



Evaluation and uncertainty analysis of regional-scale CLM4.5 net carbon flux estimates

Hanna Post^{1,2,3}, Harrie-Jan Hendricks Franssen^{1,3}, Xujun Han^{1,3}, Roland Baatz^{1,3}, Carsten Montzka¹, Marius Schmidt¹, and Harry Vereecken^{1,3}

¹Agrosphere (IBG-3), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

²Institute of Geography, University of Cologne, Cologne, Germany

³Centre for High-Performance Scientific Computing in Terrestrial Systems: HPSC TerrSys, Geoverbund ABC/J, Leo-Brandt-Strasse, 52425 Jülich, Germany

Correspondence: Harrie-Jan Hendricks Franssen (h.hendricks-franssen@fz-juelich.de)

Received: 12 December 2016 – Discussion started: 2 January 2017

Revised: 9 November 2017 – Accepted: 17 November 2017 – Published: 10 January 2018

Abstract. Modeling net ecosystem exchange (NEE) at the regional scale with land surface models (LSMs) is relevant for the estimation of regional carbon balances, but studies on it are very limited. Furthermore, it is essential to better understand and quantify the uncertainty of LSMs in order to improve them. An important key variable in this respect is the prognostic leaf area index (LAI), which is very sensitive to forcing data and strongly affects the modeled NEE. We applied the Community Land Model (CLM4.5-BGC) to the Rur catchment in western Germany and compared estimated and default ecological key parameters for modeling carbon fluxes and LAI. The parameter estimates were previously estimated with the Markov chain Monte Carlo (MCMC) approach DREAM_(ZS) for four of the most widespread plant functional types in the catchment. It was found that the catchment-scale annual NEE was strongly positive with default parameter values but negative (and closer to observations) with the estimated values. Thus, the estimation of CLM parameters with local NEE observations can be highly relevant when determining regional carbon balances. To obtain a more comprehensive picture of model uncertainty, CLM ensembles were set up with perturbed meteorological input and uncertain initial states in addition to uncertain parameters. C₃ grass and C₃ crops were particularly sensitive to the perturbed meteorological input, which resulted in a strong increase in the standard deviation of the annual NEE sum ($\sigma_{\sum \text{NEE}}$) for the different ensemble members from ~ 2 to $3 \text{ g C m}^{-2} \text{ yr}^{-1}$ (with uncertain parameters) to $\sim 45 \text{ g C m}^{-2} \text{ yr}^{-1}$ (C₃ grass) and $\sim 75 \text{ g C m}^{-2} \text{ yr}^{-1}$ (C₃

crops) with perturbed forcings. This increase in uncertainty is related to the impact of the meteorological forcings on leaf onset and senescence, and enhanced/reduced drought stress related to perturbation of precipitation. The NEE uncertainty for the forest plant functional type (PFT) was considerably lower ($\sigma_{\sum \text{NEE}} \sim 4.0\text{--}13.5 \text{ g C m}^{-2} \text{ yr}^{-1}$ with perturbed parameters, meteorological forcings and initial states). We conclude that LAI and NEE uncertainty with CLM is clearly underestimated if uncertain meteorological forcings and initial states are not taken into account.

1 Introduction

Net ecosystem CO₂ exchange (NEE), the difference of CO₂ release via soil and plant respiration and photosynthetic CO₂ uptake, is an important indicator for the net carbon sink or source function of terrestrial ecosystems. The understanding of factors controlling the spatial and temporal variability of carbon fluxes like respiration is still limited (Reichstein and Beer, 2008). Eddy covariance (EC) net carbon flux measurements are limited to a relatively small area. Chen et al. (2012) showed that the 90 % cumulative annual footprint area of 12 EC towers located at Canadian sites (with different land cover including grassland and forest) varied from about 1.1 to 5.0 km², and that the spatial representativeness of the EC flux measurements depends on the degree of the land surface heterogeneity. Biogeochemical fluxes are spatially and temporally highly variable and nonlinear due to the spatial het-

erogeneity of soil properties, vegetation and fauna, and to the temporal variability of the environmental drivers (e.g., meteorological conditions, management schemes; Chen et al., 2009; Stoy et al., 2009; Borchard et al., 2015). Therefore, conventional interpolation methods, such as kriging or inverse distance weighting, are not suitable for upscaling EC carbon flux measurements to larger areas.

Land surface models like CLM (Oleson et al., 2013) simulate the coupled carbon, nitrogen, water and energy cycle of the land surface, and are essential to understand interactions between the climate and the terrestrial carbon cycle and to predict climate–ecosystem feedbacks (e.g., Le Quéré et al., 2012; Arora et al., 2013; Brovkin et al., 2013; Todd-Brown et al., 2014). In this study, CLM version 4.5 in the biogeochemistry (BGC) mode (CLM4.5-BGC) was applied. The prognostically calculated leaf area index (LAI) is a major indicator for the model representation of plant phenology. Moreover, it affects photosynthesis and transpiration. Therefore, the LAI is a key state variable for carbon flux predictions and land surface–atmosphere exchange fluxes of water and carbon. Thus, a correct representation of the simulated LAI in terms of magnitude and timing is highly desirable. The representation of plant phenology (timing of plant emergence, length in growing season) has been shown to be seriously flawed in land surface models (LSMs; e.g., Richardson et al., 2012) including CLM (Dahlin et al., 2015), which can also affect carbon flux estimates (Baldocchi and Wilson, 2001; Richardson et al., 2012). Simulated carbon fluxes and the prognostic LAI in CLM4.5-BGC are closely linked, because they depend on common ecological key parameters and plant phenology schemes.

Commonly, LSMs are applied at global or continental scale (e.g., Stöckli et al., 2008; Bonan et al., 2011; Lawrence et al., 2012) with grid sizes of ~ 0.25 – 1.5° . At this coarse scale with such a high degree of spatial aggregation, and given the non-linearity of the governing equations, the modeled values for output variables like NEE can deviate strongly from those that would have been calculated with a fine-resolution model using fine-resolution input. Moreover, reliable calibration and validation of global LSMs is difficult, because observed data including soil carbon stocks and EC fluxes are only available for single locations. When applying a LSM for a small region or catchment with high spatial resolution (e.g., 1 km^2 , as in this study) the error of simulated fluxes is expected to be smaller due to the lower degree of spatial aggregation (Anderson et al., 2003). In addition, the land cover within a 1 km^2 grid cell more likely matches with the land cover at the EC site, which enables grid-based evaluation of modeled NEE. High spatial resolution can better represent the land surface heterogeneity and regional weather variability than a coarse spatial resolution. Thus, regional or catchment-scale applications of LSMs allow for investigating spatial patterns of model states, biogeochemical fluxes and interactions with the regional climate and catchment hydrology. Accordingly, quantification of carbon fluxes at re-

gional scales can enhance the understanding of CO_2 dynamics and their drivers (Desai et al., 2008). This has been shown in various studies, for example, for western Africa (Bonan et al., 2002; Li et al., 2007) and the Alaskan Arctic (Fisher et al., 2014). However, to our knowledge, studies like Xiao et al. (2011), who optimized a simple ecosystem model to up-scale measured EC carbon fluxes to the regional scale, do not exist yet for more complex LSMs like CLM. This is because (i) high-resolution input data are often not available, (ii) the implementation of a new model setup to a specific region is relatively time consuming and (iii) careful parameter estimation is required to allow for meaningful predictions.

LSM predictions of carbon, water and energy fluxes are still subject to a high degree of uncertainty due to (i) model structural deficits related to an imperfect and incomplete model representation of the biogeochemical processes (Todd-Brown et al., 2012; Foereid et al., 2014), (ii) poorly constrained model parameters (Abramowitz et al., 2008; Beven and Freer, 2001; Todd-Brown et al., 2013), (iii) errors in the representation of initial model conditions generated via a spin-up (Carvalhais et al., 2010; Kuppel et al., 2012; Xia et al., 2012; Exbrayat et al., 2014) and (iv) errors in both atmospheric and land surface input data. Some studies estimate the uncertainty of terrestrial carbon flux predictions based on an ensemble of many different LSMs (e.g., Fisher et al., 2014; Huntzinger et al., 2012; Piao et al., 2013; Zhao et al., 2016). These studies focus on differences between models and therefore model structural deficits. Those studies highlight that (i) carbon flux predictions are generally highly uncertain, which contributes to the uncertainty in climate change predictions, (ii) interactions of the different processes and drivers are not understood satisfactorily and (iii) models require structural improvement to produce more consistent predictions. In order to improve LSM model structure and thus model–data and inter-model consistency, a more comprehensive understanding of model functionality and the contribution and link of the different model error sources is required. However, as highlighted by Xiao et al. (2014), the uncertainty of carbon fluxes obtained by ecosystem and land surface models has not been analyzed and quantified enough, particularly in regional-scale studies.

Whereas the uncertainty of land surface model parameters has been subject to intensive investigations (e.g., Ren et al., 2013; Xiao et al., 2014), and several works are dedicated to reducing this uncertainty, for example by parameter estimation with data assimilation methods (e.g., Safta et al., 2015), the other sources of uncertainty are less studied. Earlier work has concluded that parameter uncertainty alone cannot explain observed deviations between measured and simulated NEE (e.g., Pridhoko et al., 2008; Wang et al., 2011), and the remaining deviations are often attributed to model structural errors. However, uncertainty in atmospheric forcings and initial conditions could also contribute to unexplained deviations between simulated and measured NEE. From global-scale studies it is known that NEE is sensitive to

the climate scenario. For example, in a study the LSM LPJ-GUESS was forced with output from 18 different coupled atmosphere–ocean global circulation models or Earth system models, and showed very different NEE responses depending on the climate scenario. Of the 18 models, 10 project that the land will become a carbon source in the 21st century, while the other 8 indicate that the land will act as a carbon sink (Ahlstrom et al., 2012). Very few studies analyzed the impact of uncertainty in meteorological forcings on NEE at the plot, catchment or regional scale and for shorter time periods in the recent past. Studies which analyzed interannual variability in NEE detected the important role of temperature and precipitation as drivers of this variability (e.g., Keppel-Aleks et al., 2014). Gu et al. (2016) found that for the Missouri Ozark AmeriFlux forest site, the interannual NEE variability is smaller in simulations by CLM than in the data. On the other hand, some field studies for pairs of monitoring sites (Kwon et al., 2006) or experimentation sites (with temperature increase and precipitation increase or decrease; Xu et al., 2016) found a limited impact of temperature and precipitation differences (or changes) on NEE. Zhang et al. (2012) is one of the few studies that analyzed the role of uncertain meteorological forcings (together with parameter uncertainty) for calculating NEE with a process-based ecosystem model in more detail. Their simulations for a Korean pine mixed forest site with two different meteorological input data sets revealed clear differences in model response. Spadavecchia et al. (2011) performed a rigorous uncertainty analysis for the combination of parameter and forcing uncertainty with a simple LSM for a pine stand in Oregon, USA. They found that the contribution of parameter uncertainty to NEE uncertainty is larger than the contribution of forcing uncertainty, although it also has a significant impact. If meteorological stations are located more than 100 km away from the study site, forcing uncertainty starts to dominate parameter uncertainty. The relative contribution of initial state uncertainty to prediction uncertainty of LSMs has not been thoroughly studied, although it has been recognized as one of the major challenges in model–data fusion studies (Williams et al., 2009). Richardson et al. (2010) performed Monte Carlo inverse modeling to assimilate many different data types in a forest carbon cycle model for the Howland AmeriFlux site in Maine, USA. They estimated jointly model parameters and initial carbon pools. Although they found that model performance improved strongly, they also concluded that carbon pools are more difficult to constrain than ecosystem parameters. Peylin et al. (2016) updated jointly parameters and carbon pools with the land surface model ORCHIDEE for a large-scale application. These last two studies explicitly considered initial state uncertainty together with other uncertainty sources, but did not focus on the relative contribution of the different uncertainty sources.

As atmospheric forcings and initial state uncertainty, except parameter uncertainty, have not received much consideration in the literature until now, we investigate the uncer-

tainty of model predictions in relation to uncertain model parameters, initial conditions and atmospheric forcings in this work.

Accordingly, this study has two main objectives. The first objective is to investigate to what extent can the parameter estimates based on the Markov chain Monte Carlo (MCMC) method enhance the model–data consistency of NEE and LAI for a small European catchment. The model used was CLM4.5 and compared with EC measurements and LAI from the RapidEye satellite. We applied successfully validated parameter estimates from Post et al. (2016) for C₃ grass, evergreen coniferous forest and broadleaf deciduous forest. For C₃ crops, we estimated a new set of parameters as the estimated parameters by Post et al. (2016) did not improve the model–data fit in verification experiments, using the Differential Evolution Adaptive Metropolis (DREAM; Vrugt, 2015). Main advantages of a multi-chain MCMC approach like DREAM are its robustness to find the global minimum and that it is not limited to Gaussian distributed states and parameters. The model performance with the new parameter estimates was then compared to a reference run with global default parameters. The second objective was to investigate the uncertainty of modeled NEE and LAI resulting from uncertain model parameters, atmospheric forcings and initial states. We set up three different ensembles with different combinations of perturbed model input data in order to obtain insight into the contribution of these three main sources of model uncertainty to the uncertainty of the final model output.

2 Data and methods

2.1 The Rur catchment

The Rur catchment (Fig. 1) is located around the Belgian–Dutch–German border region and covers an area of 2354 km². It is characterized by two distinctly different areas of land use and climate. In the northern lowland part, precipitation amounts are lower (650–850 mm yr^{−1}), and potential evapotranspiration is higher (580–600 mm yr^{−1}) compared to the mountainous Eifel region in the south where annual precipitation is 850–1300 mm yr^{−1} and potential evapotranspiration is 450–550 mm yr^{−1} (Montzka et al., 2008a, b). The annual mean temperature in the catchment ranges from about 7.5 °C in the south of the catchment to about 10.3 °C in the north (Baatz et al., 2014). The northern part is dominated by fertile loess soils and is intensely used for agriculture. Sugar beet and cereals (winter wheat, barley) are the most cultivated crops in the catchment (Fig. 1). In the mountainous southern part shallow, less fertile soils dominate. It is mainly covered by meadows and forests. The Rur catchment is one of four central research regions of the TERENO project (Zacharias et al., 2011). The main goal of TERENO is to determine global change impacts across different ter-

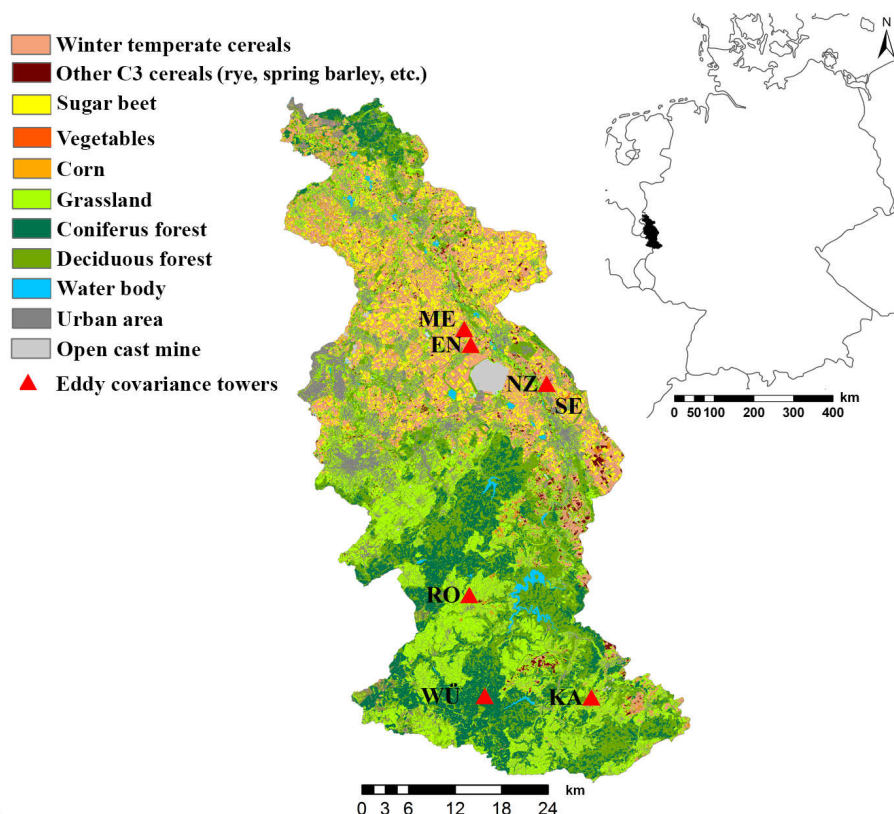


Figure 1. Land cover (Waldhoff, 2010) and eddy covariance tower sites in the Rur catchment.

restrial compartments at the regional level. Therefore, comprehensive input and evaluation data are available for the catchment, including information on land use (Lussem and Waldhoff, 2013), LAIs (Ali et al., 2015; Reichenau et al., 2016) and EC data (Schmidt et al., 2012; Graf et al., 2014; Kessomkiat et al., 2013; Post et al., 2015).

2.1.1 Eddy covariance data

Eddy covariance (EC) data in the Rur catchment were measured for four C₃ crop sites for the evaluation periods summarized in Table 1. The EC crop sites Merzenhausen (ME), Niederzier (NZ), Selhausen (SE) and Engelskirchen (EN) are located in the northern lowland region of the catchment. In ME and SE winter wheat was grown during the measurement period, and in EN and NZ sugar beet. The EC towers in ME and SE were permanently installed. For those sites EC data were available for more than 1 year. EC data at EN and NZ were measured by a roving station, which was installed for 2 to 3 months at each of the three sites (Table 1). The complete processing of the raw data was performed with the TK3.1 software (Mauder and Foken, 2011), using the quality flagging and uncertainty estimation scheme by Mauder et al. (2013) as outlined in Post et al. (2015). Only non-gap-filled data with quality flag 0 (high-quality data) and 1 (moderate-quality data) were included in this study.

Rollesbroich (RO) is an extensively used grassland site (Korres et al., 2010; Post et al., 2015). Wüstebach (WÜ; Graf et al. 2014) is located in the Eifel National Park and is largely covered by evergreen coniferous forest, particularly spruce. The parameters adopted for the current study were previously estimated and validated at RO and WÜ in a previous study (Post et al., 2016).

2.1.2 RapidEye-based leaf area index

RapidEye is a commercial satellite mission initiated by RapidEye AG (Tyc et al., 2005) and consists of five identical satellites, which were launched in August 2008. RapidEye provides multi-spectral images of five spectral bands (blue, green, red, red edge and near-infrared). The nominal temporal resolution is daily. The ground sampling distance is 6.5 m and the pixel size is 5 m for the orthorectified Level 3A data used here. The LAI data derived from satellite images are determined based on the NDVI (normalized difference vegetation index), which is related to the chlorophyll content in leaves. The NDVI is calculated based on the reflectances at near-infrared (NIR) and red (RED). NDVI-based LAI data (LAI_{NDVI}) are affected by various error sources, which can result in high uncertainty of the LAI estimate. The major error sources are summarized in Garrigues et al. (2008) and include (i) uncertainties in surface reflectance measure-

Table 1. Eddy covariance tower sites in the Rur catchment.

EC site name	Latitude (° N)	Longitude (° E)	Altitude (m)	Land use	NEE time series for parameter estimation ^a & model evaluation	Tower ^b
Merzenhausen (ME)	50.9298	6.2970	93.3	agriculture	1 Dec 2011–30 Nov 2012 ^a 1 Dec 2012–30 Nov 2013	Per.t
Selhausen (SE)	50.8658	6.4474	103.0	agriculture	1 Dec 2012–30 Nov 2013	Per.t
Niederzier (NZ)	50.8795	6.4499	102.9	agriculture	5 Apr 2013–10 Jul 2013	Rov.st.
Engelskirchen (EN)	50.9115	6.3090	108.9	agriculture	18 Aug 2012–24 Oct 2012	Rov.st.

^a Parameter estimation^b Per.t.: Permanent EC tower; Rov.st.: roving station

ments resulting, for example, from calibration errors or cloud contamination and (ii) deficiencies in the representation of canopy architecture in the algorithms applied for the LAI retrieval, for example, negligence of foliage clumping. This can lead to gross underestimation of actual LAI, especially for needleleaf forests (Chen et al., 1997). Clumping effects on landscape scale are also related to the fact that LAI algorithms have been calibrated at the plot scale, but are applied over larger heterogeneous pixels, which can induce substantial scaling biases on the LAI estimates (Garrigues et al., 2006). The latter error source is assumed relatively small for the LAI retrieval from RapidEye due to the high spatial resolution (5 m) of the images. Studies on verification or uncertainty quantification of LAI data derived from satellite images are very rare (Garrigues et al., 2008), but they are very important for land surface model applications. Ali et al. (2015) used orthorectified and radiometrically corrected Level 3A data to generate 5 m resolution LAI data for the Rur catchment, with the same methodology previously applied to MODIS data (e.g., Propastin and Erasmi, 2010). Those LAI data were validated for two crop sites (Merzenhausen and Selhausen) in the Rur catchment using in situ data measured with a destructive, ground-based method (Bréda, 2003) at several equally distributed points within the fields at 6 and 11 days during the growing season. The results indicate that LAI and LAI_{NDVI} measure in site, derived from RapidEye, are highly consistent (Ali et al., 2015). Because only the two crop sites were included in this evaluation approach, the LAI data for crop sites (winter wheat) are considered most reliable. For this study, the LAI_{NDVI} data for the Rur catchment obtained according to Ali et al. (2015) were aggregated from the 5 m² to the 1 km² grid of the CLM Rur catchment domain by arithmetic averaging.

2.2 Community Land Model setup

The Community Land Model (CLM) version 4.5 (Oleson et al., 2013) with the active biogeochemistry model (CLM4.5-BGC) is fully prognostic with respect to the seasonal timing of vegetation growth and litter fall. The day length, soil and air temperature and soil water content are main determinants of plant phenology. Plant phenology representation

follows three different schemes depending on the particular plant functional type (PFT): (1) evergreen phenology, (2) seasonal deciduous phenology and (3) stress deciduous phenology. The four most widespread CLM PFTs in the Rur catchment are (1) needleleaf evergreen temperate trees, (2) broadleaf deciduous temperate trees, (3) C₃ non-Arctic grass and (4) C₃ crops. The average PFT coverage of the vegetated land in Rur catchment was ~ 34 % C₃ crops, ~ 32 % grassland, ~ 17 % broadleaf deciduous forest and ~ 14 % coniferous forest. Evergreen coniferous trees follow phenology scheme 1, deciduous broadleaf trees follow scheme 2 and C₃ grass and C₃ crops follow scheme 3 (Oleson et al., 2013). The LAI and all carbon and nitrogen state variables in the vegetation, litter and soil organic matter are calculated prognostically. NEE in CLM is calculated as the sum of gross primary production (GPP) and total ecosystem respiration (ER). Ecosystem respiration includes heterotrophic respiration (HR) and autotrophic respiration, the sum of maintenance respiration (MR) and growth respiration (GR; Oleson et al., 2013). Photosynthesis is determined at leaf scale (Dai et al., 2004; Thornton and Zimmermann, 2007) and is up-scaled by means of the LAI.

The definition of land use cover in CLM follows a nested sub-grid hierarchy structure (Oleson et al., 2013). The main land units, which are defined as percentage coverage per grid cell are glacier, wetland, vegetated land, lake and urban area. Each land unit follows a different sub-model scheme to calculate the respective carbon, water and energy fluxes for a certain grid cell. Each vegetated land unit has 15 soil columns and can include different plant functional types. The PFTs are also defined as percentage area of the vegetated area within the grid cell.

To apply CLM for the Rur catchment domain, a land surface input data set was generated with a spatial resolution of 1 km². The land unit for each grid cell and the PFT distribution of each vegetated land unit were defined based on the land use classification derived from supervised, multi-temporal remote sensing data analysis using RapidEye and ASTER data (Waldhoff et al., 2012; Lussem and Waldhoff, 2013). In addition to the land use coverage, CLM requires information on the percentages of clay and sand for each of the 15 soil columns of the vegetated area per grid cell. For

each soil layer, the soil texture was defined based on the German soil map (BK50) provided by the Geological Survey of North Rhine-Westphalia. Mean topographic slope, mean elevation and maximum fractional saturated area were determined for the 1 km² grid from a 10 m resolution digital elevation model (scilands GmbH, 2010). Additional land surface data required to run CLM4.5 such as soil color were adopted from the default CLM4.5 0.9° × 1.25° resolution global land surface data file of year 2000 (surfdata_0.9 × 1.25_simyr2000_c110921.nc).

The atmospheric forcing data used to run CLM are hourly time series of precipitation (mm s⁻¹), incoming shortwave radiation (W m⁻²), incoming longwave radiation (W m⁻²), atmospheric pressure (Pa), air temperature (K), specific humidity (kg kg⁻¹) and wind speed (mm s⁻¹) at the lowest atmospheric level. The data were obtained for the years 2008–2013 from the reanalysis COSMO-DE data set provided by the German Weather Service (DWD) in 2.8 km × 2.8 km resolution (Baldauf et al., 2009). The COSMO-DE data were downscaled to 1 km² using natural neighbor interpolation based on Delaunay triangulation.

To generate the initial state variables such as the carbon and nitrogen pools, CLM was spun up over a period of 1200 years, using COSMO-DE data of the years 2008–2010. According to DWD (http://www.dwd.de/DE/klimaumwelt/klimaatlas/klimaatlas_node.html), the annual average temperature in the years 2008 and 2009 was ~ 0.5–1.0 °C higher than the long-term average (1961–1990) and the mean annual temperature in 2010 was ~ 0.5–1.0 °C lower. Mean precipitation amounts and freezing days were representative for the long-term average. We also studied the effect of the forcing data used for the model spin-up. For example, we tested using a longer time series (1998–2004) of global climate data. However, we found that using more recent, regional forcing data during the spin-up resulted in carbon and energy fluxes in better agreement with the observations. The model states obtained after the 1200-year spin-up were then used as input for a second 3-year “exit spin-up” also using the meteorological data for the years 2008–2010. The exit spin-up in CLM is necessary for technical reasons and switches the CLM settings from the (accelerated) spin-up mode to the “normal” mode in terms of the calculated carbon–nitrogen cycling. A longer exit spin-up period of 100 years was tested, but results and model states differed only minimally from the case with a 3-year exit spin-up.

2.3 Parameter estimation and evaluation of model performance

2.3.1 Parameter estimation with DREAM_(zs)

In this study, we estimated the posterior probability density functions (pdfs) of five PFT-specific CLM4.5 parameters for C₃ crops using the adaptive Markov chain Monte Carlo (MCMC) method DREAM_(zs) (Ter Braak and Vrugt, 2008;

Laloy and Vrugt, 2012; Vrugt, 2015), according to the approach presented in Post et al. (2016). DREAM_(zs) is based on the Bayes’ theorem (A1). In multi-chain MCMC methods like DREAM_(zs), different (in our case three) Markov chains generate random walks through the parameter space and successively visit solutions with stable frequencies stemming from a stationary distribution. It is assumed herein that the prior distribution is uniform (non-informative) using ranges of the predefined upper and lower bounds for each parameter according to Post et al. (2016). The use of multiple chains offers robust protection against premature convergence and thus the probability of becoming stuck in local minima is considerably reduced compared to single-chain methods (Vrugt, 2015). The convergence of the parameters is monitored with the $\hat{R}$ convergence diagnostic of Gelman and Rubin (1992) which is computed for each dimension as the ratio of variance within one chain and the variance between different chains. Convergence is achieved if $\hat{R}$ is smaller than 1.2 for all parameters. Because every point in the parameter space is hit with a frequency proportional to its probability, the random walk allows one to iteratively find a stable posterior distribution. Hence, after convergence, the density of the acceptance points in the parameter space approximates the posterior pdf according to Bayes’ theorem (see Appendix). DREAM exhibits excellent sampling efficiencies for multimodal and high-dimensional posterior distributions (Vrugt, 2015), which is an important advantage if applied to complex models like CLM. A full description of the DREAM_(zs) algorithms can be found in Vrugt (2015).

To constrain the C₃ crop parameters, a 1-year time series of non-gap-filled, half-hourly NEE data from the EC tower at the ME site (1 December 2011–30 November 2012) was used. The five key parameters are (1) the fraction of leaf N in Rubisco enzyme (f_{N_R}), (2) the growth respiration factor (g_R), (3) the rooting distribution parameter (1 m⁻¹) (r_b), (4) the specific leaf area at top of canopy (m² g C; sla_{top}) and (5) the soil water potential at full stomatal closure (mm; ψ_c). This selection is based on Post et al. (2016) and previous studies which highlighted the sensitivity of these parameters (Foereid et al., 2014; Göhler et al., 2013).

Parameter estimates for C₃ grass, evergreen coniferous forest and broadleaf deciduous forest have been successfully estimated and validated for central European sites by Post et al. (2016) with DREAM_(zs)-CLM4.5 and are therefore adopted in this study. Post et al. (2016) used NEE data from the RO and WÜ sites to estimate parameters for C₃ grass and coniferous forest, as well as FLUXNET data from the Fontainebleau site (FR-Fon) in France, located about 300 km southwest of the Rur catchment (48.4763° N, 2.7801° E; e.g., Migliavacca et al., 2015). The validation was based on NEE data from EC sites of corresponding PFTs that were located ~ 600 km away from the parameter estimation sites. Post et al. (2016) estimated the parameters f_{N_R} , g_R , r_b , sla_{top} and ψ_c jointly with three more parameters: the temperature coefficient (Q_{10}), the base rate for maintenance respiration (mr_b)

and the Ball–Berry slope of conductance–photosynthesis relationship (b_s). Q_{10} quantifies the fractional change of the respiration rate in response to a 10 °C temperature rise. Q_{10} , mr_b and b_s in CLM4.5 are by default hard-wired in the CLM4.5 source code and not defined separately for each PFT like the other five parameters. This implies that one single value of those parameters is defined globally, which is then applied to all PFTs. However, Post et al. (2016) found that values of the estimated hard-wired parameters, especially Q_{10} , varied for different PFTs or sites. The fact that Q_{10} , mr_b and b_s are hard-wired in CLM imposes a challenge when applying the jointly estimated parameter sets on regional scale because it requires that Q_{10} , mr_b and b_s be defined separately for each PFT. Therefore, we modified CLM in order to input the PFT-specific values for Q_{10} , mr_b and b_s of the jointly estimated parameter sets by Post et al. (2016). Post et al. (2016) showed that the estimated parameter values for C_3 crops did not enhance model performance compared to the default parameter values, if applied to an evaluation year or another C_3 crop site. This applies to the parameter values that were estimated and applied using a 1-year time series of NEE data, as done in this study. Therefore, we estimated a new set of C_3 crop parameters.

2.3.2 Evaluation of parameter estimates with eddy covariance NEE data

In order to evaluate the performance of the estimated CLM C_3 -crop-specific parameters in terms of the consistency of modeled and measured data, two CLM cases were defined and compared: (i) a reference run (CLM-Ref), which is a forward run for the years 2011–2013 with default parameters and the atmospheric input data and initial conditions as described in Sect. 2.2, and (ii) an ensemble run with 60 realizations, for the same time period, same initial conditions and same atmospheric input data as CLM-Ref, but with parameters sampled randomly from the DREAM_(ZS) multivariate posterior pdfs (CLM-Ensp). Thus, we did not perform a separate model spin-up for each of the estimated set of parameters, as this was computationally too expensive. The parameter estimates for ME were evaluated with eddy covariance NEE data from four different C_3 crop sites in the Rur catchment (ME, SE, NZ and EN). For ME, the NEE observation time series of the year that followed the parameter estimation year were used for evaluation (Table 1).

The evaluation was conducted with time series of half-hourly NEE data measured at the EC tower sites. First it was verified that the PFT coverage of one CLM grid cell coincided with the dominant PFT at the respective EC tower site. More than 80 % of each of the four grid cells was covered by C_3 crops. This was considered sufficient for grid-based model evaluation. In the following, the subscript “gc” is used to refer to the grid cell in which one of the EC towers is located; e.g., ME_{gc} refers to the respective grid cell of the ME site. The length of the available NEE time series differed

among the EC sites, ranging from 2 to 12 months (including data gaps). See Fig. 1. Accordingly, only the model output that coincides with the observed NEE data was used.

The calculated NEE output was evaluated based on the following indices:

The root mean square error (RMSE_m):

$$\text{RMSE}_m = \sqrt{\frac{1}{n} \sum_{i=1}^n (m_i - y_i)^2}, \quad (1)$$

where y is measured half-hourly NEE ($\mu\text{mol m}^{-2} \text{s}^{-1}$) at time step i for the given time series of length n and m is the modeled equivalent.

The mean absolute difference of the mean diurnal NEE cycle (MAD_{diur}):

$$\text{MAD}_{\text{diur}} = \frac{1}{48} \sum_{i=1}^{48} |\bar{m}_d - \bar{y}_d|, \quad (2)$$

where $\bar{y}$ is average measured NEE at a given time d during the day ($\mu\text{mol m}^{-2} \text{s}^{-1}$) and $\bar{m}$ is the modeled equivalent. This performance measure is evaluated half-hourly (48 times per day). For the sites ME and SE, where a complete year of NEE data was available, the mean diurnal NEE cycle and the index MAD_{diur} calculated separately for each of the four seasons within the evaluation year (winter: December–February; spring: March–May; summer: June–August; autumn: September–November). In order to summarize the four seasonal MAD_{diur} indices to one MAD_{diur} index representative for the whole evaluation year, the four seasonal indices were averaged. For the other sites where only NEE data for ~2–3 months were available, the indices including MAD_{diur} were calculated for this shorter time series.

The relative difference (%) of the NEE sum ($\sum \text{NEE}$) was calculated for all half-hourly data available in the respective evaluation period (RD _{$\sum \text{NEE}$}):

$$\text{RD}_{\sum \text{NEE}} = \frac{\sum_{i=1}^n (m_i) - \sum_{i=1}^n (y_i)}{\sum_{i=1}^n (y_i)} \times 100, \quad (3)$$

where y is measured NEE (non-gap-filled) and m is modeled equivalent.

All evaluation indices were determined for both CLM-Ref and for the ensemble mean of CLM-Ensp. Post et al. (2016) showed that parameters estimated for the forest sites WÜ and FR-Fon are well transferable to other FLUXNET sites of corresponding PFTs located more than 600 km from WÜ and FR-Fon: Tharandt in Germany (DE-Tha; e.g., Grünwald and Bernhofer, 2007) and Hainich in Germany (DE-Hai; e.g., Knohl et al., 2003). Because WÜ is the only forested EC tower site in the Rur-catchment and no additional EC towers were available for the validation of the estimated parameters, we assume the parameter estimates provide more reliable NEE data for a large part of the forested area in the

Rur catchment. We also adopted successfully estimated and validated parameters for the C₃ grass site RO from Post et al. (2016).

2.3.3 Evaluation of LAI predictions

Various studies highlight current deficits to accurately estimate and validate LAI derived from multispectral satellite images for deciduous and needleleaf forest (Ganguly et al., 2012; Tillack et al., 2014; Härkönen et al., 2015). In addition, the LAI data derived from RapidEye (LAI_{RapidEye}) used herein have not yet been validated for the forest PFTs. Thus, the model performance in terms of the LAI was only evaluated for grid cells with more than 80 % C₃ grass or C₃ crop coverage, not for the forest PFTs. The effect of the parameter estimates on LAI was evaluated for the 1 km² grid of the Rur catchment domain. LAI_{RapidEye} data of about 18 days (depending on the location) between 1 November 2011 and 16 September 2012 were used for the LAI evaluation and compared with modeled LAI data on those days and those grid cells for which RapidEye data were available. The LAI was evaluated separately for C₃ grass and C₃ crops, and both for the winter half-year (November–April) and summer half-year (May–October). To evaluate and compare modeled LAI for CLM-Ens_P and CLM-Ref, the mean absolute difference between simulated LAI and measured LAI (MAD_{LAI}) was calculated over all grid cells with more than 80 % C₃ grass ($n_{\text{pft}} = 224$) or C₃ crop ($n_{\text{pft}} = 404$) coverage:

$$\text{MAD}_{\text{LAI}} = \frac{1}{n_{\text{days}} \cdot n_{\text{pft}}} \sum_{i=1}^{n_{\text{days}} \cdot n_{\text{pft}}} |m_i - y_i|, \quad (4)$$

where m is the modeled daily LAI ($\text{m}^2 \text{m}^{-2} \text{day}^{-1}$) and y is the measured equivalent, n_{days} is number of days and n_{pft} is number of grid cells for which LAI_{RapidEye} data were available at a particular day. RMSE_{LAI} was calculated according to Eq. (1).

The mean LAI for each PFT was calculated by

$$\text{LAI}_{\text{PFT}} = \frac{1}{n_{\text{days}} \cdot n_{\text{pft}}} \sum_{i=1}^{n_{\text{days}} \cdot n_{\text{pft}}} \text{LAI}_i, \quad (5)$$

where n_{days} is the number of days RapidEye data for a given PFT were available, and LAI _{i} is the LAI observed or modeled at a particular day ($\text{m}^2 \text{m}^{-2} \text{day}^{-1}$). For each LAI_{PFT} value, a respective standard deviation was calculated.

2.4 Uncertainty estimation

2.4.1 Perturbation of atmospheric input data

In order to take the uncertainty of the meteorological input data into account, a 60-member ensemble of perturbed meteorological forcings was generated for the years 2008–2012 using hourly COSMO-DE data. The approach used to generate the perturbation fields has previously been applied in soil

moisture data assimilation studies (Reichle et al., 2007, 2010; Kumar et al., 2012; Han et al., 2012, 2013, 2014). Perturbation fields were applied to air temperature (K), incoming longwave radiation (W m^{-2}), incoming shortwave radiation (W m^{-2}) and precipitation (mm s^{-1}). Normally distributed additive perturbations were applied to longwave radiation (LW) and air temperature (Temp). Log-normally distributed multiplicative perturbations were applied to precipitation (Prec) and shortwave radiation (SW). The parameters used for the perturbations were adapted from Han et al. (2014) and are listed in Table 2. The perturbations for the different atmospheric variables are cross-correlated in order to generate physically plausible perturbations of the atmospheric forcings. We considered spatial correlation in addition to temporal correlation of the meteorological variables. The multiplicative perturbations are truncated by a defined maximum of 2.5 standard deviations, to remove outliers from the generated perturbation fields (Reichle et al., 2007). The spatially correlated noise is calculated first using the fast Fourier transform approach (Park and Xu, 2013) with a 10 km spatial correlation scale. Next, the temporally correlated noise is added to the spatially correlated noise. The temporal correlation for all perturbed variables is imposed using a first-order AR(1) autoregressive model (Reichle et al., 2007). The AR(1) temporal correlation coefficient for a time lag of 1 day was 0.368 for all variables.

2.4.2 Generation of perturbed initial state input files and the perturbed forward run for the Rur catchment

As shown in various studies, carbon fluxes predicted by land surface models strongly depend on the carbon–nitrogen pools generated during the model spin-up (Carvalhais et al., 2010; Kuppel et al., 2012). In order to take the uncertainty of initial states into account, a 60-member ensemble of perturbed initial states was generated. This was done via a 15-year spin-up using perturbed atmospheric forcings for the years 2008–2010 (Sect. 2.4.1) and parameter values sampled randomly from the joint DREAM posterior pdfs. The initial conditions used at the beginning of the 15-year perturbed spin-up were taken from the end of the main 1200-year spin-up plus 3-year exit spin-up. This perturbation takes into account the uncertainty of the less stable carbon and nitrogen pools, but uncertainty with respect to the stable and resistant carbon and nitrogen pools is not taken into account. As the stable nitrogen pool is not perturbed, the availability of N from mineralization is very similar across the ensemble members. This might be an unproblematic assumption given N fertilization by agriculture and atmospheric deposition, but it also implies that the uncertainty of the net N availability at the sites is likely underestimated.

Table 2. Parameters applied for the perturbation of the meteorological input data, adapted from Han et al. (2014).

Variables [unit]	Additive (A) or multiplicative (M)	Standard deviation	Cross correlation coefficients		
			T	SW	LW
Air temperature (T) [K]	A	1	–	0.4	0.4
Incoming shortwave radiation (SW) [W m^{-2}]	M	0.3	0.4	–	–0.5
Incoming longwave radiation (LW) [W m^{-2}]	A	20	0.4	–0.5	–
Precipitation (P) [mm s^{-1}]	M	0.5	0	–0.8	0.5

2.4.3 Evaluation of model uncertainty

In order to evaluate the effect of uncertain CLM parameters, meteorological input data and initial model states, we set up four CLM ensembles with 60 ensemble members each:

1. CLM-Ensp: ensemble model runs for 1 December 2012–30 November 2013 with deterministic (non-perturbed) initial states and non-perturbed input data from COSMO-DE (Sect. 2.2). The initial states are the outcome of the default 1203-year spin-up, without perturbations. The parameter values were sampled randomly from the DREAM_(zs) multivariate posterior pdfs (Table 3), which were estimated for C₃ crops or adopted from Post et al. (2016) for C₃ grass, evergreen coniferous forest and broadleaf deciduous forest. For C₃ crops, the setup of CLM-Ensp was identical to the one used for the evaluation of the estimated C₃ crop parameters in terms of the NEE model–data consistency (Sect. 2.3.2).
2. CLM-Ensp_{AI}: ensemble model runs for 1 December 2012–30 November 2013 with parameter values sampled according to CLM-Ensp and perturbed atmospheric forcings, but using deterministic initial states.
3. CLM-Ensp_{AI}: ensemble model runs according to CLM-Ensp_{AI}, but starting from perturbed initial states.
4. CLM-Ensp+Q₁₀: a forward run for the period 1 December 2012–30 November 2013 with deterministic forcings, deterministic initial states and parameters sampled according to CLM-Ensp, except for Q₁₀. Q₁₀ values for C₃ crop and forest PFTs were sampled from a normal distribution with a mean of 2.0 and a standard deviation of 0.5 with a minimum and maximum bound set to 1.3 and 3.0. We choose Q₁₀ because of its central role in carbon stock and flux predictions and the respective uncertainties in most land surface models (Post et al., 2008; Hararuk et al., 2014), including CLM (Post et al., 2016). The objective of ensemble CLM-Ensp+Q₁₀ was to estimate the uncertainty of the CLM output induced by the prior uncertainty of one single key parameter, and compare it to the posterior uncertainty from CLM-Ensp.

The modeled NEE sum ($\text{g C m}^{-2} \text{yr}^{-1}$) for the period 1 December 2012–30 November 2013 was calculated for each

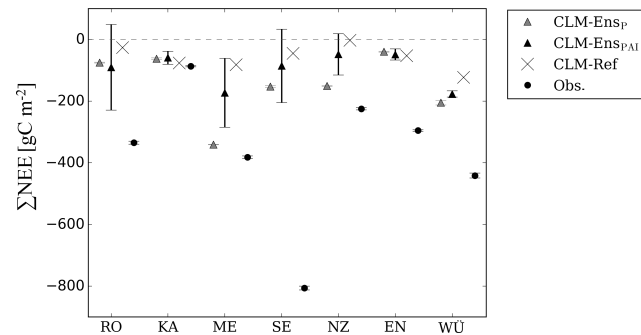


Figure 2. NEE sum (ΣNEE) determined for the evaluation periods (see Table 1) including all half-hourly time steps where eddy covariance data (Obs) were available, for the CLM ensembles Ensp with estimated parameters and Ensp_{AI} with additional perturbed atmospheric forcings and perturbed initial states (including standard deviations of the 60 ensemble members), in comparison to a reference run with default parameters (CLM-Ref). In all CLM cases time series were subsetting according to available Obs.

ensemble member using the complete time series of daily outputs. The PFT-specific NEE sum ($\Sigma\text{NEE}_{\text{PFT}}$) was then calculated separately for each of the four main PFTs by averaging the simulated NEE sum over all grid cells with more than 80 % coverage by the particular PFT. The analysis of model uncertainty was then based on the standard deviation (σ) of $\Sigma\text{NEE}_{\text{PFT}}$ between the different ensemble members, $\sigma(\Sigma\text{NEE}_{\text{PFT}})$. Accordingly for GPP and ER, $\sigma(\Sigma\text{GPP}_{\text{PFT}})$ and $\sigma(\Sigma\text{ER}_{\text{PFT}})$ were calculated.

3 Results

Evaluation of simulated NEE and LAI with estimated parameter values

As shown in Table 4, CLM parameter estimates notably reduced the mismatch between modeled and measured NEE data in comparison with the reference run (Ref), where global default parameter values were used.

MAD_{diur} was reduced by 7 % (SE_{gc}), 25 % (NZ_{gc}) and 35 % (ME_{gc}) and increased by 7 % for EN_{gc}. Thus, the mean diurnal NEE cycles were closer to the observed ones for all

Table 3. Lower and upper bound of 95 % confidence intervals of CLM4.5 parameter estimates estimated with DREAM_(zs) (C₃ crop) or adapted from Post et al. (2016).

	Year	f_{NR}	sl_{top}	gR	r_b	ψ_c	Q_{10}	mr_b	b_s
CLM4.5 default values		0.14	0.030	0.30	2.00	-2.75×10^5	1.50	2.53×10^{-6}	9.0
RO (C ₃ grass)	2011/12	0.13, 0.15	0.010, 0.010	0.39, 0.40	1.01, 1.27	-3.79×10^5 , -1.65×10^5	1.39, 1.44	4.48×10^{-6} , 4.50×10^{-6}	6.1, 6.9
CLM4.5 default values		0.18	0.030	0.30	3.00	-2.75×10^5	1.50	2.53×10^{-6}	9.0
ME (C ₃ crop)	2011/12	0.09, 0.10	0.010, 0.010	0.40, 0.40	1.00, 1.04	-3.94×10^5 , -2.62×10^5	– –	– –	– –
CLM4.5 default values		0.05	0.010	0.30	2.00	-2.55×10^5	1.50	2.53×10^{-6}	9.0
WÜ (coniferous forest)	2011/12	0.05, 0.07	0.005, 0.006	0.29, 0.40	0.75, 3.95	-3.91×10^5 , -2.07×10^5	2.50, 2.99	2.13×10^{-6} , 3.48×10^{-6}	5.0, 6.2
CLM4.5 default values		0.05	0.010	0.30	2.00	-2.55×10^5	1.50	2.53×10^{-6}	9.0
FR-Fon (deciduous forest)	2006/07	0.12, 0.12	0.010, 0.010	0.39, 0.40	1.00, 1.17	-3.89×10^5 , -2.57×10^5	1.87, 1.97	3.47×10^{-6} , 3.50×10^{-6}	5.7, 6.0

Table 4. Root mean square error $RMSE_m$ ($\mu\text{mol m}^{-2} \text{s}^{-1}$), mean absolute difference for the mean diurnal NEE cycle MAD_{diur} ($\mu\text{mol m}^{-2} \text{s}^{-1}$) and relative difference of the NEE sum over the evaluation period $RD_{\sum NEE}$ (%) for the CLM ensemble with estimated parameters (Ensp) in comparison to the reference run (Ref) with default parameters. Results are given for ME (Merzenhausen), SE (Selhausen), NZ (Niederzier) and EN (Engelskirchen). Here, n is the number of non-gap-filled half-hourly measurement data that were available to calculate these evaluation indices.

Grid cells	n	MAD_{diur} ($\mu\text{mol m}^{-2} \text{s}^{-1}$)		$RMSE$ ($\mu\text{mol m}^{-2} \text{s}^{-1}$)		$RD_{\sum NEE}$ (%)	
		Ensp	Ref	Ensp	Ref	Ensp	Ref
ME	10 157	1.5	2.30	5.95	6.74	11	79
SE	9597	3.22	3.45	8.20	9.41	81	94
NZ	2772	2.70	3.61	10.71	9.37	33	99
EN	2272	4.83	4.53	8.83	8.93	87	82

evaluation grid cells dominated by C₃ crops, except EN_{gc}, if estimated parameters were used. The $RMSE_m$ was reduced by up to 13 % (SE_{gc}) for all sites except NZ_{gc}. $RD_{\sum NEE}$ was most notably reduced compared to the other evaluation indices. The measured $\sum NEE$ was negative for each of the four EC sites and for all of the respective model outputs (Fig. 2). This would imply that all sites were net carbon sinks during the evaluation period. However, because data gaps were included in the NEE time series and because more EC data were available for summer and daytime than for winter and nighttime, those values do not represent the true NEE sum of the evaluation period. For CLM-Ref and all evaluation sites, $\sum NEE$ differed notably from the observed data (Fig. 2, Table 4). For ME_{gc}, SE_{gc} and NZ_{gc}, the predicted $\sum NEE$ was significantly closer to the observations for CLM-Ensp than for CLM-Ref, and $RD_{\sum EE}$ was reduced by a factor of 1.2 (SE_{gc}) to 7.2 (ME_{gc}). $RD_{\sum EE}$ was 13 % (SE_{gc}) to 68 % (ME_{gc}) lower for CLM-Ensp compared to CLM-Ref. For EN_{gc}, $RD_{\sum EE}$ was slightly lower for CLM-Ref, but as indicated in Fig. 2, the difference of $\sum NEE$ was not significant. Overall, results indicate that NEE sums over a time period from several months to 1 year were better represented

with estimated values than with global default values. Figure 2 highlights that the uncertainty of modeled NEE is probably underestimated if the uncertainty of meteorological input data and initial states is not taken into account. The standard deviation of CLM-Ensp_{AI} is notably higher compared to CLM-Ensp, where only parameter uncertainty is considered.

Table 5 summarizes the catchment average LAIs and the corresponding mean absolute differences (MAD_{LAI}) and RMSE between the observed and modeled LAI_{PFT}. In the summer half-year, LAI_{Ens} was closer to LAI_{RapidEye} than LAI_{Ref}, both for C₃ grass and C₃ crops. $RMSE_{LAI}$ was on average 1.5 for CLM-Ensp and 3.4 for CLM-Ref. MAD_{LAI} for summer was 1.2 for CLM-Ensp and 3.4 for CLM-Ref on average. In winter, the difference was only 0.2 ($RMSE_{LAI}$) and 0.3 (MAD_{LAI}) and thus the improvement not significant.

While CLM-Ref overestimated LAI_{PFT}, ENSP underestimated LAI_{PFT}. Overall, the standard deviation of LAI_{PFT}, i.e., the LAI variation over all included grid cells n throughout the half-year period, was very high for both the observed and the modeled data. The mean annual LAI cycles for ME_{gc} and SE_{gc} (Fig. 3) indicate that the uncertainty of LAI was un-

Table 5. Mean LAI values (LAI_{PFT}) for the RapidEye data (Obs), the CLM ensembles with estimated parameters (Ensp) and the CLM reference run (Ref) with default parameters, as well as the mean absolute differences (MAD_{LAI}) and root mean square error (RMSE_{LAI}). N is the number of grid cells in the catchment with more than 80 % coverage of one particular PFT.

PFT			LAI _{PFT}				RMSE _{LAI}		MAD _{LAI}	
			<i>N</i>	Obs	Ensp	Ref	Ensp	Ref	Ensp	Ref
C ₃ grass	w	1667	2.7 ± 1.6	1.0 ± 0.8	2.6 ± 2.2	2.6	3.0	2.0	2.5	
C ₃ grass	s	1271	3.0 ± 1.4	1.8 ± 0.6	5.5 ± 1.4	1.9	3.2	1.5	2.6	
C ₃ crop	w	4392	2.3 ± 1.3	0.7 ± 0.5	2.4 ± 1.7	2.1	1.9	1.7	1.6	
C ₃ crop	s	2887	2.1 ± 0.9	1.5 ± 0.5	5.5 ± 0.8	1.0	3.6	0.8	3.5	

w: winter half-year (November–April); s: summer half-year (May–October); $\pm$ standard deviation; Ensp: CLM ensemble with estimated parameters.

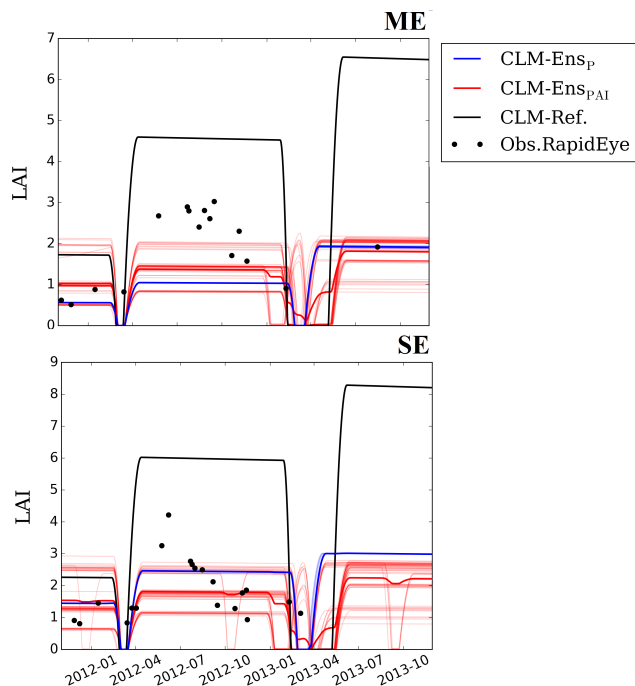


Figure 3. Daily LAIs for the evaluation period 1 December 2011–30 November 2013 for grid cells where the sites Merzenhausen (ME) and Selhausen (SE) are located. Results are shown for the 60 ensemble members of the CLM cases Ensp_{PFT} with estimated parameters, and Ensp_{AI} with additional perturbed atmospheric forcings and perturbed initial states, in comparison to a reference run with default parameters (CLM-Ref) and RapidEye data (Obs.RapidEye; bold lines: ensemble mean).

derestimated for CLM-Ensp. The uncertainty of simulated LAI was much higher for CLM-EnspAI.

Figure 3 highlights that for C₃ crops, the simulated yearly LAI cycle did not match well the observed and expected annual LAI course. The delay of the plant emergence indicated by these LAI courses is related to the strong underestimation of daytime NEE in spring 2013 (Figs. 4, 5). This underestimation of NEE can mainly be attributed to an underestimation of GPP, which is too low in spring because the simulated

plant onset was about 2 weeks later than observed. For CLM-Ref and for most of the CLM-EnspAI ensemble members, leaf onset started in May. For those model realizations, the underestimation of daytime NEE in spring was highest. In contrast, for CLM-Ensp and a small proportion of CLM-EnspAI, leaf onset started in March. For those cases, the underestimation of daytime NEE in spring was notably lower. This elucidates the close link of modeled NEE and LAI and highlights that errors in the timing of leaf onset can lead to substantial errors in simulated NEE. The evaluation of simulated LAI showed that modeled and observed LAI are closer for simulations with estimated CLM parameters than for simulations with default parameters and that for C₃ crops, simulated leaf onset was better represented.

Uncertainties of simulated carbon fluxes on catchment scale

In this section, results of modeled NEE, GPP and ER are summarized and compared for the CLM ensembles CLM-Ensp, CLM-EnspAI, CLM-EnspAI and CLM-Ensp+Q10 with focus on the model uncertainty. The uncertainty is evaluated with the standard deviations $\sigma(\sum \text{NEE}_{\text{PFT}})$, $\sigma(\sum \text{GPP}_{\text{PFT}})$ and $\sigma(\sum \text{ER}_{\text{PFT}})$ of the 60 ensemble members. In this case the time series from 1 December 2012 to 30 November 2013 of the simulated fluxes was used to calculate the PFT-specific, catchment-average sums $\sum \text{NEE}_{\text{PFT}}$, $\sum \text{GPP}_{\text{PFT}}$ and $\sum \text{ER}_{\text{PFT}}$ ($\text{g C m}^{-2} \text{ yr}^{-1}$).

Figure 6 shows the means and standard deviations of $\sum \text{NEE}_{\text{PFT}}$ for the four CLM ensembles and for CLM-Ref. The sign of $\sum \text{NEE}_{\text{PFT}}$ indicates that for CLM-Ref, all PFTs in the catchment act as net carbon sources. $\sum \text{NEE}_{\text{PFT}}$ changed significantly if estimated parameter values were applied (CLM-Ensp). In that case, all PFTs except C₃ grass acted as carbon sink. Figure 7 illustrates how estimated parameters affected the modeled annual NEE sum ($\text{g C m}^{-2} \text{ yr}^{-1}$) for the Rur catchment. With default parameter values, the annual NEE was positive for most of the grid cells, particularly in the northern part of the catchment which is dominated by agriculture. In terms of NEE, the catchment is a net CO₂ source. With estimated parameters from CLM-

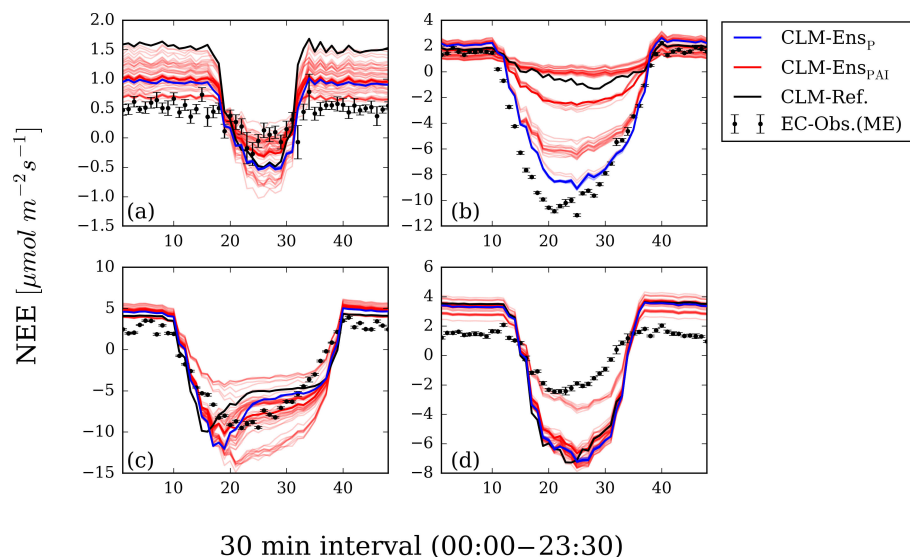


Figure 4. Mean diurnal course of half-hourly NEE for winter 2012/13 (a), spring 2013 (b), summer 2013 (c) and autumn 2013 (d) for the Merzenhausen site (ME). Results are shown for the 60 ensemble members of the CLM cases Ens_p with estimated parameters, and Ens_{pAI} with additional perturbed atmospheric forcings and perturbed initial states, in comparison to a reference run with default parameters (CLM-Ref) and EC data (EC-Obs.; bold lines: ensemble mean).

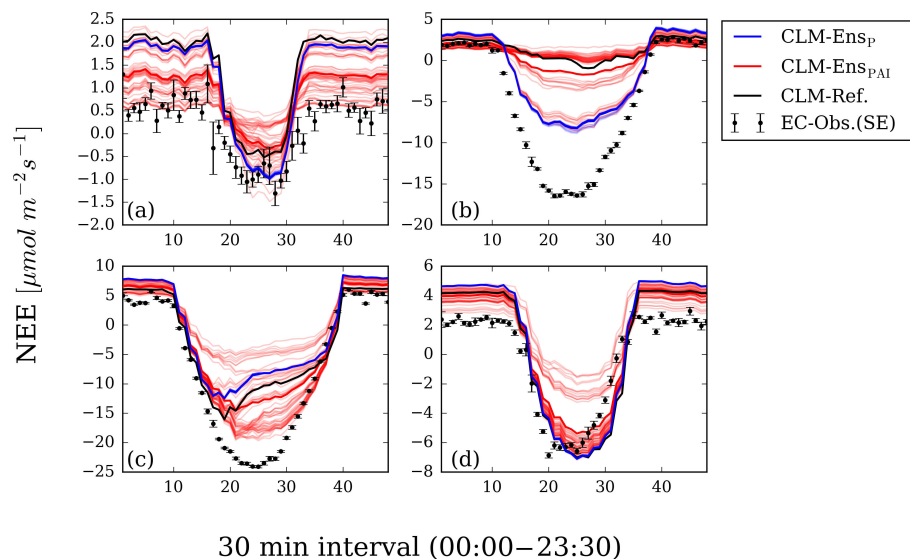


Figure 5. As Fig. 4, but for the Selhausen site (SE).

Ens_p , the NEE sum became negative for most grid cells in the catchment including a large part of the northern lowland area (Fig. 7). Thus, with estimated parameters the catchment became a net CO_2 sink (disregarding CO_2 fluxes due to harvesting, land use change or fossil fuel combustion). Figure 7 indicates that both GPP and ER increased with estimated parameters. However, GPP increased more than ER. Since we found that at verification sites simulations with estimated parameters gave NEE sums closer to the observed NEE sums than simulations with default parameters, we as-

sume the catchment-scale NEE sum (Fig. 7) and $\sum \text{NEE}_{\text{PFT}}$ (Fig. 6) are more reliable for CLM- Ens_p than for CLM-Ref.

As indicated in Fig. 6, for C_3 grass and C_3 crops, $\sigma(\sum \text{NEE}_{\text{PFT}})$ was significantly higher for CLM- Ens_{pAI} compared to CLM- Ens_p . For CLM- Ens_{pAI} , $\sigma(\sum \text{NEE}_{\text{PFT}})$ was $\sim 45 \text{ g C m}^{-2} \text{ yr}^{-1}$ (C_3 grass) and $\sim 75 \text{ g C m}^{-2} \text{ yr}^{-1}$ (C_3 crops) compared to $\sim 2\text{--}3 \text{ g C m}^{-2} \text{ yr}^{-1}$ for CLM- Ens_p . Thus, applying perturbed forcings led to a very strong increase in the uncertainty of simulated $\sum \text{NEE}_{\text{PFT}}$ by a factor of ~ 14 (C_3 grass) and ~ 42 (C_3 crops). This finding reveals that C_3 grass and C_3 crops in CLM were extremely sensitive

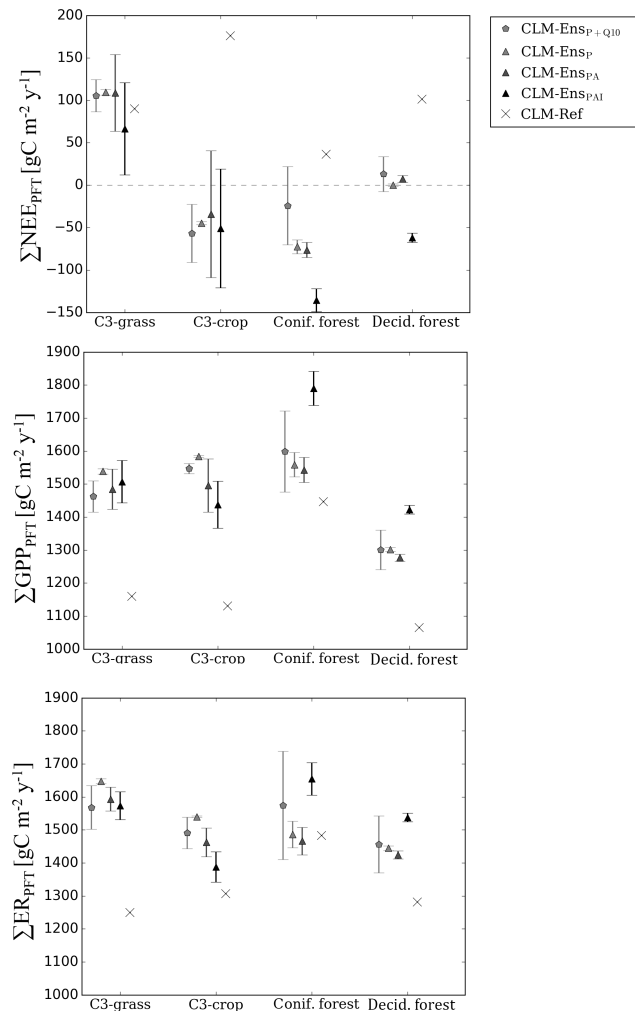


Figure 6. Ensemble mean and standard deviation of carbon fluxes simulated by four different CLM ensembles: (i) with estimated parameters (Ens_P) and (ii) additional perturbation of Q_{10} (Ens_{P+Q10}) or atmospheric forcings (Ens_P) or atmospheric forcings and initial states (Ens_{PAI}), in comparison to the reference run with default parameters (Ref). Plotted are the PFT-specific, catchment-average sums of NEE, GPP and ER.

to minor differences in the meteorological input data. In comparison, the differences of $\sigma(\sum \text{NEE}_{\text{PFT}})$ between CLM- Ens_{PA} and CLM- Ens_{PAI} were minor. Thus, the much higher $\sigma(\sum \text{NEE}_{\text{PFT}})$ for CLM- Ens_{PAI} compared to CLM- Ens_P can mainly be ascribed to the perturbed meteorological input data, rather than the perturbed initial states. Due to the large uncertainty of $\sum \text{NEE}_{\text{PFT}}$ for both CLM- Ens_{PA} and CLM- Ens_{PAI} , differences of the ensemble mean $\sum \text{NEE}_{\text{PFT}}$ were not significant between the different two CLM ensembles. For C_3 crops and C_3 grass, $\sum \text{GPP}_{\text{PFT}}$ and $\sum \text{ER}_{\text{PFT}}$ were strongly sensitive to the perturbed forcings such that the strong increase in $\sigma(\sum \text{NEE}_{\text{PFT}})$ for CLM- Ens_{PA} compared to CLM- Ens_P is related to changes in both ER and GPP (Fig. 6).

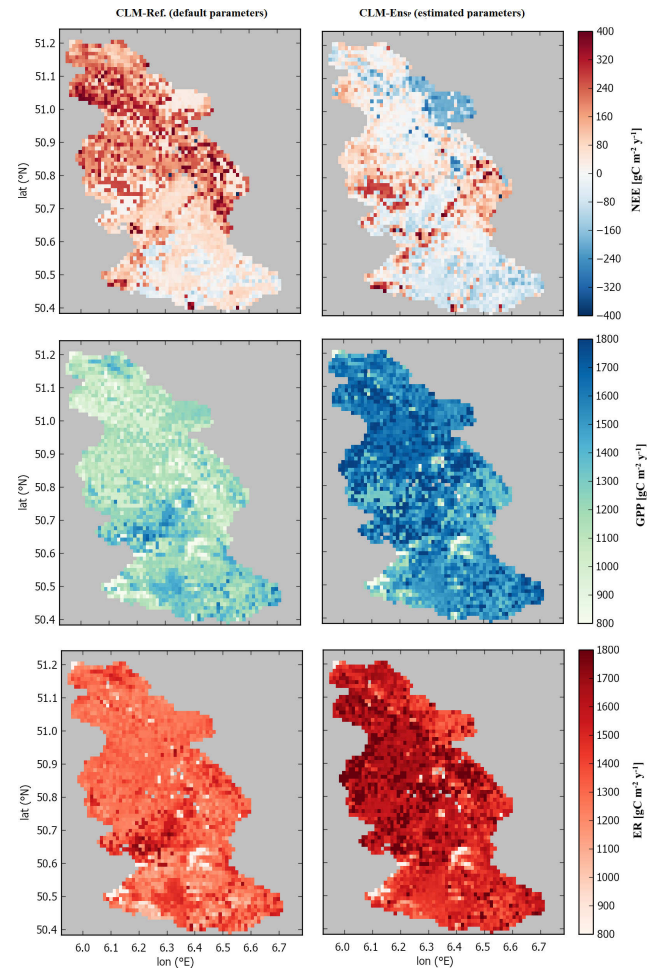


Figure 7. Annual sum of net ecosystem exchange (NEE), gross primary production (GPP) and ecosystem respiration (ER) determined with CLM4.5-BGC for the Rur catchment (December 2012–November 2013) with default parameters (CLM-Ref.) and with estimated parameters (CLM-Ens).

For coniferous forest and deciduous forest, $\sigma(\sum \text{NEE}_{\text{PFT}})$ for CLM- Ens_{PA} and CLM- Ens_{PAI} were ~ 4.0 – $13.5 \text{ gC m}^{-2} \text{ yr}^{-1}$ and thus considerably lower in comparison to C_3 grass and C_3 crops. The increase in $\sigma(\sum \text{GPP}_{\text{PFT}})$ and $\sigma(\sum \text{ER}_{\text{PFT}})$ with perturbed forcings and perturbed initial states was also minor in comparison to CLM- Ens_P . This indicates that the carbon cycle of the forest PFTs in CLM is much less sensitive to the atmospheric input data compared to C_3 grass and C_3 crops. However, when comparing CLM- Ens_{PA} and CLM- Ens_{PAI} , the additional effect of the perturbed initial states was notably higher for forest compared to C_3 grass and C_3 crops. For the forest PFTs, the ensemble mean $\sum \text{NEE}_{\text{PFT}}$ changed significantly if initial states were perturbed (Fig. 6). This was related to the fact that estimated parameters and perturbed atmospheric forcings disrupted the steady state of the forest carbon pools. Hence, carbon pools increased relatively rapidly during the

perturbed spin-up (Sect. 2.4.2), which was not the case for C_3 grass and C_3 crop. As already shown in the previous section, the spread of CLM-Ens_P was low, indicating that the uncertainty of simulated NEE induced by the posterior parameter values that were sampled from the estimated pdfs was low. This is, however, related to the fact that parameters were already conditioned to NEE data, which reduced the parameter uncertainty. The additional perturbation of Q_{10} (CLM-Ens_{P+Q10}) increased $\sigma(\sum NEE_{PFT})$ by a factor of ~ 6 (C_3 grass, coniferous forest) to ~ 19 (deciduous forest) in comparison to CLM-Ens_P. This highlights the strong effect of an uncertain Q_{10} parameter on the uncertainty of predicted carbon fluxes in CLM. As shown in Fig. 6, the perturbed Q_{10} parameter strongly affected both $\sum ER_{PFT}$ and $\sum GPP_{PFT}$. For the forest PFTs, the effect of Q_{10} (Ens_{P+Q10}) on the uncertainty of the regional-scale $\sum GPP_{PFT}$, $\sum ER_{PFT}$ and $\sum NEE_{PFT}$ was much stronger than the effect of both uncertain meteorological forcings and initial states (CLM-Ens_{PAI}). The ensemble mean $\sum ER_{PFT}$ did not change significantly, except for C_3 grass. However, for coniferous forest, the ensemble means for CLM-Ens_{P+Q10} of $\sum ER_{PFT}$ and $\sum NEE_{PFT}$ were significantly different from CLM-Ens_P. This is not surprising, since the mean of $Q_{10} = 2.0$ for CLM-Ens_{P+Q10} was considerably lower than the Q_{10} values of CLM-Ens_P sampled from the estimated parameter sets (Table 3).

4 Discussion

The results of the study showed that the model–data consistency was enhanced with estimated parameters for RO_{gc} , ME_{gc} , WU_{gc} , SE_{gc} and NZ_{gc} . For KA_{gc} and EN_{gc} , not all evaluation indices indicated an improvement of modeled NEE or the improvement was not significant. For the roving station sites NZ, KA and EN, NEE time series of only 2–3 months were available for evaluation in contrast to the other sites, where time series of a whole year were available. Accordingly, evaluation results were not directly comparable between those sites. Post et al. (2016) showed that the evaluation runs with parameter estimates had a strongly varying performance over the year. Thus, if a complete year of NEE data had been available for evaluation of the roving station sites, evaluation results for those sites would have been more informative. Particularly the transfer of parameters estimated for C_3 grass and C_3 crops to the catchment domain was found to be critical. As expected, parameter values estimated for the winter wheat site ME did not enhance the model performance at the grid cell EN_{gc} which was dominated by sugar beet. Accordingly, results showed that parameter values estimated for one C_3 crop site are not necessarily transferable to other C_3 crop types characterized by different physiology and management. This highlights the limitations of large-scale modeling with land surface models that only distinguish between very broad groups of PFTs. For SE_{gc} , where winter wheat

was grown during the evaluation period like in ME, parameter estimates clearly improved the model performance in terms of NEE. Several studies have emphasized that LSM parameters can vary within one group of PFT such that the transfer from single site estimates to other sites with the same PFT is not trivial (Groenendijk et al., 2011; Xiao et al., 2011). In addition to deficits in the representation of plant phenology, management (harvesting, cutting, etc.) is not explicitly considered for C_3 crops and C_3 grass in CLM although it has a significant impact on NEE (Borchard et al., 2015). Therefore, observed temporal LAI variations in the growing season were not well represented in the default model simulations. However, estimated parameters by CLM-Ens_P provided more reliable estimates of the NEE sums for ME and SE than the default run because the modeled plant emergence was shifted ahead. Therefore, we expect that the performance of CLM can be improved if it is coupled to a crop model which can treat specific plant traits and include management practices. Li et al. (2011) found that crop variety and management factors like irrigation, fertilization and planting date have a significant impact on NEE (and evapotranspiration) for the coupled LSM–crop model ORCHIDEE-STICS. Wu et al. (2016) showed that a crop model coupled to the LSM ORCHIDEE, combined with assimilation of many different data types, was able to reproduce many measurement data up to the measurement uncertainty, whereas other remaining errors could be related to management activities which were not correctly represented in the model.

For the forest PFTs, the uncertainty of the Q_{10} parameter had the largest impact on NEE, followed by uncertainty of the initial states and uncertainty of meteorological forcings. The crucial role of the Q_{10} parameter in carbon stock and flux predictions and the respective uncertainties in most land surface models has been highlighted in previous studies (Post et al., 2008, 2016; Hararuk et al., 2014). Post et al. (2016) showed that Q_{10} correlates strongly with other CLM4.5 key parameters like fl_{NR} and the Ball–Berry slope of stomatal conductance as well as with the initial carbon–nitrogen pools. This may explain why the spread of the regional-scale $\sum GPP_{PFT}$, $\sum ER_{PFT}$ and $\sum NEE_{PFT}$ in Ens_{P+Q10} was notably higher for coniferous forest than for the other PFTs (Fig. 6).

It is important to stress again that uncertainty in initial states was considered by an additional 15-year spin-up with perturbed parameters and meteorological forcings. This short additional spin-up does not consider uncertainty of the most stable carbon and nitrogen pools. In addition, carbon and nitrogen pools are influenced by changing land use in recent centuries and by management practices, which were not considered here. Therefore, the true uncertainty related to the initial states will be larger than in these simulation experiments. The simulation results for C_3 grass and C_3 crops were surprising as uncertainty in the meteorological forcings had a much larger impact on uncertainty of NEE than the uncertainty of initial states and Q_{10} had. We argue that the

applied meteorological perturbations (standard deviations of 1.0 K for air temperature, 30 % of the incoming shortwave radiation and 50 % of precipitation amount) are realistic for meteorological reanalysis products, and in correspondence with other studies. The perturbations (which are also temporally correlated) result in considerable variations, between ensemble members, of leaf onset and senescence. The stress deciduous phenology scheme (which determines plant onset and offset for C₃ crops and C₃ grass in CLM as outlined in Sect. 2.2) is strongly determined by various arbitrary thresholds such as “crit_offset_swi” (water stress days for offset trigger) or “crit_onset_fdd” (critical number of freezing degree days to trigger onset). The phenology representation of C₃ crops and C₃ grass could be too sensitive to those thresholds (Dahlin et al., 2015). Verheijen et al. (2015) argue that the use of PFTs instead of plant traits in LSMs reduces the adaptive response of vegetation to environmental drivers.

A second possible reason for the large impact of the perturbation of meteorological forcings on NEE uncertainty is the preferential location of the C₃ crops and C₃ grass in the drier, northern part of the catchment, which sometimes experiences drought stress during summer. The perturbation of the precipitation amounts affects drought stress and can enhance the variability in simulated NEE and LAI for C₃ crops and C₃ grasses in drier areas. Thus, the impact of the uncertainty of meteorological forcings depends on the local conditions. Jung et al. (2011) already showed that interannual variability of NEE is larger in semi-arid and semi-humid areas than in other regions. In summary, the large impact of uncertainty of meteorological input on NEE uncertainty is likely primarily model specific, but also related to the semi-humid conditions in the agricultural part of the region. High impact of uncertainty of meteorological conditions can be expected in other semi-humid and semi-arid regions and should be taken into account in simulation studies.

In order to better estimate the uncertainty of the initial carbon and nitrogen pools, it is necessary to include the spin-up (for each of the ensemble members) in the data assimilation experiments where parameters are estimated. As model runs have to be repeated thousands of times, this would be extremely CPU expensive. The spin-up should also consider uncertainty of historical land use and uncertainty of historical meteorological conditions. In particular, land use and meteorological conditions of the last few centuries impact initial states. We believe that it is important to jointly consider the uncertainty of initial conditions and parameters in LSM forward and inverse modeling runs. Pinnington et al. (2016) already pointed towards the importance of considering correlations between initial states and parameters in inverse modeling studies.

5 Conclusions

This study evaluated ecosystem parameters estimated with DREAM_(zs) for the Community Land Model (CLM4.5) and analyzed the uncertainty of modeled NEE and LAI at regional scale for a catchment in the border region of western Germany. The ecosystem parameters were estimated for EC sites with different plant functional types (PFTs) in western Germany and northern France by conditioning them to time series of NEE data. These parameters were assigned to a high-resolution (1 km²) CLM4.5 setup for the Rur catchment in Germany. It was evaluated whether the distributed land surface model reproduced measured NEE and LAI better with estimated parameters than with default parameters. Moreover, a comprehensive model uncertainty analysis was done for the four main PFTs in the catchment, comparing the effect of different model error sources (parameters, atmospheric forcings and initial states) on the CLM ensemble spread.

Parameter estimation using NEE data was found suitable to improve LAI predictions. For C₃ crop, the timing of plant emergence in spring was more accurate. This resulted in an improved characterization of NEE, which by default was highly overestimated in spring in cases where the plant emergence was delayed. This highlights the potential of DREAM-CLM parameter estimates for utilizing individual observations at a minimal number of grid cells to improve catchment-wide predictions of LAI and carbon fluxes.

Estimated parameters reduced the relative difference between the observed and modeled NEE sum significantly for most evaluation sites compared to a reference run with default parameters. For all PFTs except C₃ grass, the sign of the mean annual NEE sum over the catchment was reversed from positive to negative with estimated parameters compared to default parameters. This implies that forest and C₃ crop areas in the catchment were predicted as net CO₂ sources with global default parameter values and as net CO₂ sinks with the successfully validated parameter estimates. This elucidates the potential and relevance of parameter estimation in terms of obtaining more reliable estimates of regional carbon balances.

If estimated parameters were sampled without additional consideration of uncertain atmospheric input data and uncertain initial conditions, the uncertainty of predicted NEE and LAI was underestimated. Constraining CLM parameters with DREAM_(zs) resulted in a very low uncertainty of the predicted NEE and LAI. However, with additional consideration of uncertainty in the initial model conditions and atmospheric input data, the uncertainty of NEE was significantly higher. Thus it is essential to take into account the uncertainty of parameters, atmospheric input data and initial states for predictions of carbon fluxes or stocks with CLM4.5.

The effect of uncertainty of atmospheric forcing data on the LAI and NEE uncertainty was considerably higher for C₃ grass and C₃ crops than for the forest PFTs. This is re-

lated to the specific C₃ crop/C₃ grass phenology representation and internal model thresholds in CLM. Many of these thresholds relate to temperature and strongly control leaf onset and senescence. The strong effect of the perturbed forcing data on modeled carbon fluxes in CLM4.5 is closely linked to the effect of uncertain model parameters like the temperature coefficient Q_{10} .

The model uncertainty resulting from uncertain atmospheric forcing data for C₃ grass and C₃ crops is additionally enhanced because these crops are located in the northern part of the catchment, which is prone to some drought stress in summer. In combination with soil-moisture-related model thresholds, the effect of the perturbed precipitation on modeled NEE and LAI can be regionally different due to differences in the regional climate. This stresses the importance of regional-scale modeling.

This study demonstrated that if different crop types are grown in a region, it is particularly important to parameterize them independently in order to obtain reliable carbon flux estimates with a land surface model like CLM. Separate differentiation of different crop types including a crop-specific representation of phenology and management is very important for obtaining more reliable carbon flux and stock estimates in the future, and it is foreseen for future CLM versions (personal communication with collaborators from the CLM development team at NCAR, 2016).

Code and data availability. Data sets or codes can be made available upon request (hanna.post@t-online.de).

Appendix A: The DREAM_(zs) parameter estimation approach

The adaptive Markov chain Monte Carlo (MCMC) method DREAM_(zs) (Ter Braak and Vrugt, 2008; Laloy and Vrugt, 2012; Vrugt, 2015) estimates the posterior pdf of model parameters based on Bayes' theorem:

$$p(\mathbf{x}|\tilde{\mathbf{Y}}) = \frac{p(\mathbf{x})p(\tilde{\mathbf{Y}}|\mathbf{x})}{p(\tilde{\mathbf{Y}})}, \quad (\text{A1})$$

where $\mathbf{x}$ is the model parameters to be estimated, $\tilde{\mathbf{Y}} = \{\tilde{y}_1, \dots, \tilde{y}_n\}$ is a n vector of measured data, $p(\mathbf{x}|\tilde{\mathbf{Y}})$ is the posterior probability density function (pdf), $L(\mathbf{x}|\tilde{\mathbf{Y}}) \equiv p(\tilde{\mathbf{Y}}|\mathbf{x})$ the likelihood function, $p(\mathbf{x})$ the prior distribution and $p(\tilde{\mathbf{Y}})$ the normalizing constant. In practice, statistical inferences about $p(\mathbf{x}|\tilde{\mathbf{Y}})$ are made from its unnormalized density, $p(\mathbf{x}|\tilde{\mathbf{Y}}) \propto p(\mathbf{x})L(\mathbf{x}|\tilde{\mathbf{Y}})$.

Whether a proposal point of chain i at iteration t is accepted and thus

$$\mathbf{X}_p^i = \mathbf{X}^i + d\mathbf{X}^i \quad (\text{A2})$$

is determined with the Metropolis acceptance ratio:

$$P_{\text{accept}}(\mathbf{x}_{t-1}^i \rightarrow \mathbf{X}_p^i) = \min \left[1, \frac{p(\mathbf{X}_p^i)}{p(\mathbf{x}_{t-1}^i)} \right]. \quad (\text{A3})$$

If the candidate point is accepted, then the i th chain moves to the new position, that is, $\mathbf{x}_{t-1}^i = \mathbf{X}_p^i$, otherwise $\mathbf{x}_t^i = \mathbf{x}_{t-1}^i$ (Vrugt, 2015).

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This work was carried out with the funding of the Transregional Collaborative Research Centre 32 (TR32) and the EU FP7 project ExpeER (grant agreement no. 262060), which is supported by the European Commission through the Seventh Framework Programme for Research and Technical Development. We gratefully acknowledge the TERENO and the Transregional Collaborative Research Centre 32 (TR32), for providing eddy covariance and additional data for the Rollesbroich, Merzenhausen, Selhausen, Niederzier, Kall and Wüstebach sites. The authors gratefully acknowledge the computing time granted on the supercomputer JUROPA at the Jülich Supercomputing Centre (JSC). RapidEye data were provided by Blackbridge with the RapidEye Science Archive (RESA). We also thank Pramod Kumbhar and Guowei He for their support with high-performance computing. We kindly thank Karl Schneider and Tim Reichenau (Department of Geography, University of Cologne, Germany) for their support. Moreover, we thank Tim Hoar from the National Atmospheric Research Center (NCAR, Boulder) for his support on setting up ensemble CLM modeling as well as the Terrestrial Science Team (Gordon Bonan's group) at NCAR for their cooperation. Finally, the authors acknowledge and thank one anonymous referee and Sönke Zaehle for providing constructive comments and guiding the revision process.

The article processing charges for this open-access publication were covered by a Research Centre of the Helmholtz Association.

Edited by: Sönke Zaehle

Reviewed by: Sönke Zaehle and one anonymous referee

References

- Abramowitz, G., Leuning, R., Clark, M., and Pitman, A.: Evaluating the performance of land surface models, *J. Climate*, 21, 5468–5481, 2008.
- Ahlstrom, A., Schurgers, G., Arneth, A., and Smith, B.: Robustness and uncertainty in terrestrial ecosystem carbon response to CMIP5 climate change projections, *Environ. Res. Lett.*, 7, 4, <https://doi.org/10.1088/1748-9326/7/4/044008>, 2012.
- Ali, M., Montzka, C., Stadler, A., Menz, G., Thonfeld, F., and Vereecken, H.: Estimation and Validation of RapidEye-Based Time-Series of Leaf Area Index for Winter Wheat in the Rur Catchment (Germany), *Remote Sens.*, 7, 2808–2831, <https://doi.org/10.3390/rs70302808>, 2015.
- Anderson, M. C., Kustas, W. P., and Norman, J. M.: Up-scaling and Downscaling – A Regional View of the Soil–Plant–Atmosphere Continuum, *Agron. J.*, 95, 1408–1423, <https://doi.org/10.2134/agronj2003.1408>, 2003.
- Arora, V. K., Boer, G. J., Friedlingstein, P., Eby, M., Jones, C. D., Christian, J. R., Bonan, G., Bopp, L., Brovkin, V., Cadule, P., Hajima, T., Ilyina, T., Lindsay, K., Tjiputra, J. F., and Wu, T.: Carbon–Concentration and Carbon–Climate Feedbacks in CMIP5 Earth System Models, *J. Climate*, 26, 5289–5314, <https://doi.org/10.1175/JCLI-D-12-00494.1>, 2013.
- Baatz, R., Bogen, H. R., Hendricks Franssen, H.-J., Huisman, J. A., Qu, W., Montzka, C., and Vereecken, H.: Calibration of a catchment scale cosmic-ray probe network: A comparison of three parameterization methods, *J. Hydrol.*, 516, 231–244, <https://doi.org/10.1016/j.jhydrol.2014.02.026>, 2014.
- Baldauf, M., Förstner, J., Klink, S., Reinhardt, T., Schraff, C., Seifert, A., Stephan, K., and Wetterdienst, D.: Kurze Beschreibung des Lokal-Modells Kurzestfrist COSMO-DE (LMK) und seiner Datenbanken auf dem Datenserver des DWD, Deutscher Wetterdienst, Geschäftsbereich Forschung und Entwicklung, Offenbach, Germany, 2009.
- Baldocchi, D. D. and Wilson, K. B.: Modeling CO₂ and water vapor exchange of a temperate broadleaved forest across hourly to decadal time scales, *Ecol. Model.*, 142, 155–184, [https://doi.org/10.1016/S0304-3800\(01\)00287-3](https://doi.org/10.1016/S0304-3800(01)00287-3), 2001.
- Beven, K. and Freer, J.: Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, *J. Hydrol.*, 249, 11–29, [https://doi.org/10.1016/S0022-1694\(01\)00421-8](https://doi.org/10.1016/S0022-1694(01)00421-8), 2001.
- Bonan, G. B., DeFries, R. S., Coe, M. T., and Ojima, D. S.: Land Use and Climate, in: *Land Change Science; Observing, Monitoring and Understanding Trajectories of Change on the Earth's Surface*, *Remote Sens. Digital Image Process.*, 301–314, 2002.
- Bonan, G. B., Lawrence, P. J., Oleson, K. W., Levis, S., Jung, M., Reichstein, M., Lawrence, D. M., and Swenson, S. C.: Improving canopy processes in the Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data, *J. Geophys. Res.*, 116, G02014, <https://doi.org/10.1029/2010JG001593>, 2011.
- Borchard, N., Schirrmann, M., von Hebel, C., Schmidt, M., Baatz, R., Firbank, L., Vereecken, H., and Herbst, M.: Spatio-temporal drivers of soil and ecosystem carbon fluxes at field scale in an upland grassland in Germany, *Agric. Ecosyst. Environ.*, 211, 84–93, <https://doi.org/10.1016/j.agee.2015.05.008>, 2015.
- Bréda, N. J. J.: Ground-based measurements of leaf area index: a review of methods, instruments and current controversies, *J. Exp. Bot.*, 54, 2403–2417, <https://doi.org/10.1093/jxb/erg263>, 2003.
- Brovkin, V., Boysen, L., Arora, V. K., Boisier, J. P., Cadule, P., Chini, L., Claussen, M., Friedlingstein, P., Gayler, V., van den Hurk, B. J. J. M., Hurtt, G. C., Jones, C. D., Kato, E., de Noblet-Ducoudré, N., Pacifico, F., Pongratz, J., and Weiss, M.: Effect of Anthropogenic Land-Use and Land-Cover Changes on Climate and Land Carbon Storage in CMIP5 Projections for the Twenty-First Century, *J. Climate*, 26, 6859–6881, <https://doi.org/10.1175/JCLI-D-12-00623.1>, 2013.
- Carvalhais, N., Reichstein, M., Ciais, P., Collatz, G. J., Mahecha, M. D., Montagnani, L., Papale, D., Rambal, S., and Seixas, J.: Identification of vegetation and soil carbon pools out of equilibrium in a process model via eddy covariance and biometric constraints, *Global Change Biol.*, 16, 2813–2829, 2010.
- Chen, B., Black, T. A., Coops, N. C., Hilker, T., Trofymow, J. A., and Morgenstern, K.: Assessing tower flux footprint climatology and scaling between remotely sensed and eddy covariance measurements, *Bound.-Lay. Meteorol.*, 130, 137–167, 2009.
- Chen, B., Coops, N. C., Fu, D., Margolis, H. A., Amiro, B. D., Black, T. A., Arain, M. A., Barr, A. G., Bourque, C. P.-A., Flanagan, L. B., Lafleur, P. M., McCaughey, J.

- H., and Wofsy, S. C.: Characterizing spatial representativeness of flux tower eddy-covariance measurements across the Canadian Carbon Program Network using remote sensing and footprint analysis, *Remote Sens. Environ.*, 124, 742–755, <https://doi.org/10.1016/j.rse.2012.06.007>, 2012.
- Chen, J. M., Rich, P. M., Gower, S. T., Norman, J. M., and Plummer, S.: Leaf area index of boreal forests: theory, techniques, and measurements, *J. Geophys. Res.-Atmos.*, 102, 29429–29443, 1997.
- Dahlin, K. M., Fisher, R. A., and Lawrence, P. J.: Environmental drivers of drought deciduous phenology in the Community Land Model, *Biogeosciences*, 12, 5061–5074, <https://doi.org/10.5194/bg-12-5061-2015>, 2015.
- Dai, Y., Dickinson, R. E., and Wang, Y.-P.: A Two-Big-Leaf Model for Canopy Temperature, Photosynthesis, and Stomatal Conductance, *J. Climate*, 17, 2281–2299, 2004.
- Desai, A. R., Noormets, A., Bolstad, P. V., Chen, J., Cook, B. D., Davis, K. J., Euskirchen, E. S., Gough, C., Martin, J. G., Ricciuto, D. M., Schmid, H. P., Tang, J., and Wang, W.: Influence of vegetation and seasonal forcing on carbon dioxide fluxes across the Upper Midwest, USA: Implications for regional scaling, *Agr. Forest Meteorol.*, 148, 288–308, <https://doi.org/10.1016/j.agrformet.2007.08.001>, 2008.
- Exbrayat, J.-F., Pitman, A. J., and Abramowitz, G.: Response of microbial decomposition to spin-up explains CMIP5 soil carbon range until 2100, *Geosci. Model Dev.*, 7, 2683–2692, <https://doi.org/10.5194/gmd-7-2683-2014>, 2014.
- Fisher, J. B., Sikka, M., Oechel, W. C., Huntzinger, D. N., Melton, J. R., Koven, C. D., Ahlström, A., Arain, M. A., Baker, I., Chen, J. M., Ciais, P., Davidson, C., Dietze, M., El-Masri, B., Hayes, D., Huntingford, C., Jain, A. K., Levy, P. E., Lomas, M. R., Poulter, B., Price, D., Sahoo, A. K., Schaefer, K., Tian, H., Tomelleri, E., Verbeeck, H., Viovy, N., Wania, R., Zeng, N., and Miller, C. E.: Carbon cycle uncertainty in the Alaskan Arctic, *Biogeosciences*, 11, 4271–4288, <https://doi.org/10.5194/bg-11-4271-2014>, 2014.
- Foereid, B., Ward, D. S., Mahowald, N., Paterson, E., and Lehmann, J.: The sensitivity of carbon turnover in the Community Land Model to modified assumptions about soil processes, *Earth Syst. Dynam.*, 5, 211–221, <https://doi.org/10.5194/esd-5-211-2014>, 2014.
- Ganguly, S., Nemani, R. R., Zhang, G., Hashimoto, H., Milesi, C., Michaelis, A., Wang, W., Votava, P., Samanta, A., Melton, F., Dungan, M., Vermote, F., Gao, Y., Knyazikhin, Y., and Myneni, R. B.: Generating global leaf area index from Landsat: Algorithm formulation and demonstration, *Remote Sens. Environ.*, 122, 185–202, 2012.
- Garrigues, S., Allard, D., Baret, F., and Weiss, M.: Influence of landscape spatial heterogeneity on the non-linear estimation of leaf area index from moderate spatial resolution remote sensing data, *Remote Sens. Environ.*, 105, 286–298, 2006.
- Garrigues, S., Lacaze, R., Baret, F., Morissette, J. T., Weiss, M., Nickeson, J. E., Fernandes, R., Plummer, S., Shabanov, N. V., Myneni, R. B., Knyazikhin, Y., and Yang, W.: Validation and intercomparison of global Leaf Area Index products derived from remote sensing data, *J. Geophys. Res.*, 113, G02028, <https://doi.org/10.1029/2007JG000635>, 2008.
- Gelman, A. and Rubin, D. B.: Inference from iterative simulation using multiple sequences, *Statist. Sci.*, 7, 457–472, 1992.
- Göhler, M., Mai, J., and Cuntz, M.: Use of eigendecomposition in a parameter sensitivity analysis of the Community Land Model, *J. Geophys. Res.-Biogeosci.*, 118, 904–921, <https://doi.org/10.1002/jgrg.20072>, 2013.
- Graf, A., Bogen, H. R., Drüe, C., Hardelauf, H., Pütz, T., Heineemann, G., and Vereecken, H.: Spatiotemporal relations between water budget components and soil water content in a forested tributary catchment, *Water Resour. Res.*, 50, 4837–4857, 2014.
- Groenendijk, M., Dolman, A. J., Van der Molen, M. K., Leuning, R., Arneth, A., Delapierre, 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, 2011.
- Grünwald, T. and Bernhofer, C.: A decade of carbon, water and energy flux measurements of an old spruce forest at the Anchor Station Tharandt, *Tellus B*, 59, 387–396, 2007.
- Gu, L. H., Pallardy, S. G., Yang, B., Hosman, K. P., Mao, J. F., Ricciuto, D., Shi, X. Y., and Sun, Y.: Testing a land model in ecosystem functional space via a comparison of observed and modeled ecosystem flux responses to precipitation regimes and associated stresses in a Central US forest, *J. Geophys. Res.-Biogeosciences*, 121, 1884–1902, 2016.
- Han, X., Li, X., Hendricks Franssen, H. J., Vereecken, H., and Montzka, C.: Spatial horizontal correlation characteristics in the land data assimilation of soil moisture, *Hydrol. Earth Syst. Sci.*, 16, 1349–1363, <https://doi.org/10.5194/hess-16-1349-2012>, 2012.
- Han, X., Hendricks Franssen, H.-J., Li, X., Zhang, Y., Montzka, C., and Vereecken, H.: Joint assimilation of surface temperature and L-band microwave brightness temperature in land data assimilation, *Vadose Zone Journal*, 12, 3, 2013.
- Han, X., Franssen, H.-J. H., Montzka, C., and Vereecken, H.: Soil moisture and soil properties estimation in the Community Land Model with synthetic brightness temperature observations, *Water Resour. Res.*, 50, 6081–6105, <https://doi.org/10.1002/2013WR014586>, 2014.
- Hararuk, O., Xia, J., and Luo, Y.: Evaluation and improvement of a global land model against soil carbon data using a Bayesian Markov chain Monte Carlo method, *J. Geophys. Res.-Biogeosci.*, 119, 403–417, <https://doi.org/10.1002/2013JG002535>, 2014.
- Härkönen, S., Lehtonen, A., Manninen, T., Tuominen, S., and Peltoniemi, M.: Estimating forest leaf area index using satellite images: comparison of k-NN based Landsat-NFI LAI with MODIS-RSR based LAI product for Finland, *Boreal Environ. Res.*, 20, 181–19, 2015.
- Huntzinger, D. N., Post, W. M., Wei, Y., Michalak, A. M., West, T. O., Jacobson, A. R., Baker, I. T., Chen, J. M., Davis, K. J., Hayes, D. J., Hoffman, F. M., Jain, A. K., Liu, S., McGuire, A. D., Neilson, R. P., Potter, C., Poulter, B., Price, D., Raczka, B. M., Tian, H. Q., Thornton, P., Tomelleri, E., Viovy, N., Xiao, J., Yuan, W., Zeng, N., Zhao, M., and Cook, R.: North American Carbon Program (NACP) regional interim synthesis: Terrestrial biospheric model intercomparison, *Ecol. Model.*, 232, 144–157, <https://doi.org/10.1016/j.ecolmodel.2012.02.004>, 2012.
- Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Altaf Arain, M., Arneth, A., Bernhofer, C., Bonal, D., Chen, J., Gianelle, D., Gobron, N., Kiely, G., Kutsch, W., Lasslop, G., Law, B. E., Lindroth, A., Merbold, A., Montagnani, L., Moors, E. J., Papale, D., Sottocornale, M., Vaccari, F., and Williams, C.: Global patterns of land-

- atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations, *J. Geophys. Res.-Biogeosci.*, 116, G00J07, <https://doi.org/10.1029/2010JG001566>, 2011.
- Keppel-Aleks, G., Wolf, A. S., Mu, M. Q., Doney, S. C., Morton, D. C., Kasibhatla, P. S., Miller, J. B., Dlugokencky, E. J., and Randerson, J. T.: Separating the influence of temperature, drought, and fire on interannual variability in atmospheric CO₂, *Global Biogeochem. Cycles*, 28, 1295–1310, 2014.
- Kessomkiat, W., Hendricks-Franssen, H.-J., Graf, A., and Vereecken, H.: Estimating random errors of eddy covariance data: An extended two-tower approach, *Agric. Forest Meteorol.*, 171, 203–219, 2013.
- Korres, W., Koyama, C. N., Fiener, P., and Schneider, K.: Analysis of surface soil moisture patterns in agricultural landscapes using Empirical Orthogonal Functions, *Hydrol. Earth Syst. Sci.*, 14, 751–764, <https://doi.org/10.5194/hess-14-751-2010>, 2010.
- Knohl, A., Schulze, E.-D., Kolbe, O., and Buchmann, N.: Large carbon uptake by an unmanaged 250-year-old deciduous forest in Central Germany, *Agric. Forest Meteorol.*, 118, 151–167, 2003.
- Kumar, S. V., Reichle, R. H., Harrison, K. W., Peters-Lidard, C. D., Yatheendradas, S., and Santanello, J. A.: A comparison of methods for a priori bias correction in soil moisture data assimilation, *Water Resour. Res.*, 48, W03515, <https://doi.org/10.1029/2010WR010261>, 2012.
- Kuppel, S., Peylin, P., Chevallier, F., Bacour, C., Maignan, F., and Richardson, A. D.: Constraining a global ecosystem model with multi-site eddy-covariance data, *Biogeosciences*, 9, 3757–3776, <https://doi.org/10.5194/bg-9-3757-2012>, 2012.
- Kwon, H. J., Oechel, W. C., Zulueta, R. C., and Hastings, S. J.: Effects of climate variability on carbon sequestration among adjacent wet sedge tundra and moist tussock tundra ecosystems, *J. Geophys. Res.-Biogeosci.*, 111, G03014, <https://doi.org/10.1029/2005JG000036>, 2006.
- Laloy, E. and Vrugt, J. A.: High-dimensional posterior exploration of hydrologic models using multiple-try DREAM(ZS) and high-performance computing, *Water Resour. Res.*, 48, W01526, <https://doi.org/10.1029/2011WR010608>, 2012.
- Lawrence, D. M., Oleson, K. W., Flanner, M. G., Fletcher, C. G., Lawrence, P. J., Levis, S., Swenson, S. C., and Bonan, G. B.: The CCSM4 land simulation, 1850–2005: Assessment of surface climate and new capabilities, *J. Climate*, 25, 2240–2260, 2012.
- Le Quéré, C., Andres, R. J., Boden, T., Conway, T., Houghton, R. A., House, J. I., Marland, G., Peters, G. P., van der Werf, G. R., Ahlström, A., Andrew, R. M., Bopp, L., Canadell, J. G., Ciais, P., Doney, S. C., Enright, C., Friedlingstein, P., Huntingford, C., Jain, A. K., Jourdain, C., Kato, E., Keeling, R. F., Klein Goldewijk, K., Levis, S., Levy, P., Lomas, M., Poulter, B., Raupach, M. R., Schwinger, J., Sitch, S., Stocker, B. D., Viovy, N., Zaehle, S., and Zeng, N.: The global carbon budget 1959–2011, *Earth Syst. Sci. Data*, 5, 165–185, <https://doi.org/10.5194/essd-5-165-2013>, 2013.
- Li, K. Y., Coe, M. T., Ramankutty, N., and De Jong, R.: Modeling the hydrological impact of land-use change in West Africa, *J. Hydrol.*, 337, 258–268, 2007.
- Li, L., Vuichard, N., Viovy, N., Ciais, P., Wang, T., Ceschia, E., Jans, W., Wattenbach, M., Béziat, P., Gruenwald, T., Lehuger, S., and Bernhofer, C.: Importance of crop varieties and management practices: evaluation of a process-based model for simulating CO₂ and H₂O fluxes at five European maize (*Zea mays* L.) sites, *Biogeosciences*, 8, 1721–1736, <https://doi.org/10.5194/bg-8-1721-2011>, 2011.
- Lussem, U. and Waldhoff, G.: Land use classification 2012 of the Rur Catchment, CRC/TR32 Database (TR32DB), available at: <http://tr32db.uni-koeln.de/data.php?dataID=801>, last access: 12 September 2013.
- Mauder, M. and Foken, T.: Documentation and instruction manual of the Eddy covariance software package TK3, Univ. Bayreuth, Abt. Mikrometeorologie, 2011.
- Mauder, M., Cuntz, M., Drüe, C., Graf, A., Rebmann, C., Schmid, H. P., Schmidt, M., and Steinbrecher, R.: A strategy for quality and uncertainty assessment of long-term eddy-covariance measurements, *Agr. Forest Meteorol.*, 169, 122–135, 2013.
- Migliavacca, M., Reichstein, M., Richardson, A. D., Mahecha, M. D., Cremonese, E., Delpierre, N., Galvagno, M., Law, B. E., Wohlfahrt, G., Andrew Black, T., Carvalhais, N., Ceccherini, G., Chen, J., Gobron, N., Koffi, E., William Munger, J., Perez-Priego, O., Robustelli, M., Tomelleri, E., and Cescatti, A.: Influence of physiological phenology on the seasonal pattern of ecosystem respiration in deciduous forests, *Glob. Change Biol.*, 21, 363–376, <https://doi.org/10.1111/gcb.12671>, 2015.
- Montzka, C., Canty, M., Kreins, P., Kunkel, R., Menz, G., Vereecken, H., and Wendland, F.: Multispectral remotely sensed data in modelling the annual variability of nitrate concentrations in the leachate. *Environ. Modell. Softw.*, 23, 1070–1081, <https://doi.org/10.1016/j.envsoft.2007.11.010>, 2008a.
- Montzka, C., Canty, M., Kunkel, R., Menz, G., Vereecken, H., and Wendland, F.: Modelling the water balance of a mesoscale catchment basin using remotely sensed land cover data, *J. Hydrol.*, 353, 322–334, <https://doi.org/10.1016/j.jhydrol.2008.02.018>, 2008b.
- Oleson, K., Lawrence, D. M., Bonan, G. B., Drewniak, B., Huang, M., Koven, C. D., Levis, S., Li, F., Riley, W. J., Subin, Z. M., Swenson, S., Thornton, P. E., Bozbiyik, A., Fisher, R., Heald, C. L., Kluzek, E., Lamarque, J.-F., Lawrence, P. J., Leung, L. R., Lipscomb, W., Muszala, S. P., Ricciuto, D. M., Sacks, W. J., Sun, Y., Tang, J., and Yang, Z.-L.: Technical description of version 4.5 of the Community Land Model (CLM), NCAR, Boulder, 2013.
- Park, S. K. and Xu, L.: Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. II), Springer Science & Business Media, Berlin Heidelberg, 2013.
- Peylin, P., Bacour, C., MacBean, N., Leonard, S., Rayner, P., Kuppel, S., Koffi, E., Kane, A., Maignan, F., Chevallier, F., Ciais, P., and Prunet, P.: A new stepwise carbon cycle data assimilation system using multiple data streams to constrain the simulated land surface carbon cycle, *Geosci. Model Dev.*, 9, 3321–3346, <https://doi.org/10.5194/gmd-9-3321-2016>, 2016.
- Piao, S., Sitch, S., Ciais, P., Friedlingstein, P., Peylin, P., Wang, X., Ahlström, A., Anav, A., Canadell, J. G., Cong, N., Huntingford, C., Jung, M., Levis, S., Levy, P. E., Li, J., Lin, X., Lomas, M. R., Lu, M., Luo, Y., Ma, Y., Myneni, R. B., Poulter, B., Sun, Z., Wang, T., Viovy, N., Zaehle, S., and Zeng, N.: Evaluation of terrestrial carbon cycle models for their response to climate variability and to CO₂ trends, *Glob. Change Biol.*, 19, 2117–2132, <https://doi.org/10.1111/gcb.12187>, 2013.
- Pinnington, E. M., Casella, E., Dance, S. L., Lawless, A. S., Morison, J. I. L., Nichols, N. K., Wilkinson, M., and Quaife, T. L.: Investigating the role of prior and observation error correlations in

- improving a model forecast of forest carbon balance using four-dimensional variational data assimilation, *Agr. Forest Meteorol.*, 228, 299–314, 2016.
- Post, H., Hendricks Franssen, H. J., Graf, A., Schmidt, M., and Vereecken, H.: Uncertainty analysis of eddy covariance CO₂ flux measurements for different EC tower distances using an extended two-tower approach, *Biogeosciences*, 12, 1205–1221, <https://doi.org/10.5194/bg-12-1205-2015>, 2015.
- Post, H., Vrugt, J. A., Fox, A., Vereecken, H., and Hendricks Franssen, H.-J.: Estimation of Community Land Model parameters with DREAM(zs) for an improved assessment of net carbon fluxes at European sites, *J. Geophys. Res.-Biogeosci.*, 122, 661–689, <https://doi.org/10.1002/2015JG003297>, 2016.
- Post, J., Hattermann, F. F., Krysanova, V., and Suckow, F.: Parameter and input data uncertainty estimation for the assessment of long-term soil organic carbon dynamics, *Environ. Modell. Softw.*, 23, 125–138, <https://doi.org/10.1016/j.envsoft.2007.05.010>, 2008.
- Prihodko, L., Denning, A. S., Hanan, N. P., Baker, I., and Davis, K.: Sensitivity, uncertainty and time dependence of parameters in a complex land surface model, *Agr. Forest Meteorol.*, 148, 268–287, 2008.
- Propastin, P. and Erasmi, S.: A physically based approach to model LAI from MODIS 250 m data in a tropical region, *Int. J. Appl. Earth Obs.*, 12, 47–59, 2010.
- Reichenau, T. G., Korres, W., Montzka, C., Fiener, P., Wilken, F., Stadler, A., Waldhoff, G., and Schneider, K.: Spatial Heterogeneity of Leaf Area Index (LAI) and Its Temporal Course on Arable Land: Combining Field Measurements, Remote Sensing and Simulation in a Comprehensive Data Analysis Approach (CDAA), *PLOS ONE*, 11, e0158451, <https://doi.org/10.1371/journal.pone.0158451>, 2016.
- Reichle, R. H., Koster, R. D., Liu, P., Mahanama, S. P. P., Njoku, E. G., and Owe, M.: Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SMMR), *J. Geophys. Res.*, 112, D09108, <https://doi.org/10.1029/2006JD008033>, 2007.
- Reichle, R. H., Kumar, S. V., Mahanama, S. P., Koster, R. D., and Liu, Q.: Assimilation of Satellite-Derived Skin Temperature Observations into Land Surface Models, *J. Hydrometeorol.*, 11, 1103–1122, <https://doi.org/10.1175/2010JHM1262.1>, 2010.
- Reichstein, M. and Beer, C.: Soil respiration across scales: The importance of a model–data integration framework for data interpretation, *J. Plant Nutr. Soil Sc.*, 171, 344–354, <https://doi.org/10.1002/jpln.200700075>, 2008.
- Ren, X. L., He, H. L., Moore, D. J. P., Zhang, L., Liu, M., Li, F., Yu, G. R., and Wang, H. M.: Uncertainty analysis of modeled carbon and water fluxes in a subtropical coniferous plantation, *J. Geophys. Res.-Biogeosci.*, 118, 1674–1688, 2013.
- Richardson, A. D., Williams, M., Hollinger, D. Y., Moore, D. J. P., Dail, D. B., Davidson, E. A., Scott, N. A., Evans, R. S., Hughes, H., Lee, J. T., Rodrigues, C., and Savage, K.: Estimating parameters of a forest ecosystem C model with measurements of stocks and fluxes as joint constraints, *Oecologia*, 164, 25–40, 2010.
- Richardson, A. D., Anderson, R. S., Arain, M. A., Barr, A. G., Bohrer, G., Chen, G., Chen, J. M., Ciais, P., Davis, K. J., Desai, A. R., Dietze, M. C., Dragoni, D., Garrity, S. R., Gough, C. M., Grant, R., Hollinger, D. Y., Margolis, H. A., McCaughey, H., Migliavacca, M., Monson, R. K., Munger, J. W., Poulter, B., Raczka, B. M., Ricciuto, D. M., Sahoo, A. K., Schaefer, K., Tian, H., Vargas, R., Verbeeck, H., Xiao, J., and Xue, Y.: Terrestrial biosphere models need better representation of vegetation phenology: results from the North American Carbon Program Site Synthesis, *Glob. Change Biol.*, 18, 566–584, <https://doi.org/10.1111/j.1365-2486.2011.02562.x>, 2012.
- Safta, C., Ricciuto, D. M., Sargsyan, K., Debusschere, B., Najm, H. N., Williams, M., and Thornton, P. E.: Global sensitivity analysis, probabilistic calibration, and predictive assessment for the data assimilation linked ecosystem carbon model, *Geosci. Model Dev.*, 8, 1899–1918, <https://doi.org/10.5194/gmd-8-1899-2015>, 2015.
- Schmidt, M., Reichenau, T. G., Fiener, P., and Schneider, K.: The carbon budget of a winter wheat field: An eddy covariance analysis of seasonal and inter-annual variability, *Agr. Forest Meteorol.*, 165, 114–126, <https://doi.org/10.1016/j.agrformet.2012.05.012>, 2012.
- scilands GmbH: Digital Elevation Model 10 without anthropogenic landforms, Göttingen, Germany, 2010.
- Spadavecchia, L., Williams, M., and Law, B. E.: Uncertainty in predictions of forest carbon dynamics: separating driver error from model error, *Ecol. Appl.*, 21, 1506–1522, 2011.
- Stöckli, R., Lawrence, D. M., Niu, G.-Y., Oleson, K. W., Thornton, P. E., Yang, Z.-L., Bonan, G. B., Denning, A. S., and Running, S. W.: Use of FLUXNET in the Community Land Model development, *J. Geophys. Res.*, 113, 19, <https://doi.org/10.1029/2007JG000562>, 2008.
- Stoy, P. C., Williams, M., Disney, M., Prieto-Blanco, A., Huntley, B., Baxter, R., and Lewis, P.: Upscaling as ecological information transfer: a simple framework with application to Arctic ecosystem carbon exchange, *Landscape Ecol.*, 24, 971–986, 2009.
- Ter Braak, C. J. F. and Vrugt, J. A.: Differential Evolution Markov Chain with snooker updater and fewer chains, *Stat. Comput.*, 18, 435–446, <https://doi.org/10.1007/s11222-008-9104-9>, 2008.
- Thornton, P. E. and Zimmermann, N. E.: An improved canopy integration scheme for a land surface model with prognostic canopy structure, *J. Climate*, 20, 3902–3923, 2007.
- Tillack, A., Clasen, A., Kleinschmit, B., and Förster, M.: Estimation of the Seasonal Leaf Area Index in an Alluvial Forest Using High-Resolution Satellite-Based Vegetation Indices, *Remote Sens. Environ.*, 141, 52–63, 2014.
- Todd-Brown, K. E., Hopkins, F. M., Kivlin, S. N., Talbot, J. M., and Allison, S. D.: A framework for representing microbial decomposition in coupled climate models, *Biogeochemistry*, 109, 19–33, 2012.
- Todd-Brown, K. E. O., Randerson, J. T., Post, W. M., Hoffman, F. M., Tarnocai, C., Schuur, E. A. G., and Allison, S. D.: Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations, *Biogeosciences*, 10, 1717–1736, <https://doi.org/10.5194/bg-10-1717-2013>, 2013.
- Todd-Brown, K. E. O., Randerson, J. T., Hopkins, F., Arora, V., Hajima, T., Jones, C., Shevliakova, E., Tjiputra, J., Volodin, E., Wu, T., Zhang, Q., and Allison, S. D.: Changes in soil organic carbon storage predicted by Earth system models during the 21st century, *Biogeosciences*, 11, 2341–2356, <https://doi.org/10.5194/bg-11-2341-2014>, 2014.

- Tyc, G., Tulip, J., Schulten, D., Kruschke, M., and Oxford, M.: The RapidEye mission design. *Acta Astronautica*, 4th IAA International Symposium on Small Satellites for Earth Observation, 56, 213–219, <https://doi.org/10.1016/j.actaastro.2004.09.029>, 2005.
- Verheijen, L. M., Aerts, R., Brovkin, V., Cavender-Bares, J., Cornelissen, J. H. C., Kattge, J., and van Bodegom, P. M.: Inclusion of ecologically based trait variation in plant functional types reduces the projected land carbon sink in an earth system model, *Glob. Change Biol.*, 21, 3071–3086, 2015.
- Vrugt, J. A.: Markov Chain Monte Carlo Simulation Using the DREAM Software Package: Theory, Concepts, and MATLAB Implementation, *Environ. Model. Softw.*, 75, 273–316, 2015.
- Waldhoff, G.: Land use classification of 2009 for the Rur catchment, TR31.1, 2010.
- Waldhoff, G., Curdt, C., Hoffmeister, D., and Bareth, G.: Analysis of multitemporal and multisensor remote sensing data for crop rotation mapping, *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, I-7, 177–182, 2012.
- Wang, Y. P., Kowalczyk, E., Leuning, R., Abramowitz, G., Raupach, M. R., Pak, B., van Gorsel, E., and Luhar, A.: Diagnosing errors in a land surface model (CABLE) in the time and frequency domains, *J. Geophys. Res.-Biogeosci.*, 166, G01034, <https://doi.org/10.1029/2010JG001385>, 2011.
- Williams, M., Richardson, A. D., Reichstein, M., Stoy, P. C., Peylin, P., Verbeeck, H., Carvalhais, N., Jung, M., Hollinger, D. Y., Kattge, J., Leuning, R., Luo, Y., Tomelleri, E., Trudinger, C. M., and Wang, Y.-P.: Improving land surface models with FLUXNET data, *Biogeosciences*, 6, 1341–1359, <https://doi.org/10.5194/bg-6-1341-2009>, 2009.
- Wu, X., Vuichard, N., Ciais, P., Viovy, N., de Noblet-Ducoudré, N., Wang, X., Magliulo, V., Wattenbach, M., Vitale, L., Di Tommasi, P., Moors, E. J., Jans, W., Elbers, J., Ceschia, E., Tallec, T., Bernhofer, C., Grünwald, T., Moureaux, C., Manise, T., Ligne, A., Cellier, P., Loubet, B., Larmanou, E., and Ripoche, D.: ORCHIDEE-CROP (v0), a new process-based agro-land surface model: model description and evaluation over Europe, *Geosci. Model Dev.*, 9, 857–873, <https://doi.org/10.5194/gmd-9-857-2016>, 2016.
- Xia, J. Y., Luo, Y. Q., Wang, Y.-P., Weng, E. S., and Hararuk, O.: A semi-analytical solution to accelerate spin-up of a coupled carbon and nitrogen land model to steady state, *Geosci. Model Dev.*, 5, 1259–1271, <https://doi.org/10.5194/gmd-5-1259-2012>, 2012.
- Xiao, J. F., Davis, K. J., Urban, N. M., Keller, K., and Salendra, N. Z.: Upscaling carbon fluxes from towers to the regional scale: Influence of parameter variability and land cover representation on regional flux estimates, *J. Geophys. Res.-Biogeosci.*, 116, G00J06, <https://doi.org/10.1029/2010JG001568>, 2011.
- Xiao, J. F., Davis, K. J., Urban, N. M., and Keller, K.: Uncertainty in model parameters and regional carbon fluxes: A model-data fusion approach, *Agr. Forest Meteorol.*, 189–190, 175–186, <https://doi.org/10.1016/j.agrformet.2014.01.022>, 2014.
- Xu, X., Shi, Z., Chen, X. C., Lin, Y., Niu, S. L., Jiang, L. F., Luo, R. S., and Luo, Y. Q.: Unchanged carbon balance driven by equivalent responses of production and respiration to climate change in a mixed-grass prairie, *Glob. Change Biol.*, 22, 1857–1866, 2016.
- Zacharias, S., Bogen, H., Samaniego, L., Mauder, M., Fuß, R., Pütz, T., Frenzel, M., Schwank, M., Baessler, C., Butterbach-Bahl, K., Bens, O., Borg, E., Brauer, A., Dietrich, P., Hajsek, I., Helle, G., Kiese, R., Kunstmann, H., Klotz, S., Munch, J. C., Papen, H., Priesack, E., Schmid, H. P., Steinbrecher, R., Rosenbaum, U., Teutsch, G., and Vereecken, H.: A Network of Terrestrial Environmental Observatories in Germany, *Vadose Zone J.*, 10, 955–973, <https://doi.org/10.2136/vzj2010.0139>, 2011.
- Zhang, L., Yu, G. R., Gu, F. X., He, H. L., Zhang, L. M., and Han, S. J.: Uncertainty analysis of modeled carbon fluxes for a broad-leaved Korean pine mixed forest using a process-based ecosystem model, *J. Forest Res.-Jpn.*, 17, 268–282, 2012.
- Zhao, F., Zeng, N., Asrar, G., Friedlingstein, P., Ito, A., Jain, A., Kalnay, E., Kato, E., Koven, C. D., Poulter, B., Rafique, R., Sitch, S., Shu, S., Stocker, B., Viovy, N., Wiltshire, A., and Zaehle, S.: Role of CO₂, climate and land use in regulating the seasonal amplitude increase of carbon fluxes in terrestrial ecosystems: a multimodel analysis, *Biogeosciences*, 13, 5121–5137, <https://doi.org/10.5194/bg-13-5121-2016>, 2016.