ExpeER
Distributed Infrastructure for EXPErimentation in Ecosystem Research

Grant Agreement Number: 262060

SEVENTH FRAMEWORK PROGRAMME

CAPACITIES
INTEGRATING ACTIVITIES: NETWORKS OF RESEARCH INFRASTRUCTURES (RIS)
THEME: ENVIRONMENT AND EARTH SCIENCES

DELIVERABLE D10.3

Deliverable title: Report on upscaling methods for biogeochemical and ecological processes

Abstract: This research focused on upscaling of local measurements made by ecosystem infrastructure, like for example net ecosystem exchange, to larger scales like catchments and even continents. Results show that with help of data assimilation techniques an upscaling can be made by estimating ecosystem parameters which are typical for a specific plant functional type. Measurements made by ecosystem infrastructure allow better estimates of ecosystem parameters and are therefore able to improve estimates of net water and carbon fluxes between the land and the atmosphere for larger areas.

Due date of deliverable: M48 Actual submission date: M54

Start date of the project: December 1st, 2010 Duration: 54 months

Organisation name of lead contractor: JULICH

Contributors: Harrie-Jan Hendricks Franssen, Hanna Post, Philippe Peylin, Natasha Mc Bean, Albert Porcar-Castell, Frank Berninger, Harry Vereecken

Revision N°: V1

Dissemination level:

PU Public (must be available on the website)	**X**
PP Restricted to other programme participants (including the Commission Services)	**[ ]**
RE Restricted to a group specified by the consortium (including the Commission Services) (precise to whom it should be addressed)	**[ ]**
CO Confidential, only for members of the consortium (including the Commission Services)	**[ ]**

Table of Content

1. Executive summary

This research focused on upscaling of local measurements made by ecosystem infrastructure, like for example net ecosystem exchange (NEE), to larger scales like catchments and even continents. The work focused on two different approaches for doing the upscaling. One approach was a hybrid methodology, where ecosystem parameters first were estimated with help of measured time series of NEE using the Markov Chain Monte Carlo methodology of DREAM$_{(ZS)}$ (Laloy and Vrugt, 2012; Ter Braak and Vrugt, 2008). The advantage of this methodology is that it is not limited to non-Gaussian distributions and can also be applied for very non-linear simulation models. However, this methodology is very CPU-intensive and could only be applied for single sites. Selected were four sites with different plant function types (PFT). NEE-time series obtained by eddy covariance measurements at those sites were used for parameter estimation. Verification of parameter estimates made for a certain PFT with NEE-time series measured at other sites of the same PFT, showed that the characterization of NEE was considerably improved. In a next step, parameter estimates were applied to improve NEE simulations for the entire Rur catchment. Therefore, updated parameters were assigned to all grid cells under consideration of the parameter uncertainty.

The second approach was variational data assimilation in combination with the land surface model ORCHIDEE. Variational data assimilation adjust model simulation results to measured data by minimizing an objective function that takes into account the uncertainty of initial values for the states and parameters to be adjusted, and differences between model simulations and measured values. The objective function was minimized with respect to model parameters (mainly ecosystem parameters) and initial conditions in this work. Values were estimated for single sites and also multiple sites, which shared the same PFT. These parameter estimates were used at other sites with the same PFT and it was evaluated to what degree simulation results improved. This approach is very similar to the upscaling approach from the plot to the catchment scale, but in this case the evaluation sites were sometimes located in other continents. It was found that this upscaling methodology improved the reproduction of exchange fluxes of carbon dioxide, water and energy between the land surface and atmosphere. The parameter estimates were also used for projections of changes in net terrestrial carbon storage as function of global temperature change. These projections were compared with simulations with default parameter values. It was found that whereas default parameter values resulted in an almost linear increase in terrestrial carbon storage as function of temperature increase up to 7K, the estimated parameter values resulted in a slower increase of terrestrial carbon storage as function of temperature and even a decrease for global temperature increases larger than 5K-6K. Therefore, projected changes in terrestrial carbon storage were sensitive to the adopted ecosystem parameters.

Finally, it was investigated whether chlorophyll fluorescence (ChF) measurements and photochemical reflectance index (PRI) are promising data to further constrain simulation results with land surface models. At leaf level a high correlation was found between these variables and light use efficiency. However, for the eddy covariance tower footprint and MODIS the correlation between PRI and light use efficiency was much weaker. Surprisingly, for MODIS the PRI-LUE correlation was even slightly higher than for the flux tower.

2. UPSCALING METHODOLOGY

2.1. UPSCALING FROM PLOT SCALE TO LANDSCAPE SCALE

The upscaling of water and net ecosystem exchange fluxes from plot scale (level of eddy covariance tower footprint) to landscape scale was done following a hybrid approach. We argue that the main sources of uncertainty for this upscaling are ecosystem parameter uncertainty, uncertainty with respect to initial conditions (e.g, magnitude of carbon pools), forcing uncertainty and model structure uncertainty. The hybrid approach consists of estimating plant functional type (PFT)-specific ecosystem parameters and initial conditions at the plot scale, on the basis of NEE-fluxes measured by EC. Estimation was done with help of a Markov Chain Monte Carlo Method (MCMC), implemented in the software DREAM ("DiffeRential Evolution Adaptive Metropolis") (Vrugt et al., 2009). An important advantage of this methodology is that the full posterior probability density function (pdf) is obtained, without linearization or restrictions like Gaussianity. Parameters were estimated based on single sites for four different PFTs. In a second step, the multivariate pdf´s of ecosystem parameters were assigned to the landscape scale, and model states are actualized using sequential data assimilation with updating of leaf area index from remote sensing information. Below we give a short summary of the methodology including some specifics relevant for the work later discussed in this deliverable. We refer to deliverable D10.2 and listed references for more details about the methodology.

In contrast to local methods, global methods take into account correlations among parameters. Compared to other inverse methods, an important advantage is that estimated posterior pdf´s are not limited to Gaussianity and not affected by linearization. The latest version DREAM$_{(zs)}$ was particularly designed to solve high-dimensional search problems (Laloy and Vrugt, 2012; Ter Braak and Vrugt, 2008) more efficiently.

In DREAM$_{(zs)}$ the Gelman-Rubin convergence diagnostic (Gelman and Rubin, 1992) is computed for each dimension j using at last 50% of samples in each chain (Ter Braak and Vrugt, 2008). If the Gelman-Rubin convergence diagnostic <1.2 for all j, the chains have converged in the same area of the parameter space.

Two different likelihood functions can be tested with DREAM to constrain the model:

(1) The log-density function with heteroscedastic measurement error:

$$\log P(i,1) = -\left(\frac{N}{2}\right) * \log(2 * pi) - \sum \log(v_n) - \frac{1}{\sum \frac{Err^2}{v_n}} \qquad \textit{Eq. 1}$$

with *N* being the number of measurement data for calibration (the length of measurement vector m_n) and v_n being the measurement error vector with length (*n*) corresponding to m_n.
(2) The sum of squared errors (SSE):

$$SSSE = \sum_{t=1}^{T}\sum_{i=1}^{j}(m_t - y_{t,i})^2 \qquad \textit{Eq. 2}$$

The model error (Err) was determined by:

$$Err = m_n - y_n(i) \qquad \textit{Eq. 3}$$

with y_n being model output according to v_n and i is chain (sequence) number. In this work only the latter objective function (Eq. 2) was used.

The Ensemble Kalman Filter (EnKF) (Burgers et al., 1998; Evensen, 2003, 1994) was applied using the Data Assimilation Research Testbed DART (Anderson et al., 2009). These runs considered uncertainty input from atmospheric forcings, initial states and ecosystem parameters (on the basis of DREAM) in a 80 member ensemble run.

2.2. UPSCALING TO CONTINENTAL SCALE

The method used to upscale ecosystem fluxes (carbon and water) at continental to global scale from an ensemble of in-situ and satellite observations has been described in details in the deliverable D10.2. We thus only briefly remind below the main components and different steps associated to the upscaling procedure:

- **Ecosystem Model:**
  The principle of the approach is to use a global process-based terrestrial ecosystem model that is calibrated/optimized at an ensemble of sites representative of all ecosystems that will be considered at the continental/global scale. In our case we used the ORCHIDEE land surface model and calibrate the most uncertain model parameters controlling the carbon flow in the soil-plant-atmosphere continuum as well as the initial soil carbon pool sizes. The underlying assumption is that uncertain and poorly calibrated model parameters represent a significant share of the total model uncertainty.
- **Observations:**
  An ensemble of observations are used to calibrate the model, from in-situ measurements representative of the plot scale (Eddy-covariance fluxes, biomass,,…) to more integrated observations such as atmospheric CO_2 concentrations, going though satellite observation of the vegetation activity. The assimilation of CO_2 concentrations implies the use of an atmospheric transport model to relate measured concentration to the surface fluxes.
- **Data assimilation procedure:**
  The principle relies on the minimization of a cost function that quantifies the differences between modeled and observed quantities (given uncertainties linked to

both terms) and the differences between the parameters to be optimized and their prior knowledge. In our case, we further use the hypothesis of Gaussian errors and a least square cost function, which simplifies the derivation of the optimal parameters and their associated errors. Note finally that the optimization is done in a step-wise approach where all data stream are considered sequentially (see D10.2).

- **Upscale fluxes/stocks and associated uncertainties:**
  The final upscaling step consist of running the optimized version of the ecosystem model at continental or global scale, using prescribed climate forcing (from re-analysis or from model simulations for future projections). In this case, we applied ORCHIDEE globally over the past two decades and also using climate projection up to 2100. We also propagated the estimated parameter errors on the simulated carbon fluxes and stocks

3. RESULTS

3.1. ROLLESBROICH/RUR

The hybrid methodology as detailed in section 2.1 (further details in deliverable 10.2) was applied to upscale NEE-fluxes from the plot scale to the landscape or catchment scale.

3.1.1. Site information

The Rur catchment is located in the Belgian-Dutch-German border region. Different eddy covariance (EC) towers are located in the Rur-catchment (**Erreur ! Source du renvoi introuvable.**). The eddy covariance (EC) raw data provide plot scale estimates of the net ecosystem exchange of CO_2 between the land surface and the atmosphere (*NEE*). Half hourly NEE data used for parameter estimation and model evaluation were available for several months (e.g. Kall-Sistig) up to >2.5 years (mainly from Mai 2011 to end of 2013).

The EC tower sites “Rollesbroich”, “RurAue” and “Kall-Sistig” are grassland sites. “Wüstebach” is a conifer forest site (spruces mainly). “Merzenhausen” and “Selhausen” are crop sites where winter wheat is grown, in Selhausen in a rotation with sugar beets and potatoes. Kall-Sistig is not a long-term measurement site but a roving station was setup here for several month e.g. for uncertainty estimation (Post et al., 2015) and model evaluation purposes. The extensively managed grassland site Rollesbroich (RO) is located in the Eifel region (50.6219142 N / 6.3041256 E) in 515 meters above sea level. The coniferous forest site Wüstebach (WUE, 50.5049024 N/ 6.3313825 E) in 607 m MASL. The mountainous Eifel region is dominated by sedimentary rocks and shallow soils. Merzenhausen (ME) is located in an agriculturally used lowland region (93 m MASL; 50.9297879 N / 6.2969924 E) with fertile loamy soils. The annual precipitation in the lowland region is lower (~ 690 mm/a) and the annual mean temperature higher (~9.8°C) compared to the Eifel (Rollesbroich: ~ 7.7°C ; ~ 1033mm/a) (Post et al., 2014).

EC data from three sites (Rollesbroich, Merzenhausen and Wüstebach) were extensively used for parameter estimation with DREAM while EC data from other sites were only used for evaluation purposes.

Fig. 3.1 The Rur catchment with the different land use and vegetation types. Indicated are also the positions of the eddy covariance towers.

Because lower weights are assigned to observations with higher uncertainty during data assimilation, reliable NEE uncertainty estimates are essential for a successful data assimilation experiment. The uncertainty of the measured NEE was determined with a statistical approach described in Mauder et al., (2013) using the auto- and cross- covariances of the measured raw-data to determine the instrumental noise σ_{cov}^{noise} and the stochastic error σ_{cov}^{stoch}. Those two error terms were summed up to calculate the total uncertainty estimates for each half-hourly measurement value. These uncertainty estimates were also compared with uncertainty estimates based on an extended two-tower approach using the NEE data measured at RO as well as the second tower sites ME and KA showing good correspondence (Post et al., 2015). Measured raw-data were pre-processed with the TK3.1 software (University of Bayreuth, Department of Micrometeorology, Germany; Mauder and Foken, 2011), which includes a comprehensive quality control and flagging system. Only data

with quality flag 0 (high quality data) and 1 (moderate quality data) were used for parameter estimation and data assimilation.

The gaps were not filled to avoid introducing additional uncertainty.

3.1.2. CLM4.5 set-up

For NEE simulations and parameter estimation CLM was first setup in a single point "PTCLM" mode for the sites Rollesbroich, Merzenhausen and Wüstebach. In addition, a regional CLM domain was setup for the entire Rur catchment with 1km^2 grid resolution. The main difference of one PTCLM site and the corresponding 1km^2 grid cell is that for single sites only a single plant functional type (PFT) is defined, while 1km^2 grid cells usually contain different plant functional types represented as percentage coverage area in CLM. However, because the respective land use type is widespread around the single EC sites in the Rur catchment, the percentage of PFTs different from the site PFT was <15% for all sites except for Rur-Aue where only ~50% of the 1km^2 grid cell is grassland, the other is broadleaf deciduous temperate tree (~20%) and c3-group (~30%). For both the regional and the point simulation we used reanalyzed regional meteorological input data (COSMO_DE) provided by the German meteorological service (Baldauf et al., 2009). The COSMO_DE data includes hourly time series of air temperature, incoming short wave radiation, incoming long wave radiation, precipitation, atmospheric pressure, specific humidity and wind speed. For RO gap-filled atmospheric site measurements were available and used instead of the COSMO_DE data for point simulations. However, a comparison between forward runs with site data and COSMO_DE showed only very minor differences in estimated NEE-fluxes so that the quality of COSMO_DE was considered acceptable for NEE simulations. The soil texture (percent clay and sand) data was obtained from the German soil map (BK50). Before running CLM4.5 in the CN mode, the initial state variables such as the carbon and nitrogen pools need to be spun up. First, a single instance spin-up was done by running CLM over a period of 1200 years, driven by the COSMO_DE data for the years 2008-2010 in rotation. To generate the 80 member ensemble of CLM initial states for the data assimilation, CLM was spun up a second time for 12 years using the initial state file obtained after the first 1200-year spin-up as input. For the second spin-up a 80 member perturbed forcing ensemble for the years 2008-2010 was used.

The 80 member ensemble of perturbed meteorological forcings was generated for the years 2008-2012 using the COSMO_DE data. Fig. 3.1 shows the spread of the perturbed meteorological variables at the Rollesbroich site for the first two days in 2009. Perturbed forcings are important to take uncertainty in the meteorological input data into account and were used to generate perturbed initial conditions during the second spin up. According to (Kumar et al., 2012; Reichle et al., 2010) perturbation fields were only applied to the variables air temperature (K), long wave radiation (W/m^2), short wave radiation (W/m^2) and precipitation (mm/s). Spatial correlated noise was considered using the Fast Fourier Transform approach (Park and Xu, 2013) with a 10 km spatial correlation scale according to (Han et al., 2014).

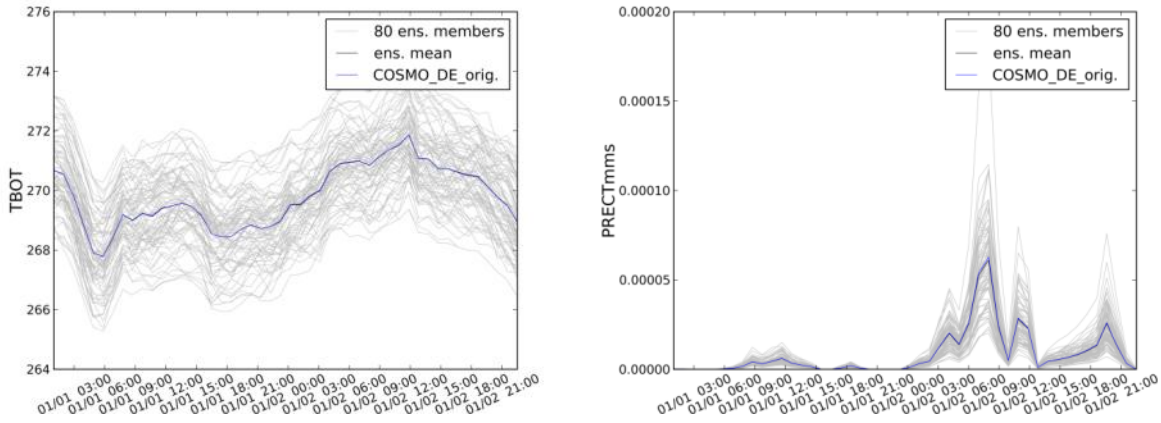

Fig. 3.1. Perturbed meteorological data: air temperature TBOT [K] (left) and precipitation PRECTmms [mm/s] for the 1st and 2nd of January 2009

3.1.3. Forward runs

Fig. 3.2 - Fig. 3.4 show NEE simulated for the sites Rollesbroich, Merzenhausen and Wüstebach for a forward CLM run (without calibrating parameters). A main discrepancy between NEE observations and CLM predictions for all sites is the underestimation of NEE magnitudes during most of the vegetation period. In addition, the time of the vegetation period (e.g. timing of vegetation period in March) sometimes shifted a few days. The simulated timing of the vegetation period and LAI development were improved by minor changes in the CLM4.5 stress deciduous phonology module and differed much more severe from the observations before those modifications (**Erreur ! Source du renvoi introuvable.**). After this modification, simulated NEE were relatively close to measurements for the entire period 2011-2012 (Fig. 3.4).

In case of Merzenhausen, measured NEE decreases abruptly in the first half of July in both years 2011 and 2012 (Fig. 3.2) which does not agree with the simulated daily NEE magnitude. This abrupt decrease of measured NEE which is a result of the ongoing senescence of the canopy (winter wheat) is not represented by CLM. The effect of senescence is larger than the effect of grassland and crop management (fertilization, grass cutting/harvesting) for both sites which does not result in such high discrepancies.

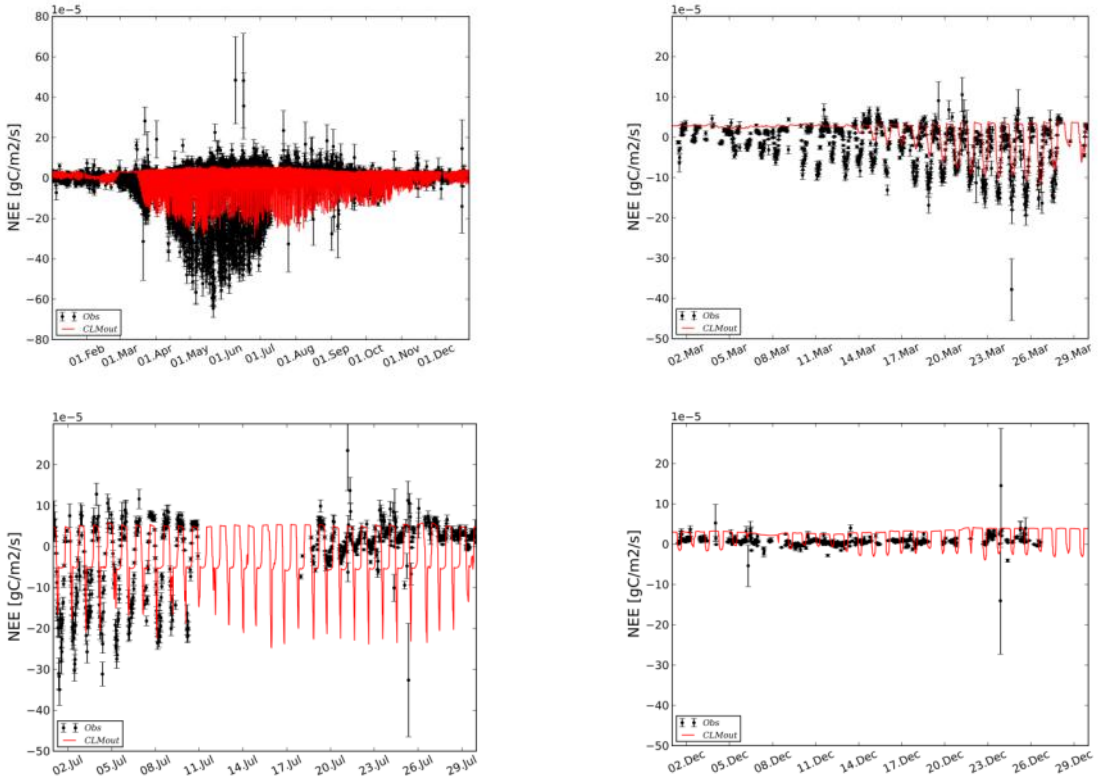

Fig. 3.2 Half hourly NEE observations (black) versus CLM outputs for Merzenhausen 2012

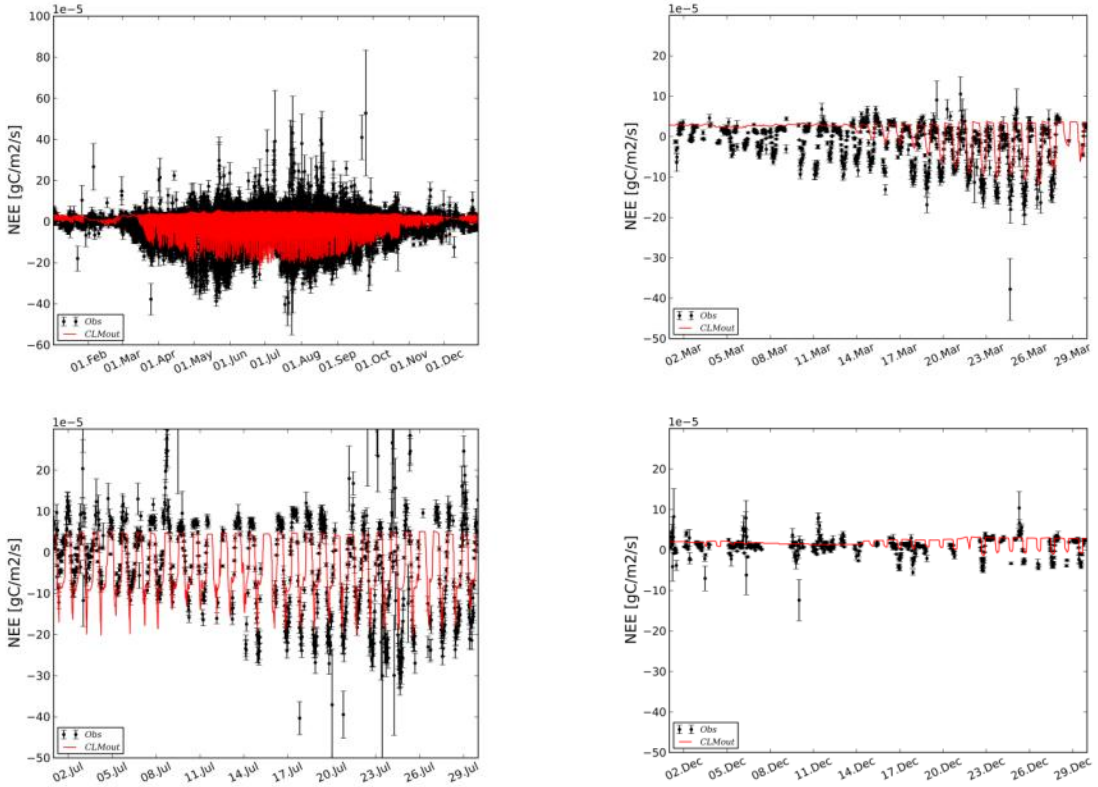

Fig. 3.3 Half hourly NEE observations (black) versus CLM outputs for Rollesbroich 2012

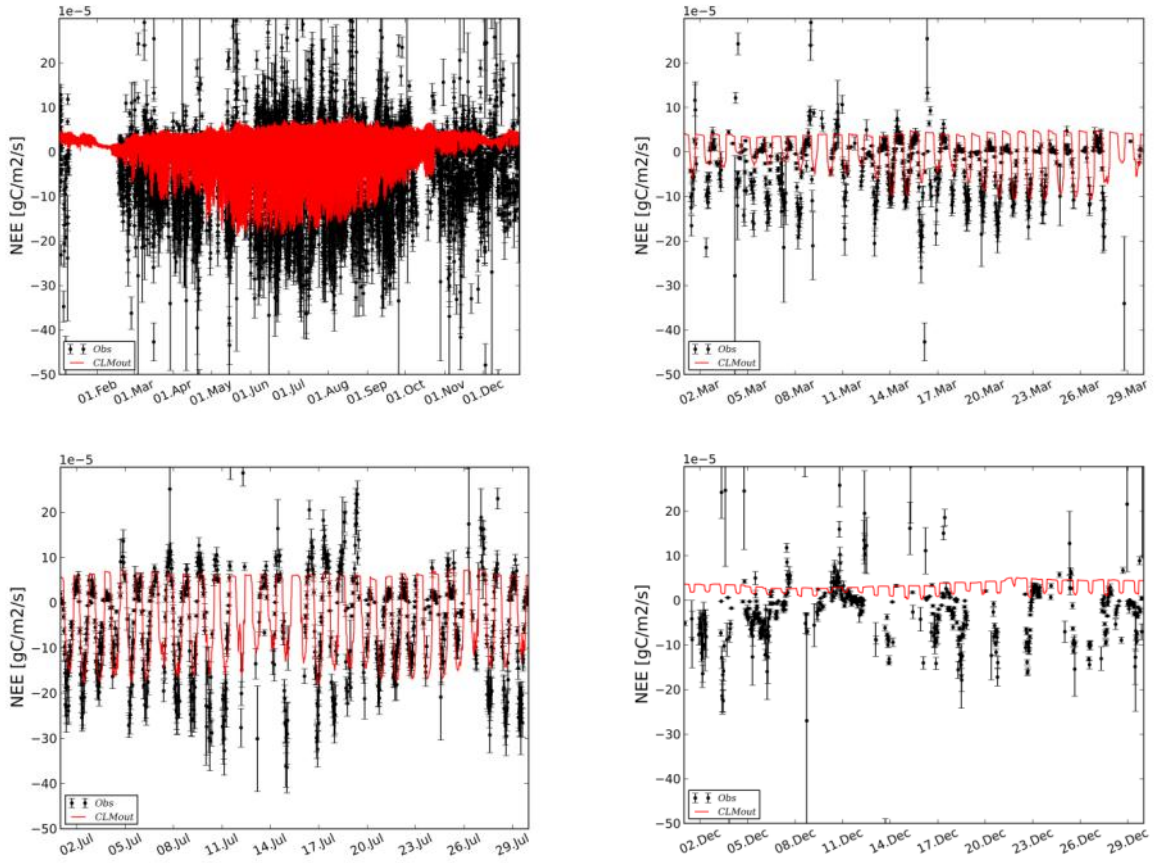

Fig. 3.4 Half hourly NEE observations (black) versus CLM outputs for Wüstenbach 2012

Tab. 3.1: Parameters estimated with DREAM

Short name	Long Name	Unit
pft-parameters		
flnr	Fraction of leaf N in Rubisco enzyme	-
grperc	Growth respiration factor	-
rootb_par	CLM rooting distribution parameter	1 m^{-1}
slatop	Specific Leaf Area (SLA) at top of canopy	m^2/gC
smpsc	Soil water potential at full stomatal closure	mm
hard-wired parameters		
q10	temperature coefficient	-
br	base rate for maintenance respiration	-
mb	Ball-Berry slope of conductance-photosynthesis relationship	-

3.1.4 Parameter estimation with MCMC

The eight parameters (

Tab. 3.1) estimated with DREAM$_{(zs)}$ were selected through a local sensitivity study (analysis of scatter plots) including 32 CLM parameters. The selection of those 32 parameters was based on previous studies on CLM parameter sensitivity studies (e.g. Göhler et al., 2013; Bonan et al., 2011; Wang et al., 2007). DREAM$_{(zs)}$ was then used to gain information about parameter uncertainty and to determine (marginal) probability distribution functions as well as maximum a posteriori (MAP) estimates for each of the 8 parameters.

The selected ecosystem parameters were estimated with help of DREAM$_{(zs)}$ using NEE-time series for the year 2012. Other experiments were performed where only NEE-data from single seasons in 2012 were used to estimate the ecosystem parameter. A verification experiment was carried out for the year 2013. The estimates of ecosystem parameters were used as input for the simulations in 2013. The parameters which were estimated for single seasons of 2012 were evaluated for the same seasons in 2013. Results for Rollesbroich are displayed in Fig. 3.6 and illustrate that for the verification period the estimated ecosystem parameters outperform the simulations with the default parameter settings. This is especially the case for the parameter estimates based on NEE-time series for a season only. The improvement achieved with parameter estimates based on data for the complete year is smaller. This indicates that ecosystem parameters might be a function of time, or that season dependent parameters compensate other model structural errors. In case of the RO site, the estimated ecosystem parameters especially improved NEE for the spring season. For the needle leaf forest site Wüstebach estimation of ecosystem parameters improves the annual NEE-cycle in the verification year, and in this case both season-based parameter estimates and estimates based on data of a complete year give similar results. See also Figure 3.7. The improvements are especially noticeable in spring and summer. Also for the cropsite Merzenhausen (not shown) improvements were achieved, which were however smaller than for the other sites. Finally, for the broadleaf forest site Fointainebleau in France (not shown) the improvement was considerable, and larger than for the other sites.

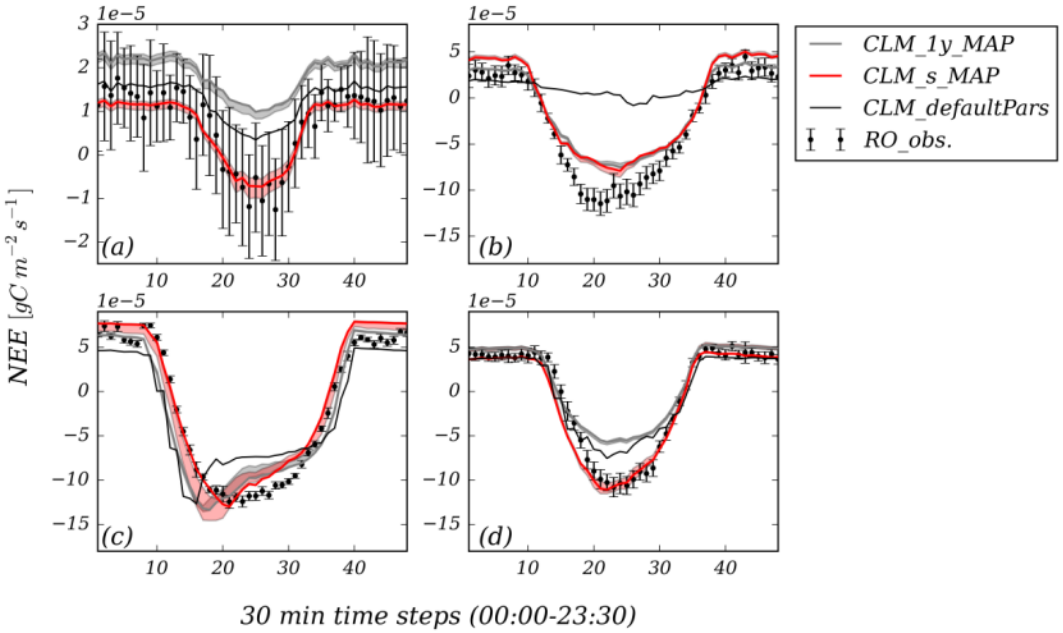

Fig. 5.6: Daily course of NEE for winter '12/'13 (a), spring 2013 (b), summer 2013 (c) and autumn 2013 (d) for the Rollesbroich site. Individual lines indicate observed NEE (RO_obs), NEE simulated with CLM using default parameters and NEE simulated with estimated parameters for the one year period and for single seasons (_s). The 95% confidence intervals were determined by sampling from DREAM posterior distributions.

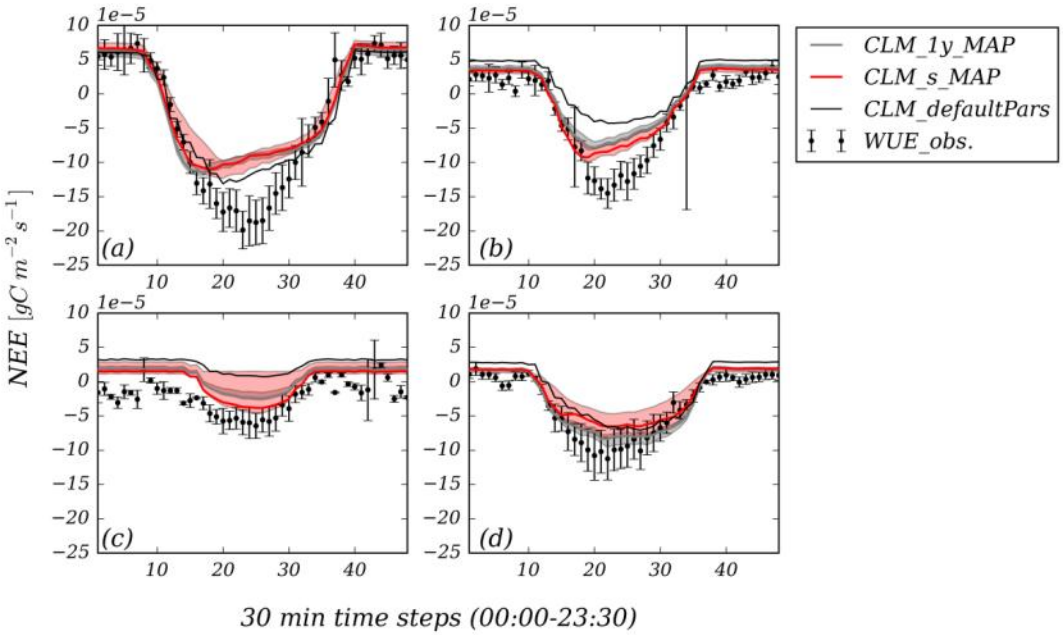

Fig. 3.7: Like Figure 3.6, but for the Wüstebach site.

The estimated ecosystem parameters for four sites were first validated for another time period independent from the calibration period. In addition parameter estimates were validated based on other sites with the same similar plant functional types which were separated more than 500km from the sites parameter estimation was applied to. The evaluation sites were Grillenburg (Germany) for grassland, Hainich (Germany) for broadleaf forest, Tharandt (Germany) for needleleaf forest and Klingenberg (Germany) for cropland. In all cases, the estimated parameters resulted in model estimation results which were closer to the measured data than the simulations with default parameters. This was the case for the average daily NEE-cycle, the yearly NEE-cycle, the total NEE sum over the year and statistical performance measures like root mean square error (RMSE) and absolute mean error (AME) evaluated over all time points. Fig. 3.8 shows as an example the evaluation for the daily NEE-course for the site Hainich.

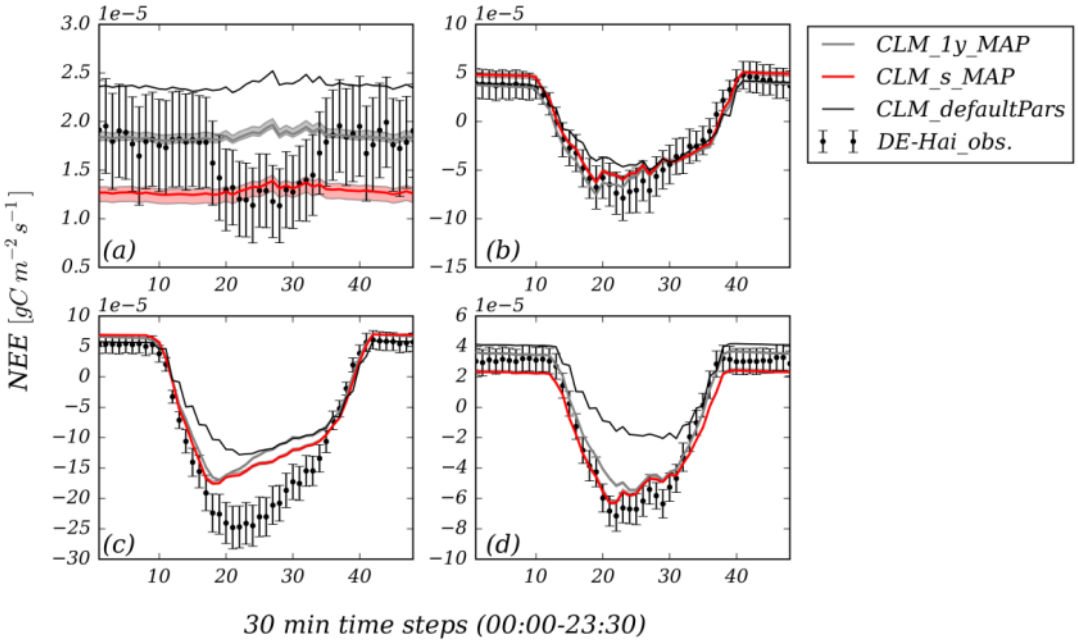

Fig. 3.8: Daily course of NEE for winter '06/'07 (a), spring 2007 (b), summer 2007 (c) and autumn 2007 (d) for the FLUXNET site DE-Hai. The lines shown are observed NEE with the EC method (DE-Hai_obs.), NEE simulated with CLM validation runs using default parameters, NEE simulated with estimated parameters from the Fontainebleau site in France (same PFT: C3-crop) for the one year period (_1y) and for the single seasons (_s). The 95% confidence intervals were determined by sampling from DREAM posterior distributions.

3.2. Upscaling carbon fluxes at continental scale using ORCHIDEE model

3.2.1. Optimization of ORCHIDEE with multiple data streams

We present the results of a sequential optimization of the main parameters of ORCHIDEE using three data streams: MODIS-NDVI data, in situ carbon and water flux measurements

(from FluxNet network) and atmospheric CO_2 concentrations. Figure 3.9 summarizes the approach and illustrates the typical model data fit for each data stream:

- In the first step, we used MODIS-NDVI to correct the phenology parameters of the deciduous Plant Functional Types in ORCHIDEE (4 or 5 parameters for each PFT). The main result is a shortening of the growing season length for most PFT compared to the prior model simulation (see MacBean et al., 2015).
- In a second step, we further correct a larger ensemble of parameters (around 15 per PFT) using more than 70 FluxNet sites with several years of NEE and LE flux measurements (assimilation of daily mean values). The improvement of the mean seasonal cycle (see figure 3.9) for each PFT is significant both in terms of amplitude and phase of the mean seasonal cycle. The results are discussed in detail in Kuppel et al. 2014.

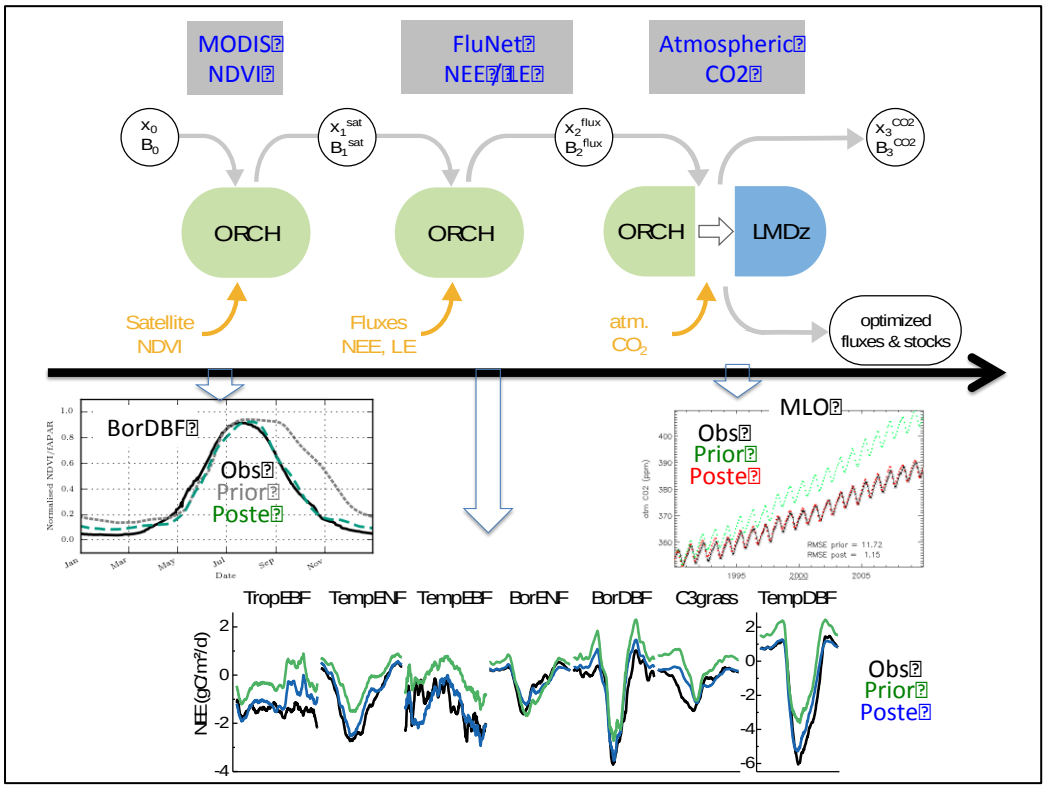

Figure 3.9: Overall scheme of the step-wise optimization of the ORCHIDEE parameters using MODIS-NDVI (step 1), FluxNet carbon and water measurements at more than 70 sites (step 2) and atmospheric observations at 72 stations (step 3). For each step the typical model – data fit is illustrated (see text). The estimated parameters (Xi) and their covariance error matrix (Bi) from each step "i" are further used as prior for the next step. After step3 the final parameters X3 and B3 will be used for the global "upscale simulation".

- Finally the atmospheric CO_2 observations at 70 sites are used to fine tune all parameters previously considered, including additional ones to optimize the initial soil carbon pools per region (around 30 regions globally). The global CO_2 trend that was too strong in the prior simulation is successfully corrected in order to match the observed trend (see Mauna Loa station in figure 3.9) by fitting the global land ecosystem carbon uptake.

One crucial point is that at each step we use the posterior parameter values and error variance-covariance matrix derived from the previous step. The method and results will be described in Peylin et al. (in preparation).

3.2.2. Spatial upscaling of carbon fluxes and stocks and associated uncertainties

Following the optimization of the land surface model parameters described above, we then run the model globally to estimate the space-time distribution of carbon fluxes and stocks.

Upscaled carbon fluxes:

Figure 3.10 illustrates the spatial distribution of the land ecosystem net annual carbon fluxes for a particular year after the optimization as well as the flux changes due to the optimization (posterior minus prior). The yearly mean optimized fluxes show significant uptakes for most ecosystems except central east Europe, south of the Amazon basin, middle US and India. The corrections through the optimization scheme (parameter optimization) were mostly to reduce the prior ORCHIDEE carbon sink at mid to high latitudes and to increase the sink over the Tropics.

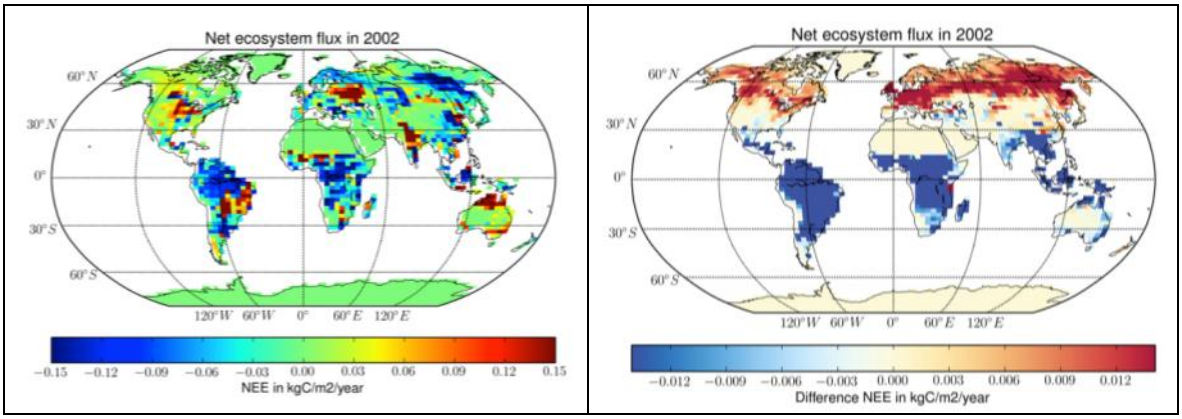

Figure 3.10: Left: estimated global net ecosystem carbon flux for a particular year, 2002 (in kgC/m^2/yr). Right: difference between the estimated (posterior) net carbon fluxes and the prior fluxes (before optimization).

If we now consider the aggregated fluxes for the Northern, Tropical and Southern land (see figure 3.11) we clearly see large year-to-year flux variations that are not in phase between the different latitude bands. The tropical fluxes present large anomalies during marked El-nino events, with lower carbon uptake (see for instance the 1998 fluxes) while the Northern

flux anomalies are more associated to large-scale temperature anomalies. If we consider the Gross Primary Production (GPP) for the same latitudinal split, the optimization significantly reduces the GPP for the tropical ecosystems (more than 20 PgC/yr) and only slightly reduces it for the North and South regions (around 5 PgC/yr). Such reduction is coherent with previous model inter-comparisons (see for instance TRENDY inter-comparison, Sitch et al. 2014) where ORCHIDEE's GPP was among the highest of an ensemble of 10 land surface models. Changes in GPP were also accompanied with similar changes in total ecosystem respiration so that the net flux changes remain compatible with the global land carbon uptake inferred from the atmospheric CO_2 growth rate.

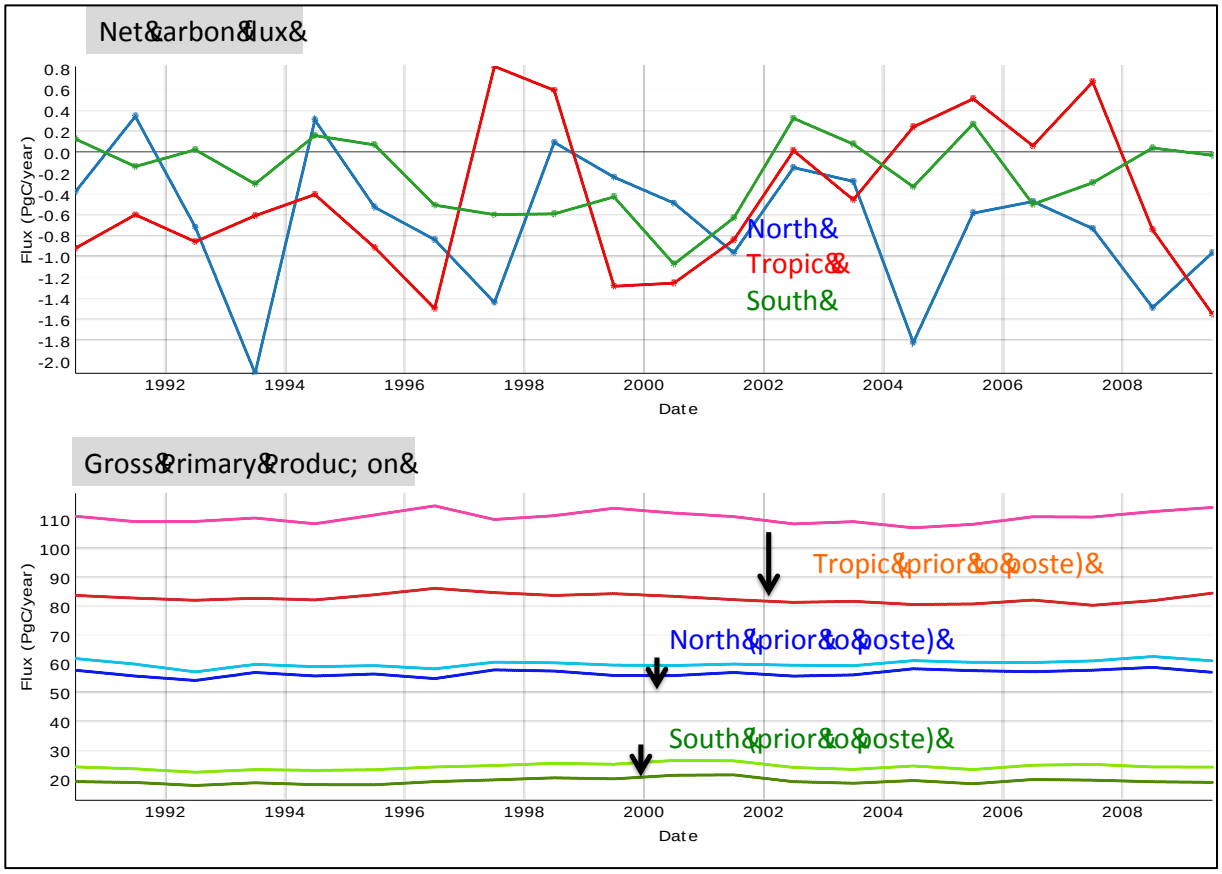

Figure 3.11: (Upper) Net annual carbon flux for three latitudinal bands: North of 30°N, Tropic (30°S-30°N) and South of 30°S. (Lower) Gross Primary Production for the same latitudinal bands for the prior and posterior (after optimization) simulations.

If we consider the continental "break-down" of the northern latitudinal band we obtain different mean fluxes and year-to-year variations for Europe, North America and North Asia. See also Figure 3.12. The optimized version of ORCHIDEE leads to a larger uptake in North Asia than North America and Europe on average. Rescaled by the respective surface areas, the carbon uptake becomes however the highest for Europe. The inter-annual variations (IAV) of the net carbon fluxes are not exactly in phase between the three regions. Europe has the highest flux IAV among the three regions. No clear trend is visible over this period (1990-2009), except for North Asia where the land sink slightly increases after the late 1990s. A more in depth analysis of these results and in particular the impact of the

optimization of the model with in situ data to estimate continental C budget (process-based model upscaling approach) is underway.

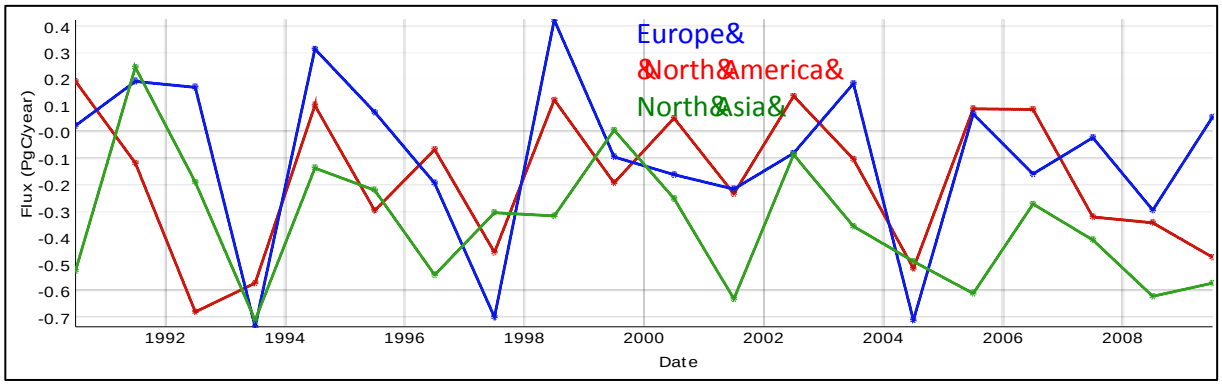

Figure 3.12: Net annual carbon fluxes from ORCHIDEE after optimization for Europe, North America and North Asia.

***Upscaled above ground carbon stocks**:*

We compare below the forest standing biomass. The optimized model with optimized GPP (input of carbon) as well as optimized subsequent growth allocation and turn-over of carbon in each pool, produces lead to different above ground forest biomasses. Observations of biomass are still scarce, but we provide a first comparison to independent biomass estimates from the FAO (global), from Thurner et al. (2014) (mid to high latitudes), and from Saatchi et al. (2011) (Tropic). Figure 3.13 shows the different estimate as separated by latitudinal bands, with the prior and posterior values simulated by ORCHIDEE. On average the above ground biomass reduction through the optimization brings the simulated values in closer agreement to the observations. However, missing processes such as modeling of age effects hamper a proper evaluation of the current model parameterizations. A new version of ORCHIDEE that accounts for management and age effects has been derived and will thus be used in future upscaling exercises.

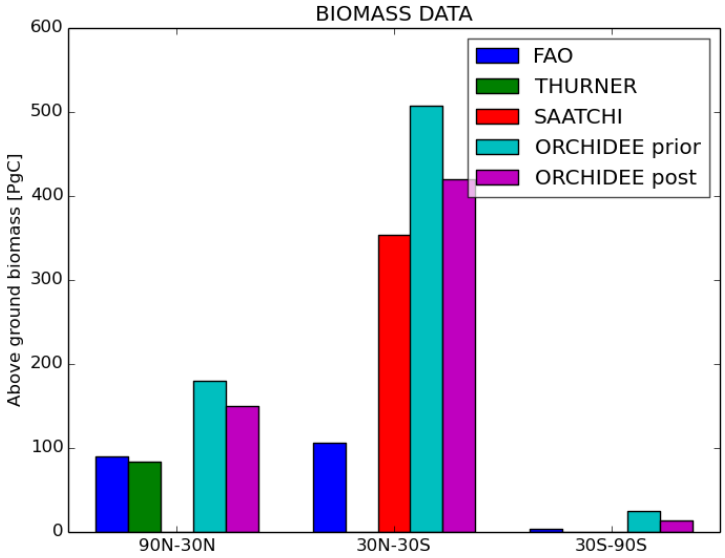

Figure 3.13: Comparison of forest above-ground biomass from the ORCHIDEE prior model and after optimisation (up-scaling of local information) to global products: Thurner et al. 2014 for the northern extra-tropics, FAO data, and Saatchi et al. 2011 tropical data (see text).

3.2.3. Temporal upscaling of carbon fluxes: future predictions

One objective is also to assess the impact of the model parameter optimization on the fate of the global carbon cycle, and in particular the evolution of the land carbon uptake in the horizon 2100. Differences in land carbon uptake will indeed be one of the major drivers of climate change in 50 to 100 years. We thus performed two forward simulations of ORCHIDEE over the next century using a given climate scenario, one with the standard parameters of the model that were used for the CMIP5 climate simulations, and one with the parameters optimized with the three data streams as described above. We then compared the global carbon budget of the two simulations in terms of cumulated fluxes over the next century as well as continental partition. The overall approach relies on the Inter-Sectoral Impact Model Intercomparison Project (ISI-MIP) protocol, with downscaled climate forcing; we took the HadGem2 climate simulation with an RCP8.5 scenario (i.e. atmospheric CO_2 concentration leading to an increase of surface temperature of 8.5°C in 2100) for this experiment. Three simulations were performed:

- A reference one that uses the version of ORCHIDEE coupled to the Earth System Model for the AR5 inter-comparison (hereafter referred to as the “AR5” simulation).
- A second one with a more recent version of ORCHIDEE (compared to the AR5 version) and the standard (non-optimized) parameters (hereafter referred to as the “PRIOR” simulation).

- A third one with the same version of ORCHIDEE as in the second simulation but with optimized parameters (hereafter referred to as the "POST" simulation).

Spin-up simulations were performed in order to set the carbon pools at the equilibrium for each case. In order to facilitate the comparison between the different simulations the changes in stocks are expressed as a function of the change in temperature and atmospheric CO_2 concentration. If we consider the sum of vegetation and soil carbon stock changes (Fig. 3.11), the optimization leads to a decrease of 170 PgC in the capacity of the terrestrial biosphere to absorb carbon (ORCHIDEE POST – ORCHIDEE PRIOR) compared to ORCHIDEE AR5. This decrease is mainly due to changes in the responses of soil and litter carbon contents to the change in environmental conditions during the 21st century (not shown). It takes place over most regions of the globe. Note that the difference between POST and PRIOR total carbon stocks only starts for temperatures above 3°C (or CO_2 concentration above 500 ppm). A detailed study of which soil parameter change is responsible for such strong reduction of the soil carbon content with increased temperature is underway.

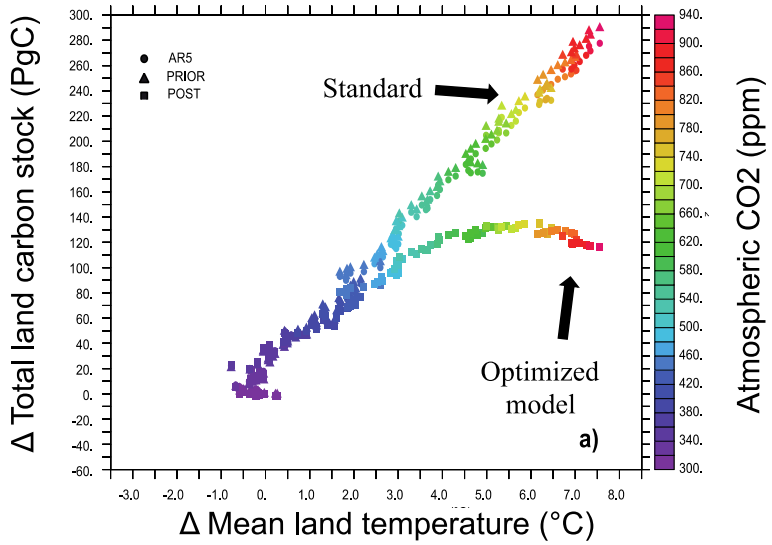

Figure 3.14: Change in total vegetation+soil carbon stock as a function of temperature (and CO_2 concentration, color code) for three cases (AR5-simulation, PRIOR, and POST);

3.3. Improved representation of ecosystem fluxes by RSI

In this study, we evaluated the potential of the PRI and ChlF as a proxy of light use efficiency of photosynthesis in a Scots pine forest in Hyytiälä at the leaf, canopy and landscape levels. In addition, the possibility of integrating chlorophyll fluorescence data into a land-surface model was preliminary explored (Kato et al 2014).

3.3.1 Materials and Methods

We progressed in the study on three different levels, centering around the Expeer flagship sites Hyytiälä, SMEAR II. The study was conducted using the measurements of the SMEAR II station of the University of Helsinki, located in southern Finland, 61◦51'N and 24◦17'E. The forest site surrounding the measuring mast is dominated by Scots pine (*Pinus sylvestris* L.), sown in 1962. The understorey is mainly composed of dwarf shrubs lingonberry (*Vaccinium vitis-idaea* L.) and blueberry (*Vaccinium myrtillus* L.), and mosses (*Pleurozium* sp., *Dicranum* sp.). In 2010, the basal area of pine on a 200 m radius around the measuring mast was 17.85 m^2/ha, that of spruce (*Picea abies* (L.) Karsten) 3.7 m^2/ha and of deciduous trees 2.74 m^2/ha. The arithmetic mean diameter at breast height (dbh) and tree height were 17.6 cm and 16.5 m for pine, 5.6 cm and 5.4 m for spruce and 3.9 cm and 7.4 m for deciduous trees. There was however both spruces and deciduous trees reaching the canopy top. The average effective leaf area index (LAI) was determined in July 2011 from hemispherical photographs taken systematically at permanent biomass sampling plots (see Ilvesniemi et al., 2009) within a radius of 200 m from the downward looking pyranometer. The effective LAI estimated from the hemispherical photographs was 2.2 using a GLA (Gap Light Analyzer) 2.0 software (SFU, New York).

The chlorophyll flurescence measurements were conducted at the leaf scale at the top of the canopy from August 2008 to August 2009 by the Multi-Channel Chlorophyll Fluorometer (MONITORING-PAM Heinz Walz GmbH, Effeltrich, Germany) with time steps between 5 minutes and 30 minutes (Porcar-Castell et al., 2008, Porcar-Castell, 2011). The yield of the chlorophyll fluorescence was measured. The photochemical reflectance index was measured at the leaf level (Porcar-Castell et al., 2012) and canopy level using a dual-field of view spectroradiometer system (Drolet et al., 2005), and a pair of PRI- broadband sensors (Skye, UK).

For pigment analysis frozen needle material was manually ground in liquid nitrogen and lyophilised. Ca. 50 mg of freeze dried material was suspended with 0.7 mL of 99.8 % acetone, mixed in vortex, and extracted for 2 h at 4 C. Extracts were centrifuged for 15 min at 15,000 rpm and 4 C. Supernatant was collected and the pellet was re-suspended with 0.7 mL acetone, mixed in vortex and extracted for 15 min in an ultrasound bath. The latter process was repeated once more and extracts were stored at -80 C. Pigment separation was performed by HPLC with a reverse phase C18 column (Waters Spherisorb ODS1, 4.6 9 250 mm; MA, USA), Identification and quantification was carried out with a photodiode array detector.

A spectroradiometer (FieldSpec-HH; ASD, Boulder, CO, USA) equipped with a fiber optic and plant probe (ASD) were used to measure leaf reflectance spectra. Spot size was 10 mm, sampling interval was 1.6 nm, and spectral resolution 3.5 nm. In order to measure the leaf reflectance from exactly the same sample and immediately after fluorescence, we used a modified dark-acclimation clip (Hansatech) for the measurements. The external diameter of the dark-acclimation clip fitted perfectly with the front measuring window of the plant probe so that the measuring clip could be easily and repeatedly placed in exactly the same position time after time. At the start of the measurements, and following every four samples, a calibrated Spectralon® reference panel was placed in the dark acclimation clip and a dark

current and white reference measurement were registered. Integration time was 540 ms, and, to avoid accumulation of NPQ during measurements, spectra were averaged over 5 measurements only. Typically, the third spectral average was stable and the sample leaf reflectance was recorded in less than 10 s.

We measured carbon dioxide fluxes using the EC technique is a direct way to determine the exchange rate of a compound across the interface between the atmosphere and a plant canopy. It measures the covariance between fluctuations in vertical wind velocity and fluctuations in compound concentration. Here, wind speed was measured by two sonic anemometers, Solent Research 1012R2. The concentrations were measured by an infrared absorption gas analyzer Li-6262 (LI-COR Inc., Lincoln, Nebraska) for CO_2 /H_2O. All these measurements were performed at the height of 23 m, approximately 10 m above the forest canopy. The collected data were quality controlled and corrected for frequency losses and sensor separation in a standard way [Aubinet et al. , 2000; Rannik , 1998]. The CO_2 storage term (the accumulation of CO_2) was estimated by means of concentration profile measurements (at 4, 8, and 17 m) using an infrared absorption analyzer (URAS 4, Hartmann and Braun, Frankfurt am Main, Germany). Data collected during low friction velocity (u*) episodes (below 0.3 m/s) were omitted due to possible underestimation of the fluxes.

The study comprised an area of 90×160km^2 centered at ca. 62°17'N and 25°05'E in middle Finland. Forests in the area are dominated by evergreen needle-leaved species Scots pine (*Pinus sylvestris* (L.)) and Norway spruce (*Picea abies* (L.) Karsten), while broadleaved deciduous species (mainly birches, *Betula pendula* Roth. and Betula pubescens Ehrh., and aspen (*Populus tremula* L.)) are mainly found as a mixture in conifer stands. The main soil types include sandy and silty podzols and peatland soils. Information on land cover type, forest age, and species specific growing stock (m^3/ha) was obtained from the product available at the data service of the Finnish Forest Research Institute (Metla, 2013). The data is in raster format on 20 × 20 m spatial resolution and is produced by the multisource national forest inventory (MS-NFI) project of Finland at the Finnish Forest Research Institute (Tomppo & Halme, 2004; Tomppo et al., 2008). The estimates of forest age and growing stock per species are produced using field measurements of the 10th and 11th (2007–2011) National Forest Inventory, digital map data and Landsat TM or ETM + images with an improved version of the nonparametric k-NN method, namely ik-NN (Tomppo & Halme, 2004; Tomppo et al., 2008). Satellite images used in the ik-NN estimation were from 2010 and 2011.

For landscape level studies we focused on the area around SMEAR 2. The study area was mainly covered by forests, but due to sparse scattered settlement and lakes, there were only 2186 MODIS albedo grid cells (about 4% of all data points) covered solely by the land cover class "forest land". These data were further filtered to remove pixels that were adjacent to water bodies or extensive cropland areas with very low or high albedo, after which 2180 pixels comprising only of productive forest land either on mineral or peatland soil were left for the analysis. The land cover data was converted to the sinusoidal projection of the MODIS albedo products. Each land cover pixel (20 × 20 m) containing the information on land cover type, forest age or species specific growing stock (pine, spruce, broadleaved species) was assigned to the overlapping MODIS grid cell. The data were divided into training and test data sets, comprising of 1744 and 436 randomly chosen MODIS pixels, respectively.

3.3.2 Results: LUE correlations with pigment and optical signals across different levels

Light use efficiencies can be measured at different scales, leaf, canopy and landscape using different tools. At the leaf level there was a close correlation between the light use efficiency and the yield of fluorescence as well as the non-photochemical quenching (Table 3.2). This indicates that the seasonal connections between PRI and LUE function are good at the leaf level.

Table 3.2. Correlation coefficients (r) at the leaf-level

	PRI	LUE
LUE	0.895	
PSII Yield	0.874	0.895
TotCAR/CHL	-0.408	-0.229
NPQ	-0.836	-0.880

At the canopy level seasonal correlations between PRI values and light use efficiencies where notably lower and the two methods to estimate PRI (Ocean Optics) and the Skye values were only moderately correlated. These are partially due to different view values and since PRI values will depend on viewing angle of the sensor. The Ocean optics PRI values are more tightly correlated to the light use values.

Table 3.3. Canopy-Level correlation coefficients (r) for snow free (growth) period (May-September) 2010-2011.

	PRI skye	PRI OO	LUE
PRI OO	0.401		
LUE	0.486	0.490	
Night min T	0.373	0.502	0.372
Day mean T	0.200	0.428	0.185
NDVI skye	0.152	0.379	0.287

Table 3.4. Canopy-Level correlation coefficients (r) for whole years 2010 or 2010-2011.

	PRI skye	PRI OO	LUE
PRI OO	0.427		
LUE	-0.204	0.401	
Night min T	-0.216	0.477	0.797
Day mean T	-0.316	0.410	0.752
NDVI	-0.638	0.022	0.678

For MODIS derived PRI (Table 3.3) the correlation values were slightly better than values from tower measurements. This is strange, since MODIS derived values are subject to non-perfect atmospheric corrections and there are known location errors attached to MODIS values.

Table 3.5. Ecosystem-level correlation coefficients between PRI derived and eddy covariance based LUE.

	PRI	LUE
LUE	0.541	
Night min T	0.734	0.521
Day mean T	0.669	0.436
MODIS NDVI	0.688	0.541

The dilemma of present satellites is that they either have a coarse spatial or a coarse temporal resolution. Both NDVI and PRI are derived from MODIS satellite which has a good temporal resolution but only a moderate spatial resolution (25 ha) which is significantly larger than the average size of a forest stand or landscape unit. Using albedo as the variable to be explained and forest maps as other input variables we were able to develop new un-mixing algorithms. The algorithms are described more in detail in Kuusinen et al. (2012).

We conclude that the observed correlation between MODIS PRI and tower level LUE is strongly controlled by measurement geometry. Decoupling the directional structural effects (shade fraction, BRDF) from the physiological ones (pigments pools, LUE) will be the next step towards the implementation of the PRI in the upscaling of LUE and assimilation into land-surface models.

4. DISCUSSION

Upscaling results illustrated the feasibility of the data assimilation approach. However, there are some main concerns. First, we might get better results not because parameter values really improved, but because the estimated parameter values cover model structural errors and incorrect initial states. The upscaling results for the continental scale on the other hand showed the transferability of parameter values across continents. This gives us additional confidence in the upscaling approach. Therefore, we have the feeling that the estimated parameter values are indeed better and can be used over a broad range of climatic conditions. However, future work should still be more focused on joint estimation of parameter values and initial states (magnitudes of carbon and nitrogen pools). Second, the parameter values cover different vegetation types. The used land surface models CLM and ORCHIDEE still have very broad definitions of plant functional types and cropland includes for example such different crop types as sugar beet, wheat, potatoes and maize. These crops have different regulations of transpiration and different root systems, but are nevertheless characterized by the same vegetation parameters. This is a limitation and points to potential improvements, which would consist of defining more plant functional types, especially for the cropland type. We believe that this is important for predicting NEE with earth system

models. Third, only a limited number of parameters was selected, but under different weather and environmental conditions, other parameters could be sensible as well. Ideally, the parameters would be selected with a global sensitivity analysis which would be repeated for many sites and for many different years. This is often not feasible, but further global sensitivity studies are still useful to gain more insight in the role of all different model parameters.

We see further potential for improving estimates of carbon budgets of earth system models. Besides model improvement which includes the inclusion of additional processes and including more interactions between different system compartments and different processes, additional conditioning to data types can improve estimates of carbon budgets. In this work the conditioning to multiple data types was analysed. Often net ecosystem exchange measurements contain information on some parameters, but not all ecosystem parameters. If in addition also biomass is measured, this provides additional constraints on ecosystem parameters. In addition, information from remote sensing is important for improving estimates between single sites. It is therefore important to continue work with the integration of multiple data streams in land surface models. Promising is for example remotely sensed fluorescence, which contains information on photosynthetic activity. The value of in situ measurements is increased if besides NEE, also soil respiration, biomass production, hydrological data like soil moisture content and other data are recorded.

5. Conclusion / Next steps

Biogeochemical fluxes measured at single sites provide important information on ecosystem functioning. However, in order to estimate water and carbon balances for larger units like catchments or continents, upscaling is needed. This work focused on upscaling with help of two different approaches: a hybrid Markov Chain Monte Carlo- sequential data assimilation approach and a variational data assimilation approach. The first methodology was tested for upscaling from the plot scale to the catchment scale. The second methodology was used for upscaling from the plot scale to the continental scale. Both methods estimated ecosystem parameters for plant functional types, which were in the upscaling step used for all sites without in situ measurements. Independent verification tests were carried out with simulations using default ecosystem parameters and simulations using estimated ecosystem parameters. It was found that the estimated ecosystem parameters reproduced the annual and daily net ecosystem exchange (NEE) cycles better than the default parameters. The improvement was often considerable. It can also be concluded that using multiple data streams for estimating ecosystem parameters is of advantage for better constraining the parameter values. It is concluded that future work should therefore focus on integrating many different data types in the upscaling framework, and this would include several data types measured in situ at observation sites (e.g., NEE, respiration, biomass, soil moisture, amongst others) as well as remote sensing information. Fluorescence, a proxy for photosynthetic activity, is of special interest for constraining carbon and water balances, but also other available remotely sensed information like leaf area index or soil moisture content still needs to be integrated better with land surface models/earth system models. Further points of interest for future research should be the joint estimation of model parameters and initial states like carbon pools, an increase of the number of plant functional types to

better represent the different vegetation and plant types and an improved selection of sensitive model parameters.

6. References

Aubinet, M., et al., 2000. Estimates of the annual net carbon and water exchange of forests: The EUROFLUX methodology, Adv. Ecol. Res., 30, 113– 175.

Anderson, J., Hoar, T., Raeder, K., Liu, H., Collins, N., Torn, R., Avellano, A., 2009. The data assimilation research testbed: A community facility. Bulletin of the American Meteorological Society, 90, 1283–1296.

Baldauf, M., Förstner, J., Klink, S., Reinhardt, T., Schraff, C., Seifert, A., Stephan, K., Wetterdienst, D., 2009. Kurze Beschreibung des Lokal-Modells Kürzestfrist COSMO-DE (LMK) und seiner Datenbanken auf dem Datenserver des DWD. Deutscher Wetterdienst, Geschäftsbereich Forschung und Entwicklung, Offenbach, Germany.

Bonan, G.B., Lawrence, P.J., Oleson, K.W., Levis, S., Jung, M., Reichstein, M., Lawrence, D.M., Swenson, S.C., 2011. Improving canopy processes in the Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data. Journal of Geophysical Research, 116, G02014, doi:10.1029/2010JG001593.

Burgers, G., van Leeuwen, P.J., Evensen, G., 1998. Analysis scheme in the ensemble Kalman filter. Monthly Weather Review, 126, 1719-1724.

Drolet, G.G., Huemmrich, K.F., Hall, F.G., Middleton, E.M., Black, T.A., Barr, A.G., Margolis, H.A., 2005. A MODIS-derived photochemical reflectance index to detect inter-annual variations in the photosynthetic light-use efficiency of a boreal deciduous forest. Remote Sensing of Environment, 98, 212–224.

Evensen, G., 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. Journal of Geophysical Research, 99(C5), 10143–10162.

Evensen, G., 2003. The ensemble Kalman filter: Theoretical formulation and practical implementation. Ocean dynamics, 53, 343–367.

Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. Statistical science, 7(4), 457–472.

Göhler, M., Mai, J., Cuntz, M., 2013. Use of eigendecomposition in a parameter sensitivity analyisis of the Community Land Model. Journal of Geophysical research-biogeosciences, 118(2), 904-921

Han, X., Franssen, H.-J.H., Montzka, C., Vereecken, H., 2014. Soil moisture and soil properties estimation in the Community Land Model with synthetic brightness temperature observations. Water Resour. Res. 50, 6081–6105. doi:10.1002/2013WR014586

Kato, T., Maignan, F., MacBean, N., Peylin, P., Goulas, Y., Daumard, F., Moya, I., Porcar-Castell. A., Nichol, C., 2014, Relationship between a model-simulated productivity and in-situ fluorescence measurements at two flux sites. Proceedings of the 5th International Workshop on Remote Sensing of Vegetation Fluorescence, 22-24 April 2014, Paris. http://www.congrexprojects.com/2014-events/14c04/proceedings

Kumar, S.V., Reichle, R.H., Harrison, K.W., Peters-Lidard, C.D., Yatheendradas, S., Santanello, J.A., 2012. A comparison of methods for a priori bias correction in soil moisture data assimilation. Water Resour. Res. 48, W03515. doi:10.1029/2010WR010261

Kuppel, S., P. Peylin, F. Maignan, F. Chevallier, G. Kiely, L. Montagnani, and A. Cescatti (2014), Model–data fusion across ecosystems: from multi-site optimizations to global simulations, *Geosci. Model Dev. Discuss.*, 7, 2961-3011, doi:10.5194/gmdd-7-2961-2014

Kuusinen, N., Tomppo, E., Berninger, F., 2012. Linear unmixing of MODIS albedo composites to infer subpixel land cover type albedos. International Journal of Applied Earth Observation and Geoinformation 23, 324-333.

Laloy, E., Vrugt, J.A., 2012. High-dimensional posterior exploration of hydrologic models using multiple-try DREAM(ZS) and high-performance computing. Water Resources Research 48 (1). doi:10.1029/2011WR010608

MacBean N., F. Maignan, P. Peylin, C. Bacour, F.-M. Bréon, and P. Ciais (2015), Using satellite data to improve the leaf phenology of a global Terrestrial Biosphere Model: impact on regional carbon budgets, Global Biogeochemical Cycles, *submitted*

Mauder, M., Cuntz, M., Drüe, C., Graf, A., Rebmann, C., Schmid, H.P., Schmidt, M., Steinbrecher, R., 2013. A strategy for quality and uncertainty assessment of long-term eddy-covariance measurements. Agricultural and Forest Meteorology, 169, 122–135.

Mauder, M., Foken, T., 2011. Documentation and instruction manual of the Eddy covariance software package TK3. Univ.Bayreuth, Abt. Mikrometeorologie.

Metla, 2013. File service for publicly available data. http://kartta.metla.fi/index-en.html (Accessed 2013).

Park, S.K., Xu, L., 2013. Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. II). Springer Science & Business Media.

Porcar-Castell, A., Pfündel, E., Korhonen, J.F.J., Juurola, E., 2008. A new monitoring PAM fluorometer (MONI-PAM) to study the short and long-term acclimation of photosystem II in field conditions. Photosynthesis Research, 96, 173–179.

Porcar-Castell A. 2011. A high-resolution portrait of the annual dynamics of photochemical and non-photochemical quenching in needles of Pinus sylvestris. Physiologia Plantarum, 143, 139–153.

Porcar-Castell, A., Garcia-Plazaola, J. I., Nichol, C. J., Kolari, P., Olascoaga, B., Kuusinen, N., Nikinmaa, E., 2012. Physiology of the seasonal relationship between the photochemical reflectance index and photosynthetic light use efficiency. Oecologia, 170(2), 313-323.

Post, H., Hendricks Franssen, H.-J., Graf, A., Schmidt, M., Vereecken, H., 2015. Uncertainty analysis of eddy covariance CO2 flux measurements for different EC tower distances using an extended two-tower approach. Biogeosciences, 12, 1205-1221.

Rannik, Ü., 1998. Turbulent atmosphere: Vertical fluxes above a forest and particle growth, Ph.D. thesis, Dep. of Phys., Univ. of Helsinki, Helsinki, Finland.

Reichle, R.H., Kumar, S.V., Mahanama, S.P., Koster, R.D., Liu, Q., 2010. Assimilation of Satellite-Derived Skin Temperature Observations into Land Surface Models. Journal of Hydrometeorology, 11, 1103-1122.

Saatchi S., Harris, N. L., Brown, S., Lefsky, M., Mitchard, E. T., Salas, W., Zutta, B. R., Buermann, W., Lewis, S. L., Hagen, S., Petrova, S., White, L., Silman, M., and A. Morel, 2011. Benchmark map of forest carbon stocks in tropical regions across three continents. Proc. Natl. Acad. Sci. U. S. A., 108, 9899-904.

S. Sitch, P. Friedlingstein, N. Gruber, S. D. Jones, G. Murray-Tortarolo, A. Ahlström, S. C. Doney, H. Graven, C. Heinze, C. Huntingford, S. Levis, P. E. Levy, M. Lomas, B. Poulter, N. Viovy, S. Zaehle, N. Zeng, A. Arneth, G. Bonan, L. Bopp, J. G. Canadell, F. Chevallier, P. Ciais, R. Ellis, M. Gloor, P. Peylin, S. L. Piao, C. Le Quéré, B. Smith, Z. Zhu, and R. Myneni, Recent trends and drivers of regional sources and sinks of carbon dioxide, Biogeosciences, 12, 653-679, 2015, doi:10.5194/bg-12-653-2015

Ter Braak, C.J.F., Vrugt, J.A., 2008. Differential Evolution Markov Chain with snooker updater and fewer chains. Stat. Comput., 18, 435–446. doi:10.1007/s11222-008-9104-9.

Thurner, M., Beer, C., Santoro, M., Carvalhais, N., Wutzler, T., Schepaschenko, D., Shvidenko, A., Kompter, E., Ahrens, B., Levick, S. R. and Schmullius, C.: Carbon stock and density of northern boreal and temperate forests, Glob. Ecol. Biogeogr., 23, 297–310, 2014.

Tomppo, E., Olsson, H., Ståhl, G., Nilsson, M., Hagner, O., Katila, M., 200

Tomppo, E., Olsson, H., Ståhl, G., Nilsson, M., Hagner, O., Katila, M., 2008. Combining national forest inventory plots and remote sensing data for forest databases. Remote Sensing of Environment, 112, 1982–1999.

Tomppo, E., Halme, M., 2004. Using coarse scale forest variables as ancillary information and weighting of variables in k-NN estimation: A genetic algorithm approach. Remote Sensing of Environment, 92, 1–20.

Vrugt, J.A., Ter Braak, C.J.F., Diks, C.G.H., Robinson, B.A., Hyman, J.M., Higdon, D., 2009. Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling. International Journal of Nonlinear Sciences and Numerical Simulation, 10, 273–290.

Wang, Y.P., Baldocchi, D., Leuning, R., Falge, E., Vesala, T., 2007. Estimating parameters in a land-surface model by applying nonlinear inversion to eddy covariance flux measurements from eight FLUXNET sites. Global Change Biology, 13, 652–670. doi:10.1111/j.1365-2486.2006.01225.x