (2013)
BE-Lon	Lonzee	CRO	50.55	4.74	Moureaux et al. (2006)
BE-Vie	Vielsalm	MF	50.31	6.00	Aubinet et al. (2001)
CH-Cha	Chamau	GRA	47.21	8.41	Zeeman et al. (2010)
CH-Dav	Davos	ENF	46.82	9.86	Zweifel et al. (2010)
CH-Fru	Frebel	GRA	47.12	8.54	Zeeman et al. (2010)
CH-Lae	Laegern	MF	47.48	8.37	Etzold et al. (2010)
CH-Oe1	Oensingen grassland	GRA	47.29	7.73	Ammann et al. (2009)
CH-Oe2	Oensingen crop	CRO	47.29	7.73	Dietiker et al. (2010)
CZ-BK1	Bily Kriz forest	ENF	49.50	18.54	Taufarova et al. (2014)
DE-Geb	Gebesee	CRO	51.10	10.91	Kutsch et al. (2010)
DE-Gri	Grillenburg	GRA	50.95	13.51	Prescher et al. (2010)
DE-Hai	Hainich	DBF	50.79	10.45	Knohl et al. (2003)
DE-Kli	Klingenberg	CRO	50.89	13.52	Prescher et al. (2010)
DE-Tha	Tharandt	ENF	50.96	13.57	Prescher et al. (2010)
DK-Lva	Rimi	GRA	55.68	12.08	Soussana et al. (2007)
ES-Agu	Aguamarga	OSH	36.94	−2.03	Rey et al. (2012)
ES-ES2	El Saler-Sueca (Valencia)	CRO	39.28	−0.32	–
ES-LMa	Las Majadas del Tietar (Caceres)	SAV	39.94	−5.77	Casals et al. (2011)
FI-Hyy	Hyytiälä	ENF	61.85	24.30	Suni et al. (2003)
FR-Aur	AuradeŽ	CRO	43.55	1.11	Tallec et al. (2013)
FR-Avi	Avignon	CRO	43.92	4.88	Garrigues et al. (2014)
FR-Fon	Fontainebleau	DBF	48.48	2.78	Delpierre et al. (2009)
FR-Hes	Hesse	DBF	48.67	7.07	Longdoz et al. (2008)
FR-LBr	Le Bray	ENF	44.72	−0.77	Jarosz el al. (2008)

BGD
12, 9393–9441, 2015

Table 1. Continued.

Site code	Site name	Land cover classification	Latitude	Longitude	Citation
FR-Lq1	Laqueuille intensive	GRA	45.64	2.74	Klumpp et al. (2011)
FR-Lq2	Laqueuille extensive	GRA	45.64	2.74	Klumpp et al. (2011)
FR-Mau	Mauzac	GRA	43.39	1.29	Albergel et al. (2010)
FR-Pue	Puechabon	EBF	43.74	3.60	Allard et al. (2008)
HU-Mat	Matra	CRO	47.85	19.73	Nagy et al. (2007)
IT-Amp	Amplero	GRA	41.90	13.61	Barcza et al. (2007)
IT-BCi	Borgo Cioffi	CRO	40.52	14.96	Kutsch et al. (2010)
IT-Cas	Castellaro	CRO	45.07	8.72	Meijide et al. (2011)
IT-Col	Collelongo	DBF	41.85	13.59	Guidolotti et al. (2013)
IT-Cpz	Castelporziano	EBF	41.71	12.38	Garbulsky et al. (2008)
IT-Lav	Lavarone	ENF	45.96	11.28	Marcolla et al. (2003)
IT-Lec	Lecceto	EBF	43.30	11.27	Chiesi et al. (2011)
IT-LMa	Malga Arpaco	GRA	46.11	11.70	Soussana et al. (2007)
IT-MBo	Monte Bondone	GRA	46.01	11.05	Marcolla et al. (2011)
IT-Ren	Renon	ENF	46.59	11.43	Marcolla et al. (2005)
IT-Ro2	Roccarespampani 2	DBF	42.39	11.92	Wei et al. (2014)
IT-SRo	San Rossore	ENF	43.73	10.28	Matteucci et al. (2014)
NL-Dij	Dijkgraaf	CRO	51.99	5.65	Jans et al. (2010)
NL-Loo	Loobos	ENF	52.17	5.74	Elbers et al. (2011)
NL-Lut	Lutjewad	CRO	53.40	6.36	Moors et al. (2010)
PT-Esp	Espirra	EBF	38.64	−8.60	Gabriel et al. (2013)
PT-Mi2	Mitra IV (Tojal)	GRA	38.48	−8.02	Jongen et al. (2011)
SE-Kno	Knottœsen	ENF	61.00	16.22	–
SE-Nor	Norunda	ENF	60.09	17.48	–
SE-Sk1	Skyttorp 1	ENF	60.13	17.92	–
SK-Tat	Tatra	ENF	49.12	20.16	–
UK-AMo	Auchencorth Moss	GRA	55.79	−3.24	Helfter et al. (2015)
UK-EBu	Easter Bush	GRA	55.87	−3.21	Skiba et al. (2013)

BGD
12, 9393–9441, 2015

Table 2. Annual temporal autocorrelation times in days, from model-data and model-model residuals. The number within the brackets shows the correlation times when excluding sites with large model-data bias from the analysis.

Reference	VPRM10 [days]	VPRM1 [days]	ORCHIDEE [days]	5PM [days]
OBSERVATION	32 (27)	33 (29)	26 (24)	70 (34)
VPRM10	–	47 (46)	30 (31)	131 (100)
VPRM1	–	–	28 (28)	116 (85)
ORCHIDEE	–	–	–	38 (32)
5PM	–	–	–	–

BGD
12, 9393–9441, 2015

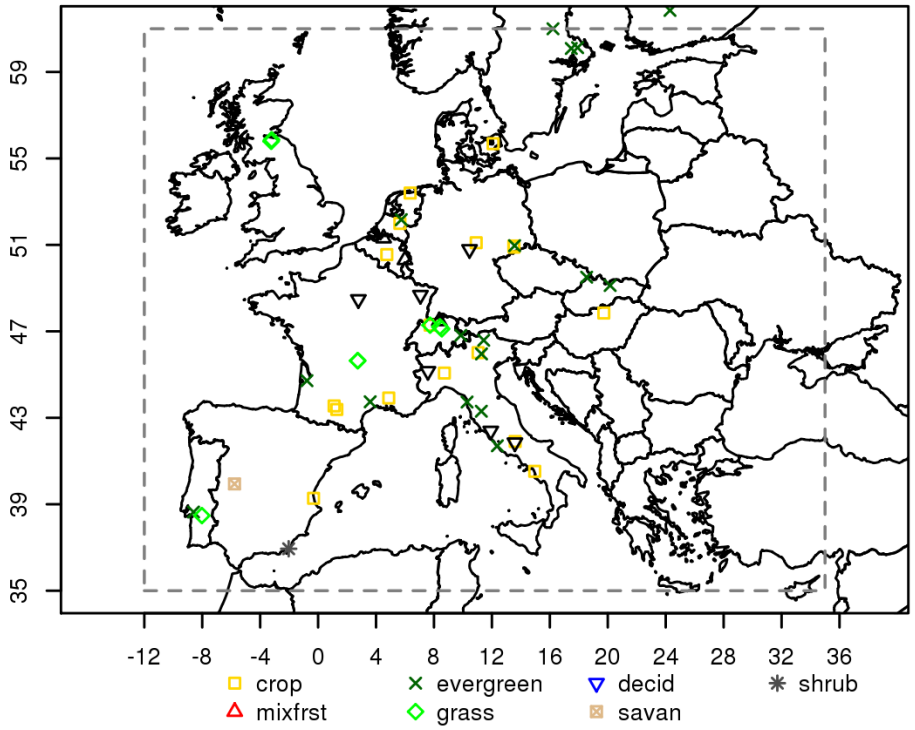

Figure 1. Eddy covariance sites used in the study. The dashed line delimits the exact domain used to calculate the aggregated fluxes.

BGD
12, 9393–9441, 2015

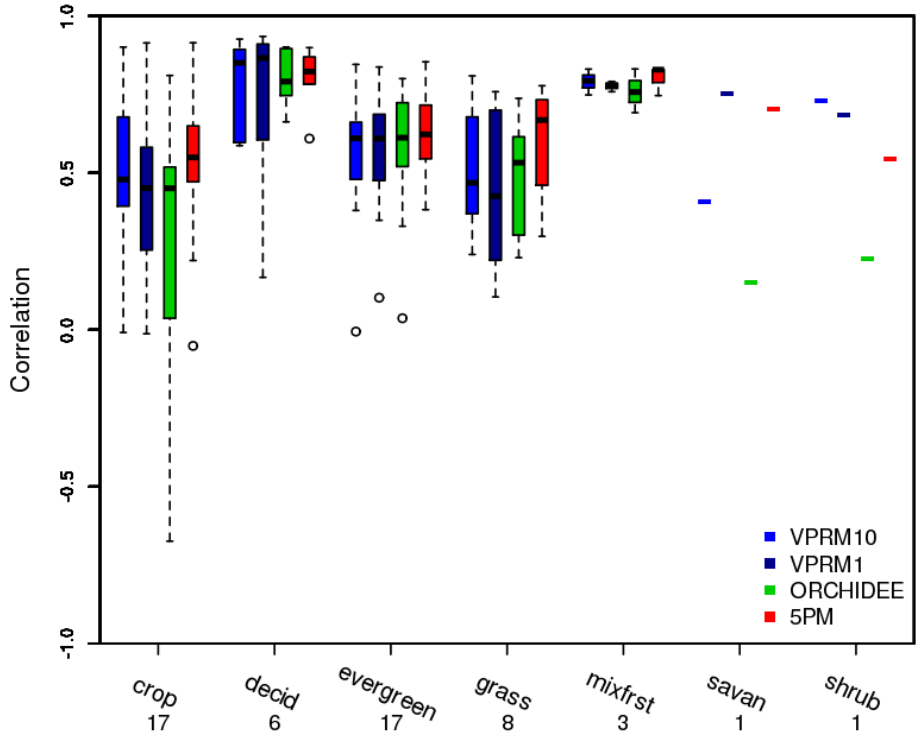

Figure 2. Box and whisker plot for correlation coefficients between modeled and observed daily fluxes as a function of the vegetation type. The numbers beneath the *x* axis indicate the number of sites involved. The bottom and the top of the box denote the first and the third quartiles. The band inside the box indicates the central 50 % and the line within is the median. Upper and lower line edges denote the maximum and the minimum values excluding outliers. Outliers are shown as circles.

BGD
12, 9393–9441, 2015

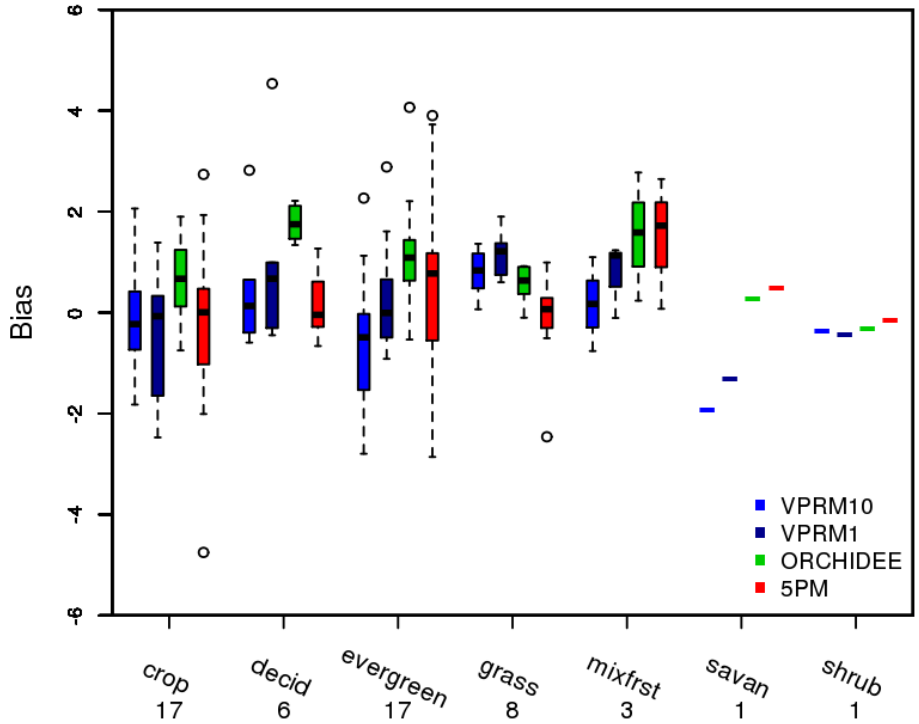

Figure 3. Box and whisker plot for the annual site specific biases of the models differentiated by vegetation type. Units at y axis are in $\mu mol\, m^{-2}\, s^{-1}$.

BGD
12, 9393–9441, 2015

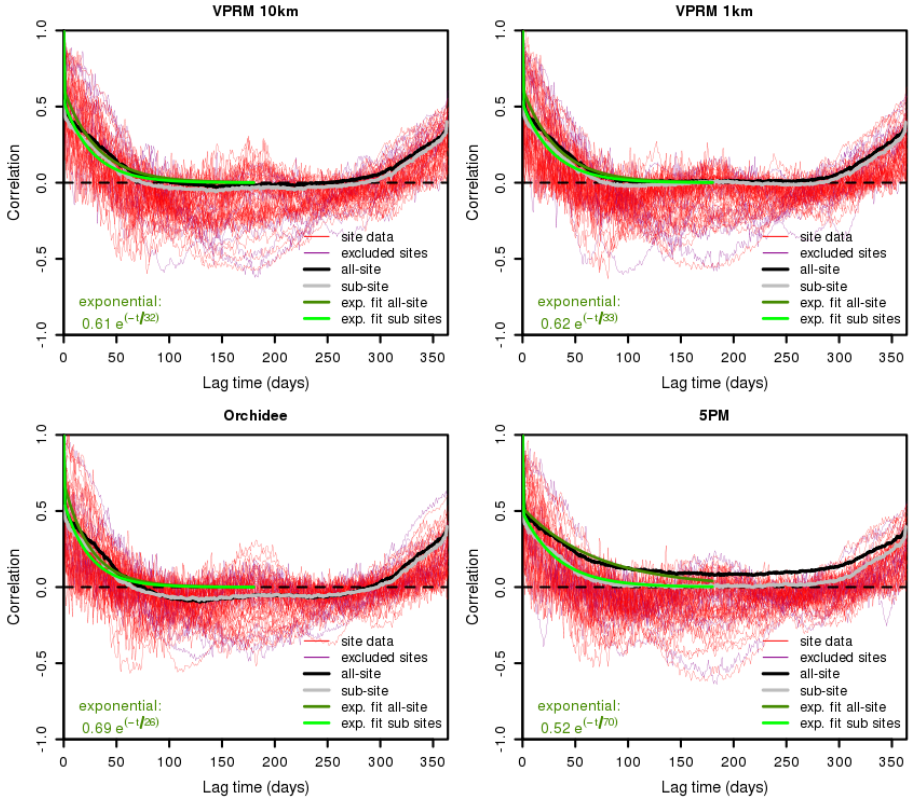

Figure 4. Temporal lagged autocorrelation from model-data daily averaged NEE residuals for all models. Red lines correspond to different sites while the dark magenta color reveals the sites with a bias larger than $\pm 2.5\,\mu mol\,m^{-2}\,s^{-1}$. Black line shows the all-site autocorrelation, and the grey line indicates the autocorrelation excluding sites with large model-data bias ("sub-site"). The dark green line is the all-site exponential fit using lags up to 180 days, and the light green line shows the all-site autocorrelation excluding the sites with large bias.

BGD
12, 9393–9441, 2015

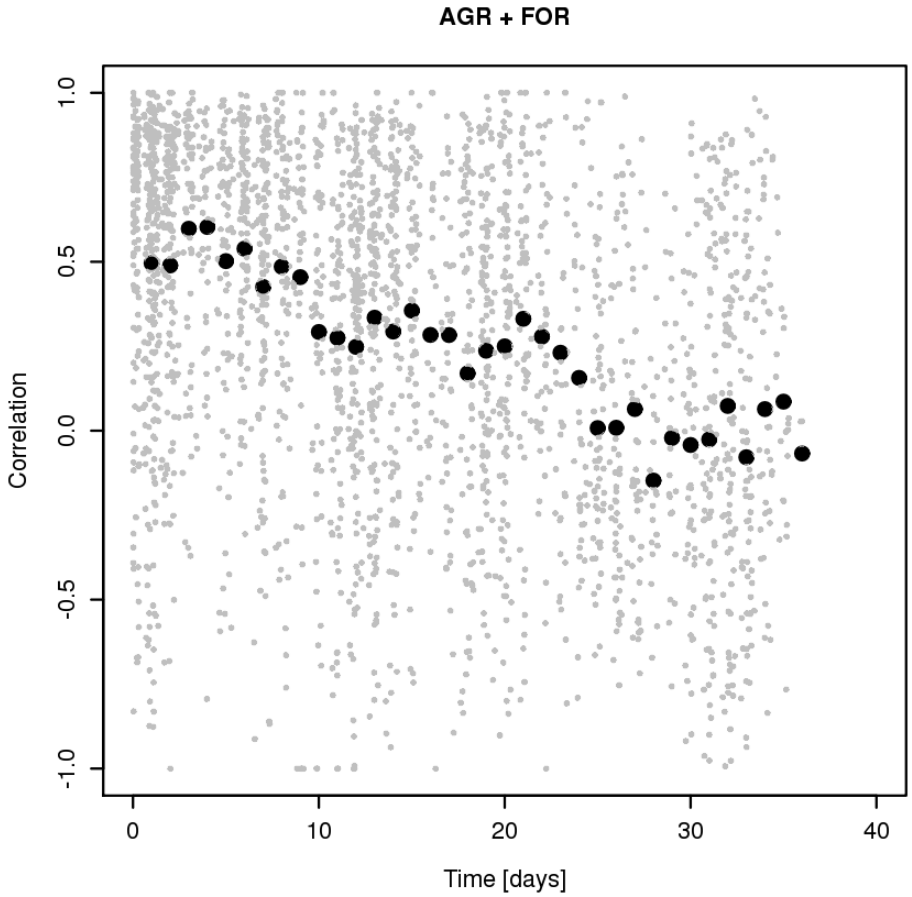

Figure 5. Temporal autocorrelation for VPRM10 – aircraft NEE residuals. Black dots represent individual flux transects pairs sampled at different times as function of time separation. Black circles represent daily scale binned data.

BGD
12, 9393–9441, 2015

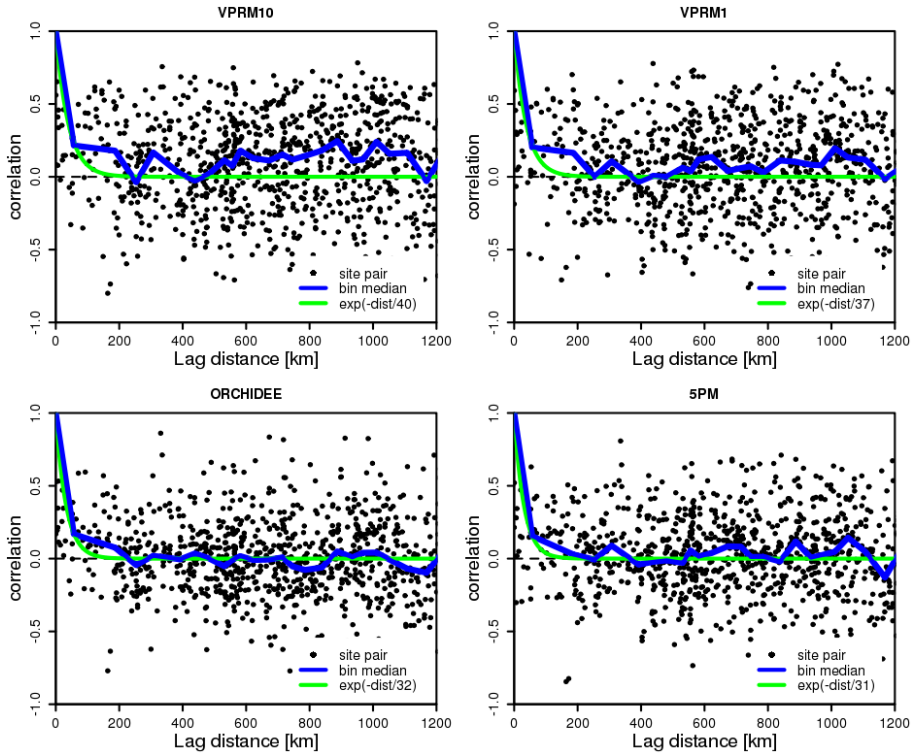

Figure 6. Distance correlogram for the daily net ecosystem exchange (NEE) residuals using all sites. Black dots represent the different site pairs; the blue line represents the median value of the points per 100 km bin and the green an exponential fit. Results are shown for residuals of VPRM at a resolution of 10 km (top left) and 1 km (top right), ORCHIDEE (bottom left), 5PM (bottom right).

BGD
12, 9393–9441, 2015

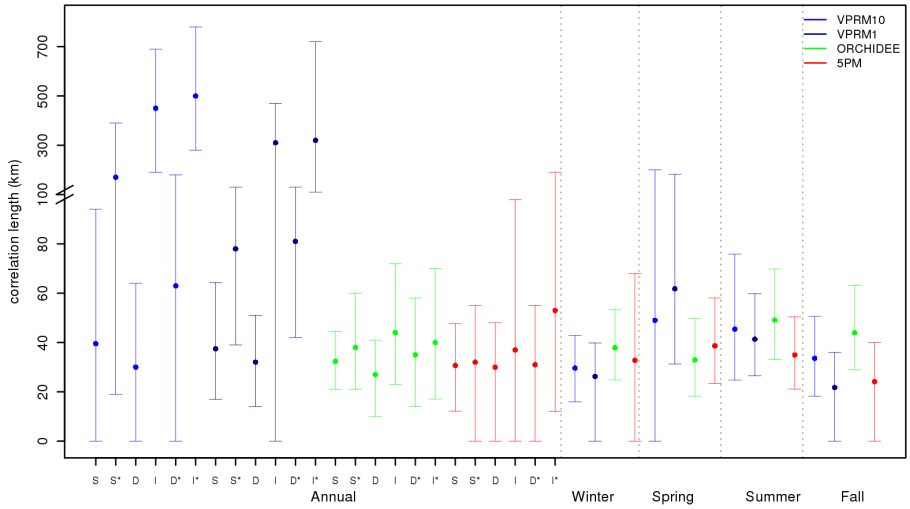

Figure 7. Annual and seasonal e-folding correlation length of the daily averaged model-data NEE residuals for VPRM at 10 and 1 km resolution, ORCHIDEE and 5PM. "S" refers to the standard case where all pairs were used, "D" refers to the case where only pairs with different vegetation types were used, "I" denotes the case in which only pairs with identical vegetation type were considered, and "*" denotes that in addition 150 days of common non-missing data are required for each pair of sites. The dot represents the best-fit value when fitting the exponential model. The upper and the lower edge of the error bars show the 2.5 and 97.5 % of the length value. Note the scale change in the *y* axis at 100 km.

BGD
12, 9393–9441, 2015

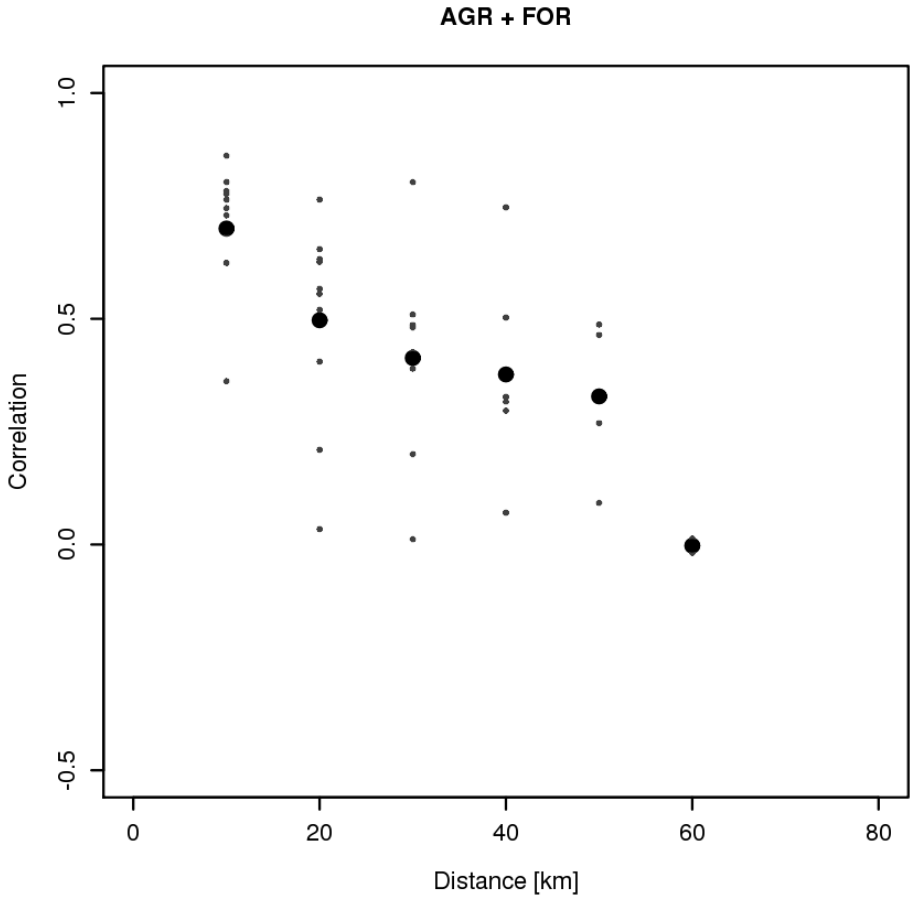

Figure 8. Distance correlogram between VPRM10 and aircraft NEE measurements. Black dots represents the different aircraft grid points pairs; black circles represent 10 km scale binned data.

BGD
12, 9393–9441, 2015

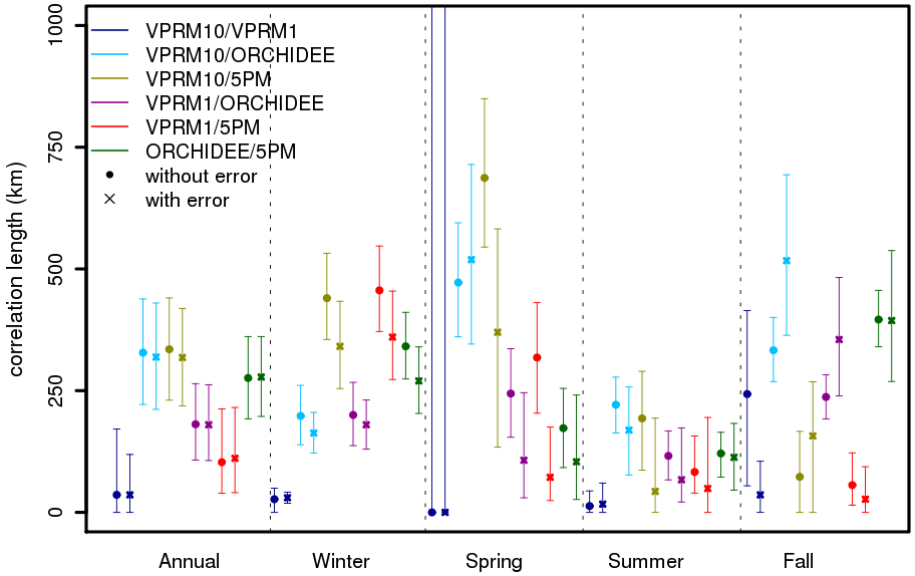

Figure 9. Annual and seasonal e-folding correlation length for an ensemble of daily averaged NEE differences between two models without (filled circle) and with random measurement errors added to the modeled fluxes used as reference (crosses). The symbols represents the best fit value when fitting the exponential model, and the upper and lower edge of the error bars show the 2.5 and 97.5 % of the correlation length. The first acronym at the legend represents the model used as reference and the second the model which was compared with. Note that for the VPRM10/VPRM1 case during spring (with and without random error), the 97.5 % of the length value exceeds the *y* axis and has a value of 1073, 1626 km respectively.

BGD
12, 9393–9441, 2015

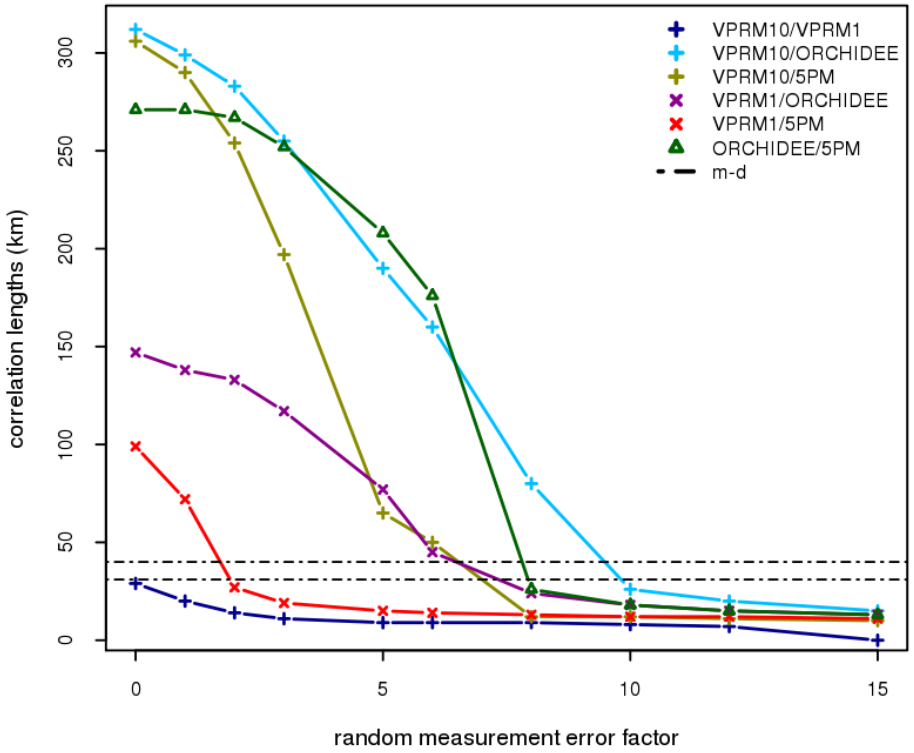

Figure 10. Annual correlation lengths as a function of the factor used for scaling the random measurement error, for all model-model combinations. The black dot-dash lines reveal the range of the spatial correlation lengths generated from the model-data comparisons.

BGD
12, 9393–9441, 2015